International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 4, August 2024, pp. 4068~4077
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i4.pp4068-4077

A light-weight and generalizable deep learning model for the prediction of COVID-19 from chest X-ray images

Md Jakaria Zobair[1], Refat Tasfia Orpa[2], Mahir Ashef[3], Shah Md Tanvir Siddiquee[1], Narayan Ranjan Chakraborty[1], Ahsan Habib[4]
[1]Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh
[3]Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh
[4]School of Information Technology, Deakin University, Geelong, Australia

Article Info

Article history:

Received Jan 12, 2024
Revised May 9, 2024
Accepted May 12, 2024

Keywords:

Chest X-ray images
COVID-19
COVID-19 detection
Neural network
Transfer learning

ABSTRACT

The detection of coronavirus disease (COVID-19) using standard laboratory tests, such as reverse transcription polymerase chain reaction (RT-PCR), is time-consuming. Complex medical imaging problems are currently being solved using machine learning and deep learning techniques. Our proposed solution utilizes chest radiography imaging techniques, which have shown to be a faster alternative for detecting COVID-19. We present an efficient and lightweight deep learning architecture for identifying COVID-19 using chest X-ray images which achieve 99.81% accuracy in intra-database testing and 100% accuracy in cross-validation testing on a separate data set. The results demonstrate the potential of our proposed model as a reliable tool for COVID-19 diagnosis using chest X-ray images, which can have a significant impact on improving the efficiency of COVID-19 diagnosis and treatment.

Corresponding Author:

Md Jakaria Zobair
Department of Computer Science and Engineering, Daffodil International University
Dhaka-1216, Bangladesh
Email: zobair.cse@diu.edu.bd

1. INTRODUCTION

The severe acute respiratory syndrome coronavirus (SARS-CoV-2, COVID-19 in short) is a new virus causing lung sickness which spread worldwide rapidly including Bangladesh [1] and was declared a world pandemic and crisis by the World Health Organization (WHO). There are two types of tests to diagnose COVID-19, antigen test and molecular/polymerase chain reaction (PCR) test. Antigen test is quick and less expensive but is comparatively less accurate [2]. On the other hand, the molecular test requires expensive machines, highly trained personnel, and specific reagents. Both tests need to transport samples from the hospital to the lab, which may require some time. Therefore, it is necessary to initiate an automated detection system to identify the presence of anomalies in collected samples (near) real-time.

Chest X-ray and computed tomography (CT) image-based COVID-19 classification methods are relatively easily accessible, cost-effective, and faster clinical methods [3], [4]. These images are being used in detecting various respiratory diseases including COVID-19 using machine learning and deep learning algorithms [5]–[9]. AlexNet was found to achieve 98% and 94.1% accuracy in COVID-19 detection from X-ray and CT images respectively [10]. Kang *et al.* [11] used a neural network-based approach which achieved accuracy and sensitivity of 95.5% and 96.6% respectively. Xu *et al.* [12] used a summation of location-attention technique and 3D ResNet-18 to identify COVID-19 from CT images and got an overall 86.7% accuracy. The tempered inception architecture was proposed by [13] to detect COVID-19 from CT images which yield 79.3% accuracy and 0.67 sensitivity. A model based on a multi-layer perceptron followed by an long short-term memory (LSTM) layer is used by Fang *et al.* [14] and acquired a 0.95 area under the curve (AUC) score. In another paper [15] 2D convolutional neural network (CNN) is applied to classify the COVID-19 cases using the CT chest images and achieved an accuracy and AUC score of 94.98% and 97.91% respectively. Jin *et al.* [16] combined ResNet-50 and 3D UNet++ from ten thousand chest CT volumes and achieved a 97.77% AUC score for the 1943 testing data set. Afshar *et al.* [17] proposed the COVID-CAPS model, a variant of CapsuleNet, which was pre-trained on natural image data and fine-tuned using X-ray images and gained accuracy, sensitivity, and AUC score of 98%, 80%, and 0.97 respectively. Song *et al.* [18] worked with the dynamic radius-encoding neural network (DRE-Net) method on chest CT images and obtained 86% accuracy. Zheng *et al.* [19] proposed UNet+3D deep network and obtained 90% accuracy on 630 CT images where 499 images were used for training and 131 images were for testing. Since X-ray images are of particular interest in this study, the below section describes COVID-19 detection studies based on X-ray images. Many researchers use deep learning models for COVID-19 detection from X-ray images. Hemdan *et al.* [20] used seven types of pre-trained CNN models namely VGG16, DenseNet201, ResNetV2, InceptionV3, InceptionResNetV2, Xception, and MobileNetV2 and they used only 50 images where there were 50% normal and 50% COVID-19 positive images. Among all the models, they got the highest 90% accuracy for both VGG16 and DenseNet201. Another study [21] used VGG-19 models along with binary robust invariant scalable key-points (BRISK) algorithm and achieved 96.60% detection accuracy. Apostolopoulos *et al.* [22] designed a CNN model to group COVID-19 cases and gain a better testing accuracy of 93.48% for 3 class classifications (such as COVID, pneumonia, and normal) and 98.75% for binary class classification (such as COVID and normal). Ucar and Korkmaz [23] proposed a model, COVIDiagnosis-Net, based on SqueezeNet and taking the Bayesian optimizer and getting 98.3% testing accuracy for 3-class classification. In another study [24], ResNet-50 which is a pre-trained architecture has been applied on a small X-ray chest data set with an accuracy of 98% on a data set of 100 images where there are 50 normal images and 50 COVID-19 positive images. Sethy *et al.* [25] extracted features using ResNet50 and used support vector machine (SVM) as a classifier which yields 95.38% detection accuracy on a limited dataset of 50 images. Ozturk evaluated a network using deep learning (DL), called DarkCOVIDNet which has 17 convolution layers that can diagnose COVID-19 automatically by analyzing X-ray images [26]. The model obtained a better accuracy that is 87.02% and 98.08% for 2 types of class, one is multi-class having COVID-positive, pneumonia-affected, and normal images and another is two-class having COVID and normal images. Wang *et al.* [27] proposed COVIDNet, a ResNet-based 19-layer architecture, which acquired 93.3% accuracy on the chest X-ray data set. Togacar *et al.* [28] utilized social mimic optimization-based feature extraction using two deep learning models - MobileNetV2 and SqueezeNet and later used SVM for classification which achieved 99.27% accuracy. In [29] a 22-layer CNN model was proposed for detecting 2, 3, and 4-class classification and obtained accuracy values 99.1%, 94.2%, and 91.2% respectively.

