

# Comparative Analysis of State-of-the-Art Deep Learning Models for Detecting COVID-19 Lung Infection from Chest X-Ray Images

HIKMAT KHAN<sup>1</sup>, PIR MASOOM SHAH<sup>2</sup>, ABDULLAH GANI<sup>3,4,\*</sup>, ZEBA GHAFAR<sup>1</sup>, IZAZ AHMAD KHAN<sup>2</sup>, MUNAM ALI SHAH<sup>1</sup>, SAIF UL ISLAM<sup>5</sup>,

<sup>1</sup>Department of Computer Science, COMSATS University Islamabad, Islamabad 45000, Pakistan(e-mail:

hikmat.khan179,zebaghaffar05@gmail.com,mshah@comsats.edu.pk

<sup>2</sup>Department of Computer Science, Bacha Khan University, Charsadda (BKUC),KPK Charsadda, Pakistan(e-mail:pirmasoomshah,azaz@bkuc.edu.pk)

<sup>3</sup>Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia(e-mail:abdullahgani@ums.edu.my)

<sup>4</sup>Faculty of Computing and Informatics, University Malaysia Sabah, Labuan 88400, Malaysia(e-mail:abdullahgani@ums.edu.my)

<sup>5</sup>Department of Computer Science, Institute of Space Technology, Islamabad 44000, Pakistan(e-mail:saiflu2004@gmail.com)

Corresponding author: Abdullah Gani (abdullahgani@ums.edu.my), University Malaysia Sabah, Labuan 88400, Malaysia

**ABSTRACT** The ongoing COVID-19 pandemic has already taken millions of lives and damaged economies across the globe. Most COVID-19 deaths and economic losses are reported from densely crowded cities. It is comprehensible that the effective control and prevention of epidemic/pandemic infectious diseases is vital. According to WHO, testing and diagnosis is the best strategy to control pandemics. Scientists worldwide are attempting to develop various innovative and cost-efficient methods to speed up the testing process. This paper comprehensively evaluates the applicability of the recent top ten state-of-the-art Deep Convolutional Neural Networks (CNNs) for automatically detecting COVID-19 infection using chest X-ray images. Moreover, it provides a comparative analysis of these models in terms of accuracy. This study identifies the effective methodologies to control and prevent infectious respiratory diseases. Our trained models have demonstrated outstanding results in classifying the COVID-19 infected chest x-rays. In particular, our trained models MobileNet, EfficientNet, and InceptionV3 achieved a classification average accuracy of 95%, 95%, and 94% test set for COVID-19 class classification, respectively. Thus, it can be beneficial for clinical practitioners and radiologists to speed up the testing, detection, and follow-up of COVID-19 cases.

**INDEX TERMS** COVID-19 detection, Chest X-rays, Deep learning models

## I. INTRODUCTION

The novel coronavirus known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was initially found in China and spread worldwide, infecting millions of people and causing a large number of deaths. It has become one of the biggest threats to human health. This disease has various symptoms, including coughing, fever, headache, loss of taste or smell, and shortness of breath. COVID-19 has affected not only human health but also imposed long-term economic, social, demographical, and cultural effects [1]. The COVID-19 spread rapidly and got the status of a pandemic in a short time as the virus can be transmitted to others through the air or by touching

contaminated surfaces. Based on the investigations, it is found that human was not the primary initiator of this disease. Instead, they acquired it from other animals such as bats and pangolin.

Due to the unavailability of vaccines, appropriate medication, and lack of health facilities on a large scale, the early detection of pandemic disease is a key to fight against its widespread infection [2]. Hence, a sustainable diagnosis methodology needs time; it will help flatten the curve of new cases, which will help the governments provide the proper and timely health facilities to the patients. The tests can be used to prevent further transmission by identifying individuals who have been

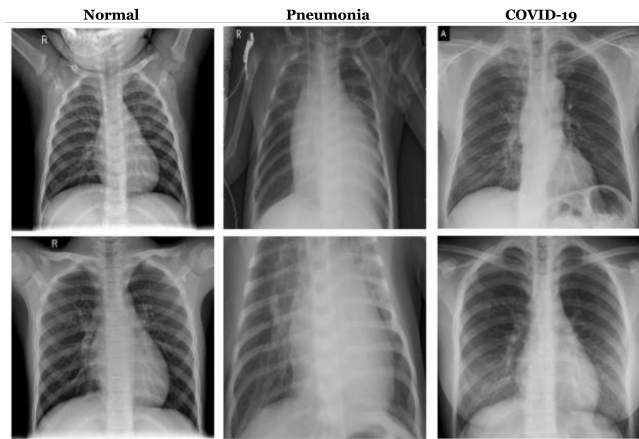


FIGURE 1: From left to right column-wise: Presents normal, pneumonia and COVID-19 CXR scans.

exposed to an infected person or animal but are not showing symptoms. This information could enable health workers to provide appropriate care to those who have tested positive for COVID-19.

The current outbreak is one of the most serious ever reported. It has highlighted the need for improved sustainable diagnostic tools and techniques that can be used earlier before patients are admitted into intensive care units. In addition, a great emphasis is required on public awareness and control measures such as washing hands, social distancing, and avoiding contaminated food or water supplies. As a result, it will help reduce the spread and transmission from person to person. However, many people infected with this virus still remain asymptomatic and undiagnosed despite these efforts. The medical scientists, practitioners, and doctors continue their tireless efforts to find the vaccine and help the patients get out of this disease. Similarly, the researchers belonging to computer science are also playing their role by proposing and developing novel techniques for the early and timely detection of COVID-19.

The virus can be tested using various methods, including manual, laboratory, genetic, and virus testing. Among these methods, virus testing is majorly used to diagnose COVID-19. In this type of testing, specimens are collected from the upper or lower respiratory tract of the patient and then are tested for possible viral infection [3]. However, some of these tests are uncomfortable for the patients and risky for the laboratory staff due to possible transmission in the air during the testing process, e.g., BAL (Bronchoalveolar lavage) sample testing. Moreover, some tests require a few days to produce the results. RNA-based testing is widely used as a standard testing approach by many laboratories worldwide; however, it involves the collection of swabs from the patients a day before testing. Moreover, it is costly, and all the laboratories cannot afford it. The efficacy or safety of

RNA sequencing as part of routine clinical diagnostic testing has also not been investigated on a large scale. Moreover, there is a lack of availability of viral transportation procedures.

Alternative to the testing mentioned above techniques, the Chest X-Ray can provide a basis to develop a more sustainable, fast, and practical testing technique. COVID-19 badly affects the lungs, leading to pneumonia in critical patients. To assess the effects of COVID-19 on the lungs, a Chest X-Ray (CXR) can be very useful. There might be other causes of lung infections causing pneumonia; however, some specific features of the X-Ray test indicate the lung damage due to COVID-19. There have been several studies on chest X-rays as a screening tool against COVID-19. However, these tests were not designed for this purpose, and they do not provide prominent information about the presence or absence of COVID-19. In addition, the CXR cannot be accurately used to diagnose pneumonia because some bacteria that live inside the body (e.g., *Pseudomonas*) can only grow when exposed to oxygen. Therefore, the CXR scans should always be performed when patients present with signs and symptoms suggestive of pneumonia. For example, chest X-rays might show elevated white blood cells (leukocytes) such as neutrophils and eosinophils. These findings can help doctors identify any infection within their patients related to COVID-19 [4]. The CXR has the advantage of its availability as many hospitals have x-ray scanners installed already on their premises, so they do not need to purchase additional equipment [5]. Moreover, the X-ray procedure is low dangerous for medical staff involved in testing as compared to the other techniques where the samples are taken from the respiratory stem of the patients. In this context, an X-ray test is considered one of the potential tools to diagnose COVID-19, as shown in Figure 1. Moreover, the physicians recommend the CXR as a quick and reliable imaging tool for detecting COVID-19. In this perspective, several deep learning techniques have been proposed for the detection of COVID-19 using radiography imaging [1, 6, 7]. These techniques show encouraging results in terms of accuracy.

