



FocusCovid: automated COVID-19 detection using deep learning with chest X-ray images

Tarun Agrawal¹ · Prakash Choudhary¹

Received: 14 January 2021 / Accepted: 29 April 2021 / Published online: 9 May 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

COVID-19 is an acronym for coronavirus disease 2019. Initially, it was called 2019-nCoV, and later International Committee on Taxonomy of Viruses (ICTV) termed it SARS-CoV-2. On 30th January 2020, the World Health Organization (WHO) declared it a pandemic. With an increasing number of COVID-19 cases, the available medical infrastructure is essential to detect the suspected cases. Medical imaging techniques such as Computed Tomography (CT), chest radiography can play an important role in the early screening and detection of COVID-19 cases. It is important to identify and separate the cases to stop the further spread of the virus. Artificial Intelligence can play an important role in COVID-19 detection and decreases the workload on collapsing medical infrastructure. In this paper, a deep convolutional neural network-based architecture is proposed for the COVID-19 detection using chest radiographs. The dataset used to train and test the model is available on different public repositories. Despite having the high accuracy of the model, the decision on COVID-19 should be made in consultation with the trained medical clinician.

Keywords Convolutional neural network · Deep learning · COVID-19 classification · Chest X-ray · Medical imaging · FocusCovid

1 Introduction

Coronaviruses are responsible for respiratory infections lining up from the common cold to severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). MERS is a zoonotic virus that has spread in humans with direct or indirect contact with camels. The first human infection was reported in the Arabian Peninsula in 2012, and since then, 27 countries had been affected with a total of 2494 cases and a 34.4% fatality rate. SARS was first reported in Guangdong province of China in 2003 and spread to 28 countries with 8000 infected and 778 deaths. SARS can be transmitted from person to person with close contact. COVID-19 also belongs to the same family of coronaviruses. The first

human COVID-19 case surfaced in December 2019 in the Wuhan province of China. Unknown pneumonia clusters were reported from Wuhan that quickly spread to the whole province and the world. WHO declared it pandemic on 30th January 2020 after 7818 confirmed cases reported in 19 countries worldwide. Till December 2020, 86,950,284 persons are confirmed with COVID-19 including 1,878,504 deaths, reported to WHO.

SARS-coronavirus (SARS-CoV) uses human Angiotensin-Converting Enzyme 2 (ACE2) as its receptor. There is a similarity between the spike proteins of SARS-CoV and SARS coronavirus 2 (SARS-CoV-2), for this reason, it is assumed that SARS-CoV-2 also uses ACE2 as its receptor (Wan et al. 2020). COVID-19 is an infectious disease caused by SARS-CoV-2 (Huang et al. 2020) with average 3.4% fatality rate. Massive alveolar damage and progressive respiratory failure can result in death (Xu et al. 2020). The virus is highly contagious and can spread by human to human transmission with micron-size droplets from nose/mouth or in close contact. The rate of reproduction (R_0) is greater than 3 for COVID-19, which means each person on average can infect three other (Grech 2020). Common symptoms associated with COVID-19 are fever, dry cough,

✉ Prakash Choudhary
pc@nith.ac.in
Tarun Agrawal
tarun20octagarwal@gmail.com; tarunagrawal@nith.ac.in

¹ Department of Computer Science and Engineering,
National Institute of Technology Hamirpur, Hamirpur,
Himachal Pradesh 177005, India

sore throat, shortness of breath, pneumonia. The fatality rate is high with comorbidities such as heart and lung diseases. With the studies available, it is reported that people aged ≥ 65 with medical history are highly prone to infection than others. Fong et al. (2020) have given a detailed description of COVID-19 and its comparison with other pandemics.

Real-Time Reverse Transcription-Polymerase Chain Reaction (RT-PCR) and Rapid Anti-body Test (RAT) are commonly used tests for COVID-19 detection. In RAT, the blood samples are tested to detect the presence of antibodies. It is not a direct method to detect the virus but shows the response of the immune system. The immune system produces antibodies to counter the virus. An antibody can take 9–20 days to show up, so it is not an efficient method to test COVID-19. In RT-PCR, a throat swab is taken from the patient, and ribonucleic acid (RNA) is extracted. If it shares the same genetic sequence as SARS-CoV-2, then the patient is positive for COVID-19. RT-PCR test takes 4–6 h to detect the presence of virus and is expensive. With the sudden rise in cases, medical infrastructure has started collapsing worldwide. The medical community needs already available diagnosis techniques to detect the COVID-19. Radiography can be quite helpful because the symptoms of COVID-19 are similar to those of pneumonia. In COVID-19 cases, several lung abnormalities are identified in chest radiographs such as ground-glass opacity, lung consolidation, and others (Guan et al. 2020). Radiography is cheap and readily available so it can help in detecting COVID-19 in suspected cases. Lungs are the primary target of the virus. The presence of a virus brings changes in the lung field that can be visualized in a chest radiograph. A trained radiologist is needed to read the COVID-19 biomarkers and differentiate them from other pulmonary diseases. With limited trained radiologists available, a reliable and fully automatic system is needed to detect the cases. Author(s) have shown in their study that radiology can also be used as an alternative approach for detecting COVID-19 (Ai et al. 2020).

In the 1960s, radiographs were getting analyzed with digital computers, and after 2 decades, computer-aided detection (CAD) was the area of focus for the research community to assist radiologists (Giger et al. 2008). Computer vision and recently deep learning models have helped the radiologist for better insight into radiographs. Deep learning models are state-of-the-art models for image classification, and they have gained popularity after successfully classifying the 1000 classes in the ImageNet dataset (Deng et al. 2009). Deep learning has played an essential role in the medical field. Recently, the release of Chest X-ray 14 (CXR-14) (Wang et al. 2017) and Chexpert (Irvin et al. 2019) dataset, has accelerated the lung segmentation and disease classification work. Wang et al. (2017) introduced the 112,120 frontal chest radiographs of more than 30,000 patients for the first time. They labeled the dataset with natural language

processing (NLP) techniques and presented that the disease can be detected by weakly supervised learning. They labeled presence or absence of fourteen conditions but Long Short Term Memory (LSTM) network also labeled the same with a better area under the curve (AUC) (Yao et al. 2017). CheXNet (Rajpurkar et al. 2017) proposed the 121 layers deep network for classification on CXR-14 dataset. CheXNet classified and localized the disease better than the trained radiologists on average F1 metric. So deep learning, in particular, can be used to identify the disease automatically. It helps in extracting features automatically and then in the disease classification/detection.

The purpose of this paper is to provide a deep convolutional neural network-based model for automatic COVID-19 detection from chest radiographs with the help of a minimal dataset. Convolutional neural network (CNN) is the most popular deep learning approach with better results for classification or detection of diseases (LeCun et al. 2015). The model proposed in the paper is end-to-end CNN architecture with no handcraft feature extraction techniques. Overfitting, vanishing gradient, and degradation are some of the many technical problems to address while training the deep learning model for COVID-19 detection. The COVID-19 radiography dataset publicly available is small and to train deep CNN models, a large dataset is required. Training deep learning models with the small dataset can lead to overfitting even after applying the overfitting prevention methods. The main challenge in designing the architecture is to keep the trainable parameters minimum to avoid overfitting. Further, the data augmentation techniques can be applied to deal with small dataset issues. The vanishing gradient problem is also an issue to address while training the deep learning models. In the worst case, it can stop the model from further training. Additionally, degradation of accuracy while training the deeper network is also an issue to address.

The main contributions of the paper are summarized below:

- An efficient deep CNN model is proposed for the COVID-19 detection using chest radiographs. The proposed model uses the residual learning and squeeze-excitation network to improve the performance of the model.
- Proposed model is validated on two datasets for the different parametric values. A weighted F1-score is used for the evaluation of the proposed model on the imbalanced test dataset. The achieved results by the model show that it detects the COVID-19 accurately.
- To the author's best knowledge, this is the only study that has used the largest COVID-19 X-ray images.

The rest of the paper is organised as follows: Sect. 2 presents the research work published for the classification of pulmonary diseases including COVID-19. Section 3

discusses the datasets and the proposed architecture. Performance metrics, experimentation details, and obtained results are presented in Sect. 4 and finally, in Sect. 5, discussion and in Sect. 6, the conclusion of the research is presented.

