

# Detection of COVID-19 from chest x-ray images using transfer learning

Jenita Manokaran<sup>1</sup>,<sup>a,\*</sup> Fatemeh Zabihollahy<sup>2</sup>,<sup>b</sup>  
Andrew Hamilton-Wright,<sup>c</sup> and Eranga Ukwatta<sup>1</sup>,<sup>a</sup>

<sup>a</sup>University of Guelph, School of Engineering, Biomedical Engineering, Guelph, Ontario, Canada

<sup>b</sup>The Johns Hopkins University, School of Medicine, Department of Radiation Oncology and Molecular Radiation Sciences, Baltimore, Maryland, United States

<sup>c</sup>University of Guelph, School of Computer Science, Guelph, Ontario, Canada

## Abstract

**Purpose:** The objective of this study is to develop and evaluate a fully automated, deep learning-based method for detection of COVID-19 infection from chest x-ray images.

**Approach:** The proposed model was developed by replacing the final classifier layer in DenseNet201 with a new network consisting of global averaging layer, batch normalization layer, a dense layer with ReLU activation, and a final classification layer. Then, we performed an end-to-end training using the initial pretrained weights on all the layers. Our model was trained using a total of 8644 images with 4000 images each in normal and pneumonia cases and 644 in COVID-19 cases representing a large real dataset. The proposed method was evaluated based on accuracy, sensitivity, specificity, ROC curve, and  $F1$ -score using a test dataset comprising 1729 images (129 COVID-19, 800 normal, and 800 pneumonia). As a benchmark, we also compared the results of our method with those of seven state-of-the-art pretrained models and with a lightweight CNN architecture designed from scratch.

**Results:** The proposed model based on DenseNet201 was able to achieve an accuracy of 94% in detecting COVID-19 and an overall accuracy of 92.19%. The model was able to achieve an AUC of 0.99 for COVID-19, 0.97 for normal, and 0.97 for pneumonia. The model was able to outperform alternative models in terms of overall accuracy, sensitivity, and specificity.

**Conclusions:** Our proposed automated diagnostic model yielded an accuracy of 94% in the initial screening of COVID-19 patients and an overall accuracy of 92.19% using chest x-ray images.

© 2021 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.8.S1.017503](https://doi.org/10.1117/1.JMI.8.S1.017503)]

**Keywords:** COVID-19; transfer learning; chest x-ray; pretrained models; deep learning.

Paper 21005SSRR received Jan. 11, 2021; accepted for publication Aug. 6, 2021; published online Aug. 23, 2021.

## 1 Introduction

Coronavirus 2019 (SARS-CoV-2 or COVID-19) is an infectious disease affecting the respiratory system, which has caused a global pandemic and widespread morbidity and mortality in humans. It has mild to severe symptoms similar to those of severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), resulting in multiorgan damage.<sup>1</sup> Currently, two tests are available for detection of COVID-19 in the affected patients: diagnostic tests (current infection) and antibody tests (past infection). Diagnostic tests, such as reverse transcription polymerase chain reaction (RT-PCR) and antigen tests, are used for rapid diagnosis of COVID-19. Since false positives (FPs) are more common in antigen tests, RT-PCR is used as the gold standard for diagnosis of the disease. RT-PCR tests require an intensive lab work to acquire the

---

\*Address all correspondence to Jenita Manokaran, [rajaman@uoguelph.ca](mailto:rajaman@uoguelph.ca)

results,<sup>2</sup> and the cost of the test is a major concern in many countries that have a private health system. Although the PCR and antigen test can now provide a rapid diagnosis, the assessment of the lungs using medical imaging will provide information on disease burden. Also, faster and earlier detection of COVID-19 would help in isolating the affected patients sooner to alleviate the disease spread.

Chest radiography (CXR) and computed tomography (CT) images are the conventional medical imaging modalities used in lung disease diagnosis.<sup>3,4</sup> Though CT images are extensively used in the COVID-19 diagnosis,<sup>5-7</sup> cost<sup>8</sup> and radiation exposure are major concerns. CXR are preferred over CT images as they have less exposure to radiation and extensively available.<sup>9</sup> Hence, in this study, CXR images are used for automatic diagnosis of COVID-19.

Deep learning techniques are widely used in various fields, such as computer vision, machine vision, and speech recognition, among which computer vision is one of the most popular fields in which promising results to have been obtained in image classification tasks.<sup>10-12</sup> In medical image analysis, deep learning has been widely investigated for computer-aided diagnosis and treatment.<sup>13,14</sup> Several state-of-art methods have been proposed for the diagnosis of COVID-19 using CXR images<sup>15-19</sup> and CT images<sup>5,20,21</sup> based on deep learning techniques. The transfer learning approaches based on deep learning have been preferred in detection of COVID-19 due to the limited available dataset. Several studies have been implemented for COVID-19 diagnosis using the pretrained model as a feature extractor by implementing transfer learning techniques.<sup>22-24</sup>

Table 1 summarizes recent studies describing methods developed for detecting COVID-19 from CXR images; the overall datasets used for training and testing the model, accuracy, and the sensitivity are provided. Ozturk et al.<sup>18</sup> developed the DarkCovidNet model for the detection of COVID-19 using multiclass classification. The model was developed using end-to-end architecture without adding any feature extraction techniques. Sensitivity of 85.35% was achieved by the proposed model. Wang and Wong<sup>25</sup> proposed COVID-Net, a deep convolutional neural network (DCNN) developed by adopting machine-driven design exploration strategy.