Deep Convolutional Neural Networks (DCNNs) is considered one of the promising techniques for solving a vast range of problems, including image classification, image segmentation, face and speech recognition, machine translation [1, 8], and many other vision-based applications [9]. Importantly, DCNNs have obtained above human-level accuracy on various image classification tasks [10]. Inspired by such successes, DCNNs are being increasingly used in varied medical applications [11, 12] such as brain tumor segmentation [13]. The utilization of DCNN in the medical domain has shown promising results that motivated the authors to investigate the applicability of the DCNN-based techniques to detect COVID-19. In this context, this paper evaluates the recent ten state-of-the-art DCNN-based architectures to detect

COVID-19 using CXR scans. These models include:

- EfficientNet and its variants [14]
- DenseNet [15]
- MobileNet [16]
- Xception [17]
- InceptionV3 [18]
- ResNet [19]
- VGG and its variants [16]
- NASANetMobile [20]
- InceptionResNetV2 [21]
- NASANetLarge [20]

The paper is organized as follows—section 2 presents the literature review. In section 3, the basic building blocks of Convolutional Neural Networks (CNNs) are explained. Section 4 describes the methodology, dataset, and models investigated for COVID-19 classification, training, experimental setup, and implementation details. In section 5, we discuss generated attention maps (Grad-Cam) using our trained models. Finally, section 6 concludes the paper and discusses future research directions.

## II. LITERATURE REVIEW

Based on the recent record-shattering performance of deep learning in various fields, many deep learning researchers actively apply reliable deep learning-based algorithms that can accelerate the detection and classification of infected COVID patients. In this regard, the paper [22] proposed a technique based upon VB-Net neural network. This model is trained on the dataset, which consisted of 249 training and 300 validations patient records with symptoms of syndromes and epidemiology.

Table 1 presents the details of obtained accuracies, sensitivities, and specificities. Moreover, their method can produce the results 3-5 days after the examination. In paper [23], the authors proposed a 2D and 3D convolutional neural network-based method to detect, quantify, and track the COVID patients effectively. In [4] authors proposed deep learning-based methodology named DeTraC (Decompose, Transfer and Compose) for classification of COVID-19 using chest x-ray images.

The paper [5] has evaluated the applicability of the pre-trained deep learning architectures, i.e., Resnet50 (with residual connections, i.e.) and VGG16 detection of COVID-19 using chest X-rays of viral and bacterial pneumonia. Their dataset comprised 135 chest X-rays of COVID-19 and 320 chest X-rays of viral and bacterial pneumonia. The under-representation of the COVID-19 class has resulted in the limitation of information in COVID-19 images compared to the bacterial pneumonia class. Such imbalance in data distribution is well known in the literature, and it has adversely affected the final classification of the model biased towards the majority class.

Table 1 presents the details of obtained accuracies, sensitivities, and specificities. Following a similar line of research, The paper [1] evaluated their proposed method

on a publicly available dataset that consists of only 100 total images. The paper [3] combined the COVID-CXR and CORDA (COVID Radiographic images Data-set for AI) chest x-rays datasets to increase the size of the available dataset. Their proposed method named “COVID-Net” has obtained a comparable sensitivity and specificity 1. The paper [24] used EfficientNet for the detection of COVID-19. The EfficientNet has obtained the state of the art accuracy on a large-scale imagenet dataset [25]. Their proposed method has 30 times fewer parameters than the baseline model and 28 and 5 times fewer parameters than the popular VGG16 and ResNet50 architectures.

The authors of the paper [26] investigate the applicability of the capsule network for the detection of the COVID-19 using X-ray images. Convolutional Neural Network, however, is prone to lose spatial information between image instances and require large datasets. An alternative modeling framework based on Capsule Networks, referred to as the COVID-CAPS can handle small datasets, which is of significant importance due to the sudden and rapid emergence of COVID-19. The paper [27] proposed a ResNet-50-based solution for the detection of COVID-19 using a chest X-ray of the patients. All their experiments were based on 278 images of chest X-rays that are obtained from the present repositories by the University of Montreal and the National Institutes of Health. The dataset consisted of three groups of classes normal, pneumonia, and COVID-19.

The paper [28] used a pre-trained CNN based to perform classification schema considering the multi-class and hierarchical perspectives since pneumonia can be structured as a hierarchy. All their experiments were based on a publicly available dataset named RYDLS-20. The RYDLS-20 dataset containing CXR images of pneumonia caused by different pathogens composed of 1,144 CXR images from seven classes: Normal lungs and lungs affected by COVID-19, MERS, SARS, Varicella, Streptococcus, and Pneumocystis as well as 1,000 CXR images of the healthy lung for the COVID-19 identification in the hierarchical classification scenario. The paper [6] investigated the applicability of the Mobilenet, a deep learning model, and demonstrated that their model had obtained comparable accuracy. They tried to provide a lightweight screening solution for COVID-19 detection.

The paper [29] has proposed a 3-step technique to fine-tune a pre-trained ResNet-50 architecture to improve model performance and reduce training time; they called it COVID-ResNet. They progressively resized the input images from 128x128x3, 224x224x3, to 229x229x3 pixels and fine-tuned three different networks at each stage. They performed all their experiments on publically available COVIDx dataset details of accuracies, sensitivities, and specificities given in Table 1. The paper [30] has proposed and 2D patches-based convolutional neural network

approach with a relatively small number of trainable parameters for COVID-19 diagnosis. Their proposed method was inspired by potential imaging biomarkers of the CXR radiographs and provided comparative results with the state-of-the-art (SOTA) method. Further, they generated interpretable saliency maps and found that they strongly correlate with radiological findings.

All their experiments were performed on the dataset comprised of 247 chest posteroanterior (PA) radiographs collected from 14 different institutions of the Japanese Society of Radiological Technology (JSRT). The dataset included normal, lung nodules and COVID-19 cases. The paper [24] has evaluated the EfficientNet family of model performance on two different datasets, which are the COVIDx and COVID-19 Database/Italian Society of Medical and Interventional Radiology. They use two approaches: flat classification and hierarchical classification and to conduct efficient training on the deep neural networks. Further, they followed the transfer learning technique and performed data augmentation techniques to increase the amount of the training dataset. Using the data augmentation technique, their model has obtained improved accuracies. Details are given in Table 1.

Ezz et al.[31] Proposed a deep learning framework for diagnosing Covid-19 patients in chest x-rays. The proposed scheme consists of 7 deep learning algorithms, which are VGG19, DenseNet121, InceptionV3, ResNetV2, Inception-ResNet-V2, Xception and MobileNetV2. Collectively, these algorithms achieved some promising results; however, the dataset used for this study is minimal. This data consists of 50 x-rays images where 25 are COVID-19 confirmed patients while 25 are health subjects. The results of the applied algorithms can be seen in table 1.

CT imaging technique is considered one of the most effective techniques for detecting lung diseases. Given this, most of the deep-learning researchers worked on CT scans by applying deep-learning models to detect COVID-19. Recently, Ophir Gozes et al. [23] applied U-NET and Resnet-50 to thoracic CTs and achieved stat of the art results. The results are presented in Table 1. In this study, the authors used multiple databases, including the US and China CT Scans.

In the first step, U-net architecture was applied to segment the abnormal lung portions. This method enabled the author to obtain the region of interest (ROI) and remove irrelevant sections. Then data augmentation techniques are applied to overcome the problem of limited data. Then pre-trained model Resnet-50 (on Imagenet) is fine-tuned on the resultant dataset. For visualizing the network decision, the Grad-cam technique is used to create an activation map.

Fei Shan et al.[22] developed a deep-learning-based system to quantify lung infection in CTs. This model can estimate the percentage, shape, volume, and auto-contouring of the infected portions of the lungs. In

this study, a novel strategy of human intervention (field experts i.e., radiologists) is adopted to guide the results of the deep-learning models. The authors collected the data from two different centers. A total of 300 COVID-19 patients were obtained from Shanghai for testing, and 249 CT images of COVID-19 confirmed patients were acquired for training purposes (outside the Shanghai). The authors developed a DL model VB-NET by integrating a bottle-neck structure[32] and V-NET. By applying this strategy, VB-net had shown faster training speed than V-NET. The results of this model can be seen in Table 1.