2 Related work

Before the pandemic, researchers had already presented the models for automatic detection of pulmonary disease using the deep learning techniques with chest radiographs. Deep learning has established itself as state-of-the-art technology to provide solutions across many fields, especially in computer vision. Radiography is the most common diagnostic tool used by physicians to examine, detect and monitor pulmonary conditions such as tuberculosis (TB), consolidation, lung nodule, emphysema, pneumonia, etc (Candemir and Antani 2019). Numerous research work on the classification of pulmonary diseases such as TB, lung nodule, etc., using deep learning are available. Recently, deep learning researchers are focusing on automatic diagnosis systems for COVID-19 detection. These diagnosis systems are aimed to support the existing medical infrastructure and reduce the burden on the medical staff. In the following sub-sections, research work conducted on the classification of pulmonary diseases such as TB, lung nodule using deep learning is briefly discussed along with COVID-19.

2.1 Pulmonary disease classification

Chen et al. (2011), deployed the CNN model for detecting the lung nodule from chest radiographs. They used the canny edge detector for removing the rib crossing. Then used the SVM classifier with Gaussian kernel for classifying the nodule with better sensitivity and reduced false-positive (FP). Chen and Suzuki (2013) detected the nodule with massive training artificial neural network (MTANN), and virtual dual-energy (VDE) for rib and bone suppression in radiographs. Subsequently, by suppressing ribs and bone, they reported an increase in sensitivity for detecting the lung nodule. DeepCNets (Bobadilla and Pedrini 2016) classified the lung nodule using CNN. It detects the nodule directly from the pixels instead of extracting features. Data augmentation and ten-fold cross-validation were applied to validate the model. Lakhani and Sundaram (2017) used pre-trained and untrained CNN models like GoogleNet (Szegedy et al. 2015) and AlexNet (Krizhevsky et al. 2012) for the classification of pulmonary tuberculosis. For experimenting, they used the four publicly available tuberculosis datasets. The proposed method accurately classified the disease with 0.99 AUC.

2.2 COVID-19 classification

Apostolopoulos and Mpesiana (2020) have used pre-trained VGG-19 and MobileNet-V2 for detecting the COVID-19 from chest radiographs along with pneumonia. They have used two datasets for the training and testing of the models. First dataset contains the 1428 radiographs (224 COVID-19, 504 Healthy, 700 Pneumonia) and the other 1442 radiographs (224 COVID-19, 504 Healthy, 714 Pneumonia). VGG-19 reported 98.75% and 93.48% while MobileNet-V2 reported 97.40% and 92.85% binary and multi-class accuracy. In their another study (Apostolopoulos et al. 2020), they have evaluated three learning strategies: training from scratch, transfer learning, and fine-tuning on the MobileNet-V2 model. They used the 3905 chest images containing seven classes along with COVID-19 and achieved 87.66% seven classes and 99.18% binary classification accuracy when the training from the scratch strategy was followed. Vaid et al. (2020) proposed the deep learning method to detect the COVID-19 using fine-tuned VGG-19. VGG-19 worked as feature extraction while a classifier with three fully connected layers and softmax function for predicting the labels. Total 364 chest radiographs are used for the training, validation, and testing of the model.

Ozturk et al. (2020) have tailored designed CNN model, DarkCovidNet for the automatic detection of COVID-19 using chest radiographs. Instead of designing the model from scratch, the authors have used DarkNet-19 (Redmon and Farhadi 2017) model design as the starting point. The proposed model consists 1,164,434 parameters. Total 1127 chest images containing 127 COVID-19, 500 normal, and pneumonia each are used for the training and testing of the model. The model achieved 98.08% and 87.02% average binary and multi-class accuracy. A deep learning model for the detection of COVID-19 using the chest radiographs is proposed by Panwar et al. (2020). They proposed 24 layer nCOVNet consisting of 18 layers of pre-trained VGG-19 and the rest are part of the classifier. They have used a total of 284 chest images (142 COVID-19 and Pneumonia each) for the training and testing of the model. They used the random sampling for creating 70% training and 30% testing dataset and achieved 88.10% accuracy for the binary classification.

Toraman et al. (2020) presented an artificial neural network based on Capsule Network (Sabour et al. 2017) using the chest radiographs for COVID-19 detection. The proposed model, CapsNet used the 2331 images consisting of 231 COVID-19 and 1050 normal and pneumonia each. The 11 layer architecture achieves 97.24% and 84.22% accuracy for binary and multi-class classification. A tailored deep model for the COVID-19 detection, COVID-Net is proposed by Wang et al. (2020). They combined the dataset from five different repositories to create a COVIDx dataset. It contains 358 COVID-19 radiographs along with normal and

pneumonia. The proposed deep learning model consists of 11.75 Million parameters. It achieves an accuracy of 93.3% on the test dataset.

Apart from chest radiographs, many COVID-19 classification studies are performed using computed tomography (CT) scans. Ahuja et al. (2020) have used ResNet18, ResNet50, ResNet101 and SqueezeNet with transfer learning and data augmentation for the COVID-19 detection. ResNet18 achieved the highest accuracy, precision, F1-Score in comparison with other models. The dataset used in the experiment consists of 349 COVID-19 and 397 normal CT-Scans. Konar et al. (2020) proposed a semi-supervised model for the diagnosis of COVID-19 using the CT-Scan. The model achieved an accuracy of 93.1% for binary classification. The study included two datasets, the first containing 1252 COVID-19 and 1230 Non-COVID-19 CT scans while the second consists of only 20 COVID-19 CT scans.

2.3 Limitations

Various models have been proposed by the deep learning community for COVID-19 detection. Mainly, transfer learning and training from the scratch strategy are applied for COVID-19 detection. Most of the proposed models have used pre-trained VGG-19, ResNet, and MobileNet models. These models have large trainable parameters, and it results in a large computational cost. COVID dataset used in the above studies is taken from multiple public repositories but the instances available are very few. Almost every study has highlighted the limitation of the dataset. Limited dataset with large trainable parameters can result in overfitting. A brief summarize related work on COVID-19 detection is presented in Table 1.

3 Material and method

This section presents the dataset and proposed architecture of the FocusCovid. Section 3.1 describes the chest radiograph dataset and the pre-processing applied to it. Problem formalization and detailed description of the proposed architecture is presented in Sect. 3.2.

3.1 Dataset and pre-processing

Sample distribution in the database impacts the developing model. It is important to have an equal number of samples that cover all classes to develop an efficient model. In order to deal with the class imbalance, oversampling techniques such as SMOTE (Chawla et al. 2002) can be used. In oversampling, minority class instances are generated to balance the distribution between different classes. It helps the model not to be biased towards the majority class. Collecting and annotating medical image datasets is extremely difficult. There are very few large medical image dataset available because of privacy issues. Few large datasets of chest radiographs are available for many pulmonary disease classification but the COVID-19 related datasets are small. Datasets used in the training and testing of the proposed model are publicly available.

3.1.1 Dataset collection

For collecting the dataset, we searched the Github, Kaggle repository, other sources such as the Italian Society of Medical and Interventional Radiology (SIRM), Radiological Society of North America (RSNA), and at

Table 1 Brief summary on related work for automatic COVID-19 detection

Model	Dataset	Instances	Classification
MobileNet-V2 (Apostolopoulos et al. 2020)	X-rays	Total: 3905— COVID-19: 455 ; Rest: 3450	Binary and Seven-class
VGG-19 and MobileNet-V2 (Apostolopoulos and Mpesiana 2020)	X-rays	Dataset 1: Total: 1428— COVID-19: 224 ; Normal: 504; Pneumonia: 700 and Dataset 2— Total: 1442; COVID-19: 224 ; Normal: 504; Pneumonia: 714	Binary and Three-class
VGG-19 (Vaid et al. 2020)	X-rays	Total: 545— COVID-19: 181 ; Normal: 364	Binary
DarkCovidNet (Ozturk et al. 2020)	X-rays	Total: 1127— COVID-19: 127 ; Normal: 500; Pneumonia: 500	Binary and Three-class
nCOVNet (Panwar et al. 2020)	X-rays	Total:284— COVID-19: 142 ; Normal: 142	Binary
Capsnet (Toraman et al. 2020)	X-rays	Total: 2331— COVID-19: 231 ; Normal: 1050; Pneumonia: 1050	Binary and Three-class
COVID-Net (Wang et al. 2020)	X-rays	Total: 13,975; COVID-19: 358 ; Normal and Pneumonia: 13617	Three-class
ResNet-18 (Ahuja et al. 2020)	CT-Scan	Total: 746; COVID-19: 349 ; Normal: 397	Binary
Semi-supervised model (Konar et al. 2020)	CT-Scan	Dataset 1: Total: 2482— COVID-19: 1252 ; Non-COVID: 1230 and Dataset 2: COVID-19: 20	Binary

last Kaggle-1 (Chowdhury et al. 2020), Kaggle-2 (Asraf 2020) are selected. Kaggle-1 dataset is used by Chowdhury et al. (2020) in their study and is the most popular dataset for COVID-19 detection studies. This dataset has 13,282 downloads and is the winner of the COVID-19 Dataset Award by the Kaggle Community. It is collected by researchers from Qatar University, the University of Dhaka along medical doctors from Malaysia. It consists of 1143 COVID-19 samples and 1345 samples each of normal and pneumonia radiographs with 1024×1024 , and 256×256 -pixel resolution. The sample distribution in the dataset is even across all the classes. This dataset is itself collected from the different public repositories such as:

- **COVID-19 Dataset:** SIRM (2020), Cohen (Cohen et al. 2020), Github (2020a, b), Euro (2020)
- **Normal and Pneumonia Dataset:** Mooney (2020)

Mooney [43] dataset is another popular dataset among deep learning researchers consisting of normal and pneumonia chest radiographs. It consists 3883 pneumonia and 1,349 normal chest radiographs from 5856 patients. Kaggle-2 (Asraf 2020) dataset is freely available on the Kaggle website. It consists of 1525 samples of COVID-19, normal, and pneumonia. It is used by Gianchandani et al. (2020) for the binary classification.

