2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 978-1-6654-0126-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/BIBM52615.2021.9669767

DeepCOVIDNet: Deep Convolutional Neural Network for COVID-19 Detection from Chest Radiographic Images

Khandaker Mamun Ahmed [1,2], Taban Eslami [3] Fahad Saeed [1], and M. Hadi Amini[1,2,*]
[1]Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, Florida, USA
[2]Sustainability, Optimization, and Learning for InterDependent Networks Laboratory (Solid Lab),
Florida International University, Miami, FL, USA
[3]Department of Computer Science, Western Michigan University, Michigan, USA
*Corresponding Author: moamini@fiu.edu

***Abstract*—The novel Coronavirus Disease 2019 (COVID-19) is a global pandemic that has infected millions of people causing millions of deaths around the world. Reverse Transcription Polymerase Chain Reaction (RT-PCR) is the standard screening method for COVID-19 detection but it requires specific molecular-biology training. Moreover, the general workflow is difficult e.g. sample collection, processing time, and analysis expertise, etc. Chest radiographic image analysis can be a good alternative screening method that is faster, more efficient, and requires minimal clinical or molecular biology trained laboratory personnel. Early studies have shown that abnormalities on the chest radiographic images are likely to be the consequence of COVID-19 infection. In this study, we propose DeepCOVIDNet, a deep learning based COVID-19 detection model. Our proposed deep-learning model is a multiclass classifier that can distinguish *COVID-19*, *viral pneumonia*, *bacterial pneumonia*, and *healthy* chest X-ray images. Our proposed model classifies radiographic images into four distinct classes and achieves the accuracy of 89.47% along with a high degree of precision, recall and F1 score. On a different dataset setting (COVID-19, bacterial pneumonia, viral pneumonia) our model achieves the maximum accuracy of 98.25%. We demonstrate generalizability of our proposed method using 5-fold cross validation for COVID-19 vs pneumonia and COVID-19 vs healthy classification that also manifests promising results.**

***Index Terms*—COVID-19, neural network, pneumonia, transfer learning, RT-PCR**

I. Introduction

A. Motivation

The novel Coronavirus disease 2019 (COVID-19) is caused by SARS-CoV-2 virus that is a new member of the coronavirus family [1]. According to the CDC, COVID-19 patients exhibit a range of symptoms including but not limited to cough, shortness of breath, fever, chills, muscle pain, headache, sore throat, and loss of taste or smell [2]. Diagnosis of COVID-19 requires assessment of symptoms across settings and ruling out the presence of co-occurring and/or alternative diagnoses that share notable features with COVID-19 such as pneumonia. However, lack of uniformity in practice guidelines and our limited understanding of COVID-19 may lead to misdiagnosis, especially for similar symptoms.

This verification for COVID-19 is usually accomplished using RT-PCR test [3]. However, RT-PCR test identifies the SARS-Coronavirus-2 (SARS-CoV-2) RNA that is normally detectable in the *acute* phase of the infection. In other words, some patients may develop severe symptoms before they are tested positive in COVID-19. Further, RT-PCR is time consuming and highly sensitive, requiring complicated manual processes carried out by trained laboratory personnel. Therefore, an easy to administer, low-cost test that may require minimal personnel training to determine the presence of COVID-19 before acute symptoms is highly desirable. Recent studies have focused on data analytics for COVID-19 prediction [4]. However, there is still a gap in using medical information for COVID-19 detection.

B. Related works

In recent years, the emergence of deep learning (DL) has shown great potential in various fields such as computer vision, text processing, and medical image classification. Moreover, applications of DL have achieved state-of-the-art results in numerous areas including natural language processing, object detection, detection and diagnosis of skin cancer and detection of cerebral microhemorrhage [5]. Deep learning models especially convolutional neural networks (CNN) have become a popular choice among researchers for medical image analysis because of its remarkable success in histopathological images [6], Magnetic Resonance Imaging (MRI) [7], skin related diseases [8], and brain tumor segmentation [9].

The gradient-based learning applied to document recognition was the first successful CNN architecture developed by Yann Lecun in 1998 to recognize handwritten digits on checks [10]. He is well known for his work on optical character recognition and considered as the founding father of convolutional nets. The model he proposed is LeNet-5, a 7-layer convolutional network, that was applied by several banks to recognize hand-written numbers on checks. LeNet-5 was applied to MNIST dataset and achieved the accuracy of 95%.

After more than 12 years of LeNet, a new CNN based model named AlexNet won the ImageNet challenge in 2012 with a large margin that rose the interest in CNNs [11]. AlexNet had a very similar architecture as LeNet but it was deeper, with more filters per layer, and with stacked convolutional

layers. AlexNet was trained for 6 days simultaneously and achieved a top-5 error of 15.3% in ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012.

In 2014 ILSVRC challenge, VGG Net addressed the major shortcomings of AlexNet by replacing large kernel-sized filters with multiple smaller filters (3×3 kernel-sized filters) [12]. The architecture of the model is very uniform with 3*3 convolutional filters and 2*2 max pooling layer with a stride of 1. It achieved top-5 accuracy of 92.7% on ILSVRC 2014 challenge. The winner of ILSVRC 2014 challenge is the GoogLeNet architecture that is well known for the inception module. It achieved a top-5 accuracy rate of 93.33% [13]. In ILSVRC 2015 challenge, ResNet was able to solve the "vanishing gradient" problem and minimized error rate even if it had 152 layers and achieved state-of-the-art accuracy of 96.3% [14], [15].