Most of the deep learning studies in the literature used pre-trained models in detecting COVID-19 from the images of the chest X-ray data set. Parameters from already trained models are fixed and only the upper layers including the decision layer can be fine-tuned with currently limited samples. Although the time and space complexity of transfer learned models follow the similar complexity of the original models, sometimes the transfer learning process itself and the description and amount of limited training samples remain unclear. In addition, the selection of a suitable transfer learning model often did not consider a large number of state-of-the-art models. Thus, building a minimal (such as least time and space complexity) but efficient deep learning model has been the prime focus in this study, which performs similarly to the transfer learning models and generalizes well across unknown test data sets. In this study, we put forward three key contributions, i) we introduce an efficient lightweight model for rapid identification of COVID-affected cases, ii) the generalizability of the proposed model is judged with a 2nd data set, and iii) the performance, complexity, and generalizability of the state-of-the-art transfer learning models areevaluated.

2. METHOD

The problem formulated, in this study, as a binary classification problem. Where the input is chest X-ray images and the output is a binary decision (such as COVID-19 vs normal). Figure 1 shows the overall system flow which includes dataset collection, pre-processing, CNN model, classification, decision-making, and performance analysis.

2.1. Dataset

We collected two types of datasets from *kaggle.com* [30], [31]. Which consist of images related to chest X-rays of COVID-19-positive and normal subjects. Where the first data set has 2,159 images and the other data set has 96 images.

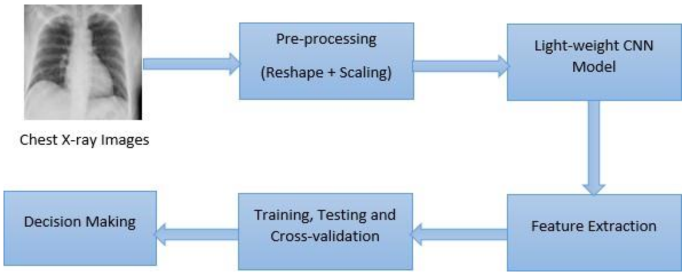

Figure 1. Overall system flow

2.2. Pre-processing

The input images used were of different sizes. That is why the first step was to resize all the images to 224×224 to unify the input shape for the model. Then we scaled the images by dividing them by 255 to get a value between 0 to 1 for all pixels.

