

Deep-learning convolutional neural networks with transfer learning accurately classify COVID19 lung infection on portable chest radiographs

Shreeja Kikkiseti, Jocelyn Zhu, Beiyi Shen, MD, Haifang Li, PhD, Tim Q Duong, PhD
Department of Radiology, Renaissance School of Medicine, Stony Brook University, New York
11794

Correspondence: Tim Duong, Ph.D., Radiology, Renaissance School of Medicine, Stony Brook
University, New York 11794 tim.duong@stonybrook.edu T 631 444-8033 F 631 444 7538

Funding: none

Manuscript Type: Original Investigation

Abbreviations

COVID-19	Coronavirus Disease 2019
CXR	chest x-ray
CNN	convolutional neural networks
AI	artificial intelligence

Acknowledgement

none

Word Count for Text: 2700

ABSTRACT

Portable chest x-ray (pCXR) has become an indispensable tool in the management of Coronavirus Disease 2019 (COVID-19) lung infection. This study employed deep-learning convolutional neural networks to classify COVID-19 lung infections on pCXR from normal and related lung infections to potentially enable more timely and accurate diagnosis. This retrospective study employed deep-learning convolutional neural network (CNN) with transfer learning to classify based on pCXRs COVID-19 pneumonia (N=455) on pCXR from normal (N=532), bacterial pneumonia (N=492), and non-COVID viral pneumonia (N=552). The data was split into 75% training and 25% testing. A five-fold cross-validation was used. Performance was evaluated using receiver-operating curve analysis. Comparison was made with CNN operated on the whole pCXR and segmented lungs. CNN accurately classified COVID-19 pCXR from those of normal, bacterial pneumonia, and non-COVID-19 viral pneumonia patients in a multiclass model. The overall sensitivity, specificity, accuracy, and AUC were 0.79, 0.93, and 0.79, 0.85 respectively (whole pCXR), and were 0.91, 0.93, 0.88, and 0.89 (CXR of segmented lung). The performance was generally better using segmented lungs. Heatmaps showed that CNN accurately localized areas of hazy appearance, ground glass opacity and/or consolidation on the pCXR. Deep-learning convolutional neural network with transfer learning accurately classifies COVID-19 on portable chest x-ray against normal, bacterial pneumonia or non-COVID viral pneumonia. This approach has the potential to help radiologists and frontline physicians by providing more timely and accurate diagnosis.

Key words:

coronavirus; machine learning; chest X-ray; computed tomography; lung infection

Introduction

Coronavirus Disease 2019 (COVID-19) is a highly infectious disease that causes severe respiratory illness (1,2). It was first reported in Wuhan, China in December 2019 (3) and was declared a pandemic on Mar 11, 2020 (4). The first confirmed case of coronavirus disease 2019 (COVID-19) in the United States was reported from Washington State on January 31, 2020.(5) Soon after, Washington, California and New York reported outbreaks. COVID-19 has already infected 10 million, killed more than 0.5 million people, and the United States has become the worst-affected country, with more than 2.4 million diagnosed cases and at least 122,796 deaths (<https://coronavirus.jhu.edu>, assessed Jun 28, 2020). There are recent spikes of COVID-19 infection cases across many states and around the world and there will likely be second waves and recurrence.

A definitive test of COVID-19 infection is the reverse transcription polymerase chain reaction (RT-PCR) of a nasopharyngeal or oropharyngeal swab specimen (6,7). Although RT-PCR has high specificity, it has low sensitivity, high false negative rate, and long turn-around time (6,7) (currently ~4 days although it is improving and other tests are becoming available (8)). By contrast, portable chest X-rays (pCXr) is convenient to perform, has a fast turnaround, and is well suited for imaging contagious patients and longitudinal monitoring of critically ill patients in the intensive care units because the equipment can be readily disinfected, preventing cross-infection. pCXr of COVID-19 infection has certain unique characteristics, such as predominance of bilateral, peripheral, and low lobes involvement, with ground-glass opacities with or without airspace consolidations as the disease progresses. These characteristics generally differ from other lung pathologies, such as bacterial pneumonia or other viral (non-COVID-19)