Trimps-Soushen, the third research institute of ministry of public security in China, won the 1st place of ILSVRC 2016 object classification task with the top-5 error rate of 2.99% [15]. In 2017 ILSVRC challenge, Hu et al. developed Squeeze-and-Excitation Network (SENet) that introduced a new architectural unit termed as "Squeeze-and-Excitation" (SE) block [16]. The SENet won the first place with state-of-the-art performance that reduced top-5 error rate to 2.25%.

The evolution of CNN models, their architectural description and performance comparison are highlighted in table I: LeNet, AlexNet, VGGNet, GoogLeNet, ResNet, Trimps-Soushen and SENet.

Considering the success of CNN based architectures for object classification, we asked the question: can we detect COVID-19 from chest radiographic images? There are some initial investigations into developing machine-learning solutions using X-ray images for diagnosing COVID-19 disease. Fang et. al. have shown that the sensitivity of CT for COVID-19 infection is 98% compared to RT-PCR sensitivity of 71% [18]. This result supports the use of chest CT for screening especially when RT-PCR is negative. Linda Wang and Alexander Wong have developed a deep neural network based method for the detection of COVID-19 from chest X-ray images [19]. A DL based segmentation system is also developed by Shan et al. that can automatically quantify the infected regions of interest as well as their volumetric ratios with respect to lungs [20]. To screen COVID-19 disease from CT images, Wang et. al. developed an Artificial Intelligence (AI) based system that has shown promising results with 82.9% accuracy, 80.5% specificity and 84% sensitivity [21].

The idea of transfer learning with the incorporation of various pre-trained deep learning models such as VGG-19, MobileNetV2, DenseNet201 and InceptionResNetV2 that are trained on ImageNet dataset is also investigated in [22]–[24] and has shown reasonable performance in COVID-19 detection. For example, the work of presented by Ioannis and Tzani achieved a good result with the accuracy of 96.78% where Brunese et al. proposed a three-fold method to detect and localise the infected areas [25]. Ismael et al. also demonstrated achievements of fine-tuning pretrained deep CNN models (ResNet18, ResNet50, ResNet101, VGG16, and VGG19) for COVID-19 classification where accuracy of the pretrained models VGG16, ResNet18, ResNet50, ResNet101 and VGG19 are 85.26%, 88.42%, 92.63% 87.37% and 89.47% respectively in normal vs COVID-19 classification setting [26]. The deep learning model for automatic COVID-19 detection introduced by Ozturk et al. presents classification accuracy of 98.08% for binary classes (COVID vs. No-Findings) and 87.02% for multi-class cases (COVID vs. No-Findings vs. Pneumonia) [27]. The CovidAID method presented by Mangal et al. obtained an accuracy of 87.2% for 4-class classification configuration [28]. In transfer learning a model developed to solve a task is reused as the starting point to solve another task in the same or similar domain. For example, a person who can play guitar can learn faster to play violin comparing others who don't have any previous experience on playing musical instruments as both of them are musical instruments that have some commonalities. But transfer learning cannot always work successfully especially when tasks have very less commonalities and this phenomenon is termed as negative transfer [29]. An example would be, having experience to play guitar can not help to learn faster to ride a bicycle. Pretrained models (transfer learning) are trained on ImageNet dataset that consists of millions of image data over thousands of different categories such as ballon, strawberry and dog. ImageNet dataset is very different than X-ray images that has different data distribution which might lead to negative transfer. Considering this issue, we built our DeepCOVIDNet model rather than applying transfer learning that shows reasonably good or similar performance compared with transfer learning techniques. In table II we have shown prior works on COVID-19 detection, their dataset settings and performance accuracy.

In this paper, we present DeepCOVIDNet, a convolutional deep learning model that can detect COVID-19 disease from chest radiographic images. Our proposed model is able to classify COVID-19, bacterial pneumonia, viral pneumonia and healthy (normal) chest X-ray images without including any further information such as demographics. DeepCOVIDNet manifests a novel CNN architecture including data preprocessing and augmentation strategy that allows training of our deep-learning model using minimal datasets. Performance of the model is evaluated by confusion matrix, precision, recall, F-1 score and accuracy. Our model has achieved promising results with more than 97% accuracy for classifying COVID-19 vs pneumonia, and COVID-19 vs healthy. An in multi-class classification, it achieved 89.47% accuracy.