3.1.2 Pre-processing on dataset

For better performance, the dataset is resized to $224 \times 224 \times 3$ (RGB). Data normalization is performed on the collected dataset by dividing the images by the number of channels (255) to ensure normalization in the range of $[0-1]$. The collected dataset is small for training deep learning models. Data augmentation can be applied as one method to deal effectively with the small dataset (Han et al. 2018; Perez and Wang 2017). It is often applied to deal with the overfitting problem and to improve the model generalization for better results. In this study also, augmentation techniques such as rotation, flipping, shear transformation, and zooming are used on the training data. Augmentation technique is not used on the test data to avoid the overfitting problem (Nour et al. 2020). Rotation between $(0^\circ, 20^\circ)$, horizontal flip, zooming, and shear range (0.2) is applied to the dataset. ImageDataGenerator function provided by Keras is used for data augmentation. The augmentation strategy used in the proposed method provides the real-time data augmentation to fit the model, not increase the size of training datasets as proposed in Nour et al. (2020); Ahuja et al. (2020); Toraman et al. (2020). Figure 1 shows the instance of chest radiographs of the collected dataset.

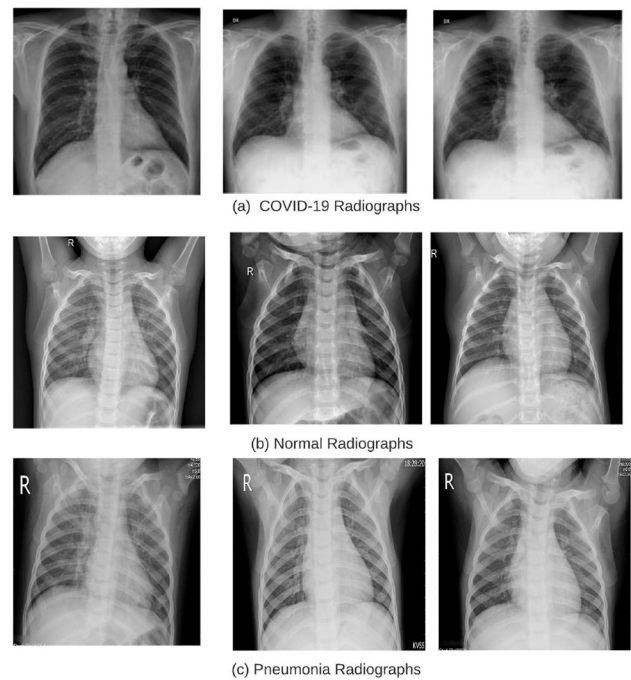


Fig. 1 Chest radiographs of COVID-19, Normal and Pneumonia

3.2 Proposed architecture of FocusCovid

Improving behavior with experience is termed learning, and it also fits for deep learning (DL). DL is a sub-category of artificial intelligence and has improved the performance of machine learning projects. GPU-based computing power and non-linearity allow the deep architecture with hidden layers to perform better than artificial neural networks (Glorot and Bengio 2010). Research studies have shown that deep learning is popular in the medical field (Li et al. 2014). A simple neural network cannot learn complex features, unlike deep learning architecture. Deep CNN extracts the local features from high layer inputs and transfers them to the lower layers for the complex pattern analysis (Islam et al. 2020). CNN has shown remarkable capability for medical image analysis tasks such as disease classification and organ segmentation (Litjens et al. 2017). Research work presented in section 2 has successfully demonstrated the capability of deep CNN for COVID-19 detection. Inspired by this, we have proposed the FocusCovid for the COVID-19 detection. In the following, first the problem formalization and then the proposed architecture is described.

3.2.1 Problem formalization

In this paper, the supervised learning technique is used for the COVID-19 detection. Let consider, the dataset has S_i training instances, $M = \{X, Y\}$, where X represents the input radiograph and Y represents the true labels. We can represent

the training samples as, $X = \{x_1, x_2, \dots, x_{S_i}\}$, and the true label associated with each sample as $Y = \{y_1, y_2, \dots, y_{S_i}\}$. The $y_i \in [1, 2, 3]$, where 1, 2 and 3 represents the COVID-19, normal and pneumonia classes, respectively. The output predicted by the classifier, $f^w : X \rightarrow Z$ parameterized by a weight w , can differ from the true label Y . This difference in the predicted label (Z) and the true label (Y) can be termed as prediction error rate. The main aim is to keep error rate minimum by finding suitable parameters w during the training process.

3.2.2 Details of FocusCovid architecture

Encoder–decoder architecture is required for image segmentation. The encoder is a CNN architecture that extracts the features and transfers them to the decoder for segmentation. The decoder uses those feature maps therefore, the better the encoder is, the better will be the segmentation results. So, instead of designing the proposed CNN architecture from scratch, the FocusNet (Kaul et al. 2019) was used as the initial point. FocusNet is the U-Net-based encoder–decoder architecture that was proposed for medical image segmentation. This architecture has successfully demonstrated the segmentation capabilities on different medical datasets. It has two branches of encoder–decoder, so instead of using the encoders of both branches, we have modified the second branch according to needs for the COVID-19 classification. In the FocusNet (Kaul et al. 2019), there are three blocks of residual and strided residual layers with two Squeeze-Excitation (SE) layer (Hu et al. 2018) in between. Having the two SE layers is the architectural requirement of FocusNet but it is not such for the FocusCovid. So, the second SE layer is removed from all the blocks. Additionally, we have increased the depth of the encoder by adding the fourth residual-strided residual block with a single SE layer. Further, the number of filters is reduced to have less trainable parameters (16,32,64,128,256). The total number of parameters in the architecture is 2,940,122. To the author's best knowledge, no study has proposed a similar architecture for the COVID-19 detection in chest radiographs. Figure 3 shows the block diagram of the proposed architecture.

Generally, increasing the depth of architecture might increase the accuracy but adding more layers can also lead to higher training error (He et al. 2016). With adding more layers, accuracy might increase but as reported in He and Sun (2015); Srivastava et al. (2015) with increasing depth, accuracy gets saturated and then degrades rapidly. This problem is termed degradation. Vanishing gradient is another problem to address while training the deep learning model, therefore, (He et al. 2016) introduces the residual learning concept to address this issue and the degradation problem. Further, in their other

paper (He et al. 2016), introduced the concept of identity mapping to improve the generalization and make training easier. In the FocusCovid, there are four blocks of residual and strided residual layers. Increasing the depth of FocusCovid with such blocks will have two disadvantages. First, it will increase the number of trainable parameters that can lead to overfitting, and second, might have resulted in a degradation problem, as discussed above. Other than residual layers, the proposed model consists one Global Average Pooling (GAP) layer, two dropout, and three fully connected layers. In the architecture, strides are used for the downsampling instead of average/max pooling.

In the following, different blocks and layers used in the architecture are discussed.