2.3. Proposed model

Size and the number of kernels is the key hyperparameters of all convolution operations. We use 1×1, 3×3, and 7×7 kernels to build our architecture. The no of kernel tells about the depth of the output feature maps and we use 64, 128, and 256 respectively. We use a 7×7 kernel as a larger kernel removes more noise from the image [32]. But, this larger kernel size has a disadvantage. It blurs out edges more than 3×3 kernel size. That means the 7×7 kernel removes more noise but can extract less information from the image. That is why a 3×3 kernel is used to extract information from the images. Before applying the 3×3 kernel, we use a 1×1 kernel, as it down-samples the input and produces smaller feature vectors for 3×3 convolution to work on. By reducing the no of parameters, we are reducing the no of unrelated features possible. This helps the model to learn features common to different situations and so to generalize better. We define padding as the same for the layer's outputs will have the same spatial dimensions as its inputs. The outputs of convolution layers are then passed through a linear activation function. We use the rectified linear unit (ReLU) that determines the function: $f(a) =$ maximum (0, a). We use max pooling and the size of the filter is 2×2 with a stride of 2 which down-samples the dimension of feature maps by a factor of 2. In fully connected layer output feature maps are converted into a 1D (one-dimension) array and each input is connected with every output by a learnable weight. Our first dense layer has 224 neurons and the final fully connected layer has two nodes as we are classifying two types of images. Dropouts reduce the chances of over-fitting by dropping neurons. Here we use 0.2 as it gives us a better performance. We use the sigmoid function in the decision-making layer which normalizes output real value from zero to one and the summation of all values is 1. If the value is greater than or equal to 0.5 then the model predicts the output as COVID positive otherwise non-COVID for a particular image. Table 1 summaries of models. Figure 2 summarizes the CNN model architecture which consists of an initial single convolution layer followed by 3 blocks of layers followed by a final dense layer. Also, the summary of our model is given below:

Table 1. Summary of our model

Layer	Output shape	Parameters
conv2d input (InputLayer)	[(None, 224, 224, 3)]	0
conv2d (Conv2D)	(None, 224, 224, 64)	9472
max-pooling2d (MaxPooling2D)	(None, 112, 112, 64)	0
conv2d 1 (Conv2D)	(None, 112, 112, 64)	4160
conv2d 2 (Conv2D)	(None, 112, 112, 64)	36928
max-pooling2d (MaxPooling2D)	(None, 56, 56, 64)	0
conv2d 3 (Conv2D)	(None, 56, 56, 128)	8320
conv2d 4 (Conv2D)	(None, 56, 56, 128)	147584
max-pooling2d (MaxPooling2D)	(None, 28, 28, 128)	0
conv2d 5 (Conv2D)	(None, 28, 28, 256)	33024
conv2d 6 (Conv2D)	(None, 28, 28, 256)	590080
max-pooling2d (MaxPooling2D)	(None, 14, 14, 256)	0
flatten (Flatten)	(None, 50176)	0
dense (Dense)	(None, 224)	11239648
dropout (Dropout)	(None, 224)	0
dense 1 (Dense)	(None, 1)	225
Total parameters: 12069441		
Trainable parameters: 12069441		
Non-trainable parameters: 0		

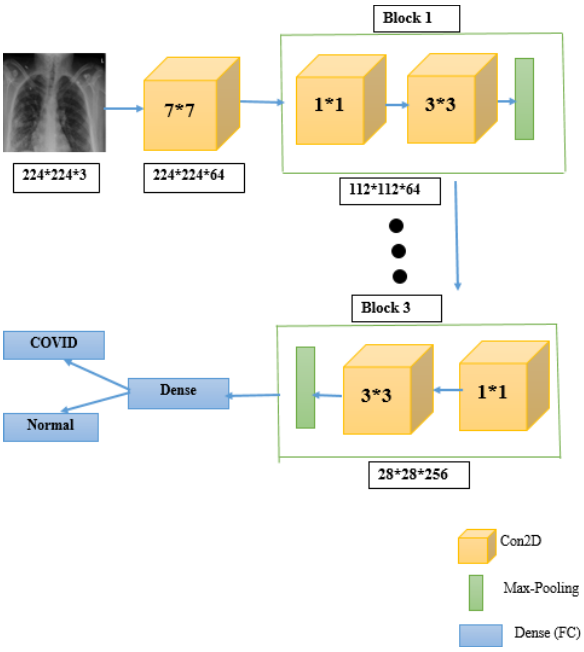

Figure 2. Proposed CNN architecture

2.4. Model training

We use 1,726 images (COVID-460 images and normal-1266 images) which is about 70% of our total dataset-1 to train our model. We also use 5-fold cross-validation to get a more accurate result with this medium range of datasets. In this respect, at first 100 epochs are used and also early stopping. We observe that 17 epochs are executed with patient=5. So, we use 20 epochs for all the models. We use the Adam optimizer for updating the weights, binary cross-entropy loss, and the default learning rate of 0.001.

2.5. Model testing

Model testing usually entails evaluating a trained model's resilience, generalization capacity, and predictive accuracy using a variety of datasets. We use the rest of the 30% images of dataset-1 as testing where there are 433 images (116 COVID images and 317 normal images). To observe how our model is generalized, we used an additional dataset (we call it dataset-2) where there are 96 images (30 COVID images and 66 normal images).