C. Contributions

As we mentioned in the related works, bacterial pneumonia and viral pneumonia chest X-rays have not been addressed while detecting COVID-19 that exhibits similar symptoms of COVID-19 infection. In order to detect COVID-19, we propose DeepCOVIDNet: a multiclass classifier that can distinguish COVID-19, viral pneumonia, bacterial pneumonia, and healthy chest X-ray images. Our contributions include:

TABLE I: Architectural description and performance comparison of various CNN models: LeNet, AlexNet, VGGNet, GoogLeNet, ResNet, Trimps-Soushen and SENet. (The error rate for ImageNet dataset: top-5 error rate)

Architecture Name	Year	Main contribution	Parameters	Error Rate	Depth	Reference
LeNet	1998	First popular CNN architecture	0.060 M	MNIST: 5.0%	5	LeCun et al., 1998 [10]
AlexNet	2012	Deeper and wider than LeNet Uses relu, droupout and pooling	60M	ImageNet: 16.4%	8	Krizhevsky et al., 2012 [11]
VGGNet	2014	Homogenous topology Uses small sized kernels	138 M	ImageNet: 7.3%	19	Simonyan and Zisserman, 2015, [12]
GoogLeNet	2014	Introduced inception module	4 M	ImageNet: 6.7%	22	Szegedy et al., 2015 [13]
ResNet	2015	Introduced residual module	25.6 M	ImageNet: 3.7%	152	He et al., 2015 [14]
Trimps -Soushen	2016	Ensembled the pretrained models of Inception-v3, Inception-v4, ResNet-200, Inception-ResNet-v2, and Wide ResNet	-	ImageNet: 2.99%	-	Zhang et al., 2016 [17]
SENet	2017	Introduced a new architectural unit termed as "Squeeze-and-Excitation" (SE) block	Additional ≈10%	ImageNet: 2.25%	154	Hu et al., 2017 [16]

TABLE II: Existing works, dataset settings and accuracy

Related works	Dataset setting	Accuracy
Brunese et al. [25]	Healthy vs disease	96%
	Disease vs COVID-19	98%
Ismael et al. [26]	Normal vs COVID-19	90.3%
Wang et al. [3]	COVID-19, pneumonia, normal	93.3%
Arpan et al. [28]	COVID-19, pneumonia, normal	90.5%
Gozes et al. [30]	COVID-19 vs normal	94%
Yang et al. [31]	COVID vs non-COVID CT	83.3%
Born et al. [32]	Bacterial pneomonia, COVID-19 and healthy	89%
Tulin et al. [27]	COVID vs No-Findings	98.08%
	COVID vs No Fildings vs Pneomonia	87.02%
Harsh et al. [33]	Bacterial pneumonia, viral pneumonia, COVID-19 and healthy	88.10%

- A deep convolutional neural network (DeepCOVIDNet) to detect COVID-19 that improved state-of-the-art results by increasing the accuracy about 2% from 87.2% [25] to 89.47%. In a large scale population the improvement of 2% could play a significant role.
- An end-to-end structure without manual feature extraction and selection methods.
- Data augmentation technique to prevent overfitting and to enhance the model's generalizability and dataset size. Moreover, a data preprocessing method to remove noises from the dataset.
- DeepCOVIDNet classifier that classifies COVID-19, viral pneumonia, bacterial pneumonia, and healthy chest X-ray images.

Moreover, in order to highlight our contributions as compared with prior works, we have compared our work with existing methods in table III.

D. Organization

The paper is organized as follows: II presents the methodology of the work including data pre-processing and augmentation. Section III presents the dataset description. Section IV represents experimental setup. Section V is dedicated for result analysis, followed by section VI that concludes the paper.

II. Proposed method

The proposed method is illustrated in Figure 1, in which chest X-ray images are used as the input data. Our proposed model DeepCOVIDNet is a 11 layer convolutional network with 6 convolutional layer, 3 max pooling layer, 1 global average pooling layer where the final layer is a fully connected dense layer. Max pooling and average pooling layers reduce variance and computational complexity, and extract low level features. Different numbers of (3,3) sized kernels are employed in the convolutional layers while (2,2) max pooling layers are used for subsampling. As it is a classification problem, softmax activation function is applied in the final layer and a (2,2) rectified linear unit (ReLU) is implemented in all other layers to introduce non-linearity. The model is optimized using the Adam optimizer, categorical crossentropy loss function is utilized for calculating loss and the learning rate is set as 0.001. The model is trained on 120 epochs on multi-class configurations and 25 epochs for pneumonia vs COVID-19 classification and pneumonia vs normal configurations, and 20 epochs for COVID-19 vs normal configurations.

A. Data Augmentation and Preprocessing

Data preprocessing is a crucial part in deep learning architectures to eliminate anomalies that may lead to inaccurate interpretation of machine-learning artifacts. It includes removing noise and redundant data, and making the data suitable so that model can learn and predict accurate results. In our pneumonia and normal dataset, there were some noises such as letter 'R' in the x-ray images. As, noise is not expected in the dataset and it may lead to inconsistent results, noise is removed from the x-ray images. The process is done by analyzing pixel value and comparing it with its neighbor pixels. Noise is in white color and based on our empirical analysis, we have found that the pixel value of it is greater than 245. So, we changed the pixels' values that have the value greater than