lung infection. Based on CXR and laboratory findings, clinicians might start patients on empirical treatment before the RT-PCR results become available or even if the RT-PCR come back negative due to high false negative rate of RT-PCR. Early treatment in COVID-19 patients is associated with better clinical outcomes. Similarly, computed tomography (CT), which offers relatively more detailed features (such as subtle ground-glass opacity (9,10)), has also been used in the context of COVID-19. However, CT suite and equipment are more challenging to disinfect, and thus it is much less suitable for examining patients suspected of or confirmed with contagious diseases in general and COVID-19 in particular. Longitudinal CT monitoring of critically ill patients in the intensive care units is also challenging. In short, pCXR has become an indispensable imaging tool in the management of COVID-19 infection, is often one of the first examinations a patient suspected of COVID-19 infection receives in the emergency room, and ideally used for longitudinal monitoring of critically ill patients in the intensive care units.

The usage of pCXR under the COVID-19 pandemic circumstances is unusual in many aspects. For instance, pCXR is preferred as it can be used at the bedside without moving the patients, but the imaging quality is not as good as conventional CXR (11). In addition, COVID-19 patients may not be able to take full inspirations during the examination, obscuring possible pathology, especially in the lower lung fields. Many sicker patients may be positioned on the side which compromises imaging quality. Thus, pCXR data under the COVID-19 pandemic circumstances are suboptimal and, thus, may be more challenging to interpret. Moreover, pCXR is increasingly read by non-chest radiologists in some hospitals due to increasing demands, resulting in reduced accuracy and efficiency.

pCXR images contain important clinical features that could be easily missed by the naked eyes. Computer-aided methods can improve efficiency and accuracy of pCXR interpretations,

which in turn provides more timely and relevant information to frontline physicians. Deep-learning artificial intelligence (AI) has become increasingly popular for analyzing diagnostic images (12,13). AI has the potential to facilitate disease diagnosis, staging of disease severity and longitudinal monitoring of disease progression.

One common machine-learning algorithm is the convolutional neural network (CNN) (14,15), which takes an input image, learns important features in the image such as size or intensity, and saves these parameters as weights and bias to differentiate types of images (16,17). CNN architecture is ideally suited for analyzing images. Moreover, the majority of machine learning algorithms to date are trained to solve specific tasks, working in isolation. Models have to be rebuilt from scratch if the feature-space distribution changes. Transfer learning overcomes the isolated learning paradigm by utilizing knowledge acquired for one task to solve related ones. Transfer learning in AI is particularly important for small sample size data because the pre-trained weights enable more efficient training and improved performance (18,19).

Many artificial intelligence (AI) algorithms based on deep-learning convolutional neural networks have been deployed for pCXR applications (20-24) and these algorithms can be readily repurposed for COVID-19 pandemic circumstances. While there are already many papers describing prevalence and radiographic features on pCXR of COVID-19 lung infection (see reviews (25,26)), there is a few peer-reviewed AI papers (27-32) and non-peer reviewed papers (33-36) to classify CXRs of COVID-19 patients from CXR of normals or related lung infections. The full potential of AI applications of pCXR under COVID-19 pandemic circumstances is not yet fully realized.

The goal of this pilot study is to employ deep-learning convolutional neural networks to classify normal, bacterial infection, and non-COVID-19 viral infection (such as influenza)

against COVID-19 infection on pCXR. The performance was evaluated using receiver-operating curve (ROC) analysis. Heatmaps were also generated to visualize and assessment the performance of the AI algorithm.