**Table 1** Methods described for diagnosis of COVID-19 from CXR and CT images.

Source	Imaging modality	Model	Classification	Total number of images	Sensitivity (%)	Accuracy (%)
Ozturk et al. <sup>18</sup>	Chest x-ray	DarkCovidNet	Three-class	1127	85.35	87.02
Wang et al. <sup>25</sup>	Chest x-ray	COVID-Net	Three-class	13975	91	93.3
Panwar et al. <sup>26</sup>	Chest x-ray	nCOVnet	Two-class	284	97.62	88
Sethy et al. <sup>23</sup>	Chest x-ray	ResNet50 plus SVM	Three-class	381	97.29	95.33
Jain et al. <sup>27</sup>	Chest x-ray	ResNet50 (stage-I)	Three-class (stage-I)	1832	97.14	93
		ResNet101 (stage-II)	Two-class (stage-II)			97.78
Narin et al. <sup>28</sup>	Chest x-ray	ResNet-50	Two-class	3141 (dataset 1)	91.8	98
				1834 (dataset 2)		
				3113 (dataset 3)		
Ioannis et al. <sup>22</sup>	Chest x-ray	MobileNetV2	Three-class	1442	98.66	94.72
Asnaoui et al. <sup>29</sup>	Chest x-ray and CT	Inception_ ResNetV2	Three-class	6087	92.11	92.18
Hemdan et al. <sup>30</sup>	Chest x-ray	COVIDX-Net	Two-class	50	83.33	90

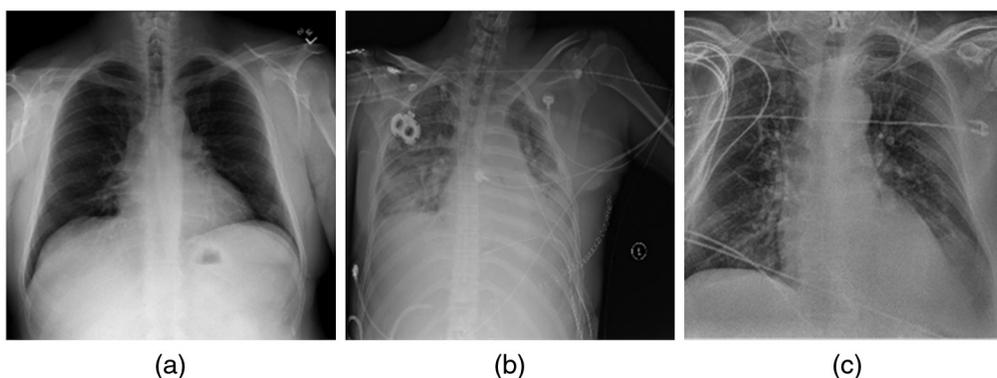
A multiclass classification was performed using CXR images in identifying COVID-19 from normal and non-COVID disease (i.e., pneumonia). Both qualitative and quantitative analysis were performed to determine the success of the model. The developed model was able to achieve a sensitivity of 91% for detecting COVID-19. Panwar et al.<sup>26</sup> introduced nCOVnet, an algorithm developed based on a transfer learning model for faster diagnosis of COVID-19 from CXR images. The proposed model is built using VGG 16 as the base model. The model was able to achieve 97.62% in detecting COVID-19. Sethy et al.<sup>23</sup> described a comparative study on Xception, ResNet18, ResNet50, Resnet 101, Inceptionv3, Inceptionresnetv2, GoogleNet, Densenet201, VGG16, VGG19, and AlexNet using the deep features extracted from the CNN layers and fed to the SVM classifier for classification. ResNet50 performed better compared with other classification models on a binary classification between viral pneumonia and COVID-19 with a classification accuracy of 95.38%. Jain et al.<sup>27</sup> utilized ResNet50 model that was used in stage-I network model for distinguishing viral (including COVID-19) from bacterial pneumonia and normal cases using CXR images. Based on the result obtained, ResNet101 was used in stage-II network model to classify COVID-19 from other viral pneumonia. The model achieved an accuracy of 97% in COVID-19 detection. Narin et al.<sup>28</sup> proposed three CNNs for the detection of COVID-19 through CXR images. The pretrained models chosen for comparative study were ResNet50, InceptionV3, and Inception-ResNetV2. A binary classification was performed between normal and COVID-19. ResNet50 model outperformed all the other models with the highest classification accuracy of 98%. Ioannis et al.<sup>22</sup> compared VGG 19, Mobile Net, Inception, Xception, and Inception ResNetv2 pretrained models for the automatic detection of COVID-19 using CXR images. They implemented multiclass classification among COVID-19, normal, and pneumonia. MobileNetV2 achieved highest sensitivity rate of 98.6%. Asnaoui et al.<sup>29</sup> put forward a comparative study on the pretrained DCNN models, namely VGG16, VGG19, Inception-ResNetV2, InceptionV3, ResNet50, DenseNet201, and MobileNetV2 for the classification of CXR into normal, bacteria, and coronavirus (multiclass classification). Inception-ResNetV2 model provided a sensitivity rate of 92.11% in detecting coronavirus. Hemdan et al.<sup>30</sup> proposed COVIDX-Net a deep learning framework based on seven DCNNs namely, VGG19, Xception, ResNetV2, InceptionV3, Inception-ResNetV2, DenseNet201, and MobileNetV2 for the diagnosis of COVID-19 using x-ray images. The VGG19 and DenseNet201 achieved an accuracy of 90% compared with other models with *F1*-score of 0.89 for normal and 0.91 for COVID-19.