TABLE III: Existing methods and our proposed work comparison

Related works	Contributions	Our proposed work
Wang et al. [19]	Provided promising results with 88.9% accuracy in COVIDx dataset (pneumonia and Covid-19) that includes COVID-19 images.	Our proposed work further classifies COVID-19, bacterial pneumonia, viral pneumonia and healthy X-ray images.
Ioannis and Tzani [23]	Implemented transfer learning that achieved good results with the accuracy of 96.78% for COVID-19 detection. The two dataset presented in this paper are imbalanced.	We present similar approach but we classified bacterial and viral pneumonia further with balanced dataset.
Shan et al. [20]	Proposed method can automatically quantify infected regions of interest as well as their volumetric ratios with respect to lungs.	Our predictive model classifies COVID-19 where Shan's model automatically segments infected regions.
Narin et al. [34]	Have used ResNet50, InceptionV3 and Inception-ResNetV2 pretrained models for prediction COVID-19 from chest X-ray images. A very small dataset (50 COVID-19 vs 50 normal) is being used in this experiment.	Created our own model. Experiment is carried out on relatively larger dataset. Data augmentation and data preprocessing are also implemented.
Brunese et al. [25]	Presented a three-fold method to detect the presence of pneumonia, classification between pneumonia vs COVID-19 and highlighted the important regions for prediction.	Demonstrated the classification between COVID-19 vs pneumonia, COVID-19 vs normal, pneumonia vs normal and multiclass.
Ozturk et al. [27]	Proposed DarkCovidNet: a deep learning model for the automatic diagnosis of COVID-19.	Demonstrated the effectiveness of DeepCOVIDNet to detect COVID-19 in multiple classification scenarios.
Mangal et al. [28]	Presented CovidAID: COVID-19 AI detector that achieved an accuracy of 87.2% for 4-class classification configuration and 90.5% for the 3-class classification.	DeepCOVIDNet model achieved 89.47% accuracy for 4-class classification where COVID-19 vs Pneumonia and COVID-19 vs normal configurations it achieved 97.44% and 97.37% accuracy.

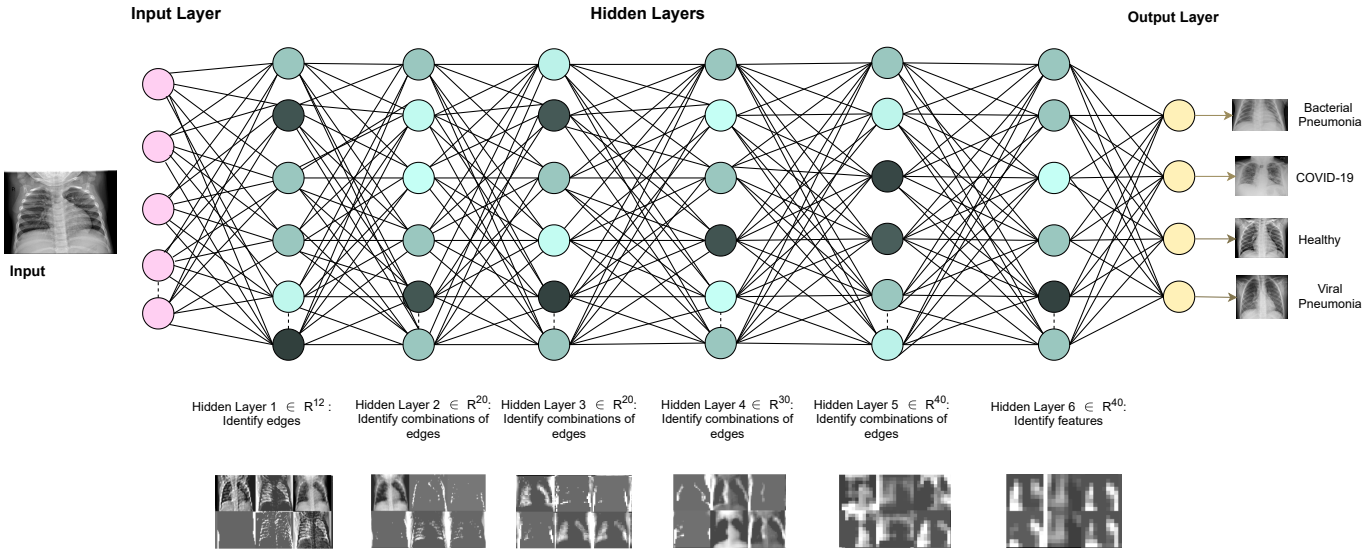


Fig. 1: DeepCOVIDNet: Deep learning model for multiclass classification: Bacterial pneumonia, COVID-19, Viral pneumonia and Healthy patient from chest X-ray images

245. Moreover, to make the change similar to its surrounding pixels, we have also considered its neighbor pixels (Figure 2). If the image size is $[p, q]$, we iterate through each pixel x_i, y_j where $i \in [0, p], j \in [0, q]$. Based on equation 1, for a random pixel $x_{i=a}, y_{j=b}$, where the pixel value $Px_{i=a}, y_{j=b}$ is more than 245 we calculate the minimum (m) of the four neighbor pixels. Then, based on equation 2, we generate a random number within the range of m and $m + 10$ that we have found from equation 1 and assign this random value in $fx_{i=a}, y_{j=b}$; otherwise we assign $Px_{i=a}, y_{j=b}$ in $fx_{i=a}, y_{j=b}$.

