

Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT

Lin Li^{*1a,1b}, Lixin Qin^{*2}, Zeguo Xu^{1a}, Youbing Yin³, Xin Wang³, Bin Kong³, Junjie Bai³, Yi Lu³, Zhenghan Fang³, Qi Song³, Kunlin Cao³, Daliang Liu⁴, Guisheng Wang⁵, Qizhong Xu⁶, Xisheng Fang^{1a}, Shiqin Zhang^{1a}, Juan Xia^{1a}, Jun Xia^{*6}

1a Department of Radiology, Wuhan Huangpi People's Hospital, Wuhan, Hubei, China 430301

1b Jiangnan University Affiliated Huangpi People's Hospital, Wuhan, Hubei, China 430301

2. Department of Radiology, Wuhan Pulmonary Hospital, Wuhan, Hubei, China 430030

3. Keya Medical Technology Co., Ltd, Shenzhen, Guangdong, China 518116

4. Department of Radiology, Liaocheng People's Hospital, Shandong, China, 252000

5. Department of CT, The Third Medical Center of Chinese PLA General Hospital, Beijing, China 100039

6. Department of Radiology, Shenzhen Second People's Hospital/the First Affiliated Hospital of Shenzhen University Health Science Center, Shenzhen, China 518035

*L.L. and L.Q. contributed equally to this manuscript.

Address correspondence to:

Jun Xia

Department of Radiology, Shenzhen Second People's Hospital/the First Affiliated Hospital of Shenzhen University Health Science Center, Shenzhen, China 518035 xiajun@email.szu.edu.cn

Article Type: Original Research; Thoracic Imaging

Abbreviations:

AUC = area under the receiver operating characteristic curve

CI = confidence interval

COVID-19 = coronavirus disease 2019

COVNet = COVID-19 detection neural network

CAP = community acquired pneumonia

DICOM = digital imaging and communications in medicine

Key Results:

A deep learning method was able to identify COVID-19 on chest CT exams (area under the receiver operating characteristic curve, 0.96).

A deep learning method to identify community acquired pneumonia on chest CT exams (area under the receiver operating characteristic curve, 0.95).

There is overlap in the chest CT imaging findings of all viral pneumonias with other chest diseases that encourages a multidisciplinary approach to the final diagnosis used for patient treatment.

Summary Statement:

Deep learning detects coronavirus disease 2019 (COVID-19) and distinguish it from community acquired pneumonia and other non-pneumonic lung diseases using chest CT.

Abstract:

Background: Coronavirus disease has widely spread all over the world since the beginning of 2020. It is desirable to develop automatic and accurate detection of COVID-19 using chest CT.

Purpose: To develop a fully automatic framework to detect COVID-19 using chest CT and evaluate its performances.

Materials and Methods: In this retrospective and multi-center study, a deep learning model, COVID-19 detection neural network (COVNet), was developed to extract visual features from volumetric chest CT exams for the detection of COVID-19. Community acquired pneumonia (CAP) and other non-pneumonia CT exams were included to test the robustness of the model. The datasets were collected from 6 hospitals between August 2016 and February 2020. Diagnostic performance was assessed by the area under the receiver operating characteristic curve (AUC), sensitivity and specificity.

Results: The collected dataset consisted of 4356 chest CT exams from 3,322 patients. The average age is 49 ± 15 years and there were slightly more male patients than female (1838 vs 1484; p -value=0.29). The per-exam sensitivity and specificity for detecting COVID-19 in the independent test set was 114 of 127 (90% [95% CI: 83%, 94%]) and 294 of 307 (96% [95% CI: 93%, 98%]), respectively, with an AUC of 0.96 (p -value<0.001). The per-exam sensitivity and specificity for detecting CAP in the independent test set was 87% (152 of 175) and 92% (239 of 259), respectively, with an AUC of 0.95 (95% CI: 0.93, 0.97).

Conclusions: A deep learning model can accurately detect COVID-19 and differentiate it from community acquired pneumonia and other lung diseases.

Introduction

Coronavirus Disease 2019 (COVID-19) has widely spread all over the world since the beginning of 2020. It is highly contagious and may lead to acute respiratory distress or multiple organ failure in severe cases (1-4). On January 30, 2020, the outbreak was declared as a “public health emergency of international concern” (PHEIC) by World Health Organization (WHO).