- **Initial block:** The first block of the architecture is Initial block. It takes the input in $224 \times 224 \times 3$ resolution. It consists of the convolutional layer, Batch Normalization (BN) layer, and the activation layer. The number of filters used in the convolution layer is 16 with 3×3 sizes. The convolutional layer is the basic layer found in all CNN architectures. It consists of filters whose parameters are updates (learned) during the model training. These filters are applied to the dataset to capture the low and high-level features. Filters convolve with input to generate the activation maps and the output is obtained by stacking all activation maps along depth dimension. The convolution process is defined in Eq. (1)

$$Y_j^l = p \left(\sum_{i \in N_j} Y_i^{l-1} * f_{ij}^l + b_j^l \right) \quad (1)$$

Y_j^l and Y_j^{l-1} represents the current and previous convolutional layer, f_{ij}^l represents the filter or kernel, b_j^l shows the bias term and N_j matches the input map. BN (Ioffe and Szegedy 2015) technique is used for the training of the deep neural network. It performs the normalization on each mini-batch during training. It reduces the internal co-variate shift and accelerates the training process. The BN layer is followed by the activation layer. It gives the non-linear feature to the model. Many activation layers are proposed by the researchers but mainly the CNN is the combination of any of Sigmoid, ReLu, LeakyReLu, and Softmax layer.

- **Residual learning block:** Proposed architecture consists of four residual and strided residual learning blocks for the feature extraction. Block diagrams of residual and the strided residual learning are shown in Fig. 2a. We have introduced the residual mapping to address the degradation and vanishing gradient problem during the training of deep learning networks. Shortcut connection

Fig. 2 Building block **a** residual identical mapping, **b** SE block

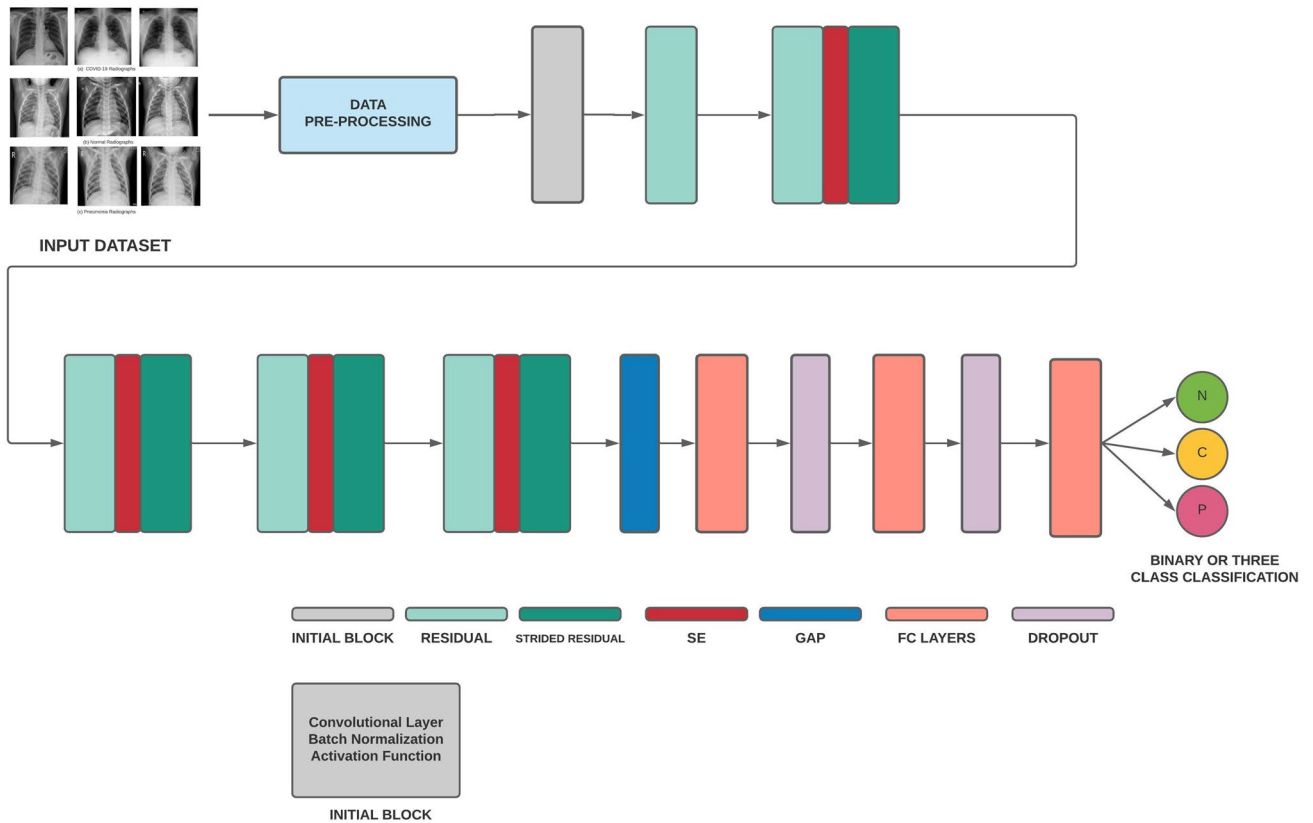
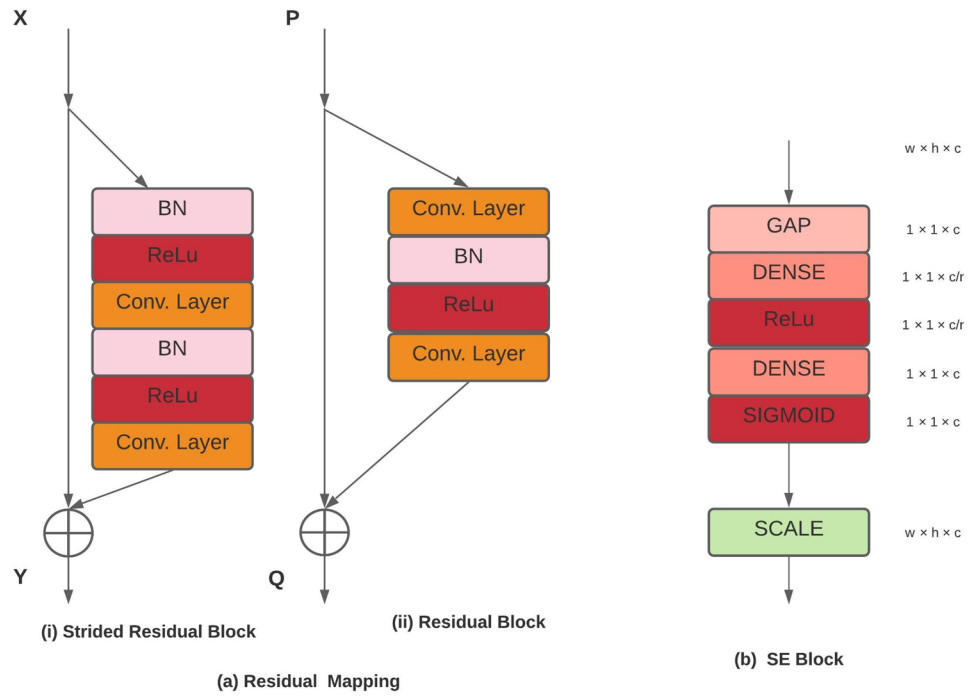


Fig. 3 Block diagram of proposed architecture

introduced in ResNet (He et al. 2016) performs the identity mapping and their output is added with stacked layers output, as shown in Fig. 2a(ii). In the strided residual block, pre-activation identity residual mapping is used. The advantage introduced to the architecture by this is twofold. First, network optimization is eased, and second, the use of the BN layer as pre-activation improves the regularization of the architecture (He et al. 2016). Generally, there is a tradition of using the max/avg pooling layers for the downsampling but in the proposed architecture, increased strides are used. As studied by Springenberg et al. (2014), max-pooling can be replaced by the strided convolutional layers. In the proposed architecture, we have used the stride = (2,2) for the downsampling. In each residual and strided residual block, there are two convolutional layers with 3×3 filter, batch normalization layer and an activation layer (ReLU). Both block layers are similar in the number of filters (32,64,128,256) except for the first residual layer where 16 filters are used. There are total 18 convolutional layers in the architecture. Kernel is initialized with ‘he normal’ and regularized with $L2(1e-4)$.

- **Squeeze-Excitation (SE) block:** Hu et al. (2018) in their paper studied the relationship between channels and proposed the SE network that performs dynamic channel-wise feature re-calibration. It helps the network to selectively emphasize informative features by learning the use of global information. SE blocks in earlier layers excite informative features, strengthening the representation of lower-level features. While at later layers, it responds in a specialized manner to different inputs. In the proposed architecture, SE blocks are used at all levels to be benefited from the feature re-calibration across the whole network. Figure 2b shows the SE block.
- **Other layers:** Global Average Pooling (GAP) (Lin et al. 2013), dropout (Srivastava et al. 2014) and fully connected layers (FC) are used after last strided residual block. GAP (Lin et al. 2013) can be used in two forms. In the first form, GAP replaces the fully connected (FC) layers completely and in another form, it feeds its output to one or more FC layers. A fully connected layer is added after the GAP layer. It has full connections among neurons. These layers are generally located at the end of CNN. Input applied to these layers is multiplied with the FC weights to produce results. In the proposed architecture, the FC layer has 128 and 64 neurons with ReLU activation function. While dropout (Srivastava et al. 2014) adds the regularization to the CNN by randomly dropping the neurons at hidden layers. Dropped neurons have no role in forwarding or backward pass during the training. During each forward pass, the architecture is different despite sharing weight. This also helps in avoiding overfitting. In the proposed architecture, it is charac-

terized by a 20% dropout rate. At last, for generating the output at the end, a dense layer with 2/3 neurons and a Softmax classifier is used.