Figure 3 shows the result before and after this operation is performed. And, there is a clear indication that noise is removed effectively from the dataset.

$$m = min((x_{i=a-10}, y_{j=b}), (x_{i=a+10}, y_{j=b}), (x_{i=a-10}, y_{j=b+10}), (x_{i=a+10}, y_{j=b+10})) \quad (1)$$

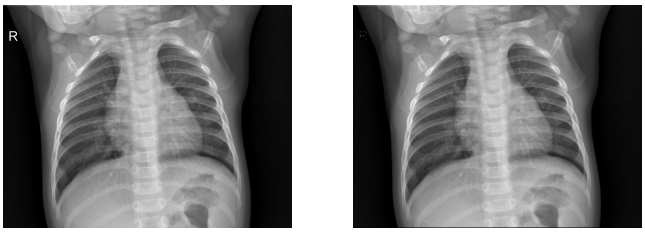
$$fx_{i=a}, y_{j=b} := \begin{cases} Px_{i=a}, y_{j=b} & if, Px_{i=a}, y_{j=b} <= 245 \\ random(m, m+10) & otherwise \end{cases} \quad (2)$$

Data augmentation improves model accuracy and robustness. It increases the number of training samples that reduces overfitting and improves model generalization [35]. It also

$x_{i=0}, y_{j=0}$	...	$x_{i=a-10}, y_{j=0}$	...	$x_{i=a}, y_{j=0}$	...	$x_{i=a+10}, y_{j=0}$	...
...	...	...	...	...	...	...	...
$x_{i=0}, y_{j=b-10}$	...	$x_{i=a-10}, y_{j=b-10}$	...	$x_{i=a}, y_{j=b-10}$	...	$x_{i=a+10}, y_{j=b-10}$	...
...	...	...	...	...	...	...	...
$x_{i=0}, y_{j=b}$	...	$x_{i=a-10}, y_{j=b}$	...	$x_{i=a}, y_{j=b}$	...	$x_{i=a+10}, y_{j=b}$	...
...	...	...	...	...	...	...	...
$x_{i=0}, y_{j=b+10}$	...	$x_{i=a-10}, y_{j=b+10}$	...	$x_{i=a}, y_{j=b+10}$	...	$x_{i=a+10}, y_{j=b+10}$	...
...	...	...	...	...	...	...	$x_{i=p}, y_{j=q}$

Fig. 2: Data preprocessing: Here $x_{i=a}, y_{j=b}$ is a random pixel (orange color) that has value more than 245. To change the pixel value, we considered four neighbor pixels (light green color)

alleviates scarcity of training data when the dataset is relatively small. After data preprocessing, training data is augmented by random cropping and, using shearing and zooming with values of 0.2 and 0.2. Moreover, input images are resized to 64x64 and rescaled by 1./255.



(a) (a) Before processing: Normal X-ray image with letter "R"

(b) (b) After processing: Normal X-ray image without letter "R"

Fig. 3: Data preprocessing: Noise Removal(e.g R letter) from X-ray images

B. Pneumonia (bacterial and viral) and Healthy classification

People affected by bacterial or viral pneumonia also show similar symptoms of COVID-19 like cough, fever, shaking or chills, fatigue, sweating and weakness. As a result, it is hard to determine whether a person is infected by COVID-19 or by pneumonia. Initially, we built a model for classifying pneumonia (bacterial and viral) and healthy patients from chest radiographic images. Our dataset contains 1583 pneumonia and 1583 normal chest X-ray images from where 1314 images of each class are used for model training, 234 for validation and 35 for testing.

C. COVID-19 and other classification

This section discusses COVID-19 vs healthy, COVID-19 vs pneumonia, and COVID-19, bacterial pneumonia, viral pneumonia and healthy classification.

In here, we classified COVID-19 and healthy chest X-ray images. But this classification result might not be enough because a person infected by pneumonia virus can also show similar symptoms of COVID-19. As a result, we have carried our work further for classifying COVID-19 vs pneumonia and COVID-19, bacterial pneumonia, viral pneumonia and normal classification. For this multiclass classification, we have used 120 epochs and for pneumonia vs COVID-19 classification 25 epochs are being used. Moreover, in both classifications 60, 19 and 10 images of each class are used as training, validation and testing data. Figure 1 shows the architecture of our multiclass classifier.

1) 5 Fold Cross Validation: Cross validation is the technique that we have used to evaluate our model performance. Model is validated by giving some data that the model didn't see before. We split the dataset into training and testing. The model learns from training dataset and validates from the unseen testing dataset. In our COVID-19 vs normal and COVID-19 vs pneumonia classification, we have used a 5-fold cross validation method. The whole dataset is divided into 80% training set and 20% testing set. Figure 4 explains the structure of our 5 fold cross validation method. After 5 iterations, average model performance (AMP) is calculated according to equation 3.

$$AMP = \frac{1}{5}\sum_{i=1}^{5} performance_i \tag{3}$$

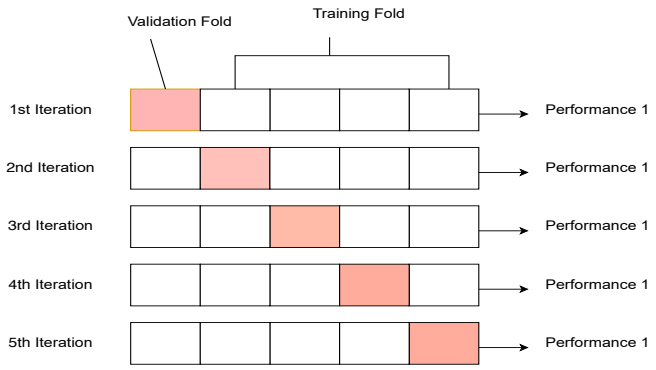


Fig. 4: 5 fold cross validation architecture

III. Dataset Description

The dataset we have used in this experiment contains bacterial pneumonia, viral pneumonia, COVID-19, and healthy chest X-ray images. Data was collected from two different sources. The COVID-19 dataset was collected from IEEE open source Github repository created by Joseph Paul Cohen, Paul Morrison and Lan Dao [36]. The repository contains 100 COVID-19 positive chest X-ray images in which 89 X-rays are taken from the frontal position and the rest of the images are taken from either top or side view. The other dataset was collected from Kaggle that contains chest X-ray images of 2780 bacterial pneumonia, 1493 viral pneumonia, and 1583 healthy patients [37].