Materials and Methods

Data sources: This retrospective study used publicly available pCXR of i) COVID-19 infection, ii) non-COVID-19 viral infection, iii) bacterial pneumonia, and iv) normal subjects. The COVID-19 pCXR were downloaded from (19) on May 27th, 2020. The original download contained 673 CT or pCXR images of COVID-19, SARS, acute respiratory distress syndromes, pneumocystis, streptococcus, legionella, Chlamydomphila, E Coli, Klebsiella, lipoid, Varicella, and influenza. The final sample size for COVID-19 patients was 455 pCXR from 197 patients. We recognized that this dataset was a public, community-driven dataset and there are potential selection biases. A radiologist (BS) evaluated all images for quality and relevance and each case was COVID-19 positive based on available data. Thus, this dataset is useful and valid for the purpose of algorithm development.

The other datasets were taken from the established Kaggle chest X-ray image (pneumonia) dataset (<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>). Although the Kaggle database has a large sample size, we randomly selected a sample size comparable to that of COVID-19 to avoid asymmetric sample size bias that could skew sensitivity and specificity. The sample sizes chosen for bacterial pneumonia, non-COVID-19 viral pneumonia, and normal pCXR were 492, 552 and 532 patients, respectively. Similarly, a chest radiologist evaluated all images for quality.

CNN: The CNN architecture was based on VGG16, a convolutional neural network (37). The VGG16 model was used because it was pretrained on the ImageNet database and properly employs transfer learning which makes the training process efficient. The data was normalized first by transforming all files into RGB images and resizing them into 224x224 pixels to make them compatible with the VGG16 framework. Next, the images were one-hot-encoded and split into 75% training and 25% testing. VGG16 implements 13 convolutional layers: 5 Max Pooling layers and 3 Dense layers which sum up to 21 layers and 16 weight layers. Conv 1 has 64 filters while Conv 2 has 128 filters, Conv 3 has 256 filters while Conv 4 and Conv 5 have 512 filters. VGG-16 also uses weights pre-trained on the ImageNet dataset. The first two layers have 2 sublayers while the 4th and 5th layers have 3 sublayers. A max-pooling layer was used after each step in the model to down sample the input and identify its important features. All convolutional layers used rectified linear units (ReLUs) as an activation function because it adds a small number of learnable parameters. Three fully connected layers were used, each having 4096 nodes. Dropout layers were used, along with the Softmax function, to prevent overfitting. For data analysis, batch sizes of 32 were used to limit computational expense and trained for 50 epochs. Several optimizers were tested however, Adams optimization function gave the lowest validation loss. The learning rate was lowered from the recommended 0.01 to 0.001 to prevent overshooting the global minimum loss. Categorical cross entropy was used as a loss function since the loss value decreases as the predicted probability converges to the actual label. The VGG16 architecture was utilized for computation efficiency and ease to implement, for immediate translation potential.

CNN analysis was performed on the whole pCXR as well as virtually segmented lungs. Lung segmentation was performed using a CNN architecture with 22 convolutional layers, 4 max-pooling layers, and 4 merged layers for connectivity. A ReLu activation function was used with the Keras library. The output consisted of a mask of the segmented lungs. The segmented lungs were then fed into the CNN model for the Covid19 classification. This model was trained on the Montgomery dataset and achieved an IoU score of 0.956 and dice score of 0.972.

Heatmaps: To visualize the spatial location on the images that the CNN networks were paying attention to, heatmaps were generated with class activation maps algorithm (38). This was done by adding global average pooling into CNN and calculating gradient backpropagation given one specific output class to obtain the class activation maps, indicating the discriminative image regions CNN paid attention to.

Statistical methods and performance evaluation: Five-cross-validation was used. Performance of the prediction model used standard ROC analysis of the area under the curve (AUC), accuracy, sensitivity, specificity, precision, recall and F1 scores. Precision was computed using true positives divided by the sum of false positives and true positives; Recall was computed using the true positives divided by the sum of true positives and false negatives; F1 scores were the mean of recall and precision rates.