4 Performance evaluation

The different evaluation metric used for the evaluation of the proposed model are described in Sect. 4.1. In Sect. 4.2, the experimentation details and results obtained for the binary and three class classifications are described.

4.1 Evaluation metrics

The confusion matrix is often used for checking the performance of the classifier on test data. It is a tabular table with actual and predicted instances of the represented classes in test data. It can be used to find out other parameters such as F1-score, specificity, sensitivity, precision, and classification accuracy. For calculating these parameters from the confusion matrix, few terms are needed to be mentioned, such as

- (a) True Positive (TP) for correctly predicting the positive class
- (b) True Negative (TN) for correctly predicting negative class.
- (c) False Negative (FN) for incorrectly predicting the positive class.
- (d) False Positive (FP) for incorrectly predicting the negative class.

As the images for the test set are imbalanced, i.e. different testing classes have different image instances. So to evaluate the model, we have used the area under curve (AUC) score for the binary classification and weighted F1 score for three class classification. For the three-class classification, precision (P) and recall (R) is calculated for all classes separately based on one-vs.-rest and then the average of P and R is taken before calculating the F1 score. F1 score is the weighted harmonic mean of P and R, and it helps in model evaluation on the imbalance dataset (Bhagat et al. 2021).

The equations mentioned below calculates the above-mentioned parameters:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (2)$$

$$Sensitivity = Recall = R = TP/(TP + FN) \quad (3)$$

$$Precision = P = TP/(TP + FP) \quad (4)$$

$$F1 - Score = F1 = \frac{2 \times P \times R}{P + R} \quad (5)$$

4.2 Results

For training and testing the models, Google Colaboratory referred to as Colab is used. It provides a 12GB NVIDIA Tesla K80 GPU for use up to 12 h. In this work, we have used Adam optimizer with initial learning rate as 0.0005. A dynamic learning approach has been used for model training. ReduceLROnPlateau method is used to reduce the learning rate when the model stops improving. Factor = 0.5, patience = 3, and the min learning rate is fixed to $1e-6$. EarlyStopping method is also used with patience = 10 to monitor the validation loss. If validation loss is not improved for continuous ten epochs, the model will stop the training. Epochs are set to 65 and the batch size to 16. Models are trained for approximately 30 min and 40 min for both binary and three class classification, respectively. Categorical crossentropy is used as a loss function.

Cross-validation method is used to estimate the general effectiveness of the models. In K-fold cross-validation, the dataset is split into K mutually exclusive subsets. The model runs K times, each time taking (K–1) set for training and the rest for validating/testing the model. In this paper, we have performed the 5-fold cross-validation to evaluate the model. The dataset is divided into 5 sets, 4 sets were used for the training and remaining is used for the testing. This process is carried out for all the sets and results are recorded. The average is calculated to evaluate the performance of the model. Further, the Kaggle-2 dataset is used to validate the proposed model.

The most common findings of COVID-19 are ground-glass opacity (GGO), thickening of the adjacent pleura, air space consolidation, bronchovascular thickening (Shi et al. 2020), while of pneumonia is GGO, vascular thickening, bronchial wall thickening (Bai et al. 2020). There are some other findings in COVID-19 but rare such as multiple tiny pulmonary nodules, pneumothorax, smoother interlobular septal thickening with pleural effusion, some of the findings are similar in pneumonia and COVID-19 (Kanne et al. 2020). Keeping this in mind, two approaches are used in the study. In the first approach, binary classification is conducted. Only two classes, normal and COVID-19 are used for the classification, and in the second, three class classifications (Normal, COVID-19, pneumonia) are conducted. COVID-19 can be detected lonesomely with normal radiographs but since it also has some findings similar to pneumonia, it is necessary to include it in classification.

4.2.1 Binary classification

The result for binary classification is presented in this subsection. 2484 chest radiographs of normal and COVID-19 are used for the training and testing of the model. Precision,

Table 2 Results for COVID-19 and normal class on Kaggle-1 dataset

Class	Fold	Precision	Sensitivity	F1 score	AUC
COVID-19	1	0.99	0.98	0.98	0.99
–	2	1.00	1.00	1.00	1.00
–	3	1.00	0.99	0.99	0.99
–	4	0.98	0.98	0.98	0.98
–	5	1.00	1.00	1.00	1.00
–	Average	0.994	0.990	0.990	0.992
Normal	1	0.98	0.99	0.99	0.99
–	2	1.00	1.00	1.00	1.00
–	3	0.99	1.00	0.99	0.99
–	4	0.98	0.98	0.98	0.98
–	5	1.00	1.00	1.00	1.00
–	Average	0.990	0.994	0.992	0.992

Table 3 Results of binary classification between COVID-19 and normal classes on Kaggle-1 dataset

Fold	Precision	Sensitivity	F1 score	Accuracy
1	0.99	0.99	0.99	0.99
2	1.00	1.00	1.00	1.00
3	0.99	0.99	0.99	0.99
4	0.98	0.98	0.98	0.98
5	1.00	1.00	1.00	1.00
Average	0.992	0.992	0.992	0.992

sensitivity, F1 score, and AUC are calculated for each fold for both classes (Table 2), and the overall average of each metric for both classes, and total accuracy is reported in Table 3.

For the COVID-19 class, lower performance values are reported at fold 4. Minimum precision, sensitivity, and F1-score reported in fold 4 is 0.98 while the higher values are reported in fold 2, fold 3 and fold 5. Precision, sensitivity, and F1-score values are 1.00 for each metric, respectively in fold 2, fold 3 and fold 5. The average precision, sensitivity, and F1-score of COVID-19 are 0.994, 0.990, and 0.990, respectively. For the normal class, a lower precision value of 0.98 is reported in fold 1 and fold 4, and a higher value of 1.00 in fold 2 and fold 5. A lower sensitivity value of 0.98 in fold 4 and a higher value of 1.00 is reported in fold 2, fold 3 and fold 5. F1-score value of 1.00 is in fold 1, fold 2 and fold 5, while 0.98 in fold 4. Precision, sensitivity, F1-score, and accuracy for both the classes are presented in Table 3. The overall average accuracy of both classes is 0.992. Further, value 0.992 is reported for precision, sensitivity, and F1-score, respectively. Further, to measure the class imbalance, AUC is also reported in Table 2

Table 4 Results for COVID-19, normal and pneumonia classification on Kaggle-1 dataset

Class	Fold	Precision	Sensitivity	F1 score
COVID-19	1	0.99	0.97	0.98
	2	0.99	1.00	0.99
	3	0.99	1.00	0.99
	4	0.98	0.97	0.98
	5	0.99	0.99	0.99
	Average	0.988	0.986	0.986
Normal	1	0.91	0.98	0.95
	2	0.90	0.97	0.93
	3	0.89	0.98	0.93
	4	0.93	0.99	0.96
	5	0.90	0.99	0.94
	Average	0.906	0.982	0.942
Pneumonia	1	0.97	0.91	0.94
	2	0.97	0.90	0.93
	3	0.98	0.87	0.92
	4	0.98	0.92	0.95
	5	0.99	0.89	0.94
	Average	0.978	0.898	0.936

Table 5 Results of three class classification between COVID-19, normal and pneumonia class classes on Kaggle-1 dataset

Fold	Precision	Sensitivity	F1 score	Accuracy
1	0.96	0.95	0.95	0.95
2	0.95	0.95	0.95	0.95
3	0.95	0.95	0.95	0.95
4	0.96	0.96	0.96	0.96
5	0.96	0.95	0.95	0.95
Average	0.956	0.952	0.952	0.952

Table 6 Results of binary classification between COVID-19 and normal class on Kaggle-2 dataset

Fold	Precision	Sensitivity	F1 score	Accuracy	AUC
1	0.97	0.97	0.97	0.97	0.97
2	0.96	0.96	0.96	0.96	0.96
3	0.96	0.96	0.96	0.96	0.96
4	0.95	0.95	0.95	0.95	0.95
5	0.96	0.96	0.96	0.96	0.96
Average	0.960	0.960	0.960	0.960	0.960

Another experiment is done on the Kaggle-2 dataset to validate the results. This dataset is used by Gianchandani et al. (2020) for the binary classification. Results are reported in Table 6. There is no major difference in the results that shows the efficiency of the proposed model.