The disease is typically confirmed by reverse-transcription polymerase chain reaction (RT-PCR). However, it has been reported that the sensitivity of RT-PCR might not be high enough for the purpose of early detection and treatment of the presumptive patients (5,6). Computed tomography (CT), as a non-invasive imaging approach, can detect certain characteristic manifestations in the lung associated with COVID-19 (7,8). Therefore, CT could serve as an effective way for early screening and diagnosis of COVID-19. Despite of its advantages, CT may share certain similar imaging features between COVID-19 and other types of pneumonia, thus making it difficult to differentiate.

Recently, artificial intelligence (AI) using deep learning technology has demonstrated great success in the medical imaging domain due to its high capability of feature extraction (9-11). Specifically, deep learning was applied to detect and differentiate bacterial and viral pneumonia in pediatric chest radiographs (12,13). Attempts have also made to detect various imaging features of chest CT (14,15).

In this study, we proposed a three-dimensional deep learning framework to detect COVID-19 using chest CT, referred to COVID-19 detection neural network (COVNet). Community acquired pneumonia (CAP) and other non-pneumonia exams were included to test the robustness of the model.

Materials and Methods

Patients

This retrospective study was approved by the ethics committees of the participating hospitals. Further consent was waived with approval. It includes 4,536 three-dimensional (3D) volumetric chest CT exams from 3,506 patients acquired at 6 medical centers between Aug 16, 2016 and Feb 17, 2020 (Figure 1). The exclusion criteria included: 1) contrast CT exam; 2) exams with slice thickness >3mm. After applying the exclusion criteria, 4,356 3D chest CT exams from 3,322 patients are selected in this study (Figure 1). The average age is 49 ± 15 years and there are slightly more male patients than female (1838 vs 1484; p-value=0.29). CT exams with multiple reconstruction kernels at the same imaging session or acquired at multiple time points were included. The final dataset consists of 1296 (30%) exams for COVID-19, 1735 (40%) for CAP and 1325 (30%) for non-pneumonia. All the COVID-19 were confirmed as positive by RT-PCR and were acquired from Dec 31, 2019 to Feb 17, 2020.

All the COVID-19 were confirmed as positive by RT-PCR and were acquired from Dec 31, 2019 to Feb 17, 2020. The median period from the symptom onset to the first chest CT exam was 7 days with the range from 0 to 20 days. The most common symptoms were fever (81%) and cough (66%). For each patient, one or multiple CT exams at several time points during the course of the disease were acquired (Average CT exams per patient was 1.8, with a range from 1 to 6).

CAP and other non-pneumonia patients were randomly selected from the participating hospitals between Aug 16, 2016 and Feb 17, 2020. The admission distribution of the patients with CAP was: outpatient (46%, 713 of 1551), inpatient (38%, 588), emergency (3%, 40 of 1551) and physical examination (3%, 46 of 1551). Out of the 1325 CAP patients, 210 received laboratory confirmation of the etiology: 112 were confirmed to be bacterial culture positive and 98 negative (65 mycoplasma and 28 viral pneumonia).

Other non-pneumonia patients have no lung disease (510), or have lung nodules (552), chronic inflammation (264), chronic obstructive pulmonary disease (COPD, 115) or other diseases (14, including pneumothorax, tracheal diverticulum, etc)

CAP and other non-pneumonia patients were randomly selected from the participating hospitals between Aug 16, 2016 and Feb 17, 2020. All the exams were randomly split with a ratio of 9:1 into a training set and an independent testing set at the patient level. The training dataset is then further split for training the model and internal validation (10% of the samples). The independent testing set was not used for training and internal validation. The patient demographic statistics are summarized in Table 1. The patient disease statistics are summarized in Table 2.

Imaging Protocol

CT examinations were performed using different manufacturers with standard imaging protocols. Each volumetric exam contains 51~1094 CT slices with a varying slice-thickness from 0.5mm to 3mm. The reconstruction matrix is 512x512 pixels with in-plane pixel spatial resolution from 0.29x0.29 mm² to 0.98x0.98 mm². Please refer to Table E1 in Appendix E1 [online].

Deep Learning Model

We develop a 3D deep learning framework for the detection of COVID-19, referred to COVNet (Figure 2). It is able to extract both 2D local and 3D global representative features. The COVNet framework consists of a ResNet50 (16) as the backbone, which takes a series of CT slices as input and generates features for the corresponding slices. The extracted features from all slices are then combined by a max-pooling operation. The final feature map is fed to a fully connected layer and softmax activation function to generate a probability score for each type (COVID-19, CAP, and non-pneumonia).