Results

Figure 2 shows examples of pCXR from a normal subject and from patients with different lung infections. COVID-19 is often characterized by ground-glass opacities with or

without nodular consolidation with predominance of bilateral, peripheral and lower lobes involvement. Non-COVID-19 viral pneumonia is often characterized by diffuse interstitial opacities, usually bilaterally. Bacterial pneumonia is often characterized by confluent areas of focal airspace consolidation.

Figure 3 shows the training and validation loss and accuracy as a function of the epoch. Loss decreases and accuracy improved with increasing epoch for both training and validation dataset. The accuracy typically reached > 0.8 .

The results of the multiclass CNN classification for the whole CXR in the form of the confusion matrix is shown in **Table 1**. The precision, recall and F1 scores for the whole pCXR are shown in **Table 2**. The overall precision, recall and F1 scores showed good to excellent performance. For CNN with transfer learning performed on the whole pCXR, the overall sensitivity, specificity, accuracy, and AUC were 0.79, 0.93, and 0.79, .84 respectively. For CNN performed on segmented lungs, the overall sensitivity, specificity, accuracy, and AUC were 0.91, 0.93, 0.88, 0.89 respectively. The performance was generally better using segmented lungs.

To visualize the spatial location on the images that the CNN networks were paying attention to for classification, heatmaps of the COVID-19 versus normal pCXR are shown in **Figure 4**. The CNN algorithm was able to localize the area of pathology on pCXR. For CNN performed on the whole pCXR, the majority of the hot spots were reasonably localized to regions of ground glass opacities and/or consolidations, but some hot spots were located outside the

lungs. For CNN performed on segmented lungs, the majority of the hot spots were reasonably localized to regions of ground glass opacities and/or consolidations, mostly as expected.

Discussion

This study developed and applied a deep-learning CNN algorithm with transfer learning to classify COVID-19 CXR from normal, bacterial pneumonia, and non-COVID viral pneumonia CXR in a multiclass model. Heatmaps showed reasonable localization of abnormalities in the lungs. The overall sensitivity, specificity, accuracy, and AUC were 0.91, 0.93, 0.88, and .89 respectively (segmented lungs).

There are a few AI studies to date using machine learning methods to classify CXRs of COVID-19, normal and related lung infections. By the time this paper is reviewed many more papers will be published. **Hurt et al.** used a U-net CNN algorithm to predict pixel-wise probability maps for pneumonia on CXR on 10 COVID-19 patients (27). No ROC analysis was performed. **Apostolopoulos and Mpesiana** used deep-learning algorithm to predict COVID-19 CXR with 98.66% sensitivity, 96.46% specificity, and 96.78% accuracy from a collection of 1427 CXRs of which 224 were COVID-19 CXRs (28). **Elaziz et al.** used an innovative feature selection algorithms and standard classifier to classify CXR between COVID-19 (N=216) and non-COVID-19 (N=1675). This method achieved accuracy rates of 96.09% and 98.09% for each of the respective datasets (29). Note that patient cohorts were highly asymmetric. **Murphy et al.** used an artificial intelligence to classify COVID-19 CXRs (N=223) from non-COVID-19 CXRs (N=231) with an 0.81 AUC and they also showed that AI outperformed expert readers (30). **Ozturk et al.** used an AI model to perform multiclass classification for COVID-19 (N=127) vs.

No-Findings (N=500) vs. Pneumonia (N=500) as well as a binary classification for COVID vs. No-Findings which achieved 87.02% and 98.08% accuracies, respectively (31). **Pereira et al.** performed a multiclass classification and a hierarchical classification for COVID-19 vs pneumonia vs no-finding using resampling algorithms, texture descriptors, and CNN. This model achieved a F1-Score of 0.65 for the multiclass approach and F1 score of 0.89 for the hierarchical classification (32). AUC and accuracy were not reported. A few non-peer reviewed pre-prints using AI to classify COVID-19 CXRs have also been reported (33-36). Our study had one of the larger cohorts, balanced sample sizes, and multi-class model. Our approach is also amongst the simplest AI models with comparable performance index, likely facilitate immediate clinical translation. Together, these studies indicate that AI has the potential to assist frontline physicians in distinguishing COVID-19 infection based on CXRs.