One of the major limitations of the previous studies is the relatively smaller test dataset employed for the classification. Moreover, an imbalanced realistic representation of the patient population was not considered in the testing phase. Furthermore, some of the methods were designed to differentiate COVID-19 from normal cases by ignoring the pneumonia cases, or by combining bacterial, viral pneumonia, and COVID-19 under one class. To address these shortcomings and to further enhance the diagnosis of COVID-19, we propose a fully automated deep learning-based method for the detection of COVID-19 from CXR images. The proposed method was based on a transfer learning approach using dense convolutional neural network 201 (DenseNet201)<sup>31</sup> pretrained model. Our model was trained and tested using a dataset that reflected realistic class imbalance in the actual setting. As a benchmark, we compared results of our method to those of several state-of-the-art CNN architectures, including VGG16,<sup>32</sup> VGG19,<sup>32</sup> DenseNet121,<sup>31</sup> ResNet50,<sup>33</sup> ResNet101,<sup>33</sup> MobileNetv2,<sup>34</sup> and Inceptionv3.<sup>35</sup>

## 2 Materials and Methods

### 2.1 Dataset

Our dataset comprised 8644 CXR images of normal, respiratory distress syndrome (ARDS), COVID-19, MERS, pneumonia, and SARS from five open-access GitHub repositories<sup>36-40</sup> used in Wang et al.<sup>25</sup> It consisted of 4000 normal cases, 4000 pneumonia cases, and 644 COVID-19 cases. Out of 644 images of COVID-19, 20% were taken from COVID-19 Image Data Collection,<sup>36</sup> 5% from the Figure 1 COVID Chest X-ray Dataset,<sup>40</sup> 25% from the Actualmed COVID-19 Chest X-ray Dataset,<sup>39</sup> and 50% of images were obtained from the COVID-19

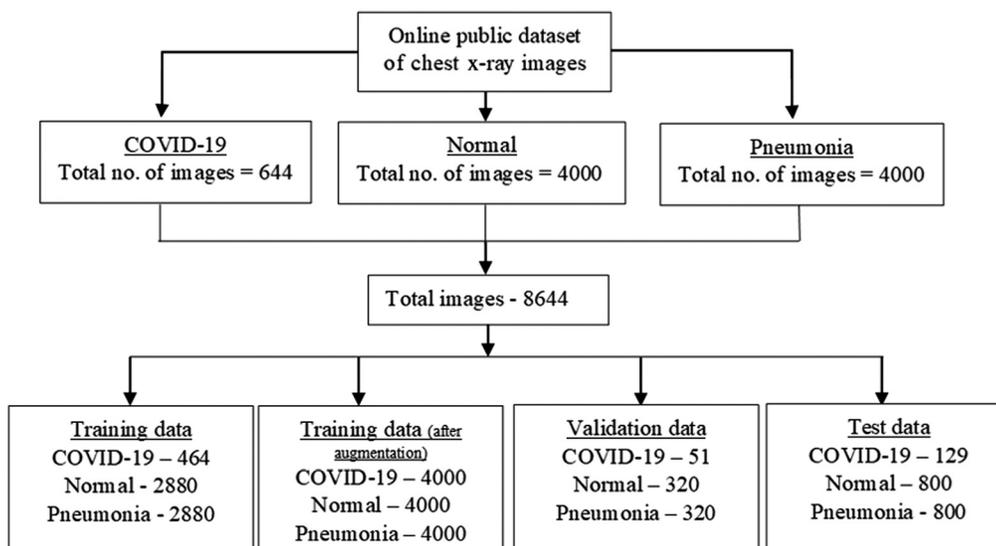


**Fig. 1** Examples of (a) normal, (b) pneumonia, and (c) COVID-19 in the dataset.

Radiography Database.<sup>38</sup> About 5% of pneumonia images were taken from COVID-19 Image Data Collection<sup>36</sup> and the remaining 95% from the RSNA Pneumonia Detection Challenge dataset.<sup>37</sup> All the normal images were obtained from the RSNA Pneumonia Detection Challenge dataset.

CXR images of COVID-19 affected patients have patchy and hazy lungs compared with normal, healthy lungs. Certain characteristic findings can be observed on the lower lobes and periphery of the lungs affected with COVID-19. Examples of normal, pneumonia, and COVID-19 are shown in Fig. 1. Since the purpose of this paper is to correctly classify COVID-19 cases, both bacterial and viral pneumonia are grouped under one label (pneumonia).

The CXR images are resized to  $224 \times 224$  to be compatible with the pretrained model used in this study. Except for Xception where the default input size of  $299 \times 299$  is used. All the images in the dataset used for training, validation, and testing are normalized between 0 and 1. Since the dataset is imbalanced, data augmentation techniques were incorporated to increase the training dataset. The training data of each class are increased to 4000 images summing up to 12,000 images in total. The images were horizontally flipped, magnified, and rotated for augmentation. The augmentation techniques were not applied on the validation and test data set. The 6223 (80%) of the dataset were used for training the model and the remaining 1729 (20%) of the dataset were used for testing the model. From the 80% of training dataset, 692 images (10%) were used for validation. A detailed description on the number of datasets used for training, validation, and test is provided in Fig. 2.