Specifically, given a 3D CT exam, we first preprocess it and extract the lung region as the region of interest (ROI) using a U-net (17) based segmentation method. The preprocessed image is then passed to our COVNet for the predictions. The code to reproduce the results in this study is made available via: <https://github.com/bkong999/COVNet.git>. Please refer to Appendix E1 for the details.

Statistical Analysis

The performance of our trained deep learning model was evaluated using an independent testing set, which was not used during model development. Statistical analysis was performed using R (version 3.6.3). ANOVA tests and χ^2 tests were used to compare differences among different groups for continuous and dichotomous variables, respectively. A two-sided p value less than 0.05 was considered to be statistically significant. The sensitivity and specificity for both COVID-19 and CAP were calculated. Receiver operating characteristic (ROC) curve was plotted and the area under the curve (AUC) was calculated with the 95% confidence intervals (CIs) based on DeLong's method (18).

Results

Study Population Characteristics

Table 1 list the study population characteristics for the training and independent testing datasets. There are slightly more male patients than females across COVID-19, CAP and non-pneumonia groups for both training dataset (age: COVID-19: 53%; CAP: 56%; Non-Pneumonia: 55%; p-value=0.55) and independent testing dataset (age: COVID-19: 52; CAP: 62; Non-Pneumonia: 52; p-value=0.17). Patients with COVID-19 and CAP are older than the ones in the non-pneumonia group for the training dataset (age: COVID-19: 53; CAP: 51; Non-Pneumonia: 41; p-value<0.001) and independent testing dataset (age: COVID-19: 52; CAP: 53; Non-Pneumonia: 41; p-value<0.001).

Model Performance

Once the deep learning model is trained, it is very fast to process a new testing exam. The average processing time for each CT exam is 4.51 seconds on a workstation (GPU NVIDIA Quadro M4000 8GB, RAM 16GB, and Intel Xeon Processor E5-1620 v4 @3.5GHz).

The performance of COVNet on the detection of COVID-19 and CAP are summarized in Table 3. The sensitivity and specificity for COVID-19 are 90% (114 of 127; p-value<0.001) with 95% confidence interval (95% CI) of [95% CI: 83%, 94%] and 96% (294 of 307; p-value<0.001) with [95% CI, 93%, 98%], respectively. For the detection of CAP, our deep learning model yields 87 % (152 of 175; p-value<0.001) [95% CI: 81%, 91%] for the sensitivity and 92% (239 of 259; p-value<0.001) [95% CI, 88%, 95%] for the specificity. The ROC curves were shown in Figure 3. The corresponding AUC values for COVID-19 and CAP are 0.96 [95% CI, 0.94,0.99] and 0.95 [95% CI: 0.93, 0.97], respectively.

To improve the interpretability of our model, we adopted the Gradient-weighted Class Activation Mapping (Grad-CAM) method (19) to visualize the important regions leading to the decision of the deep learning model. Such a localization map is fully generated by the model without additional manual annotation. Figure 4 shows the heatmap to the suspected regions for the examples of COVID, CAP and non-pneumonia. These heatmaps illustrated that our algorithm paid most attention to the abnormal regions while ignore the normal-like regions as shown in the non-pneumonia example.

In addition, Figure 5 shows one representative example of CAP case that is misclassified as COVID-19 and Figure 6 shows another example of COVID-19 case that is misclassified as CAP. The consecutive slices around the abnormality are shown from left to right on each row. It seems that these are very challenging cases. It might be useful to include the history of exposure to further improve the accuracy.

Discussion

In this study, we designed and evaluated a three-dimensional deep learning model for detecting coronavirus disease 2019 (COVID-19) from chest CT images. On an independent testing data set, we showed that this model achieved high sensitivity (90% [95% CI: 83%, 94%] and high specificity of 96% [95% CI: 93%, 98%] in detecting COVID-19. The AUC values for COVID-19 and community acquired pneumonia (CAP) were 0.96 [95% CI: 0.94, 0.99] and 0.95 [95% CI: 0.93, 0.97], respectively.

In this study, a 3D deep learning framework was proposed for the detection of COVID-19. This framework is able to extract both 2D local and 3D global representative features. Deep learning has achieved superior performance in the field of radiology (10-13). Previous studies have successfully applied deep learning techniques to detect pneumonia in pediatric chest radiographs and further to differentiate viral and bacterial pneumonia in 2D pediatric chest radiographs (12-13). We were able to collect a large number of CT exams from multiple hospitals, which included 1296 COVID-19 CT exams. More importantly, 1735 CAP and 1325 non-pneumonia CT exams were also collected as the control groups in this study in order to ensure the detection robustness considering that certain similar imaging features may be observed in COVID-19 and other types of lung diseases.