Table 7 Results of three class classification between COVID-19, normal and pneumonia class on Kaggle-2 dataset

Fold	Precision	Sensitivity	F1 score	Accuracy
1	0.95	0.95	0.95	0.95
2	0.96	0.95	0.95	0.95
3	0.95	0.94	0.94	0.94
4	0.94	0.94	0.94	0.94
5	0.94	0.94	0.94	0.94
Average	0.948	0.944	0.944	0.944

Confusion matrix and loss graph for both datasets are presented in Figure 4 and Figure 5.

4.2.2 Three class classification

In this section, three class classification results are presented. Total 3829 radiographs for all three classes are used in the experiment. Results obtained for the different metrics are reported in Table 4 and Table 5. In Table 4, precision, sensitivity, and F1-score of COVID-19, normal, and pneumonia are presented for all 5 folds, and at last, the average is reported. Table 5 presents the weighted average of precision, sensitivity, and F1-score of all three classes. Weighted F1-score enables us to evaluate the model on an imbalanced dataset.

A higher precision value of 0.99 is reported in fold 1, fold 2, fold 3 and fold 5 for the COVID-19. While the lower precision value of 0.98 in fold 4. The average precision, sensitivity and F1-score value of COVID-19 reported are 0.988, 0.986, and 0.986, respectively. For the normal class, a lower precision value of 0.89 is reported in fold 3 and higher 0.93 in fold 4. The lower sensitivity value of 0.97 in fold 2, while the higher value of 0.99 in fold 4 and fold 5. The average precision, sensitivity, and F1-score for the normal class are 0.906, 0.982, and 0.942, respectively. For the pneumonia class, a lower precision value of 0.97 at fold 1 and fold 2, and a higher value of 0.99 at fold 5. The lower sensitivity value of 0.87 is reported at fold 3, while a higher value of 0.92 at fold 4. The average precision, sensitivity, and F1-score for the pneumonia class are 0.978, 0.898, and 0.936, respectively.

The test dataset is imbalanced for the three-class classification, so to evaluate the performance of the model, we have used the weighted F1-score. To calculate the F1-score, the precision and sensitivity for each class are calculated based on one-vs.-rest. Table 5 shows the weighted F1-score for the imbalanced test classes. The weighted F1-score value of 0.95 is reported at fold 1, fold 2, fold 3 and fold 5 and 0.96 at fold 4.

As done for binary classification, results for three-class classification are also validated on the Kaggle-2 dataset.

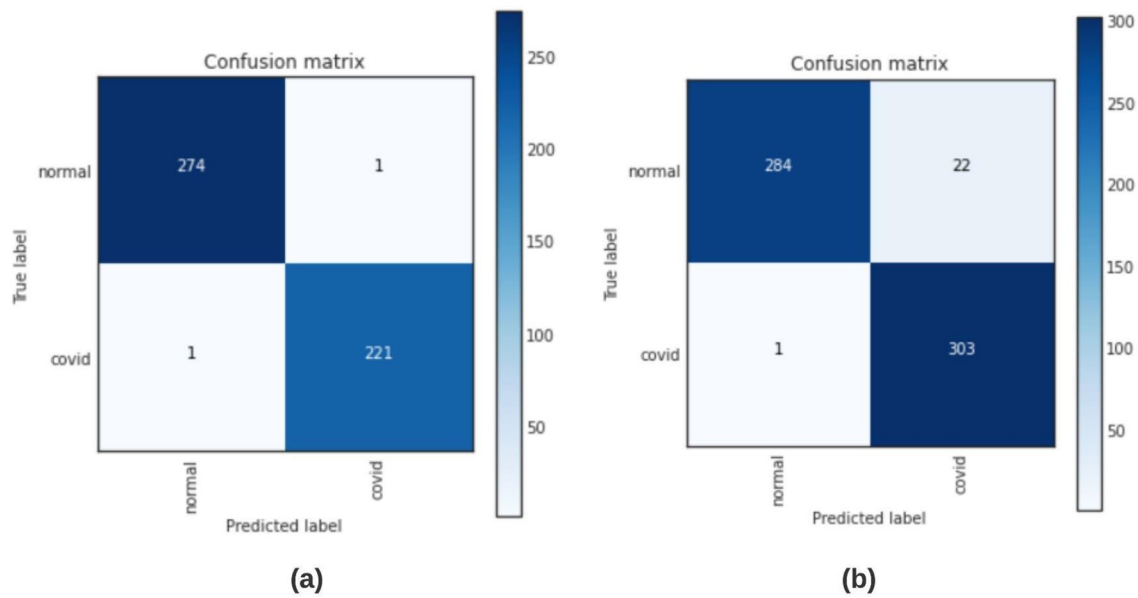


Fig. 4 Confusion matrix for the binary classification on **a)** Kaggle-1 and **b)** Kaggle-2 dataset

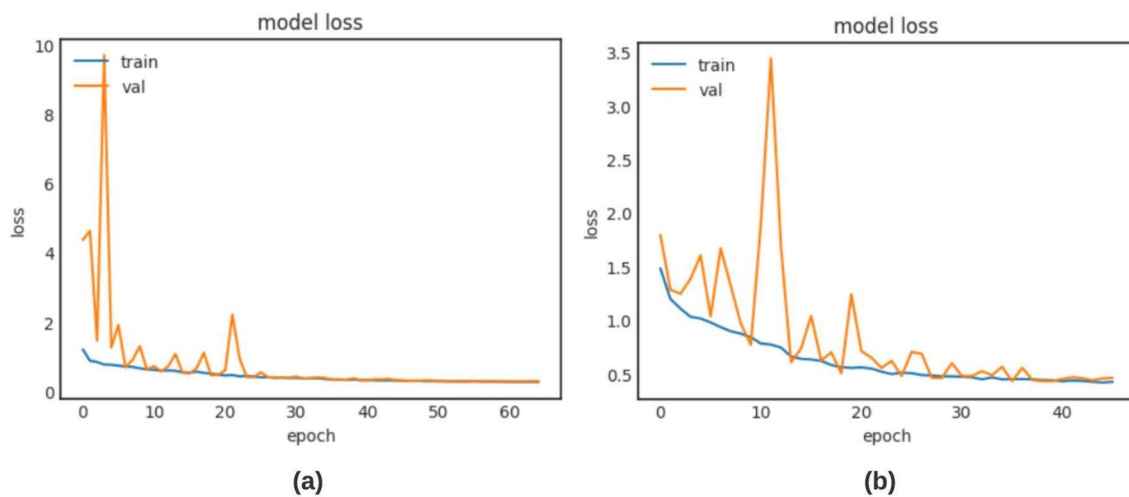


Fig. 5 Loss graph for the binary classification on **a)** Kaggle-1 and **b)** Kaggle-2 dataset

Results are reported in Table 7. Obtained results for three-class classification on the Kaggle-2 dataset validate the efficiency of the proposed model. Confusion matrix and loss graph for both datasets are presented in Figure 6 and Figure 7.

5 Discussion

The proposed model is compared with other states of the work and a brief discussion is presented in this section. This comparison is limited as the one-to-one comparison

is not possible due to differences in sample size, simulation environment, hardware, model parameters, and different methodologies. Table 8 presents all the studies included for the comparison of the proposed model with other state-of-the-art models.