In [26], the authors came up with a capsule network for the detection of COVID-19 in chest Xrays. The experiments were carried out on [33] and [34] datasets. Combining these two datasets resulted in 4 different labels ( bacterial, non-bacterial, COVID-positive, and COVID negative). Since the main purpose of this study is to develop a deep learning model that can contribute to diagnoses of COVID-19, all the labels were turned into only two labels. Bacterial, non-bacterial, COVID negative are presented as COVID-negative while COVID-positive has remained as COVID-positive class. Coming to the network architecture, the COVID-CAPS consisted of 3 capsule layers and four convolutional layers and had shown promising results on a limited dataset. When the model was pre-trained on NIH dataset [32] its results were improved significantly, which are precisely presented in Table 1.

M Farooq et al.[35] presented the COVID-ResNet model for the classification of COVID-infection in chest Xrays. The authors used 50 layered pre-trained (on imagenet) reneat architecture. These experiments were carried out on COVIDx Dataset. Since this database contained limited samples and deep-learning models require a more significant amount of training samples, a data augmentation strategy is adopted. The authors claimed that this model achieved the stat of the art results by using the transfer learning technique even after 40 epochs. The results can be seen in Table 1.

Asmaa Abbas et al.[4] fine-tuned their previously developed deep learning model for COVID-19 detection in chest X-rays. This model was named Decompose, Transfer, and Compose (DeTraC). The authors claim that this model can deal with irregularities in images. The data used for this study is the composition of two different datasets. One dataset was acquired from the Japanese Society of Radiological Technology, and the other one was obtained from [33]. The data augmentation technique is applied to increase the training data. The authors used pre-trained Alexnet with three convolutional layers for class decomposition. For classification purposes, Resnet-18, which is pre-trained on imagenet is used. Resnet-18 consisted of 18 layers. Since this network has to classify between six different classes, the last fully-connected layer was changed accordingly. The model achieved comparatively 95.12% 97.91% 91.87% and



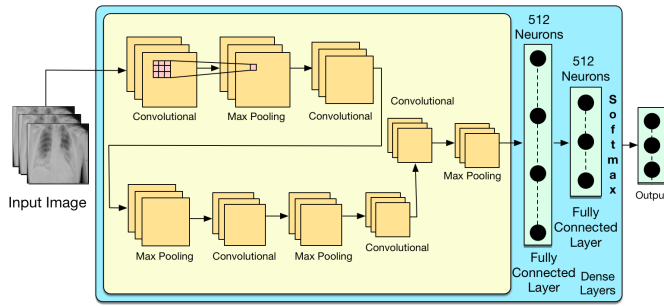


FIGURE 2: Block diagram of typical Convolutional Neural Networks (CNNs).

93.36% in terms of accuracy, sensitivity, specificity, and precision, respectively.

### III. CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks (CNNs) have obtained above human expert-level classification accuracies on various domain tasks, which include image captioning [48], image classification [14], semantic segmentation [49][50], face recognition [51][52], speech recognition [53], language translation [54], machine translation [55], medical imaging analysis [56][57][58][59] and many other vision tasks [60].

The first and simplest neural network was introduced as a feedforward network with multiple nodes arranged in a layer connected through edges with weights. In the following subsection, we will briefly explain the basic building blocks of typical CNNs networks.

#### A. CONVOLUTIONS

Using multiple convolutional filters, the convolution operator takes the image as an input and performs convolution across an image's spatial dimension. The convolutional filters perform the role of feature extractors. They extract essential features that help the network obtain minimum loss and highest attainable classification accuracy.

#### B. FILTERS

The filters are the learnable parameters of each layer. In conventional settings, each layer has a fixed number of learnable parameters, commonly called neurons. These filters learn different features from input either directly in image space (if filters are operating on the image) or in feature space (if the layer receives intermediate features maps). The ultimate objective of the filters is to learn the features that help the network attain the lowest optimization loss and highest classification accuracies. Therefore, the filters know different components, including some filters playing the role of edge detectors, other learned color-related features, etc. The early layer filters learn the abstract features. In contrast, the later layer filters (which are near the classification layer) learn more

concrete and semantic features built upon the early learn filter activation.

#### C. FEATURE MAPS

Each filter uses a kernel window of a fixed dimension for a feature detector during convolution. The windows slide over the image with a pre-defined strike from left to right to produce a dot product saved at the location of the corresponding pixels in the output feature map. The number of filters defines the channel dimension for the output features (also known as a tensor of features). The stride value represents the convolution operation movement in spatial dimensions. If we set stride to one, the convolution operation will move pixel by pixel through the adjacent spatial location. Every spatial location will be considered. Setting the stride value to higher than one value will skip the adjacent pixel and output the feature tensor in a reduced dimension. The stride value is also used as a dimensionality reduction trick.

#### D. NON LINEARITIES

The researcher proposed different types of non-linearity that help the network to learn an arbitrary complex task. The non-linearity makes the neural network a universal function approximator. Most widely used non-linearities include, sigmoidal, ReLU, parametric ReLU and tangent [61]. The detailed list of non-linearity can be found in [61]. Among all non-linearity functions, ReLU is the most widely used function. The ReLU stands for the Rectified Linear Unit for a non-linear operation. The output is calculated as follows:

$$f(x) = \max(0, x)$$

The ReLU performs an element-wise operation where all negative activations are suppressed to zero in the output feature map.

#### E. PADDING

The operation of padding extra zeros or ones to the input dimension is called padding. The padding operation meets the intermediate layers' input and dimensional output requirements.

#### F. POOLING

Pooling is the most common technique used for dimensionality reduction in deep learning. The researchers proposed a variety of pooling techniques which include average pooling, max-pooling, and global average pooling [62] [63]. Among them, max pooling is the most commonly used technique for dimensionality reduction. The below expression shows the mathematical formulation of the max pooling.

$$f_{\max}(x) = \max_i x_i$$

$$f_{\text{avg}} = \frac{1}{N} \sum_{i=1}^N x_i$$

TABLE 1: The table presents the accuracy, sensitivity, specificity, and F1-score of the existing approaches employed to detect COVID-19 using CXR scans.

| Summary of Literature Review |          |             |             |          |
|------------------------------|----------|-------------|-------------|----------|
| Method                       | Accuracy | Sensitivity | Specificity | F1 Score |
| Oh et.al[30]                 | 92%      | 100%        | -           | 77%      |
| Abbas et.al [4]              | 95%      | 98%         | 92%         | -        |
| Hall et.al[5]                | 97%      | 97%         | 94%         | -        |
| Khan et.al [1]               | 90%      | 100%        | -           | 87%      |
| Tartaglione et.al [3]        | 85%      | 90%         | 80%         | 90%      |
| Wang, Linda et.al [32]       | 93%      | 87%         | 96%         | -        |
| Gozes et.al[23]              | -        | 96%         | 98%         | -        |
| Luz et.al[24]                | 91%      | 90%         | 92%         | -        |
| Bukhari et.al[27]            | 98%      | 98%         | -           | 98%      |
| Apostolopoulos et.al[36]     | 97 %     | 99%         | 96 %        | -        |
| Afshar et.al[26]             | 96%      | 90%         | 96%         | -        |
|                              | 98% (FT) | 80% (FT)    | 98% (FT)    | -        |
| Pierira et.al[28]            | -        | -           | -           | 89%      |
| Butt et.al[37]               | 86%      | -           | -           | -        |
| Wang, S and Kang et.al[32]   | 79%      | 63%         | 83%         | -        |
| Jin,,chang et.al [38]        | 95%      | -           | -           | -        |
| Jin, Shou [39]               |          | 97%         | 92%         | -        |
| Chowdhury et.al[40]          |          | 97%         | 92%         | -        |
| Farooq et.al[35]             | 96%      | 100%        | -           | 100%     |
| Ni, Qianqian et.al [41]      | 91%      | 88%         | 95          | 92%      |
| Shety et.al [42]             | 95%      | 96%         | 91%         |          |
| Ozkaya et.al[43]             | 98%      | 98%         | 97%         | 97%      |
|                              | 96%      | 95%         | 93%         | 98%      |
| De Moraes et.al[44]          |          | 68%         | 85%         | -        |
| Minaee et.al[8]              |          | 97%         | 90%         | -        |
| Acar et.al[45]               | 99%      | 100%        |             | 100%     |
| Bai et.al[46]                | 85%      | 79%         | 88%         |          |
|                              | 90%      | 88%         | 91%         |          |
| Tabik et.al[47]              | 97%      | 88%         | 66%         |          |

#### G. FULLY CONNECTED LAYER

The softmax activation is a widely used activation function for the output of the deep learning approaches. The below expression shows the mathematical formulation of the softmax function.