COVID-19 has widely spread all over the world since the first case was detected at the end of 2019. Early diagnosis of the disease is important for treatment and the isolation of the patients to prevent the virus spread. RT-PCR is considered as the reference standard; however, it has been reported that chest CT could be used as a reliable and rapid approach for screening of COVID-19 (5,6). The abnormal CT findings in COVID-19 have been recently reported to include ground-glass opacification, consolidation, bilateral involvement, peripheral and diffuse distribution (1-5,7-8).

This study has several limitations. Firstly, COVID-19 is caused by a corona virus and may have similar imaging characteristics as pneumonia caused by other types of viruses. However, due to the lack of laboratory confirmation of the etiology for each of these cases, we were not able to select other viral pneumonias for comparison in this study. Instead, we randomly selected CAP from August 2016 to February 2020. We propose that our sampling method was sufficient for the typical distribution of various subtypes of CAP. These cases should have included non-COVID-19 viral pneumonias (such as influenza virus), bacterial pneumonia and organizing pneumonia from any cause. It would be desirable to test the performance of COVNet in distinguishing COVID-19 from other viral pneumonias that have real time polymerase chain reaction confirmation of the viral agent in a future study. Secondly, a disadvantage of all deep learning methods is the lack of transparency and interpretability (e.g. it is impossible to determine what imaging features are being used to determine the output). While we used a heatmap to visualize the important regions in the images leading to the decision of the algorithm, heatmaps are still not sufficient to visualize what unique features are used by the model to distinguish between COVID-19 and CAP. Thirdly, there is a great deal of overlap in how the lung responds to various insults and there is a significant amount of overlap in the presentation of many diseases in the lung that depend on host factors (e.g. age, drug reactivity, immune status, underlying co-morbidities). No one method will be able to differentiate all lung diseases based simply on the imaging appearance on chest CT. A multidisciplinary approach for this is recommended. Furthermore, though a large size of data was collected in this study, the test set came from the same hospitals as the training set. We plan to collect additional CT exams from different centers to evaluate its performance in near future. Finally, this study focuses on whether one exam is COVID-19 or not, but has not addressed categorizing the disease into different severities. As a next step, it would be important to not only predict the presence of COVID-19, but also the severity degree to further help monitor and treat patients.

In conclusion, a robust deep learning model is developed to differentiate COVID-19 and CAP from chest CT images. These results demonstrate that a machine learning approach using convolutional networks model has the ability to distinguish COVID-19 from community acquired pneumonia.

Author contributions: Guarantors of integrity of entire study, L.L., L.Q., Y.Y., J.X.; study concepts/study design or data acquisition or data analysis/interpretation, L.L., L.Q., Z.X., Y.Y., X.W., B.K.,

J.B., Y.L., Z.F., Q.S., K.C., D.L., G.W., Q.X., X.F., S.Z., J.X., J.X.; manuscript drafting or manuscript revision for important intellectual content, L.L., L.Q., Z.X., Y.Y., X.W., B.K., J.B., Y.L., Z.F., Q.S., K.C., D.L., G.W., Q.X., X.F., S.Z., J.X., J.X.; approval of final version of submitted manuscript, L.L., L.Q., Z.X., Y.Y., X.W., B.K., J.B., Y.L., Z.F., Q.S., K.C., D.L., G.W., Q.X., X.F., S.Z., J.X., J.X.; literature research, Y.Y., X.W., B.K., J.B., Y.L., J.X.; clinical studies, L.L., L.Q., D.L., G.W., J.X.; experimental studies, Y.Y., X.W., B.K., J.B., Y.L., Z.F., statistical analysis, Y.Y., X.W., B.K., Y.L.; and manuscript editing, L.L., L.Q., Y.Y., X.W., B.K., J.B., Y.L., Q.S., J.X.;

Reference

1. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet* 2020 Jan 30. doi: 10.1016/S0140-6736(20)30211-7.
2. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *Jama* 2020.
3. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *The New England Journal of Medicine* 2020.
4. Holshue ML, DeBolt C, Lindquist S, et al. First Case of 2019 Novel Coronavirus in the United States. *New England Journal of Medicine* 2020 Jan 31. doi:10.1056/NEJMoa2001191.
5. 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. <https://doi.org/10.1148/radiol.2020200642>.
6. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology*. 2020. <https://doi.org/10.1148/radiol.2020200432>
7. Chung M, Bernheim A, Mei X, et al. CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV). *Radiology* 2020: 200230.
8. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 2020 Jan 24. doi: 10.1016/S0140-6736(20)30183-5.
9. Xia C, Li X, Wang X, et al. A Multi-modality Network for Cardiomyopathy Death Risk Prediction with CMR Images and Clinical Information. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2019.
10. Kong B, Wang X, Bai J, et al. Learning tree-structured representation for 3D coronary artery segmentation. *Computerized Medical Imaging and Graphics* 2020, 80: 101688.
11. Ye H, Gao F, Yin Y, et al. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. *European Radiology*, 2019, 29: 6191–6201.
12. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 2018, 172, 1122–1131.