Heatmaps are informative tools to visualize regions that CNN algorithm pays attention to for detection. This is particular important given AI operates on high dimensional space. Such heatmaps enable reality checks and make AI interpretable with respect to clinical findings. Our algorithm showed that the majority of the hotspots were highly localized to abnormalities within the lungs, i.e., ground glass opacity and/or consolidation, albeit imperfect. The majority of the above-mentioned machine learning studies to classify COVID-19 CXRs did not provide heatmaps. We also noted that CNN on whole pCXR image resulted in some hot spots located outside the lungs. CNN of segmented lungs solved this problem. Another advantage of using segmented lung is reduced computational cost during training. Transfer learning also reduced computational cost, making this algorithm practical. The performance is generally better using segmented lungs.

Most COVID-19 positive patients showed significant abnormalities on pCXR (39). Some early studies have even suggested that pCXR could be used as a primary tool for COVID-19 screening in epidemic areas (39,40), which could complement swab testing which still has long turnaround time and non-significant false positive rate. In some cases, imaging revealed chest abnormalities even before swab tests confirm infection (41,42). In addition, pCXR can detect superimposed bacteria pneumonia, which necessitates urgent antibiotic treatment. pCXR can also suggest acute respiratory distress syndrome, which is associated with severe negative outcomes and necessitates immediate treatment. Together with the anticipated widespread shortage of intensive care units and mechanical ventilators in many hospitals, pCXR also has the potential to play a critical role in decision-making, especially in regards to which patients to admit to the ICU, put on mechanical ventilation, or when to safely extubate. A timely implementation of AI methods could help to realize the full potential of pCXR in this COVID-19 pandemic.

This pilot proof-of-principal study has several limitations. This is a retrospective study with a small sample size and the data sets used for training had limited alternative diagnoses. Although the Kaggle database has a large sample size for non-COVID-19 CXR, we chose the sample sizes to be comparable to that of COVID-19 to avoid asymmetric sample sizes that could skew sensitivity and specificity. Future studies will need to increase the COVID-19 sample size and include additional lung pathologies. The spatiotemporal characteristics on pCXR of COVID-19 infection and its relation to clinical outcomes are unknown. Future endeavors could include developing AI algorithms to stage severity, and predict progression, treatment response,

recurrence, and survival, to inform and advise risk management and resource allocation associated with the COVID-19 pandemic.

In conclusion, deep learning convolutional neural networks with transfer learning accurately classify COVID-19 pCXR from pCXR of normal, bacterial pneumonia, and non-COVID viral pneumonia patients in a multiclass model. This approach has the potential to help radiologists and frontline physicians by providing efficient and accurate diagnosis.

REFERENCES

1. Hui DS, E IA, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* 2020;91:264-266. DOI: 10.1016/j.ijid.2020.01.009.
2. Lu H, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *J Med Virol* 2020;92:401-402. DOI: 10.1002/jmv.25678.
3. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020;382:1199-1207. DOI: 10.1056/NEJMoa2001316.
4. <https://www.who.int>.
5. Holshue ML, DeBolt C, Lindquist S, et al. First Case of 2019 Novel Coronavirus in the United States. *N Engl J Med* 2020;382:929-936. DOI: 10.1056/NEJMoa2001191.
6. Tang YW, Schmitz JE, Persing DH, et al. The Laboratory Diagnosis of COVID-19 Infection: Current Issues and Challenges. *J Clin Microbiol* 2020. DOI: 10.1128/JCM.00512-20.
7. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* 2020. DOI: 10.1001/jama.2020.3786.
8. <https://www.cdc.gov/coronavirus/2019-ncov/lab/index.html>.
9. Li K, Wu J, Wu F, et al. The Clinical and Chest CT Features Associated with Severe and Critical COVID-19 Pneumonia. *Invest Radiol* 2020. DOI: 10.1097/RLI.0000000000000672.