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

#### IV. METHODOLOGY

In this research, we comparatively investigated the feasibility of the recent ten state-of-the-art deep learning approaches for detecting COVID-19 cases from CXR images. All experiments were performed on an open-access benchmark dataset [3, 64]. The dataset consisted of three classes, i.e., COVID-19, normal, and pneumonia CXR scans. We divided the whole dataset into two splits training and test splits. We kept the test split separate from obtaining the models' final classification accuracy (i.e., we did not include in the training set

images). We used a deep convolutional neural network (CNN) based method to classify COVID19, normal, and pneumonia images based on CXR scans. Our models were trained with hyper-parameter settings given in Table 4.

We reached optimal hyper-parameters settings for our models after performing an array of experiments [65]. We tested different combinations of hyperparameters' settings for each experiment and validated the test set. In Table 1, we compared the best-trained model performance with other proposed approaches. Finding the optimal hyperparameters of deep any deep model is a stiff challenge. In this regard, our search for optimal hyper-parameters was done through an array of experiments. We analyzed the intermediate experiment output and crafted the next experiment based on observations for the quest to find the optimal hyperparameters.

We designed the overall research steps as 1).

experiment, 2). obtain results 3). critically analyze obtained results 4). designed future experiments based on obtained observations to get better results). After repeating the experimentation-observation cycle many times, we found the optimal hyperparameters settings for the COVID-19 dataset. We validated the performance of the hyperparameter's validity on the test set. We froze the best-performing hyperparameters settings for later minor fine-tune that followed. We kept trying different reasonable combinations for the remaining hyperparameters (i.e., epoch count, stopping criterion ) until we found the optimal hyper-parameter value in their parameter space. The detail on optimum hyperparameters can be seen in Table 4.

### A. DATASET

For this research work, we have used publicly available datasets to train and evaluate deep learning models[3][64]. The Table ??, presents the overall dataset distribution. As can be seen that the COVID-19 class in the dataset is under-represented as compared to the other two classes, i.e., pneumonia and normal classes. As a consequence of such an imbalanced dataset, the deep learning model will not be able to learn the under-represented class compared to the other comparatively over-represented classes. To address this challenge, we have to introduce an intelligence strategy at the classes level which enforces the deep learning methods to treat all the classes equally despite their imbalance distribution. The strategy should especially enforce the deep learning method to learn the underrepresented detection class (i.e., Covid-19) and two classes (i.e., normal and pneumonia). For this purpose, we introduced the class-wise weightage to enforce the deep learning models to learn unbiased classification of both over and underrepresented classes.

Without such class-wise weightage, the naively trained deep learning models (i.e., without addressing the class-imbalance issue) on this dataset will produce biased predictions towards the majority classes. Therefore, we intelligently calculated class-wise weight, as mentioned in [66]. We assigned higher weightage to the minority class (i.e., COVID-19 in our case) and lower weightage to majority classes in the final optimization function of the deep learning model. The Table 3 presents the class-wise weightage distribution. Given these class-wise weights assigned, the deep learning model will be highly penalized for the misclassification of the COVID-19 example compared to the normal and pneumonia examples.

Such varying misclassification weightage will play a role in forming the weak supervisory signal that guides the training of the deep model and will eventually guide the optimization algorithm to converge to a local minimum with characters of better performance on under-represented classes. Another benefit of such a strategy is prioritizing the importance at the class level by providing such weak supervisory signals to the optimization algorithm.

Following our class-wise weightage, the deep learning

TABLE 2: Class-wise distribution of overall dataset.

| Training Distribution |           |          |
|-----------------------|-----------|----------|
| Normal                | Pneumonia | COVID-19 |
| 7966                  | 5488      | 176      |
| Test Distribution     |           |          |
| Normal                | Pneumonia | COVID-19 |
| 885                   | 608       | 24       |

TABLE 3: Class-wise weightage distribution addresses the class-imbalance issue. We gave higher weightage to the minority classes, especially COVID-19 classes. Such higher weightage will enforce the deep learning model to correctly classify the minority classes (i.e., COVID-19 in our case).

| Class-wise weightage |           |          |
|----------------------|-----------|----------|
| Normal               | Pneumonia | COVID-19 |
| 5                    | 10        | 25       |

model will receive a higher penalty for the misclassification of the minority classes compared to the majority classes. The objective of the optimization function is to obtain a lower cost. Therefore, during the optimization, the optimization algorithm (i.e., gradient descent) will be enforced to give more attention to minority class learning compared to majority class learning and minimize the final objective function by obtaining a lower cost.

### B. CNN ARCHITECTURES

In this paper, we have evaluated the applicability of the ten state-of-the-art CNN architectures to detect COVID-19. The list of the architectures includes; the famous VGG and variant [67], Xception [17], DenseNet and variants [15], Inceptions [21], Resnet50 [19], InceptionResNetV2 [67], NASNetMobile [16], NASNetLarge [16] and EfficientNet [14]. All these models have proven their performance by obtaining promising results in different years of ImageNet competitions [68]. The EfficientNet model [14], which is originally proposed by Google, is the current winner of the ImageNet competition and has obtained above human detection and classification accuracies. In all of our experiments, we initialized all architectures with ImageNet weights and then fine-tuned them on our dataset for 500 epochs. We also included the early stopping criteria with patient 15 epochs to prevent overfitting.

### C. TRAINING & EXPERIMENTAL SETUP

We performed the final ten experiments with different state-of-the-art CNNs architectures. Table 4 presents the details of hyper-parameters being used in the different experimental setups. In this hyper-parameter Table 4, for each experiment, we have provided exact hyper-parameters values that were used to obtain good test set accuracies. We initialized all models with ImageNet weights and then fine-tuned them on the COVID dataset for 500 epochs. We

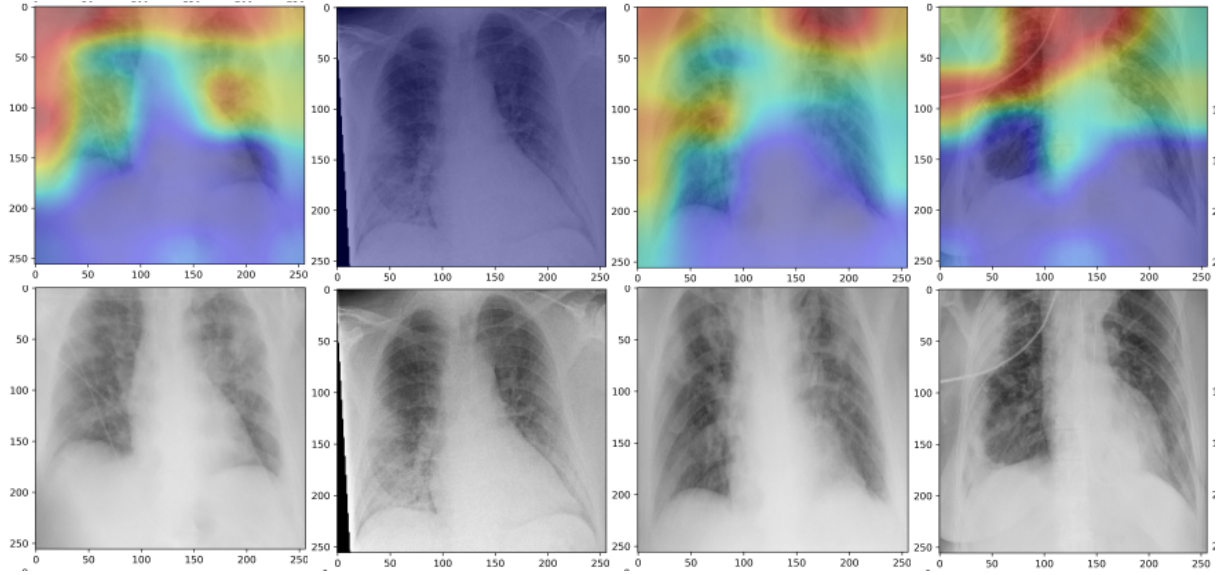


FIGURE 3: COVID-19 True Positives: correctly classified COVID-19 X-ray scans with corresponding attention maps (CAM). The figure is best viewed in color.