3. RESULTS AND DISCUSSION

This section narrates the experimental results, and performance of VGG19, ResNet50, ResNet101, NasNetMobile, and our proposed model separately and traces the best suitable model for the detection of COVID-19 performed on chest X-ray dataset. We evaluate the performances based on accuracy, precision, re-call, and F-1 score which can be computed by counting the number of accurate predictions using the terms true/false position/negative, where true positive (TP) indicates that a COVID case is truly identified as COVID, true negative (TN) indicates that a normal instance is truly identified as normal, false positive (FP) indicates that the event is normal and is identified as COVID-19, and false negative (FN) indicates that the thing is COVID and is identified as normal. Precision (P) is the ratio of TP to the summation of TP and FP. recall (R) is the ratio of TP to the summation of TP and FN F-1 score (F1) is the harmonic mean of the previous two terms which are precision and recall. The correctness of our model is assessed by performing the experiment.

Table 2 describes the results acquired by diverse popular CNN transfer learning models to diagnose COVID-19 automatically by using Chest X-ray images. It can be observed that VGG-19 outperforms in both accuracy and F1 score which are 99.54% and 98.68% for dataset 1. Whereas ResNet-101 has the lowest accuracy and F1 score of 86% and 84.50% respectively.

Figure 3 describes the training and testing accuracy when we use 100 epochs (using 11 CNN layers) and set the early stopping at patient=5 based on the validation accuracy. The X-axis denotes the number of epochs and Y-axis represents accuracy. If the model does not get higher testing accuracy than the previous one for consecutive 5 times, then the learning is stopped. For the training part, from the figure, it can be seen that at first the accuracy is below 75% but it increases dramatically at the 2nd epoch and achieves about 90% accuracy and at the end, it achieves about 99%. The testing part, it is also seen some fluctuation. Though the beginning testing accuracy is near 80%, it increases by about 14% from the previous one. After epoch 10 the model is unable to increase the accuracy for the next 5 consecutive epochs and stops. The same types of results are experienced for the 10 layers, 8 layers, 7 layers, 6 layers, and 5 layers CNN model. That is why we use 20 epochs for our next experiments instead of 100 epochs.

From Figure 4 it can be observed that all the CNN architectures performed well in terms of 5-fold average testing accuracy and 7 layers CNN architecture has the highest value which is 99.81%. On the other hand, for cross-validation, some fluctuations in accuracy were observed. At first, for 11-layer CNN, the accuracy is 98.96% then it goes down for 10-layer CNN which is about 82% and for 7-layer CNN it hits the highest point (100% accuracy). Then again it shows some downfall and reaches the lower point in 5 layers CNN which is about 72%.

Figure 5 describes the number of trainable parameters in terms of layers where the X-axis represents the number of layers and the Y-axis represents the number of trainable parameters. The curve shows some fluctuations. When the number of layers of CNN architecture is 11 then the trainable parameter is 27,507,265 but it falls dramatically when we reduce one layer. For, the 8-layer CNN architecture the number of trainable parameters increases instantly and achieves the highest value which is 46,050,881.

Table 2. Performances of different transfer learning models

Model name	Dataset 1		Dataset 2	
Accuracy	F1 score	Accuracy	F1-score	
VGG-19	99.54	98.68	98.96	98.50
ResNet-50	99.00	98.50	76.08	76.50
ResNet-101	86.00	84.50	62.50	62.00
NasNetMobile	99.00	98.50	98.96	98.50