COVID detection is the trending topic in these days. Many models are proposed for it such as Apostolopoulos et al. (2020), Apostolopoulos and Mpesiana (2020), Vaid et al. (2020), Panwar et al. (2020), Toraman et al. (2020), Wang et al. (2020), Ozturk et al. (2020). Table 8 provides the details for comparison of proposed model with others for binary and three class classification for different

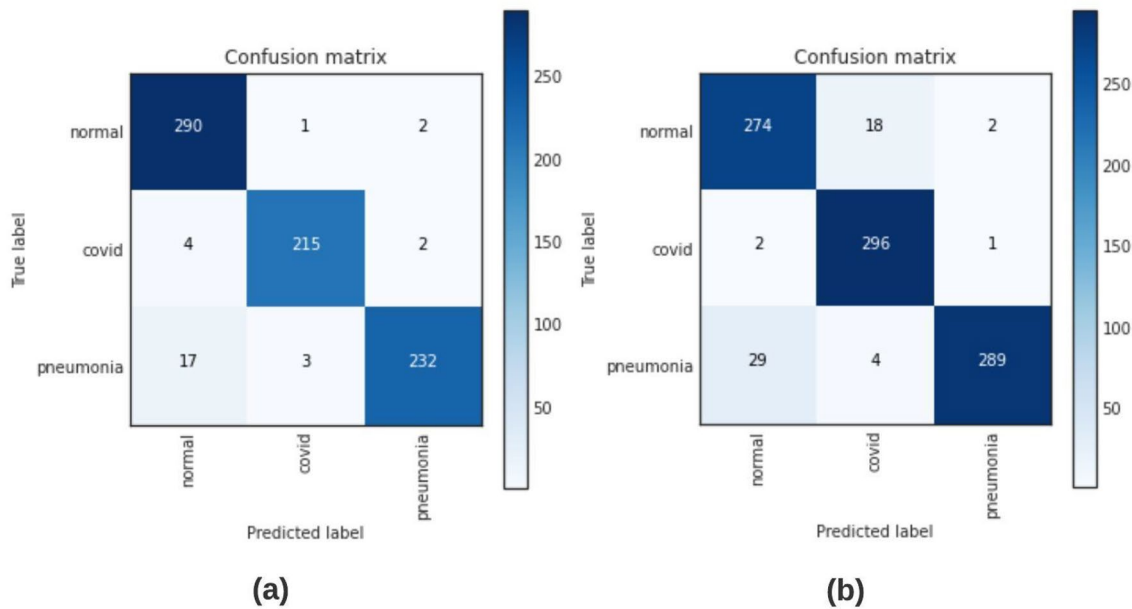


Fig. 6 Confusion matrix for the three class classification on **a)** Kaggle-1 and **b)** Kaggle-2 dataset

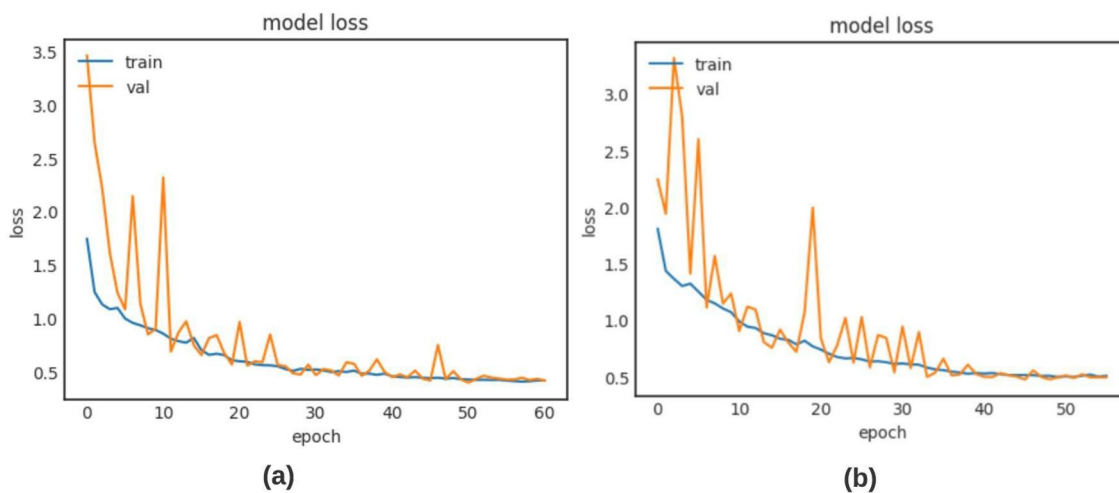


Fig. 7 Loss graph for the three class classification on **a)** Kaggle-1 and **b)** Kaggle-2 dataset

parametric values. For binary classification, proposed model has outperformed all the models for precision, sensitivity, F1 metric and accuracy. It has comparable results with MobileNet-V2 (Apostolopoulos and Mpesiana 2020) for sensitivity (99.10%) and MobileNet-V2 (Apostolopoulos et al. 2020) for accuracy (99.18%). Apostolopoulos and Mpesiana (2020) have used the pre-trained VGG-19 and MobileNet-V2 to detect the COVID-19. In comparison to these models, FocusCovid has achieved better results with fewer parameters. In (Apostolopoulos et al. 2020), MobileNet-V2 have achieved the 99.18% accuracy and 97.36% sensitivity. In comparison, FocusCovid has

achieved better sensitivity (99.20%) and comparable accuracy (99.20%).

Vaid et al. (2020) have used the pre-trained VGG-19 for COVID-19 detection. They have achieved 96.30% accuracy for the binary classification. While the proposed FocusCovid has achieved 99.20% accuracy for the binary classification on more radiographs. DarkCovidNet (Ozturk et al. 2020) is the CNN based architecture used for the binary and multi-class classification. It has less trainable parameters in comparison to FocusCovid but the proposed FocusCovid has shown superior results for both binary and three class classifications. FocusCovid have outperformed the DarkCovidNet

Table 8 Comparison of the state of the art models for binary and three class classification

Model	Dataset	Class	Precision	Sensitivity	F1 score	Accuracy
MobileNet-V2 (Apostolopoulos et al. 2020)	X-rays	2	–	97.36	–	99.18
VGG-19 (Apostolopoulos and Mpesiana 2020)	X-rays	2	–	92.85	–	98.75
MobileNet-V2 (Apostolopoulos and Mpesiana 2020)	X-rays	2	–	99.10	–	97.40
VGG-19 (Vaid et al. 2020)	X-rays	2	95.15	96.55	95.80	96.30
DarkCovidNet (Ozturk et al. 2020)	X-rays	2	98.03	95.13	96.51	98.08
nCOVNet (Panwar et al. 2020)	X-ray	2	–	97.62	–	88.10
Capsnet (Toraman et al. 2020)	X-rays	2	97.08	97.42	97.24	97.24
Proposed FocusCovid	X-rays	2	99.20	99.20	99.20	99.20
DarkCovidNet (Ozturk et al. 2020)	X-rays	3	89.96	85.35	87.37	87.02
VGG-19 (Apostolopoulos and Mpesiana 2020)	X-rays	3	–	92.85	–	93.48
MobileNet-V2 (Apostolopoulos and Mpesiana 2020)	X-rays	3	–	99.10	–	92.85
Capsnet (Toraman et al. 2020)	X-rays	3	84.61	84.22	84.21	84.22
COVID-Net (Wang et al. 2020)	X-rays	3	–	93.3	–	93.3
Proposed FocusCovid	X-rays	3	95.60	95.20	95.20	95.20

Models that have performed better on a different metric values are marked in bold along with the proposed FocusCovid

(Ozturk et al. 2020) on every parametric value. Panwar et al. (2020) performed the binary classification using the nCOVnet. nCOVnet is the pre-trained VGG-19 network with 14, 846, 530 parameters. It achieved an overall accuracy of 88.10% that is far less than the FocusCovid (99.20%). Toraman et al. (2020) proposed the convolutional capsnet for COVID-19 detection with capsule network. It achieved the accuracy of 97.24% and 84.22% for binary and three class classification. It has comparable accuracy with the FocusCovid but has not performed well for the three-class classification. Additionally, among the related studies done in this paper, FocusCovid has been evaluated on the largest COVID-19 chest radiographs.

Chest radiography is preferred by the radiologist to detect and get a glimpse of the lungs. Capturing X-rays is simple and has a low cost. When the cases of COVID-19 are increasing many folds and the RT-PCR test takes many hours to give the results, radiography can be used to detect and isolate the COVID-19 patient. The conducted experiments have given encouraging results but some limitations of the dataset need to be overcome in the future. A strongly labeled larger dataset of COVID-19 is needed for truly exploiting deep learning. Cases of patients showing mild symptoms can also be included in the dataset. Many models that are proposed uses the pre-trained models for feature extraction such as Apostolopoulos et al. (2020), Apostolopoulos and Mpesiana (2020), Vaid et al. (2020) and some has focused on training the model from scratch such as Toraman et al. (2020), Wang et al. (2020), Ozturk et al. (2020). In this study, instead of using the transfer learning strategy, a CNN model is designed and trained from scratch for classification.

Results are encouraging but need to be cross-validated with the medical radiologists as this is preliminary work.

This study aims to check the deep learning models to automatically detect the COVID-19 and to release the burden from the medical fraternity. More experiments with in-depth large data needed to be performed for further testing the proposed model. Further, segmentation and rib suppression can be used to increase the detection rate of COVID-19. Most of the studies that are performed have used the Cohen (Cohen et al. 2020) dataset directly or indirectly. More diversified data needs to be released so, experiments can be performed to differentiate the COVID-19 from other viral pneumonia such as SARS or MERS.