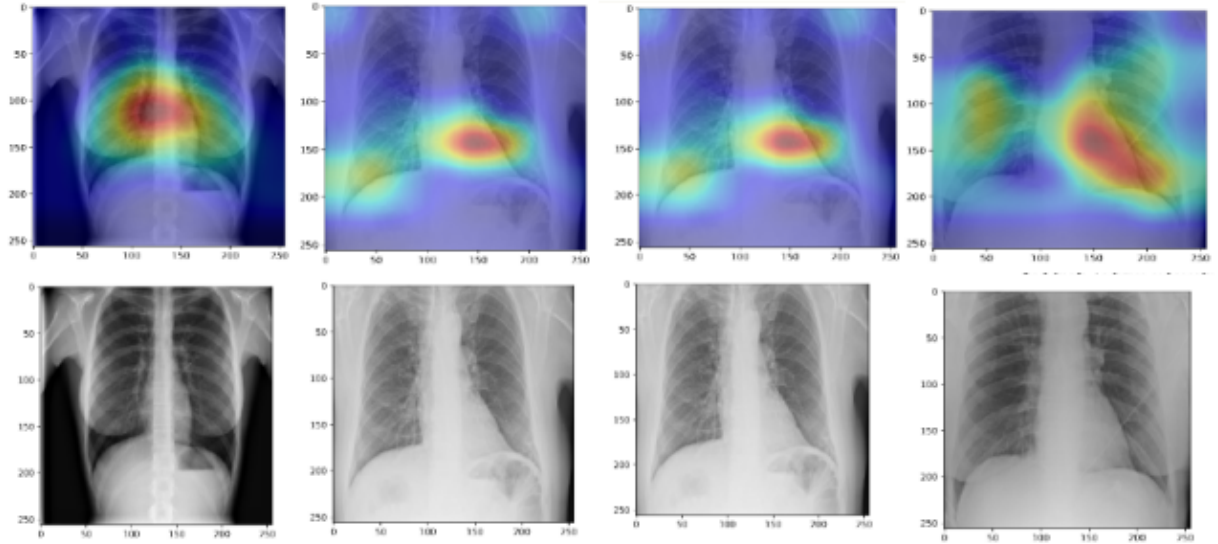


FIGURE 4: Normal True Negatives: correctly classified normal X-ray scans with corresponding attention maps (CAM). The figure is best viewed in color.

used Adam as an optimizer with a learning rate of 0.001, and all other parameters of adam were fixed to default values, as discussed in [69]. All experiments were performed with maximum epochs 500, batch size 128, a dropout layer with 0.5 dropout probability was added before the softmax classification layer, and early stopping criteria equal 25 epochs. We did not perform any data augmentation. We used the below expressions to calculate accuracy, sensitivity, specificity, precision, and F-measure.

deep learning library <sup>1</sup> with a TensorFlow backend <sup>2</sup>.

$$Accuracy = \frac{\text{No. of images correctly classified}}{\text{Total No. of images}}$$

$$Recall/Sensitivity = \frac{TP}{TP + FN}$$

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

<sup>1</sup>Keras - Deep learning framework <https://keras.io>

<sup>2</sup>TensorFlow - An end-to-end open-source machine learning platform <https://www.tensorflow.org/>

All models were trained and evaluated using the Keras



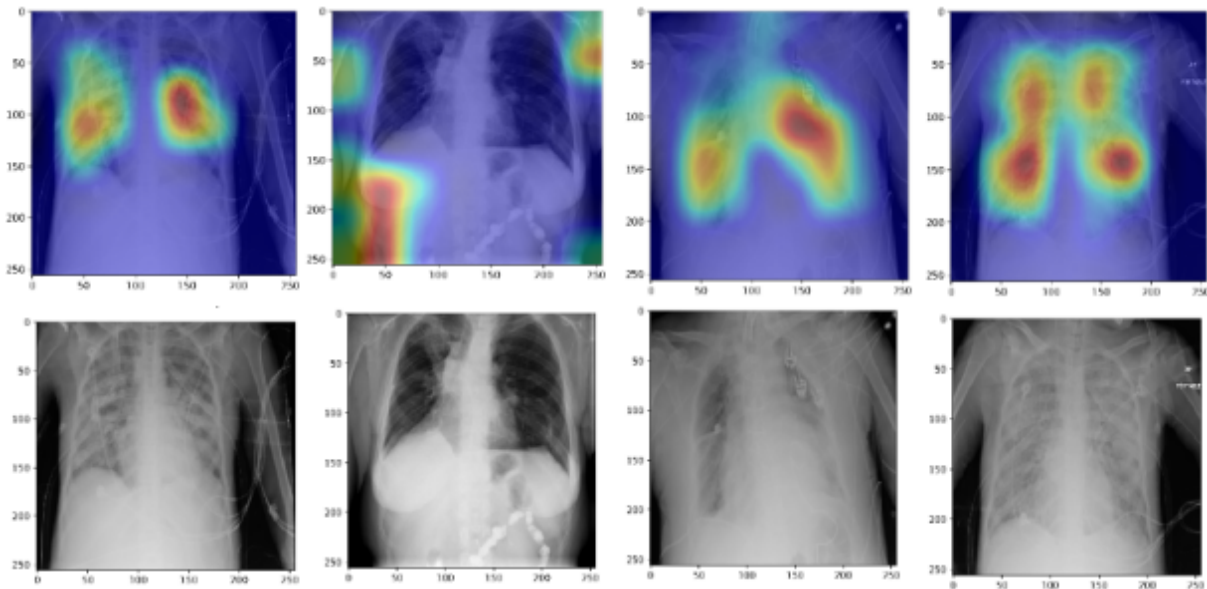


FIGURE 5: Pneumonia True Positives: correctly classified pneumonia X-ray scans with corresponding attention maps (CAM). The figure is best viewed in color.

$$Precision = \frac{TP}{TN + FP}$$

$$F - measure = \frac{(2 * Precision * Recall)}{Precision + Recall}$$

TABLE 4: The table presents the optimum hyper-parameter values found for CNN models.

| Hyper-parameters  |                |
|-------------------|----------------|
| Input Size        | (256, 256, 3)  |
| Dropout           | 0.5            |
| Optimizer         | Adam           |
| Learning rate     | 0.001          |
| Epoch             | 500            |
| Batch size        | 256            |
| Stopping criteria | Early stopping |

## V. RESULT & DISCUSSION

We designed experiments that helped us to comparatively evaluate the performance of 10 state-of-the-art CNNs models on three classes classification of COVID19, normal, and pneumonia classes. This list of considered CNNs architectures consisted of models as; the famous VGG and variant [67], Xception [17], DenseNet and variants [15], Inceptions [21], Resnet50 [19], InceptionResNetV2 [67], NASNetMobile [16], NASNetLarge [16] and EfficientNet [14]. Figure 7 presents the class-wise confusion matrix of all our ten trained models that were trained under experimental setting given in Table 4.

Amongst all models, it can be seen that MobileNet, VGG19, NASNetMobile, InceptionV3, and EfficientNet family of models have obtained consistent and promising results as compared to InceptionResNetV2 and NASNetLarge architectures. Both InceptionResNetV2 and NASNetLarge architectures suffer from overfitting problems and have received the lowest prediction accuracy, especially on COVID-19. Besides that, our other trained models demonstrate the applicability of deep learning models to detect COVID-19 disease in the CXR scan. However, It is important to mention that MobileNet, EfficientNetB0, and EfficientNetB1 have obtained 100% accuracy for COVID-19 class detection in our experiments. The accuracy of the 100% COVID-19 classification has further highlighted the efficient applicability of the deep learning models in early COVID-19 disease detection. Table 5 presents a comparison of our deep learning model with other CNNs based baseline methods. The results show that our trained models have obtained promising results and consistently performed on this benchmark and can beat most existing models in terms of classification accuracy, sensitivity, specificity, and precision.