**Fig. 2** Distribution of the dataset among different classes.

## 2.2 Transfer Learning-Based Method

In this study, we developed our model based on DenseNet201,<sup>31</sup> which is a densely connected convolutional neural network that is 201 layers deep in which each layer receives collective knowledge from all preceding layers resulting in a compact network. It was designed to address the vanishing gradient issue in deep neural networks by employing skip connection similar to those of ResNet and has been proved to predict pneumonia at a higher rate using CXR images.<sup>41</sup> The fundamental difference between ResNet and DenseNet rests upon the fact that the features are combined using summation in ResNet, whereas they are concatenated in DenseNet leading to a dense connectivity pattern. This connectivity pattern introduced in DenseNet reduces the required training parameters resulting in a parametrically efficient model that has shown promising results in ImageNet and CIFAR-100 datasets. It is also memory and computationally efficient. DenseNet and ResNet have various versions, such as DenseNet121,<sup>31</sup> DenseNet201,<sup>31</sup> ResNet50,<sup>33</sup> and ResNet101,<sup>33</sup> which are considered in our study. The numbers represent the layers in the neural network.

The proposed model was then evaluated and further compared with the other state-of-the-art pretrained models, such as VGG16,<sup>32</sup> VGG19,<sup>32</sup> DenseNet121,<sup>31</sup> ResNet50,<sup>33</sup> ResNet101,<sup>33</sup> MobileNetv2,<sup>34</sup> and Inceptionv3,<sup>35</sup> and for a quantitative justification, we compared our proposed model with a lightweight CNN model designed from scratch. Based on our previous experimental results on classification using cascaded and multiclass classification, we determined that the multiclass classification yielded higher accuracy in COVID-19 identification.<sup>42</sup> Also, the segmented lungs from the CXR images classified COVID-19 with a low sensitivity value compared with the nonsegmented lungs.<sup>43</sup> Hence, a multiclass classification is performed using the proposed model to classify COVID-19 from normal and pneumonia classes using nonsegmented CXR images. The developed model uses initial pretrained weights and the last fully connected layer is removed and replaced with the new network developed using the global averaging layer, batch normalization layer, and a dense layer with ReLU activation. The final classification layer implements the Softmax activation function for multiclass classification. This is kept consistent for all the pretrained models considered in this work.

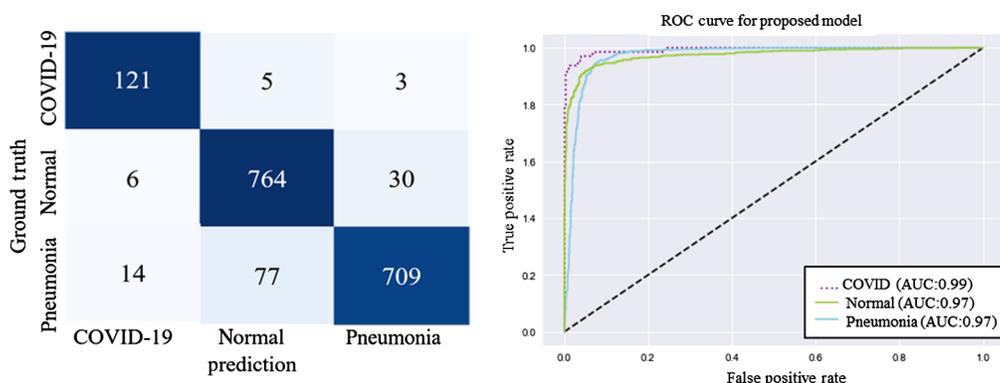
The models considered were trained for 100 epochs and the best fit was stored using the model check point that stores the model based on the maximum validation accuracy. The adaptive moment estimation (ADAM) optimizer with a learning rate of 0.0001 was used for training the model. An optimum value of 32 is considered as the batch size.<sup>44</sup> These parameters were kept the same for all the eight models used in this work. All the models were trained and tested using SHARCNET, a high-performance cluster in Canada.

## 2.3 Evaluation Metrics

The dataset considered is an imbalanced dataset that comprises an equal number of normal and pneumonia CXR images and a diminutive number of COVID-19 images. The classification accuracy cannot be determined using the overall accuracy as we are detecting the COVID-19 cases which is a minority class in the dataset considered, whereas both normal and pneumonia belong to the majority class. The minority class is regarded as the positive class, and the majority classes both normal and pneumonia are grouped under negative class. So, the confusion matrix is calculated to estimate the true positives (TP) that determines the correctly identified positive class, and the false negatives (FN) represents the positive classes that are misclassified as negative class. The true negatives (TN) determine the correctly identified negative class and FP indicates the number of negative classes misclassified as positive class. Since we are interested in determining the accuracy of the COVID-19 compared with the overall accuracy, sensitivity, specificity, and receiver operating characteristic (ROC) curve were scrutinized for model evaluation.

## 3 Results

The automated diagnostic tool developed using the DenseNet201 model for the detection of COVID-19 evaluated using the test dataset yielded an overall accuracy of 92% with sensitivity of 94%. The model was also able to classify 764 normal and 709 pneumonia out of 800 cases.