IV. Experimental Setup

The entire experiment is carried out in Windows 10 operating system using python programming language. The deep learning classifiers have been implemented using python version 3.7.3 and the Keras package with TensorFlow version 2.0.0 on Intel(R) Core(TM) i7-3.6 GHz processor. In addition, the experiment is executed using the NVIDIA Quadro k620 graphics processing unit (GPU) with 2GB GPU memory.

V. Result Analysis

In this experiment, we study COVID-19 detection from chest radiology images. Initially, we investigate COVID-19 vs normal chest X-ray classification. Then, we classify COVID-19, viral pneumonia, bacterial pneumonia, and normal chest radiology images. Moreover, normal and pneumonia classification, COVID-19 and normal classification, and COVID-19 and pneumonia classifications are also analyzed. Besides, to reduces biasness in performance evaluation we have implemented 5-fold cross validation.

Performance of the model is measured through training and validation accuracy, receiver operating characteristic (ROC) curve, confusion matrix, precision, recall and F1 score. The formula for calculating precision, recall and F1 score is presented in equations 4, 5, 6 where tp = True positive; fp = False positive; tn = True negative; fn = False negative.

$$Precision(P) = \frac{tp}{tp + fp} \tag{4}$$

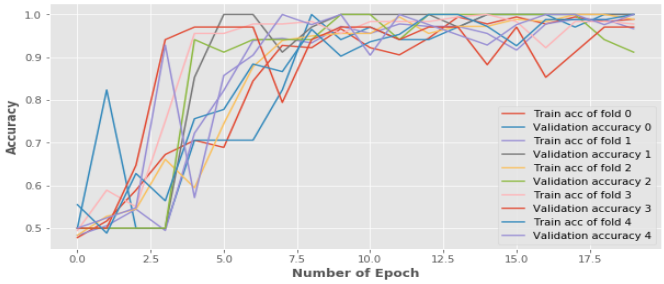
$$Recall(R) = \frac{tp}{tp + fn} \tag{5}$$

$$f - 1 = 2 * \frac{P * R}{P + R} \tag{6}$$

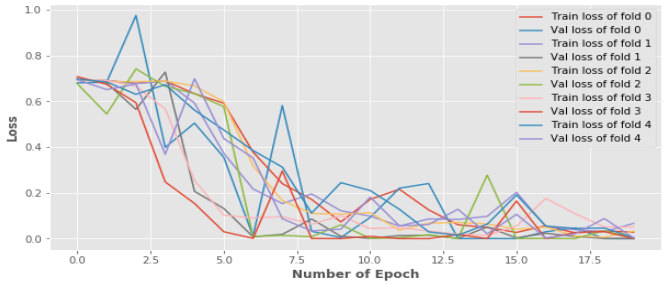
As we have a limited amount of available data, in COVID-19 vs normal classification, we have randomly selected 60 images for training, 19 images for validation and 10 images for testing for normal class. After 20 epochs, the model achieved 97.37% accuracy with the precision, recall and F1 score of 1. Figure 5 shows training and validation accuracy graph. In figure 6, we demonstrate the training and validation accuracy and loss of the 5 fold cross validation where model has achieved an average accuracy of 96.12% (table V). The 5-fold cross validation is also experimented for COVID-19 vs pneumonia classification where model achieved an average accuracy of 97.65%.

Fig. 5: COVID-19 and normal classification: Training and validation accuracy

In COVID-19, bacterial pneumonia, viral pneumonia and healthy chest X-ray classification our model has achieved an accuracy of 89.47%. Moreover, in figure 7 we analyse the ROC curve where area under the curve (AUC) value of each class is 1, 1, 0.91 and 0.9 respectively. We have also achieved high precision, recall and F-1 score for this multi-class setting (table IV). In another multi-class dataset setting (COVID-19, bacterial pneumonia, viral pneumonia) we have achieved 98.25% accuracy with a good precision, recall, f-1 scores. Figure 8 shows training loss and accuracy result of the experiment. Results of the COVID-19 vs pneumonia, COVID-19 vs normal and pneumonia vs normal X-ray images are also investigated where accuracy of these classifiers are 97.44%, 97.37% and 88.03% respectively (table IV).