Developing a better understanding of deep learning models is an active research area. Deep Convolution Neural Networks are often referred to as black-box models due to minimal knowledge of their internal actions. As an effort to develop more complex explainable deep learning models. Recently, many researchers have proposed methods to provide class activation maps (CAM) that visualize the deep learning prediction with the hope to help human experts develop understandable [70] the deep learning models. In this regard, the authors in [3], proposed methods that can generate the gradient-based

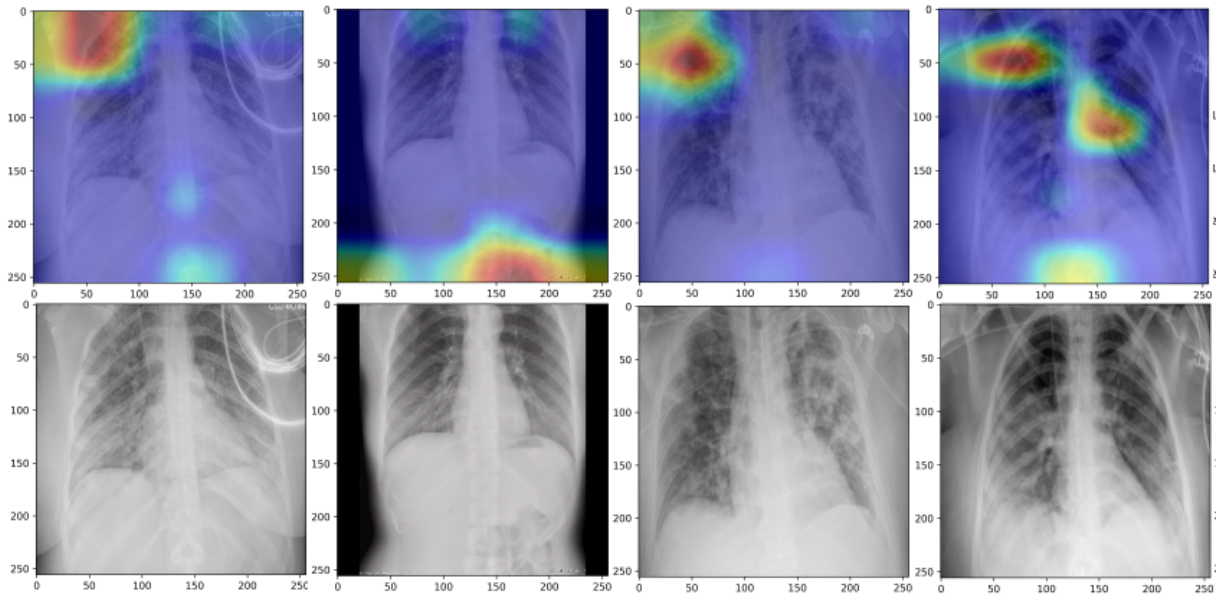


FIGURE 6: This Figure presents the attention maps for miss-classified cases. Starting from top to bottom, The first inner rectangle (i.e., borders highlighted in red color) shows attention maps for COVID-19 scans that were miss classified either into normal or pneumonia class. The second inner rectangle (i.e., borders highlighted in green) presents attention maps for normal scans that were miss classified either into COVID-19 or pneumonia class. The third inner rectangle (i.e., borders highlighted in blue) presents attention maps for pneumonia scans that were miss classified either into COVID-19 or normal class. Inside each inner rectangle, the first row shows attention maps while the second row contains the corresponding input scans (Best viewed in colors).

CAM (i.e., grad-cam), which highlights the more informative the input image concerning the final prediction of the model for each class. The availability of such information, along with the model's predictions, plays a vital role in developing trustworthiness in deep learning-based solutions. Furthermore, the availability of a grad-cam makes it possible to validate the reliability of deep learning by a human expert (i.e., a doctor). We used a grad-cam algorithm to visualize the prediction of the deep learning models.

Figures 4,3 and 5 present the input image, model prediction, and corresponding Grad-Cams for our trained models, for normal, COVID-19 and pneumonia input scans using [71][3] gradient class activate map (Grad-Cam) algorithm, respectively. The attention maps are a simple and efficient way to visualize the features that the model has based its final classification decisions on. We used the Grad-Cam algorithm to visualize trained model features for all three classes (i.e., COVID-19, normal, and pneumonia). The Figures 4,3 and 5 present the attention maps (i.e. Grad-Cam) for COVID-19, normal and pneumonia classes, respectively. In all three Figures 4,3 and 5 the second, fourth, and sixth rows represent the input CXR scans, while the first, third, and fifth rows represent the corresponding attention maps of our trained model. The attention maps explain the model predictions at pixels level in the input CXR scan space.

The availability of such attention maps facilitates the

explainability of the prediction of deep learning. Moreover, attention maps also help develop trust in the prediction of the CNNs models, especially deployed/used in health care. The attention maps of the deep learning model in Figure /reffig:map demonstrate which input pixels or features contributed most to the final prediction of the model. Interestingly, in the case of COVID-19, just like domain experts (i.e., radiologists), we can see the model has given more focus/attention to the middle of the chest (i.e., lungs locations) while making the ultimate decision. Furthermore, From the given attention maps, It can also be seen that the model has not only considered but also analyzed the repository organ (i.e., lungs) conditions given in an input x-ray scan before classifying the input scan as a COVID-19 scan. Similarly, in the case of normal patient x-rays, the model has based its classification on the left and right ribs of the chest. Finally, in the case of classifying the pneumonia patient x-ray, It can be seen that the model has based its classification decision on the upper left and right part of the chest x-rays.

To comprehensively understand the models' predictions, we also visualize the attention maps for the misclassified examples. Figure 6 presents the attention maps for misclassified COVID-19, normal, and pneumonia examples. The true class (e.g., COVID-19) is confused with anyone of the other two classes (i.e., either normal or pneumonia).