10. Xu X, Yu C, Qu J, et al. Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2. *Eur J Nucl Med Mol Imaging* 2020;47:1275-1280. DOI: 10.1007/s00259-020-04735-9.
11. In: Portable versus Fixed X-ray Equipment: A Review of the Clinical Effectiveness, Cost-effectiveness, and Guidelines. Ottawa (ON), 2016.
12. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-444. DOI: 10.1038/nature14539.
13. Ehteshami Bejnordi B, Mullooly M, Pfeiffer RM, et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod Pathol* 2018;31:1502-1512. DOI: 10.1038/s41379-018-0073-z.
14. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* 2012:1097-1105.
15. Hinton GE, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors. arxiv: 12070580v1 2012.
16. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *ICLR* 2015.
17. Arel I, Rose DC, Karnowski TP. Deep machine learning-a new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine* 2010;5:13-18.
18. Samala RK, Chan HP, Hadjiiski LM, et al. Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms. *Phys Med Biol* 2017;62:8894-8908. DOI: 10.1088/1361-6560/aa93d4.

19. Byra M, Styczynski G, Szmigielski C, et al. Transfer learning with deep convolutional neural network for liver steatosis assessment in ultrasound images. *Int J Comput Assist Radiol Surg* 2018;13:1895-1903. DOI: 10.1007/s11548-018-1843-2.
20. Mekov E, Miravittles M, Petkov R. Artificial intelligence and machine learning in respiratory medicine. *Expert Rev Respir Med* 2020;1-6. DOI: 10.1080/17476348.2020.1743181.
21. Prayer F, Rohrich S, Pan J, et al. [Artificial intelligence in lung imaging]. *Radiologe* 2020;60:42-47. DOI: 10.1007/s00117-019-00611-2.
22. Harris M, Qi A, Jeagal L, et al. A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS One* 2019;14:e0221339. DOI: 10.1371/journal.pone.0221339.
23. Li B, Kang G, Cheng K, et al. Attention-Guided Convolutional Neural Network for Detecting Pneumonia on Chest X-Rays. *Conf Proc IEEE Eng Med Biol Soc* 2019;2019:4851-4854. DOI: 10.1109/EMBC.2019.8857277.
24. Heo SJ, Kim Y, Yun S, et al. Deep Learning Algorithms with Demographic Information Help to Detect Tuberculosis in Chest Radiographs in Annual Workers' Health Examination Data. *Int J Environ Res Public Health* 2019;16. DOI: 10.3390/ijerph16020250.
25. Bao C, Liu X, Zhang H, et al. coronavirus disease 2019 (covid-19) CT findings: A systematic review and meta-analysis. *J Am Coll Radiol* 2020, in press.
26. Lomoroa P, Verdeb F, Zerbonia F, et al. COVID-19 pneumonia manifestations at the admission on chest ultrasound, radiographs, and CT: single-center study and

- comprehensive radiologic literature review. *European Journal of Radiology Open* 2020, in press.
27. Hurt B, Kligerman S, Hsiao A. Deep Learning Localization of Pneumonia: 2019 Coronavirus (COVID-19) Outbreak. *J Thorac Imaging* 2020;35:W87-W89. DOI: 10.1097/RTI.0000000000000512.
 28. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med* 2020;43:635-640. DOI: 10.1007/s13246-020-00865-4.
 29. Elaziz MA, Hosny KM, Salah A, et al. New machine learning method for image-based diagnosis of COVID-19. *PLoS One* 2020;15:e0235187. DOI: 10.1371/journal.pone.0235187.
 30. Murphy K, Smits H, Knoop AJG, et al. COVID-19 on the Chest Radiograph: A Multi-Reader Evaluation of an AI System. *Radiology* 2020:201874. DOI: 10.1148/radiol.2020201874.
 31. Ozturk T, Talo M, Yildirim EA, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 2020;121:103792. DOI: 10.1016/j.combiomed.2020.103792.
 32. Pereira RM, Bertolini D, Teixeira LO, et al. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Comput Methods Programs Biomed* 2020;194:105532. DOI: 10.1016/j.cmpb.2020.105532.
 33. Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images. *arXiv:200309871v1* 2020.