**Fig. 3** Confusion matrix and the ROC curve for the proposed model.

Considering COVID-19, it was able to classify 121 cases out of 129 cases. Only eight cases were misclassified. Most of the misclassification occurred between normal and pneumonia cases due to the early onset of pneumonia. The confusion matrix and the ROC curve for the proposed model are shown in Fig. 3. The proposed model was able to achieve area under the curve (AUC) of 0.99, 0.97 and 0.97 for COVID-19, normal and pneumonia classes, respectively.

We compared our method to seven state-of-the-art pretrained models. All the models were able to classify all the three classes (normal, pneumonia, and COVID-19) with a relatively high accuracy. Among 129 COVID-19 cases considered in the test dataset, all the models were able to classify more than 100 COVID-19 cases. Our proposed model based on DenseNet201 classified 121 cases bagging the first place, while ResNet50 was able to detect 117 cases holding the second place. When considering TN cases, VGG16, DenseNet121, DenseNet201, and ResNet101 were able to identify more than 700 cases correctly for both normal and pneumonia cases. VGG16 shows the least FP with two normal cases misidentified as COVID-19 and no FP for pneumonia cases. The least detection of TP was provided by VGG16 and highest FP was provided by the MobileNetV2 with least TN classification. The confusion matrix for all the seven pretrained models is summarized in Fig. 4.

The ROC curve for all the eight models used in this study is provided in Fig. 5, and AUC for all the three classes (normal, pneumonia, and COVID-19) were calculated and provided in the ROC curve. When the AUC for COVID-19 is taken into consideration, VGG16, VGG19, DenseNet121, and proposed model achieved the highest value of 0.99, and the least value of 0.95 was achieved by MobileNetV2. When the normal and pneumonia are examined, DenseNet121 obtained 0.98 for both the cases. A summarized AUC for all the classes provided in Table 2 shows that the DenseNet121 achieved higher AUC for all the three classes with 0.99 for COVID-19, 0.98 for normal, and 0.98 for pneumonia, and MobileNetV2 showed a lower result compared with the other models with 0.95 for COVID-19, 0.96 for normal, and 0.97 for pneumonia.

The evaluation metrics such as sensitivity, specificity, accuracy, and  $F1$ -score are summarized in Table 3. Considering the overall classification accuracy, DenseNet201 model was able to classify COVID-19 from normal and pneumonia with highest classification accuracy of 92% among all the other models considered. The DenseNet121 model was able to achieve the same classification accuracy of 92% as DenseNet201 model. VGG16, VGG19, ResNet101, and InceptionV3 obtained a classification accuracy of 90%, which is 2% less than the highest accuracy proposed.

The sensitivity and specificity rate are some of the important metrics to be considered when the dataset is imbalanced. Based on sensitivity rate, proposed model achieved highest rate of 0.94 and specificity of 0.99. DenseNet121 with less deep model compared with DenseNet201 achieved the same specificity rate but a lower sensitivity rate of 0.89. ResNet50 was able to achieve 0.91 which is second to the proposed model. Though ResNet50 achieved a higher sensitivity rate when compared with ResNet101, we can imply that the network depth can improve the overall classification accuracy but a reduction in the sensitivity is observed. VGG16 and VGG19 show higher accuracy as they are large network and provide better performance on the

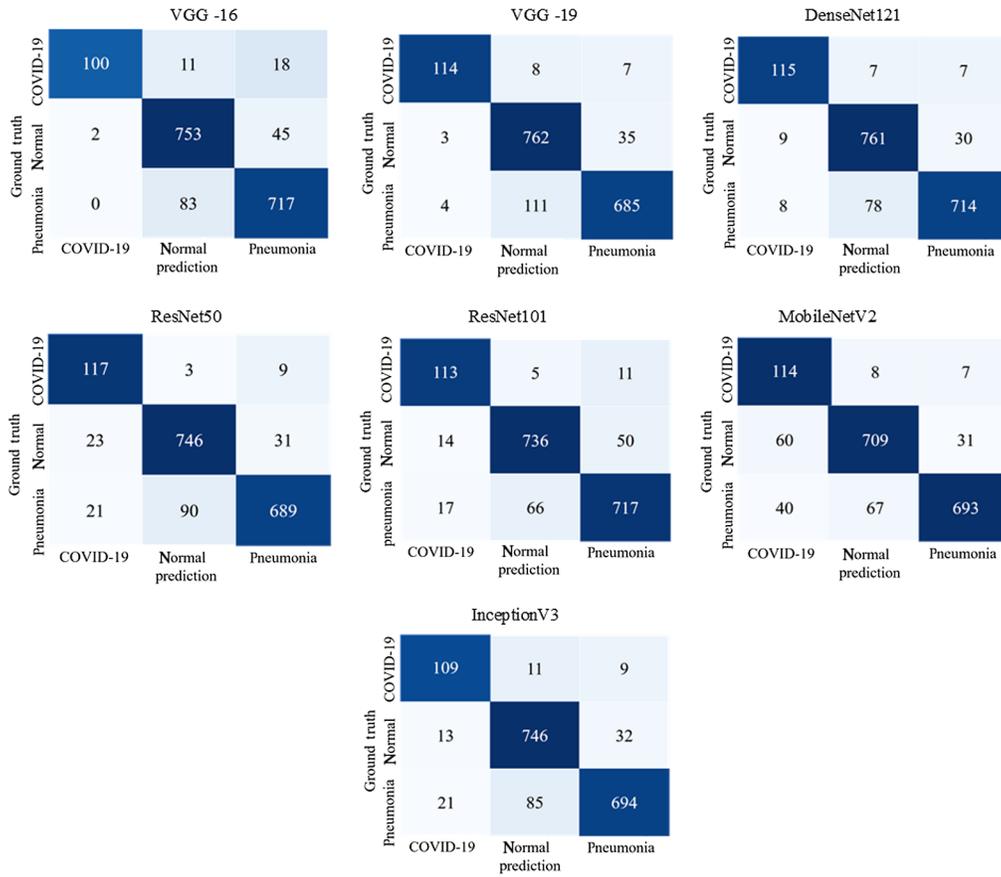


Fig. 4 Confusion matrix for all the seven models used for analytical analysis.