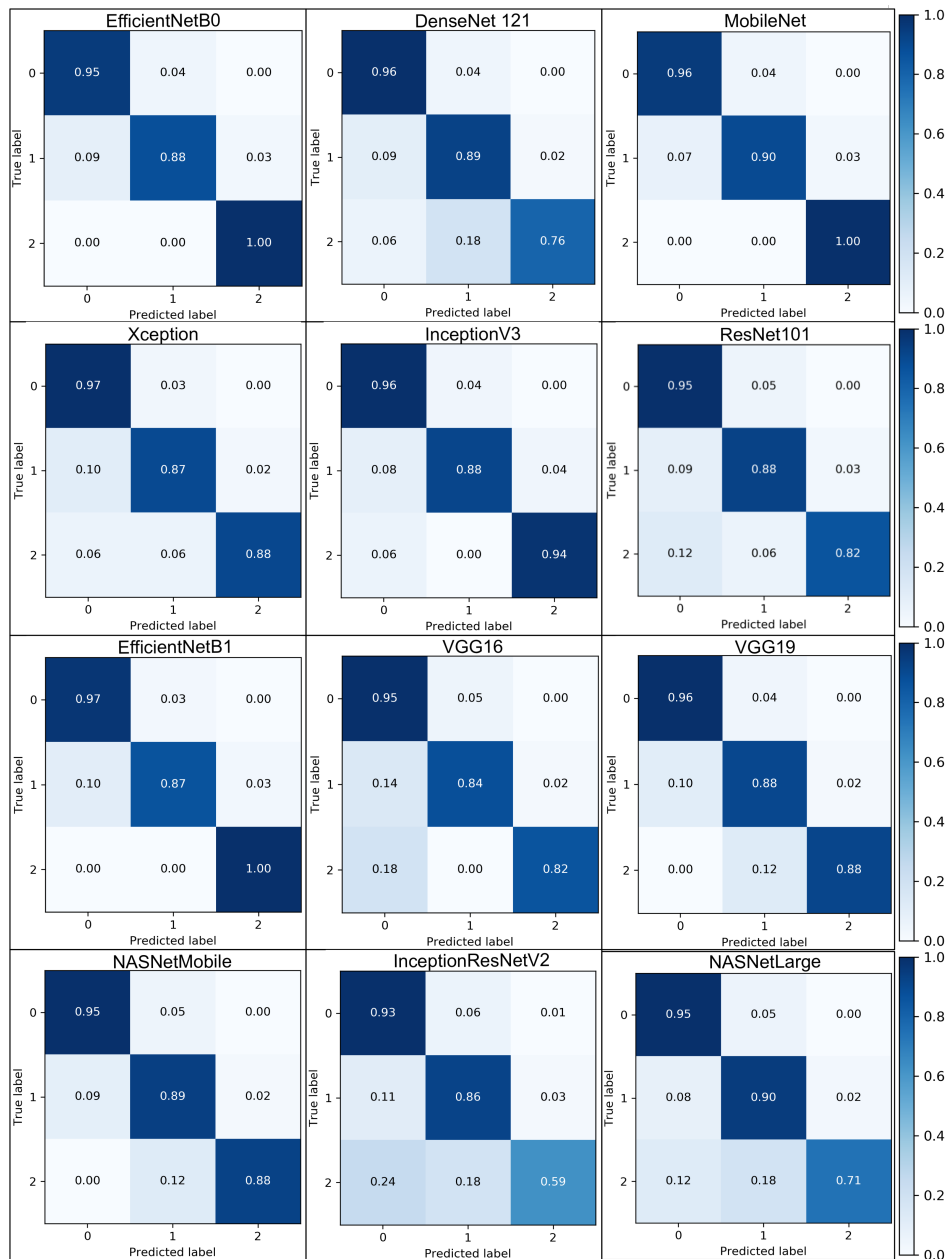


FIGURE 7: Confusion matrices for considered CNN models.

## VI. CONCLUSION

This paper comprehensively evaluated the applicability of the recent top ten state-of-the-art Deep Convolutional Neural Networks (CNNs) for automatically detecting COVID-19 infection from chest X-ray images. These models includes, EfficientNet and its variants [14] DenseNet [15], MobileNet [16], Xception [17], InceptionV3 [18], ResNet [19], VGG and its variants [16], NASNetMobile [20], InceptionResNetV2 [21] and NASNetLarge [20]. We showed that our trained models obtained outstanding results in classifying the COVID-19 infected chest x-rays. In particular, our trained models MobileNet, EfficientNet, and InceptionV3 achieved a

classification average accuracies of 95%, 95%, and 94% test set, for COVID-19 class classification, which is significantly higher than previous state-of-the-art models. Such results demonstrated how deep learning could substantially help in the early detection of COVID-19 patients using X-rays scans. Moreover, it also provided an excellent example for other researchers interested in applying this method to real-world problems. Lastly, we believe it can be a beneficial tool for clinical practitioners and radiologists to speed up testing, detection, and follow-up of COVID-19 cases.

TABLE 5: The table presents the results of our trained models. We evaluated the performance of our trained models based on five metrics: accuracy, sensitivity, specificity, and F1-score.

| Our Trained Models   |            |             |             |            |            |
|----------------------|------------|-------------|-------------|------------|------------|
| Method               | Accuracy   | Sensitivity | Specificity | F1 Score   | Precision  |
| InceptionNetV2 [18]  | 79%        | 79%         | 90%         | 81%        | 78%        |
| NasNetLarge [20]     | 85%        | 85%         | 93%         | 87%        | 85%        |
| ResNet101 [19]       | 87%        | 87%         | 94%         | 88%        | 87%        |
| VGG16 [16]           | 87%        | 87%         | 94%         | 89%        | 87%        |
| DenseNet121 [15]     | 87%        | 87%         | 94%         | 87%        | 87%        |
| NASNetMobile[20]     | 91%        | 91%         | 95%         | 91%        | 91%        |
| VGG19[16]            | 91%        | 91%         | 95%         | 91%        | 91%        |
| XceptionNet[17]      | 91%        | 91%         | 95%         | 91%        | 91%        |
| EfficientNetB0[14]   | 91%        | 91%         | 95%         | 90%        | 91%        |
| InceptionNetV3 [21]  | 94%        | 94%         | <b>97%</b>  | 94%        | 94%        |
| EfficientNetB1 [14]  | <b>95%</b> | <b>95%</b>  | <b>97%</b>  | <b>95%</b> | <b>95%</b> |
| <b>MobileNet[16]</b> | <b>95%</b> | <b>95%</b>  | <b>97%</b>  | <b>95%</b> | <b>95%</b> |

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## REFERENCES

- [1] 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," *Computer Methods and Programs in Biomedicine*, p. 105581, 2020.
- [2] M. S. Nadeem, M. A. Zamzami, H. Choudhry, B. N. Murtaza, I. Kazmi, H. Ahmad, and A. R. Shakoori, "Origin, potential therapeutic targets and treatment for coronavirus disease (covid-19)," *Pathogens*, vol. 9, no. 4, p. 307, 2020.
- [3] E. Tartaglione, C. A. Barbano, C. Berzovini, M. Calandri, and M. Grangetto, "Unveiling covid-19 from chest x-ray with deep learning: a hurdles race with small data," *arXiv preprint arXiv:2004.05405*, 2020.
- [4] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network," *arXiv preprint arXiv:2003.13815*, 2020.
- [5] L. O. Hall, R. Paul, D. B. Goldgof, and G. M. Goldgof, "Finding covid-19 from chest x-rays using deep learning on a small dataset," *arXiv preprint arXiv:2004.02060*, 2020.
- [6] X. Li and D. Zhu, "Covid-xpert: An ai powered population screening of covid-19 cases using chest radiography images," *arXiv preprint arXiv:2004.03042*, 2020.
- [7] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, and D. Shen, "Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19," *IEEE reviews in biomedical engineering*, 2020.
- [8] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning," *arXiv preprint arXiv:2004.09363*, 2020.
- [9] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multi-grade brain tumor classification using deep cnn with extensive data augmentation," *Journal of computational science*, vol. 30, pp. 174–182, 2019.
- [10] Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le, "Self-training with noisy student improves imagenet classification," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10 687–10 698.
- [11] Q. Feng and Z. Ding, "Mri radiomics classification and prediction in alzheimer's disease and mild cognitive impairment: A review," *Current Alzheimer Research*, vol. 17, no. 3, pp. 297–309, 2020.
- [12] Z. Sherkatghanad, M. Akhondzadeh, S. Salari, M. Zomorodi-Moghadam, M. Abdar, U. R. Acharya, R. Khosrowabadi, and V. Salari, "Automated detection of autism spectrum disorder using a convolutional neural network," *Frontiers in Neuroscience*, vol. 13, 2019.
- [13] X. Feng, N. J. Tustison, S. H. Patel, and C. H. Meyer, "Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic



- features,” *Frontiers in Computational Neuroscience*, vol. 14, p. 25, 2020.
- [14] M. Tan and Q. V. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” *arXiv preprint arXiv:1905.11946*, 2019.
  - [15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
  - [16] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, “Mobilenets: Efficient convolutional neural networks for mobile vision applications,” *arXiv preprint arXiv:1704.04861*, 2017.
  - [17] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
  - [18] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Thirty-first AAAI conference on artificial intelligence*, 2017.
  - [19] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
  - [20] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, “Learning transferable architectures for scalable image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8697–8710.
  - [21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.
  - [22] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, and Y. Shi, “Lung infection quantification of covid-19 in ct images with deep learning,” *arXiv preprint arXiv:2003.04655*, 2020.
  - [23] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, A. Bernheim, and E. Siegel, “Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis,” *arXiv preprint arXiv:2003.05037*, 2020.
  - [24] E. Luz, P. L. Silva, R. Silva, and G. Moreira, “Towards an efficient deep learning model for covid-19 patterns detection in x-ray images,” *arXiv preprint arXiv:2004.05717*, 2020.
  - [25] K. He, R. Girshick, and P. Dollár, “Rethinking imagenet pre-training,” in *Proceedings of the IEEE international conference on computer vision*, 2019, pp. 4918–4927.
  - [26] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, “Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images,” *arXiv preprint arXiv:2004.02696*, 2020.
  - [27] S. U. K. Bukhari, S. S. K. Bukhari, A. Syed, and S. S. H. SHAH, “The diagnostic evaluation of convolutional neural network (cnn) for the assessment of chest x-ray of patients infected with covid-19,” *medRxiv*, 2020.
  - [28] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla Jr, and Y. M. Costa, “Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios,” *Computer Methods and Programs in Biomedicine*, p. 105532, 2020.
  - [29] F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, and Y. Bengio, “Renet: A recurrent neural network based alternative to convolutional networks,” *arXiv preprint arXiv:1505.00393*, 2015.
  - [30] Y. Oh, S. Park, and J. C. Ye, “Deep learning covid-19 features on cxr using limited training data sets,” *IEEE Transactions on Medical Imaging*, 2020.
  - [31] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, “Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images,” *arXiv preprint arXiv:2003.11055*, 2020.
  - [32] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2097–2106.
  - [33] J. P. Cohen, P. Morrison, and L. Dao, “Covid-19 image data collection,” *arXiv preprint arXiv:2003.11597*, 2020.
  - [34] P. Mooney, “kaggle chest x-ray images (pneumonia) dataset,” 2020.
  - [35] M. Farooq and A. Hafeez, “Covid-resnet: A deep learning framework for screening of covid19 from radiographs,” *arXiv preprint arXiv:2003.14395*, 2020.
  - [36] I. D. Apostolopoulos and T. A. Mpesiana, “Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural @networks,” *Physical and Engineering Sciences in Medicine*, p. 1, 2020.
  - [37] C. Butt, J. Gill, D. Chun, and B. A. Babu, “Deep learning system to screen coronavirus disease 2019 pneumonia,” *Applied Intelligence*, p. 1, 2020.
  - [38] C. Jin, W. Chen, Y. Cao, Z. Xu, X. Zhang, L. Deng, C. Zheng, J. Zhou, H. Shi, and J. Feng, “Development and evaluation of an ai system for covid-19 diagnosis,” *medRxiv*, 2020.
  - [39] S. Jin, B. Wang, H. Xu, C. Luo, L. Wei, W. Zhao, X. Hou, W. Ma, Z. Xu, Z. Zheng et al., “Ai-assisted ct imaging analysis for covid-19 screening: Building

- and deploying a medical ai system in four weeks,” medRxiv, 2020.
- [40] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al-Emadi et al., “Can ai help in screening viral and covid-19 pneumonia?” arXiv preprint arXiv:2003.13145, 2020.
- [41] Q. Ni, Z. Y. Sun, L. Qi, W. Chen, Y. Yang, L. Wang, X. Zhang, L. Yang, Y. Fang, Z. Xing et al., “A deep learning approach to characterize 2019 coronavirus disease (covid-19) pneumonia in chest ct images,” *European Radiology*, pp. 1–11, 2020.
- [42] P. K. Sethy and S. K. Behera, “Detection of coronavirus disease (covid-19) based on deep features,” *Preprints*, vol. 2020030300, p. 2020, 2020.
- [43] U. Ozkaya, S. Ozturk, and M. Barstugan, “Coronavirus (covid-19) classification using deep features fusion and ranking technique,” arXiv preprint arXiv:2004.03698, 2020.
- [44] A. F. de Moraes Batista, J. L. Miraglia, T. H. R. Donato, and A. D. P. Chiavegatto Filho, “Covid-19 diagnosis prediction in emergency care patients: a machine learning approach,” medRxiv, 2020.
- [45] E. Acar, E. SAHIN, and I. Yilmaz, “Improving effectiveness of different deep learning-based models for detecting covid-19 from computed tomography (ct) images,” medRxiv, 2020.
- [46] H. X. Bai, R. Wang, Z. Xiong, B. Hsieh, K. Chang, K. Halsey, T. M. L. Tran, J. W. Choi, D.-C. Wang, L.-B. Shi et al., “Ai augmentation of radiologist performance in distinguishing covid-19 from pneumonia of other etiology on chest ct,” *Radiology*, p. 201491, 2020.
- [47] S. Tabik, A. Gómez-Ríos, J. Martín-Rodríguez, I. Sevillano-García, M. Rey-Area, D. Charte, E. Guirado, J. Suárez, J. Luengo, M. Valero-González et al., “Covidgr dataset and covid-sdnet methodology for predicting covid-19 based on chest x-ray images,” arXiv preprint arXiv:2006.01409, 2020.
- [48] P. Sharma, N. Ding, S. Goodman, and R. Soricut, “Conceptual captions: A cleaned, hypernymed, image alt-text dataset for automatic image captioning,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018, pp. 2556–2565.
- [49] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-end object detection with transformers,” arXiv preprint arXiv:2005.12872, 2020.
- [50] A. Kirillov, Y. Wu, K. He, and R. Girshick, “Pointrend: Image segmentation as rendering,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2020, pp. 9799–9808.
- [51] D. Zeng, R. Veldhuis, and L. Spreeuwers, “A survey of face recognition techniques under occlusion,” arXiv preprint arXiv:2006.11366, 2020.
- [52] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, “Deepfakes and beyond: A survey of face manipulation and fake detection,” arXiv preprint arXiv:2001.00179, 2020.
- [53] C. Geng, S.-j. Huang, and S. Chen, “Recent advances in open set recognition: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [54] R. Dabre, C. Chu, and A. Kunchukuttan, “A comprehensive survey of multilingual neural machine translation,” arXiv preprint arXiv:2001.01115, 2020.
- [55] S. Yang, Y. Wang, and X. Chu, “A survey of deep learning techniques for neural machine translation,” arXiv preprint arXiv:2002.07526, 2020.
- [56] G. Haskins, U. Kruger, and P. Yan, “Deep learning in medical image registration: a survey,” *Machine Vision and Applications*, vol. 31, no. 1, p. 8, 2020.
- [57] T. G. Debelee, F. Schwenker, A. Ibenthal, and D. Yohannes, “Survey of deep learning in breast cancer image analysis,” *Evolving Systems*, vol. 11, no. 1, pp. 143–163, 2020.
- [58] H. Khan, P. M. Shah, M. A. Shah, S. ul Islam, and J. J. Rodrigues, “Cascading handcrafted features and convolutional neural network for iot-enabled brain tumor segmentation,” *Computer Communications*, vol. 153, pp. 196–207, 2020.
- [59] P. M. Shah, A. Zeb, U. Shafi, S. F. A. Zaidi, and M. A. Shah, “Detection of parkinson disease in brain mri using convolutional neural network,” in *2018 24th International Conference on Automation and Computing (ICAC)*. IEEE, 2018, pp. 1–6.
- [60] L. Liu, W. Ouyang, X. Wang, P. W. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, “Deep learning for generic object detection: A survey,” *CoRR*, vol. abs/1809.02165, 2018. [Online]. Available: <http://arxiv.org/abs/1809.02165>
- [61] L. Datta, “A survey on activation functions and their relation with xavier and he normal initialization,” arXiv preprint arXiv:2004.06632, 2020.
- [62] N. Akhtar and U. Ragavendran, “Interpretation of intelligence in cnn-pooling processes: A methodological survey,” *Neural Computing and Applications*, pp. 1–20, 2020.
- [63] M. Lin, Q. Chen, and S. Yan, “Network in network,” arXiv preprint arXiv:1312.4400, 2013.
- [64] G. Maguolo and L. Nanni, “A critic evaluation of methods for covid-19 automatic detection from x-ray images,” arXiv preprint arXiv:2004.12823, 2020.
- [65] M. Nishio, S. Noguchi, H. Matsuo, and T. Murakami, “Automatic classification between covid-19 pneumonia, non-covid-19 pneumonia, and the healthy on chest x-ray image: combination of data augmentation methods in a small dataset,” arXiv preprint arXiv:2006.00730, 2020.
- [66] G. King and L. Zeng, “Logistic regression in rare events data,” *Political analysis*, vol. 9, no. 2, pp. 137–163, 2001.

- [67] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [68] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [69] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of machine learning research*, vol. 12, no. 7, 2011.
- [70] Z. Qin, F. Yu, C. Liu, and X. Chen, "How convolutional neural network see the world-a survey of convolutional neural network visualization methods," arXiv preprint arXiv:1804.11191, 2018.
- [71] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

...