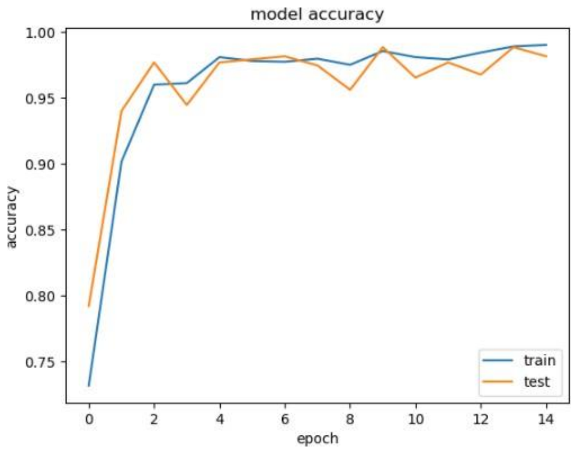

Figure 3. Training and testing accuracy with early stopping

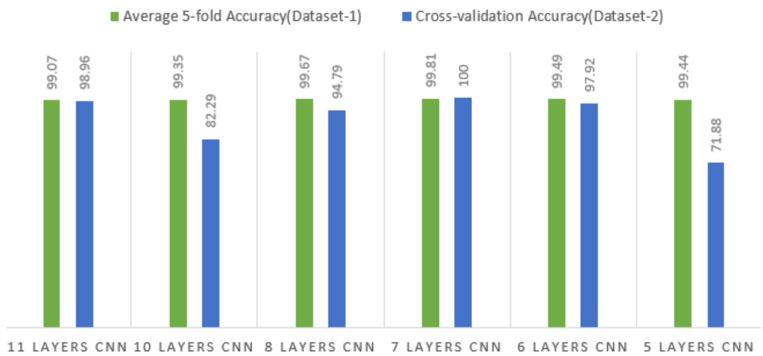

Figure 4. Accuracy in different layers CNN model for dataset-1 and dataset-2

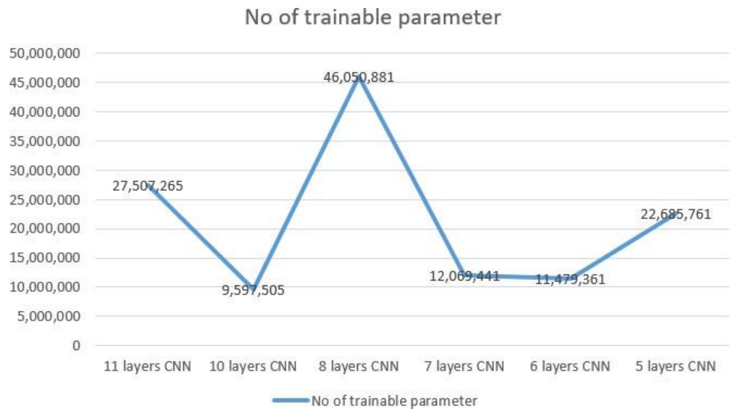

Figure 5. Number of trainable parameters in different CNN architectures

Table 3 describes the precision, recall, F1-score, and support for CNN architectures of different depths for cross-validation datasets. We get the best result for 7 layers of CNN architecture where for both true and false, we get 100% accurate results in terms of precision, recall, and f1 score. On the other hand, we get the lowest values for 5-layers of CNN architecture. 11-layer CNN architecture achieve the 2nd highest F1-score which is 0.99.

Figure 6 shows the confusion metrics of our proposed CNN architecture. Table 4 describes weighted memory and execution time for some popular transfer learning CNN models and our proposed CNN model. For the execution time, all the samples of dataset 2 are used (96 images). It can be seen that our model has the lowest weight which is 23,808 and execution time is 4 seconds which is the lowest. VGG-19 has the second-highest weight which is 154,112 but it takes the highest execution time of 42.02 seconds. NasNetMobile has the highest amount of weights but it takes less time to classify the images which is about 7 seconds.

Table 3. Precision, recall, F1 Score and support of different layers CNN architecture for dataset-2

Model	Precision	Recall	F1-score
11 layer	0.975	0.985	0.985
10 layer	0.87	0.82	0.812
8 layer	0.96	0.93	0.94
7 layer	1.00	1.00	1.00
6 layer	0.985	0.97	0.975
5 layer	0.795	0.765	0.715

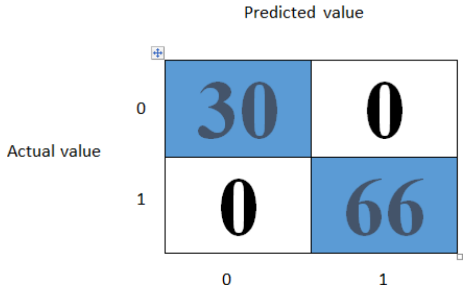

Figure 6. Confusion matrix of 7 layers CNN model

Table 4. Weighted memory and execution time of different models for dataset-2

Model	Weights (convolution layers)	Execution time (seconds)
VGG-19	154,112	42.02
ResNet-50	390,144	25.08
ResNet-101	390,144	15.12
NasNetMobile	1,013,888	7.08
Ours	23,808	4.00

4. DISCUSSION

From the result section it is seen that pre-trained deep convolution neural networks perform well for chest X-ray imaging problems. But we tried to address a question, is it possible to do better or similar by using a lightweight model with less run time? For this, we have built a lightweight CNN model that generalizes well and outperforms, or at least performs at the same level as, the pre-trained models. VGG-19 is one of the most popular transfer learning CNN models to solve real-world deep learning problems and in our COVID-19 detection problem, it performs best. Because VGG-19 uses higher convolution layers and only one kernel size (3×3) and for the non-linear multiple layers the depth of the network increases and it is possible for the architecture to learn more complicated features. In dataset 2, VGG-19 correctly identified almost all the images which is why the accuracy and F1 score are near to 100%. However, ResNet-101 cannot identify the true negative (normal images) properly and gives an accuracy of 62.50%. Because ResNet works on the basis of the skip connection technique and sometimes memorizes the training data instead of realizing the feature [33]. So, this model may perform worse for a limited dataset as used in this study.