34. Karim MR, Döhmen T, Rebholz-Schuhmann D, et al. DeepCOVIDExplainer: Explainable COVID-19 Predictions Based on Chest X-ray Images. arXiv:200404582 2020.
35. Khuzani AZ, Heidari M, Shariati SA. COVID-Classifer: An automated machine learning model to assist in the diagnosis of COVID-19 infection in chest x-ray images. medRxiv 2020 in press;<https://doi.org/10.1101/2020.05.09.20096560>.
36. Narin A, Kaya C, Pamuk Z. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv:200310849 2020
37. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:14091556 2014.
38. Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:2921-2929.
39. Zhang P, Zhu L, Cai J, et al. Association of Inpatient Use of Angiotensin Converting Enzyme Inhibitors and Angiotensin II Receptor Blockers with Mortality Among Patients With Hypertension Hospitalized With COVID-19. Circ Res 2020, in press. DOI: 10.1161/CIRCRESAHA.120.317134.
40. Ai T, Yang Z, Hou H, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology 2020:200642. DOI: 10.1148/radiol.2020200642.
41. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology 2020:200432. DOI: 10.1148/radiol.2020200432.

42. Li D, Wang D, Dong J, et al. False-Negative Results of Real-Time Reverse-Transcriptase Polymerase Chain Reaction for Severe Acute Respiratory Syndrome Coronavirus 2: Role of Deep-Learning-Based CT Diagnosis and Insights from Two Cases. *Korean J Radiol* 2020;21:505-508. DOI: 10.3348/kjr.2020.0146.

Table 1. Confusion table showing the multiclass CNN classification (whole CXR)

	Normal	COVID-19	Non-COVID-19 viral pneumonia	Bacterial pneumonia
Normal	122	3	17	2
Covid19	6	102	3	6
Non-COVID-19 viral Pneumonia	16	2	94	20
Bacterial pneumonia	4	1	30	85

Table 2 shows the precision and recall rate and F1 score (whole CXR).

	Precision	Recall	F1 - score
Normal	0.82	0.85	0.84
Covid19	0.94	0.87	0.91
Non-covid19 viral pneumonia	0.65	0.71	0.68
Bacterial pneumonia	0.75	0.71	0.73

Figure legends

Figure 1. VGG16 architecture with 16 weighted layers including 3 fully connected layers.

Figure 2: Examples of chest radiographs (a) normal, (b) COVID-19 viral pneumonia, (c) non-COVID-19 viral pneumonia, and (d) bacterial pneumonia. COVID-19 is often characterized by ground-glass opacities with or without nodular consolidation with predominance of bilateral, peripheral and lower lobes involvement. Non-COVID-19 viral pneumonia is often characterized by diffuse interstitial opacities, usually bilaterally. Bacterial pneumonia is often characterized by confluent areas of focal airspace consolidation. Arrows indicate regions of above-described characteristic features.

Figure 3: CNN training and validation loss and accuracy. Loss decreases and accuracy improved with increasing epoch for both training and validation dataset.

Figure 4: pCXR from a COVID-19 patient, the corresponding segmented lung, heatmap from CNN analysis using whole pCXR, and heatmap from CNN analysis using segmented lung overlaid on whole CXR. Arrows indicated regions of ground glass opacity and/or consolidations.