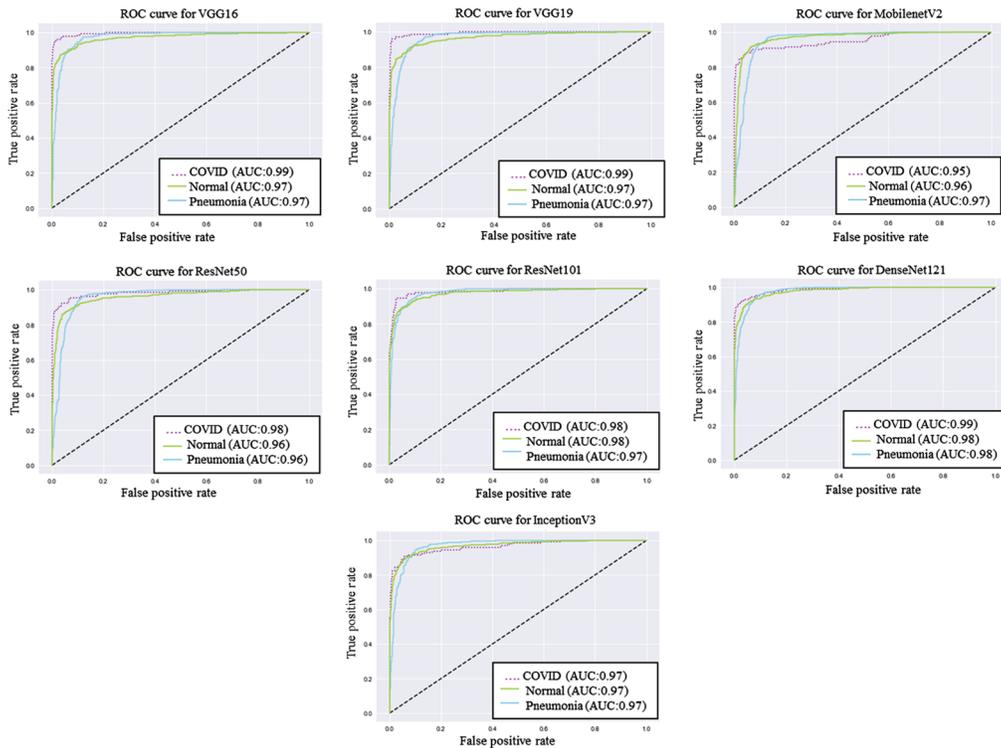


Fig. 5 ROC curve for all the seven models used for analytical analysis.

**Table 2** AUC for all three classes for all the eight models (the model with best result is highlighted in bold).

Model	COVID-19	Normal	Pneumonia
VGG16	0.99	0.97	0.97
VGG19	0.99	0.97	0.97
<b>DenseNet121</b>	<b>0.99</b>	<b>0.98</b>	<b>0.98</b>
Proposed model	0.99	0.97	0.97
ResNet50	0.98	0.96	0.96
ResNet101	0.98	0.98	0.97
MobileNetv2	0.95	0.96	0.97
Inceptionv3	0.97	0.97	0.97

larger dataset used in this work. The classification accuracy remains the same for both VGG16 and VGG19, whereas the sensitivity rate is higher for VGG19, and VGG16 has the least sensitivity rate in detecting COVID-19.

*F1*-score is included in the metrics to provide a balance between the exactness and completeness of the model. Considering the *F1*-score achieved by each model, VGG19 achieved a higher score of 0.91 and DensNet201 acquired 0.90, following VGG19. An exceptionally low *F1*-score of 0.66 was achieved by the MobileNetV2 model.

Also, to quantitatively justify, our proposed model was compared with a lightweight CNN designed from scratch. The lightweight CNN contains fewer parameters resulting in a faster CNN and provides a similar performance to the state-of-art pretrained models considered. The CNN architecture is made up of repeated blocks of convolution, batch normalization, and pooling as shown in Fig. 6. The architecture consists of five convolution and pooling layers and a fully connected classification layer. The convolutional and the hidden layers are followed by the ReLU activation function. The convolutional layer comprises 64, 128, and 256 neurons with convolution kernel of size  $3 \times 3$  with stride length of 1. The pooling layer comprises  $2 \times 2$  kernel with stride length 2. Before convolution, the inputs are padded using “same” padding in Tensorflow.

The results obtained by training and evaluating the model using the dataset are provided in Table 4. The model was able to achieve higher sensitivity and specificity value with an overall classification accuracy of 89.7%. Using the transfer learning approach, our proposed model was able to outperform in terms of overall classification accuracy, sensitivity, specificity, and *F1*-score, providing a more accurate diagnosis.

**Table 3** Evaluation metrics for eight pretrained models (model that achieved best result under each metric is highlighted in bold).