Regarding searching for a lightweight CNN model, we investigated different possible architectures and found a more efficient and generalized CNN model that performs well in terms of time and space complexity. In this respect, we use CNN architectures with six different depths including 11-layer, 10-layer, 8-layer, 7-layer, 6-layer, and 5-layer CNN models. From the experiment, we observed that the CNN model with 7-layers performed best. For the 7-layer CNN model, the average testing accuracy for dataset-1 is 99.81% and for dataset-2 this accuracy is 100%. That means the 7-layer CNN model generalizes best for the identification of COVID-19 chest X-ray images. Though we do not use more layers as compared to VGG-19 (which gives 98.96% accuracy for dataset-2), we were able to achieve the best result. We use a 7×7 kernel as our first input layer which helps us to remove the noise from the images and a 1×1 kernel which down-samples the input and helps us to extract information. When we increase the layers by adding another layer (8 layers CNN model) the number of trainable parameters increased by about 4 times but the accuracy decreased. That means it is not mandatory to get higher accuracy with the help of a higher number of trainable parameters and we can get better accuracy by using a suitable architecture. For the 11-layer CNN model, the average testing accuracy is 99.074% for dataset-1 and the total training parameters are 27,507,265 which is the 2nd highest, and achieved 98.96% cross-validation accuracy for dataset-2. For 10 CNN layers, the trainable parameters are 9,597,505 which yields similar testing performance for dataset-1 but for dataset-2 the validation accuracy decreased by about 16%. We reduced the last layer (used in the 11-layer CNN model) which contains a 1×1 kernel and this kernel is used to produce smaller feature vectors for which the performance dropped. CNN model with 7-layer depth seems suitable for the current study and any shallow or deep version yields poor accuracy, thus we finally use the 7-layer CNN model as our proposed model.

For both datasets 1 and 2, transfer learning (TL) architectures perform well as they are already pre-trained by the image-net dataset. The main favor of this TL is that the models are already generalized for the classification problems by extracting features for a target problem where there is insufficient data. On the other hand, the customized model seems to perform better for both sufficient and insufficient data. Initially, most of the customized models cannot start with higher accuracy as transfer learning but as they learn through training epochs, their learning increases and is found to achieve similar performances to the transfer learning models.

Table 5. Comparison between different state of art models and our model

Name	K-fold	Methodology	ACC (2-class)	ACC (multi-class)
Ozturk *et al.* [26]	5- fold	17 layers CNN (DarkCOVIDNet)	98.08%	87.02%
Wang *et al.* [27]	NA	19 layers CNN (ResNet variant)	-	93.30%
Togacar *et al.* [28]	5- fold	MobileNetV2-SqueezeNet & SVM	99.27%	-
Hussain *et. al.* [29]	5- fold	22 layers CNN	99.1%	94.2%
Ours	5- fold	7 layers CNN	100%	-

From Table 5 it is seen that four state-of-the-art CNN models proposed by different authors can identify COVID with $>$ 87% accuracy using chest X-ray images. Hussain *et al.* [29] got a higher accuracy of 99.1% for binary classification by using 22 22-layer CNN model. On the other hand, our model outperformed by achieving 100% accuracy with only a 7-layer deep CNN model which has a comparatively small memory footprint. Our proposed model provides similar performance to some existing transfer learning models but in terms of time and space complexity, our model performs better than the others, see in Table 4. Considering the space complexity, for a convolution layer the weighted memory can be calculated from the kernel size ($n \times n$), no of filters (f), no of channels (c), and image shape [34].

$$Total\ Weights = \sum_{i=1}^{x} (n_i * n_i) * f_i * c_i \tag{1}$$

where x is the total number of layers in the architecture. In our model, the weights of the first convolution layer is 9,472 (7×7 kernel, 64 filters, 3 channels and 224×224 image shape, yields 7×7×3×64+64=9,472), the second convolution layer is 256(=1×1×3×64+64), the third convolution layer is 1792(=3×3×3×64+64), the fourth to seventh layers are 512, 3,584, 1,024 and 7,168 respectively which yields a total weight of 23,808. The weights will be increased if the convolution layer increases. For the other models (TL models in Table 4), the no of convolution layer is higher than our proposed model, so the total weights are also higher which consumes morememory.

5. CONCLUSION

In this study, we propound a light weighted CNN model for predicting and classifying COVID-19 vs normal X-ray images. The exploration results conclude that our model is best suited for the prediction of COVID-19. Our proposed model is susceptible to redact binary classification tasks with an accuracy of 100%. The performance evaluation argues that it outsails some similar existing models. According to these results, our proposed model seems to be an applicable tool that can lend a helping hand to the medical staff to diagnose and predict COVID-19 infection cases in a measurable time which will allow them to make faster decisions about the quarantine of patients. Also, it may reduce the pressure on medical staff and quicken the process of identifying COVID-19 events. The utility of our mentioned model for coronavirus detection can also be thought to be useful for other chest-related disease detection including tuberculosis and pneumonia, however, this requires further investigation to ascertain, which forms a natural extension of current work. The main attenuation of this study is that the model is underpinned using an average number of COVID-19 samples and we plan to validate our approach using a large and diverse dataset.