6 Conclusion

The number of COVID-19 cases is rising daily. The existing medical infrastructure is collapsing and the medic is working for late hours to assist. In this study, a model has been proposed for the COVID-19 detection automatically. The automatic system can help in detection of COVID-19 cases early and help to stop the spread of the virus. No handcraft feature extracting technique is used in the model. The proposed CNN architecture is trained from scratch instead of using the transfer learning techniques. Two separate datasets are used to validate the proposed model. A drawback of this study is that a limited dataset is used. More samples of COVID-19 and other pulmonary disorders are needed to be used for validating the proposed model.

Declarations

Conflict of interest Authors do not have any conflict of interest.

References

- Ahuja S, Panigrahi BK, Dey N, Rajinikanth V, Gandhi TK (2020) Deep transfer learning-based automated detection of covid-19 from lung ct scan slices. *Applied Intelligence*. 51:571–85
- Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L (2020) Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in China: a report of 1014 cases. *Radiology* 296:200642
- Apostolopoulos ID, Aznaouridis SI, Tzani MA (2020) Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases. *Journal of Medical and Biological Engineering* 40:1
- Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine* 43:1
- Asraf A (2020) COVID dataset. <https://www.kaggle.com/amanullahaasraf/covid19-pneumonia-normal-chest-xraypa-dataset>. Accessed 02 Jan 2021
- Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TML, Pan I, Shi LB, Wang DC, Mei J et al (2020) Performance of radiologists in differentiating covid-19 from viral pneumonia on chest ct. *Radiology* 296:200823
- Bhagat P, Choudhary P, Singh KM (2021) A novel approach based on fully connected weighted bipartite graph for zero-shot learning problems. *J Ambient Intell Humaniz Comput* 1–16
- Bobadilla JCM, Pedrini H (2016) Lung nodule classification based on deep convolutional neural networks. *Iberoamerican congress on pattern recognition*. Springer, New York, pp 117–124
- Candemir S, Antani S (2019) A review on lung boundary detection in chest x-rays. *International journal of computer assisted radiology and surgery* 14(4):563–576
- Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research* 16:321–357
- Chen S, Suzuki K (2013) Computerized detection of lung nodules by means of “virtual dual-energy” radiography. *IEEE Transactions on Biomedical Engineering* 60(2):369–378
- Chen S, Suzuki K, MacMahon H (2011) Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification. *Medical physics* 38(4):1844–1858
- Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahub ZB, Islam KR, Khan MS, Iqbal A, Emadi NA, Reaz MBI, Islam MT (2020) Can ai help in screening viral and covid-19 pneumonia? *IEEE Access* 8:132665–132676. <https://doi.org/10.1109/ACCESS.2020.3010287>
- Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M (2020) Covid-19 image data collection: prospective predictions are the future. *arXiv preprint arXiv:2006.11988*
- Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, pp 248–255. Ieee
- (2020) EURO: COVID dataset. <https://www.eurorad.org>. Accessed 02 Jan 2021
- Fong SJ, Dey N, Chaki J (2020) An introduction to covid-19. *Artificial intelligence for coronavirus outbreak*. Springer, New York, pp 1–22
- Gianchandani N, Jaiswal A, Singh D, Kumar V, Kaur M (2020) Rapid covid-19 diagnosis using ensemble deep transfer learning models from chest radiographic images. *J Ambient Intell Humaniz Comput* 1–13
- Giger ML, Chan HP, Boone J (2008) Anniversary paper: history and status of cad and quantitative image analysis: the role of medical physics and aapm. *Medical physics* 35(12):5799–5820
- (2020a) GitHub: COVID dataset. <https://github.com/armiro/COVID-CXNet>. Accessed 02 Jan 2021
- (2020b) GitHub: COVID dataset. <https://github.com/ml-workgroup/covid-19-image-repository>. Accessed 02 Jan 2021
- Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp 249–256
- Grech V (2020) Unknown unknowns-covid-19 and potential global mortality. *Early Human Development* 144:105026
- Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, Liu L, Shan H, Lei CL, Hui DS et al (2020) Clinical characteristics of coronavirus disease 2019 in china. *New England journal of medicine* 382(18):1708–1720
- Han D, Liu Q, Fan W (2018) A new image classification method using CNN transfer learning and web data augmentation. *Expert Systems with Applications* 95:43–56
- He K, Sun J (2015) Convolutional neural networks at constrained time cost. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 5353–5360
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
- He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. *European conference on computer vision*. Springer, New York, pp 630–645
- Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 7132–7141
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395(10223):497–506
- Ioffe S, Szegedy C (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*
- Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, Marklund H, Haghgoo B, Ball R, Shpanskaya K et al (2019) Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol 33, pp 590–597
- Islam MZ, Islam MM, Asraf A (2020) A combined deep cnn-lstm network for the detection of novel coronavirus (covid-19) using x-ray images. *Informatics in medicine unlocked* 20:100412
- Kanne JP, Little BP, Chung JH, Elicker BM, Ketai LH (2020) Essentials for radiologists on covid-19: an update—radiology scientific expert panel. *Radiology* 296(2):E113–E114. <https://doi.org/10.1148/radiol.2020200527>
- Kaul C, Manandhar S, Pears N (2019) Focusnet: an attention-based fully convolutional network for medical image segmentation. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), pp 455–458. IEEE
- Konar D, Panigrahi BK, Bhattacharyya S, Dey N (2020) Auto-diagnosis of covid-19 using lung ct images with semi-supervised shallow learning network

- Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* 25:1097–1105
- Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* 284(2):574–582
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
- Li Q, Cai W, Wang X, Zhou Y, Feng DD, Chen M (2014) Medical image classification with convolutional neural network. In: 2014 13th international conference on control automation robotics & vision (ICARCV), pp 844–848. IEEE
- Lin M, Chen Q, Yan S (2013) Network in network. arXiv preprint [arXiv:1312.4400](https://arxiv.org/abs/1312.4400)
- Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Medical image analysis* 42:60–88
- (2020) Mooney: pneumonia dataset. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>. Accessed 02 Jan 2021
- Nour M, Cömert Z, Polat K (2020) A novel medical diagnosis model for covid-19 infection detection based on deep features and Bayesian optimization. *Applied Soft Computing* 97:106580
- Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR (2020) Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in Biology and Medicine* 121:103792
- Panwar H, Gupta P, Siddiqui MK, Morales-Menendez R, Singh V (2020) Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. *Chaos Solitons Fractals* 138:109944
- Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning. arXiv preprint [arXiv:1712.04621](https://arxiv.org/abs/1712.04621)
- Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, et al. (2017) Chestnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint [arXiv:1711.05225](https://arxiv.org/abs/1711.05225)
- Redmon J, Farhadi A (2017) Yolo9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7263–7271
- Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. *Adv Neural Inf Process Syst* 30:3856–3866
- Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, Fan Y, Zheng C (2020) Radiological findings from 81 patients with covid-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet Infectious Diseases* 20:425–34
- (2020) SIRM: COVID dataset. <https://www.sirm.org/category/senza-categoria/covid-19/>. Accessed 02 Jan 2021
- Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M (2014) Striving for simplicity: the all convolutional net. arXiv preprint [arXiv:1412.6806](https://arxiv.org/abs/1412.6806)
- Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 15(1):1929–1958
- Srivastava RK, Greff K, Schmidhuber J (2015) Highway networks. arXiv preprint [arXiv:1505.00387](https://arxiv.org/abs/1505.00387)
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
- Toraman S, Alakus TB, Turkoglu I (2020) Convolutional capsnet: a novel artificial neural network approach to detect covid-19 disease from x-ray images using capsule networks. *Chaos Solitons Fractals* 140:110122
- Vaid S, Kalantar R, Bhandari M (2020) Deep learning covid-19 detection bias: accuracy through artificial intelligence. *International Orthopaedics* 44:1
- Wan Y, Shang J, Graham R, Baric RS, Li F (2020) Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of sars coronavirus. *J Virol* 94(7), Am Soc Microbiol
- Wang L, Lin ZQ, Wong A (2020) Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports* 10(1):1–12
- Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) 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, pp 2097–2106
- Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, Liu S, Zhao P, Liu H, Zhu L et al (2020) Pathological findings of covid-19 associated with acute respiratory distress syndrome. *Lancet respiratory medicine* 8(4):420–422
- Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K (2017) Learning to diagnose from scratch by exploiting dependencies among labels. arXiv preprint [arXiv:1710.10501](https://arxiv.org/abs/1710.10501)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.