Model	Accuracy	Sensitivity	Specificity	<i>F1</i> -score
VGG16	0.908	0.78	<b>1</b>	0.87
VGG19	0.9028	0.88	0.99	<b>0.91</b>
DenseNet121	0.9196	0.89	0.99	0.88
<b>Proposed model (DenseNet201)</b>	<b>0.9219</b>	<b>0.94</b>	0.99	0.90
ResNet50	0.8976	0.91	0.97	0.81
ResNet101	0.9057	0.88	0.98	0.83
MobileNetv2	0.8768	0.88	0.93	0.66
Inceptionv3	0.9019	0.84	0.98	0.80



images were collected from various resources, a binary classification was performed using the developed model between normal and pneumonia cases obtained from the same source to eliminate the bias between the two classes (normal and pneumonia). We were able to achieve a classification accuracy of 95.75%, which shows that having data from different sources or the same sources does not influence the results, and our model is robust for images collected from different sources using different machines, but there is still a possibility for bias between COVID-19 and other classes decision boundary. Also, since we are dealing with x-ray images, they are more standardized compared with the other medical imaging modalities for detection of lung diseases.<sup>45,46</sup>

Our fully automated deep learning-based model based on DenseNet201 was able to detect COVID-19 with a higher accuracy based on a large and imbalanced dataset using CXR images. Considering the results obtained, the proposed model was able to outperform the other tested models in terms of overall accuracy, sensitivity, and specificity in identifying COVID-19 cases from normal and pneumonia cases. The analysis of the proposed model using higher number of pretrained models provided an insight on different model performance for medical image classification by implementing transfer learning approach using a dataset that is a large realistic representation of the patient population. Hence, the developed model utilizing CXR images will assist in the initial diagnosis of COVID-19 by the radiologists and will aid in subsequent analysis for clinical prognosis, including determining the severity of the disease affecting the lungs.

## Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

## Acknowledgments

The authors acknowledge the funding support from Natural Sciences and Engineering Council of Canada (NSERC) and COVID-19 seed grant by the University of Guelph.

## References

1. E. Estrada, "Fractional diffusion on the human proteome as an alternative to the multi-organ damage of SARS-CoV-2," *Chaos* **30**, 081104 (2020).
2. W. Wang et al., "Detection of SARS-CoV-2 in different types of clinical specimens," *J. Am. Med. Assoc.* **323**, 1843–1844 (2020).
3. A. K. Jaiswal et al., "Identifying pneumonia in chest x-rays: a deep learning approach," *Meas. J. Int. Meas. Confed.* **145**, 511–518 (2019).
4. M. Annarumma et al., "Automated triaging of adult chest radiographs with deep artificial neural networks," *Radiology* **291**, 272 (2019).
5. C. Zheng et al., "Deep learning-based detection for COVID-19 from chest CT using weak label," *IEEE Trans. Med. Imaging* **39**(8), 2615–2625 (2020).
6. X. Wang et al., "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Trans. Med. Imaging* **39**, 2615–2625 (2020).
7. H. Ko et al., "COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation," *J. Med. Internet Res.* **22**(6), e19569 (2020).
8. N. Garin et al., "High discordance of chest x-ray and CT for detection of pulmonary opacities in ED patients: implications for diagnosing pneumonia," *Am. J. Respir. Crit. Care Med.* **31**(2), 401–405 (2016).
9. A. Jacobi et al., "Portable chest x-ray in coronavirus disease-19 (COVID-19): a pictorial review," *Clin. Imaging* **64**, 35–42 (2020).
10. Q. Zheng et al., "Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process," *IEEE Access* **6**, 15844–15869 (2018).
11. Q. Zheng et al., "Layer-wise learning based stochastic gradient descent method for the optimization of deep convolutional neural network," *J. Intell. Fuzzy Syst.* **37**(4), 5641–5654 (2019).