REFERENCES

[1] C. Ouchicha, O. Ammor, and M. Meknassi, "CVDNet: a novel deep learning architecture for detection of coronavirus (COVID-19) from chest x-ray images," *Chaos, Solitons & Fractals*, vol. 140, Nov. 2020, doi: 10.1016/j.chaos.2020.110245.

[2] O. Filchakova, D. Dossym, A. Ilyas, T. Kuanysheva, A. Abdizhamil, and R. Bukasov, "Review of COVID-19 testing and diagnostic methods," *Talanta*, vol. 244, Jul. 2022, doi: 10.1016/j.talanta.2022.123409.

[3] N. Narayan Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays," *Irbm*, vol. 43, no. 2, pp. 114–119, Apr. 2022, doi: 10.1016/j.irbm.2020.07.001.

[4] E. Ayan and H. M. Unver, "Diagnosis of Pneumonia from chest x-ray images using deep learning," in *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, Apr. 2019, pp. 1–5, doi: 10.1109/EBBT.2019.8741582.

[5] F. Bel and S. Selvaraj, "Efficient machine learning classifier to detect and monitor COVID-19 cases based on internet of things framework," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 4, pp. 4605–4611, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4605-4611.

[6] P. H. Khotimah *et al.*, "Monitoring Indonesian online news for COVID-19 event detection using deep learning," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 1, pp. 957–971, Feb. 2023, doi: 10.11591/ijece.v13i1.pp957-971.

[7] Z. Iklima, T. M. Kadarina, and E. Ihsanto, "Realistic image synthesis of COVID-19 chest X-rays using depthwise boundary equilibrium generative adversarial networks," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 5, pp. 5444–5454, Oct. 2022, doi: 10.11591/ijece.v12i5.pp5444-5454.

[8] M. Y. Kamil, "A deep learning framework to detect Covid-19 disease via chest X-ray and CT scan images," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 1, pp. 844–850, Feb. 2021, doi: 10.11591/ijece.v11i1.pp844-850.

[9] S. Haque, M. A. Hoque, M. A. I. Khan, and S. Ahmed, "COVID-19 detection using feature extraction and semi-supervised learning from chest X-ray images," in *2021 IEEE Region 10 Symposium (TENSYMP)*, Aug. 2021, pp. 1–5, doi: 10.1109/TENSYMP52854.2021.9550977.

[10] H. Maghdid, A. T. Asaad, K. Z. G. Ghafoor, A. S. Sadiq, S. Mirjalili, and M. K. K. Khan, "Diagnosing COVID-19 pneumonia from x-ray and CT images using deep learning and transfer learning algorithms," in *Multimodal Image Exploitation and Learning 2021*, Apr. 2021, p. 26, doi: 10.1117/12.2588672.

[11] H. Kang *et al.*, "Diagnosis of Coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2606–2614, Aug. 2020, doi: 10.1109/TMI.2020.2992546.

[12] X. Xu *et al.*, "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, Oct. 2020, doi: 10.1016/j.eng.2020.04.010.

[13] B. Ghoshal and A. Tucker, "Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection," *arxiv.org/abs/2003.10769*, Mar. 2020.

[14] C. Fang *et al.*, "Deep learning for predicting COVID-19 malignant progression," *Medical Image Analysis*, vol. 72, Aug. 2021, doi: 10.1016/j.media.2021.102096.

[15] B. Wang *et al.*, "AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system," *Applied Soft Computing*, vol. 98, Jan. 2021, doi: 10.1016/j.asoc.2020.106897.

[16] C. Jin *et al.*, "Development and evaluation of an artificial intelligence system for COVID-19 diagnosis," *Nature Communications*, vol. 11, no. 1, Oct. 2020, doi: 10.1038/s41467-020-18685-1.

[17] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "COVID-CAPS: a capsule network-based framework for identification of COVID-19 cases from X-ray images," *Pattern Recognition Letters*, vol. 138, pp. 638–643, Oct. 2020, doi: 10.1016/j.patrec.2020.09.010.

[18] Y. Song *et al.*, "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *IEEE/ACM*