(a) (a) Training and validation accuracy



(b) (b) Training and validation loss

Fig. 6: 5 fold cross validation of normal and COVID-19 classification. (a) training and validation accuracy; (b) training and validation loss

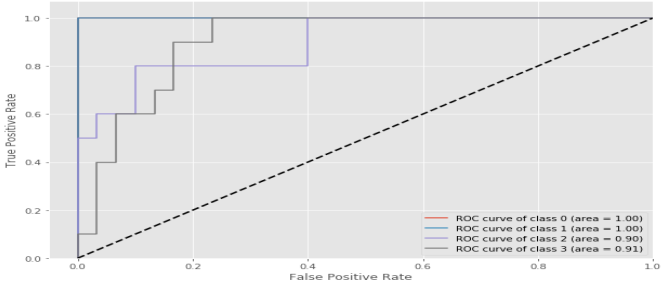


Fig. 7: Receiver Operating Characteristic (ROC) Curve for the classification of COVID-19, bacterial pneumonia, viral pneumonia and normal X-ray images. Here, Class 0 - bacterial pneumonia, class 1 - COVID-19, class 2 - normal and class 3 - viral pneumonia

It can be noted from table 4 that the model has achieved best performance in 3-class classification setting, and 2-class classification settings except pneumonia vs normal classification where it performed comparatively less. In 4-class

TABLE IV: COVID-19 vs pneumonia, COVID-19 vs normal, pneumonia vs normal and multiclass X-ray images classification: Precision, Recall and F-1 scores are calculated from the test dataset and the accuracy shown from the validation dataset

Classifier	Class	Training Dataset	Validation Dataset	Test Dataset	Precision	Recall	F1 Score	Accuracy v.(%)
COVID-19 vs Pneumonia	COVID-19	60	20	10	1	1	1	97.44
	Pneumonia	60	20	10	1	1	1	
COVID-19 vs Normal	COVID-19	60	19	10	1	1	1	97.37
	Normal	60	19	10	1	1	1	
Pneumonia vs Normal	Pneumonia	1314	234	35	0.86	0.91	0.89	88.03
	Normal	1314	234	35	0.91	0.86	0.88	
Multiclass	Bacterial pneumonia	60	19	10	1.00	1.00	1.00	89.47
	COVID-19	60	19	10	0.91	1.0	0.95	
	Normal	60	19	10	0.86	0.60	0.71	
	Viral Pneumonia	60	19	10	0.67	0.80	0.73	
Multiclass	Bacterial pneumonia	60	19	10	0.91	1	0.95	98.25
	COVID-19	60	19	10	0.9	0.9	0.9	
	Viral Pneumonia	60	19	10	0.89	0.80	0.84	

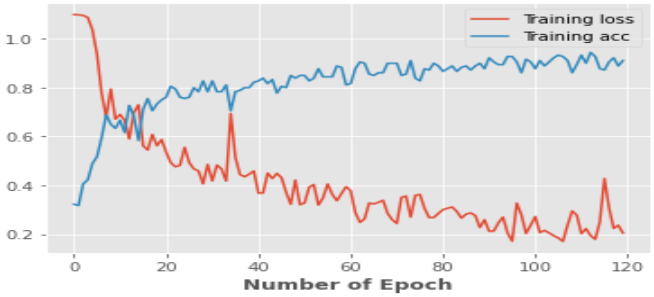


Fig. 8: COVID-19 bacterial pneumonia, viral pneumonia classification: Training accuracy and loss

classification, model classified COVID-19 better than normal and viral pneumonia and, obtained precision, recall and f-1 score values are 0.91, 1.0 and 0.95 respectively.

TABLE V: 5 fold cross validation model accuracy

Classification	Iteration Number	Accuracy(%)	Average Accuracy(%)
COVID-19 vs Pneumonia	1	100	97.65
	2	100	
	3	91.18	
	4	97.06	
	5	100	
COVID-19 vs Normal	1	100	96.12
	2	94.12	
	3	97.06	
	4	94.12	
	5	95.24	

VI. Conclusion and Future Works

In this study, we have presented DeepCOVIDNet in the detection of COVID-19, based on chest X-ray images. Our proposed model achieved notable performance for classifying COVID-19, pneumonia (bacterial and viral) and healthy X-ray images. Although COVID-19 and pneumonia exhibit similar symptoms, our muticlass classification model successfully classified them with the accuracy of 89.47%. We have also demonstrated other classification scenarios such as COVID-19 vs pneumonia, COVID-19 vs normal and pneumonia vs normal where our model achieved significant performance. Data preprocessing and augmentation is applied to remove noises and improve model's generalizability.

In future, we plan to experiment our model with large dataset so that we can have better inference before the development of clinical diagnostic models.

Acknowledgments

We are thankful to the open source community that made the dataset available for the researchers. We hope that our contribution will help the world community to overcome this pandemic situation. Research reported in this paper was partly supported by NIGMS of the National Institutes of Health (NIH) under award number R01GM134384. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Fahad Saeed was additionally supported by the NSF CAREER award OAC 1925960. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also acknowledge the

resources and support of Sustainability, Optimization, and Learning for InterDependent networks laboratory (solid lab) at the Florida International University (www.solidlab.network).

REFERENCES

[1] Dale Fisher and David Heymann. Q&a: The novel coronavirus outbreak causing covid-19. *BMC medicine*, 18(1):1–3, 2020.

[2] Coronavirus disease 2019 (covid-19): Symptoms of coronavirus. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html, 2020. [Online; accessed 6-May-2020].

[3] Wenling Wang, Yanli Xu, Ruqin Gao, Roujian Lu, Kai Han, Guizhen Wu, and Wenjie Tan. Detection of sars-cov-2 in different types of clinical specimens. *Jama*, 2020.