12. Q. Zheng et al., “PAC-Bayesian framework-based drop-path method for 2D discriminative convolutional network pruning,” *Multidimens. Syst. Signal Process.* **31**, 793–827 (2020).
13. R. Fakoor et al., “Using deep learning to enhance cancer diagnosis and classification,” in *Proc. ICML Work. Role Mach. Learn. Transform. Healthc.* (2013).
14. A. Ortiz et al., “Ensembles of deep learning architectures for the early diagnosis of the Alzheimer’s disease,” (2016).
15. B. Ghoshal and A. Tucker, “Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection,” arXiv:2003.10769 (2020).
16. R. Kumar et al., “Accurate prediction of COVID-19 using chest x-ray images through deep feature learning model with SMOTE and machine learning classifiers,” medRxiv, 2020.04.13.20063461 (2020).
17. K. Medhi and I. Hussain, “Automatic detection of COVID-19 infection from chest x-ray using deep learning,” medRxiv, 2020.05.10.20097063 (2020).
18. T. Ozturk et al., “Automated detection of COVID-19 cases using deep neural networks with x-ray images,” *Comput. Biol. Med.* **121**, 103792 (2020).
19. S. H. Yoo et al., “Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest x-ray imaging,” *Front. Med.* **7**, 427 (2020).
20. Q. Ni et al., “A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images,” *Eur. Radiol.* **30**, 6517–6527 (2020).
21. J. Chen et al., “Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography,” *Sci. Rep.* **10**, 19196 (2020).
22. I. D. Apostolopoulos and T. Bessiana, “Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks,” *Phys. Eng. Sci. Med.* **43**, 635–640 (2020).
23. P. K. Sethy et al., “Detection of coronavirus disease (COVID-19) based on deep features,” *Int. J. Math. Eng. Manage. Sci.* **5**(4), 643–651 (2020).
24. A. M. Ismael and A. Şengür, “Deep learning approaches for COVID-19 detection based on chest x-ray images,” *Expert Syst. Appl.* **164**, 114054 (2021).
25. L. Wang and A. Wong, “COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest x-ray images,” *Sci. Rep.* **10**, 19549 (2020).
26. H. Panwar et al., “Application of deep learning for fast detection of COVID-19 in x-rays using nCOVnet,” *Chaos Solitons Fractals* **138**, 109944 (2020).
27. G. Jain et al., “A deep learning approach to detect COVID-19 coronavirus with x-ray images,” *Biocybern. Biomed. Eng.* **40**, 1391–1405 (2020).
28. A. Narin, C. Kaya, and Z. Pamuk, “Automatic detection of coronavirus disease (COVID-19) using x-ray images and deep convolutional neural networks,” *Pattern Anal. Appl.* **24**, 1207–1220 (2021).
29. K. El Asnaoui and Y. Chawki, “Using x-ray images and deep learning for automated detection of coronavirus disease,” *J. Biomol. Struct. Dyn.* **39**, 3615–3626 (2020).
30. E. E.-D. Hemdan et al., “COVIDX-net: a framework of deep learning classifiers to diagnose COVID-19 in x-ray images,” arXiv:2003.11055 (2020).
31. G. Huang et al., “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 4700–4708 (2017).
32. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *Comput. Biol. Learn. Soc.*, 1–14 (2015).
33. K. He et al., “Deep residual learning for image recognition,” in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, pp. 770–778, IEEE (2016).
34. M. Sandler et al., “MobileNetV2: inverted residuals and linear bottlenecks,” in *IEEE/CVP Conf. Comput. Vision and Pattern Recognit. (CVPR)*, pp. 4510–4520 (2018).
35. C. Szegedy et al., “Rethinking the inception architecture for computer vision,” in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, Las Vegas, Nevada, pp. 2818–2826 (2016).
36. J. Paul Cohen, P. Morrison, and L. Dao, “COVID-19 Image Data Collection,” arXiv 2003.11597 (2020).
37. Radiological Society of North America, RSNA Pneumonia Detection Challenge, <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data> (2019).

38. Radiological Society of North America, COVID-19 Radiography Database, <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database> (2019).
39. A. Chung, “Actualmed COVID-19 Chest X-ray Dataset Initiative,” <https://github.com/agchung/Actualmed-COVID-chestxray-dataset> (2020).
40. A. Chung, “Figure 1 COVID-19 Chest X-ray Dataset Initiative,” <https://github.com/agchung/Figure1-COVID-chestxray-dataset> (2020).
41. P. Rajpurkar et al., “CheXNet: radiologist-level pneumonia detection on chest x-rays with deep learning,” arXiv abs/1711.05225 (2017).
42. J. Manokaran et al., “Deep learning-based detection of COVID-19 from chest x-ray images,” *Proc. SPIE* **11600**, 116000O (2021).
43. L. O. Teixeira et al., “Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest x-ray images,” arXiv:2009.09780 (2020).
44. Y. Bengio, “Practical recommendations for gradient-based training of deep architectures,” *Lect. Notes Comput. Sci.* **7700**, 437–478 (2012).
45. M. Zak and A. Krzyżak, “Classification of lung diseases using deep learning models,” *Lect. Notes Comput. Sci.* **12139**, 621–634 (2020).
46. S. T. Hwa Kieu et al., “A survey of deep learning for lung disease detection on medical images: state-of-the-art, taxonomy, issues and future directions,” *J. Imaging* **6**(12), 131 (2020).

**Jenita Manokaran** is currently pursuing her PhD under Dr. Ukwatta on image processing based on artificial intelligence-based methods. She received her bachelor’s degree in electronics and communication engineering from Karunya University, India, in 2008. She graduated with a Master of Applied Science degree from the University of Windsor, Canada, in 2015. She has also worked as a professor of electronics engineering technology at St. Clair College, Windsor, Canada.

**Fatemeh Zabihollahy** is currently a postdoctoral research fellow at the Johns Hopkins University. She received her PhD in electrical and computer engineering in 2020 from Carleton University, Canada. She is the recipient of Carleton Medal for her outstanding graduate work at the PhD level. She worked in the medical devices industry as an R&D engineer for 10 years. Her research interest is in the field of application of deep learning techniques for medical image analysis.

**Andrew Hamilton-Wright** is an associate professor in the School of Engineering at the University of Guelph, Canada. He uses machine-learning and knowledge discovery techniques in clinically relevant biomedical applications. His research focus is on the application of computer-based decision support tools for important human decision-making problems, including the improvement of human understanding of physiological data. He has a total of 49 peer-reviewed technical publications within the last 10 years.

**Eranga Ukwatta** received his PhD from Western University, Canada, in 2013. He was a multi-center postdoctoral fellow with Johns Hopkins University and University of Toronto. He is currently an assistant professor at the University of Guelph, Canada, and an adjunct professor in systems and computer engineering with Carleton University, Canada. He has more than 90 journal articles and conference proceedings. His research interests include medical image segmentation and registration, deep learning, and computational modeling.