Transactions on Computational Biology and Bioinformatics, vol. 18, no. 6, pp. 2775–2780, Nov. 2021, doi: 10.1109/TCBB.2021.3065361.
[19] X. Wang *et al.*, "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2615–2625, Aug. 2020, doi: 10.1109/TMI.2020.2995965.
[20] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images," *Sustainable Machine Intelligence Journal*, vol. 6, Mar. 2020, doi: 10.61356/smij.2024.66104.
[21] A. Bhattacharyya, D. Bhaik, S. Kumar, P. Thakur, R. Sharma, and R. B. Pachori, "A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images," *Biomedical Signal Processing and Control*, vol. 71, Jan. 2022, doi: 10.1016/j.bspc.2021.103182.
[22] I. D. Apostolopoulos and T. A. Mpesiana, "COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.
[23] F. Ucar and D. Korkmaz, "COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images," *Medical Hypotheses*, vol. 140, Jul. 2020, doi: 10.1016/j.mehy.2020.109761.
[24] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 1207–1220, May 2021, doi: 10.1007/s10044-021-00984-y.
[25] M. E. Khoudour and I. Biskri, "Detection of Coronavirus disease (COVID-19) based on deep features and support vector machine," *2023 International Conference on Machine Learning and Applications (ICMLA)*, pp. 1332–1337, Dec. 2023, doi: 10.1109/ICMLA58977.2023.00201.
[26] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103792.
[27] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Scientific Reports*, vol. 10, no. 1, Nov. 2020, doi: 10.1038/s41598-020-76550-z.
[28] M. Toğaçar, B. Ergen, and Z. Cömert, "COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches," *Computers in Biology and Medicine*, vol. 121, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103805.
[29] E. Hussain, M. Hasan, M. A. Rahman, I. Lee, T. Tamanna, and M. Z. Parvez, "CoroDet: a deep learning based classification for COVID-19 detection using chest X-ray images," *Chaos, Solitons and Fractals*, vol. 142, Jan. 2021, doi: 10.1016/j.chaos.2020.110495.
[30] USP Anupam, "Covid-chestXray," *Kaggle.com*, https://www.kaggle.com/datasets/dapy15/covidchestxray (accessed Jun. 05, 2022).
[31] A. D. Singh, "COVID-19 preprocessed x-ray dataset." *Kaggle.com*, https://www.kaggle.com/datasets/amritpal333/covid19-preprocessed-dataset (accessed Jun. 05, 2022).
[32] M. S. Bhattacharjee, "Noise removal from digital mammogram images for proper diagnosis of breast cancer," *International Journal for Research in Applied Science and Engineering Technology*, vol. 7, no. 3, pp. 331–337, Mar. 2019, doi: 10.22214/ijraset.2019.3058.
[33] "Residual networks (ResNet)–deep learning," *geeksforgeeks.org*, www.geeksforgeeks.org/residual-networks-resnet-deep-learning/ (accessed Oct. 12, 2022).
[34] M. Hollemans, "How fast is my Beowulf?" *machinethink.net*, https://machinethink.net/blog/how-fast-is-my-model/ (accessed Sep. 11, 2022).

BIOGRAPHIES OF AUTHORS

Md Jakaria Zobair completed his B.Sc. (engineering) in computer science and engineering from Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh, and M.Sc. (ongoing) in computer science and engineering at East West University, Dhaka, Bangladesh. Currently, he is working as a lecturer in the Department of Computer Science and Engineering at Daffodil International University, Dhaka, Bangladesh. Before that, he worked as a lecturer in the Department of Computer Science and Engineering at German University Bangladesh, Gazipur, Bangladesh. He is a fast learner and his research interests are in machine learning, deep learning, and medical image processing. He can be contacted at email: zobair.cse@diu.edu.bd.

Refat Tasfia Orpa received her B.Sc. (engineering) in computer science and engineering from Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh. Her research interests are in machine learning, deep learning, and artificial intelligence. She can be contacted at email: refattasfiacse16@gmail.com.

Mahir Ashef received his B.Sc. (engineering) in computer science and engineering from Bangladesh University of Business and Technology, Dhaka, Bangladesh. His research interests are in machine learning, deep learning, and artificial intelligence. Currently he is working as a software engineer in Dohatech, Bangladesh. He can be contacted at email: mahirasef30@gmail.com.

Shah Md Tanvir Siddiquee is working as an assistant professor at the Department of Computer Science and Engineering, Daffodil International University, Daffodil Smart City, Dhaka, Bangladesh. His research interests include topics such as cloud computing, networking, tax analysis, and machine learning. He has published his research outcomes in several international peer-reviewed journals and conferences. He can be contacted at email: tanvir.cse@diu.edu.bd.

Narayan Ranjan Chakraborty is working as an associate head and associate professor at the Department of Computer Science and Engineering, Daffodil International University, Daffodil Smart City, Dhaka, Bangladesh. His research interests include topics such as cyber security, information security, text mining, machine learning, and natural language processing. He has published his research outcomes in several international peer-reviewed journals and conferences. He can be contacted at: narayan@daffodilvarsity.edu.bd.

Ahsan Habib is working as a lecturer, School of Information Technology, Deakin University, Geelong, VIC 3225, Australia. He received his PhD from School of Information Technology, Deakin University, Burwood, VIC, Australia. He is a researcher, academician, entrepreneur, and software development enthusiast. The research area is emerging technologies in healthcare especially, interpretable deep learning. He has published his research outcomes in several international peer-reviewed journals and conferences. He can be contacted at email: m.habib@deakin.edu.au.