[4] Meleik Hyman, Calvin Mark, Ahmed Imteaj, Hamed Ghiaie, Shabnam Rezapour, Arif M Sadri, and M Hadi Amini. Data analytics to evaluate the impact of infectious disease on economy: Case study of covid-19 pandemic. *Patterns*, 2(8):100315, 2021.

[5] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, 2017.

[6] Zhongyi Han, Benzheng Wei, Yuanjie Zheng, Yilong Yin, Kejian Li, and Shuo Li. Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific reports*, 7(1):1–10, 2017.

[7] Taban Eslami, Vahid Mirjalili, Alvis Fong, Angela Laird, and Fahad Saeed. Asd-diagnet: A hybrid learning approach for detection of autism spectrum disorder using fmri data. *Frontiers in Neuroinformatics*, 13:70, 2019.

[8] Seung Seog Han, Myoung Shin Kim, Woohyung Lim, Gyeong Hun Park, Ilwoo Park, and Sung Eun Chang. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, 138(7):1529–1538, 2018.

[9] Heba Mohsen, El-Sayed A El-Dahshan, El-Sayed M El-Horbaty, and Abdel-Badeeh M Salem. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, 3(1):68–71, 2018.

[10] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[12] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[13] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[15] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.

[16] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[17] Trimps-Soushen: Good practices for deep feature fusion. https://www.youtube.com/watch?v=NaoVOOhVC3w, 2021. [Online; accessed 06-April-2021].

[18] Yicheng Fang, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, and Wenbin Ji. Sensitivity of chest ct for covid-19: comparison to rt-pcr. *Radiology*, page 200432, 2020.

[19] Linda Wang and Alexander Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images. *arXiv preprint arXiv:2003.09871*, 2020.

[20] Fei Shan, Yaozong Gao, Jun Wang, Weiya Shi, Nannan Shi, Miaofei Han, Zhong Xue, Dinggang Shen, and Yuxin Shi. Lung infection quantification of covid-19 in ct images with deep learning. *arXiv preprint arXiv:2003.04655*, 2020.

[21] Shuai Wang, Bo Kang, Jinlu Ma, Xianjun Zeng, Mingming Xiao, Jia Guo, Mengjiao Cai, Jingyi Yang, Yaodong Li, Xiangfei Meng, et al. A deep learning algorithm using ct images to screen for corona virus disease (covid-19). *MedRxiv*, 2020.

[22] Ezz El-Din Hemdan, Marwa A Shouman, and Mohamed Esmail Karar. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. *arXiv preprint arXiv:2003.11055*, 2020.

[23] Mpesiana Apostolopoulos, Ioannis D and Tzani A. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, page 1, 2020.

[24] Jianpeng Zhang, Yutong Xie, Yi Li, Chunhua Shen, and Yong Xia. Covid-19 screening on chest x-ray images using deep learning based anomaly detection. *arXiv preprint arXiv:2003.12338*, 2020.

[25] Luca Brunese, Francesco Mercaldo, Alfonso Reginelli, and Antonella Santone. Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. *Computer Methods and Programs in Biomedicine*, 196:105608, 2020.

[26] Aras M Ismael and Abdulkadir Şengür. Deep learning approaches for covid-19 detection based on chest x-ray images. *Expert Systems with Applications*, 164:114054, 2021.

[27] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in biology and medicine*, 121:103792, 2020.

[28] Arpan Mangal, Surya Kalia, Harish Rajgopal, Krithika Rangarajan, Vinay Namboodiri, Subhashis Banerjee, and Chetan Arora. Covidaid: Covid-19 detection using chest x-ray. *arXiv preprint arXiv:2004.09803*, 2020.

[29] Zirui Wang, Zihang Dai, Barnabás Póczos, and Jaime Carbonell. Characterizing and avoiding negative transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11293–11302, 2019.

[30] Ophir Gozes, Maayan Frid-Adar, Hayit Greenspan, Patrick D Browning, Huangqi Zhang, Wenbin Ji, Adam Bernheim, and Eliot 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.

[31] Xingyi Yang, Xuehai He, Jinyu Zhao, Yichen Zhang, Shanghang Zhang, and Pengtao Xie. Covid-ct-dataset: a ct scan dataset about covid-19. 2020.

[32] Jannis Born, Gabriel Brändle, Manuel Cossio, Marion Disdier, Julie Goulet, Jérémie Roulin, and Nina Wiedemann. Pocovid-net: automatic detection of covid-19 from a new lung ultrasound imaging dataset (pocus). *arXiv preprint arXiv:2004.12084*, 2020.

[33] Harsh Panwar, PK Gupta, Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, and Vaishnavi Singh. Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. *Chaos, Solitons & Fractals*, 138:109944, 2020.

[34] Ali Narin, Ceren Kaya, and Ziynet Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *arXiv preprint arXiv:2003.10849*, 2020.

[35] Nikita Dvornik, Julien Mairal, and Cordelia Schmid. On the importance of visual context for data augmentation in scene understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[36] Joseph Paul Cohen, Paul Morrison, and Lan Dao. Covid-19 image data collection. *arXiv 2003.11597*, 2020.

[37] Chest X-Ray Images (Pneumonia). https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia, 2020. [Online; accessed 3-May-2020].