13. Rajaraman S, Candemir S, Kim I, et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Applied Sciences* 8.10 (2018): 1715.
14. Depeursinge A, Chin AS, Leung AN, et al. Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution CT. *Investigative Radiology* 50.4 (2015): 261.
15. Anthimopoulos M, Christodoulidis S, Ebner L, et al. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Transactions on Medical Imaging* 35.5 (2016): 1207-1216.
16. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
17. Ronneberger O, Philipp F, Thomas B. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2015.
18. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988 (44), 837–845.
19. Selvaraju RR, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

Table 1. Summary of training and independent testing datasets.

	Training				Independent Testing			
	COVID-19	CAP	Non-Pneumonia	<i>p</i> Value	COVID-19	CAP	Non-Pneumonia	<i>p</i> Value
Patients	400 (14%)	1396 (47%)	1173 (40%)	-	68 (19%)	155 (44%)	130 (37%)	-
Exams	1165 (30%)	1560 (40%)	1193 (31%)	-	131 (30%)	175 (40%)	132 (30%)	-
Male patients	212 (53%)	782 (56%)	645 (55%)	0.55	30 (52%)	96 (62%)	68 (52%)	0.17
Age (year)	5±15	5±19	41±13	<0.001	52±17	5±20	4±13	<0.001

Note: Values in parentheses are the percentage. Ages are reported as mean ± standard deviation. COVID-19 = coronavirus disease 2019, CAP = community acquired pneumonia.

Table 2. Summary of the Diseases of training and independent testing datasets.

Dataset	Category	rt-PCR confirmed	rt-PCR not confirmed	Bacterial culture positive	Bacterial culture negative	Bacterial culture status unknown	Non-infectious lung disease	Other infectious lung disease	Normal Lung
Training Set	COVID-19	400 (1165 exams)	0						
	CAP			Bacterial Pneumonia 72 Bacterial & Mycoplasma Pneumonia 26	Viral Pneumonia 24 Mycoplasma Pneumonia 60 Pneumocystis Carinii Pneumonia 2	1212			
	Other						Nodule 500 COPD 107 Other 13	Chronic Inflammation 235	459
							CHF: NA Drug reaction: NA		
Testing Set	Category	rt-PCR confirmed	rt-PCR not confirmed	Bacterial culture positive	Bacterial culture negative	Bacterial culture status unknown	Normal or Non-infectious Lung disease	Other infectious Lung disease	Normal Lung
	COVID-19	68 (131 exams)	0						
	CAP			Bacterial Pneumonia 13 Bacterial & Mycoplasma Pneumonia 1	Viral Pneumonia 7 Mycoplasma Pneumonia 5	129			
	Other						Nodule 52 COPD 8 Other 1	Chronic inflammation 29	51
						CHF: NA Drug reaction: NA			

Note: Only some CAP patients received laboratory confirmation of the etiology. COVID-19 = coronavirus disease 2019, CAP = community acquired pneumonia, rt-PCR = reverse-transcription polymerase chain reaction, COPD = chronic obstructive pulmonary disease, CHF= Congestive heart failure.

Table 3. The performance of deep learning framework COVNet on the independent testing set.

	Sensitivity %	Specificity %	AUC	P-value
COVID-19	90 (114 of 127) [83, 94]	96 (294 of 307) [93, 98]	0.96 [0.94, 0.99]	<0.001
CAP	87 (152 of 175) [81, 91]	92 (239 of 259) [88, 95]	0.95 [0.93, 0.97]	<0.001
Non-Pneumonia	94 (124 of 132) [88, 97]	96 (291 of 302) [94, 98]	0.98 [0.97, 0.99]	<0.001

Note: Values in parentheses are the numbers for the percentage calculation. Values in brackets are 95% confidence intervals [95%CI, %]. AUC = area under the receiver operating characteristic curve, COVID-19 = coronavirus disease 2019, CAP = community acquired pneumonia, COVNet = COVID-19 detection neural network.

FIGURES

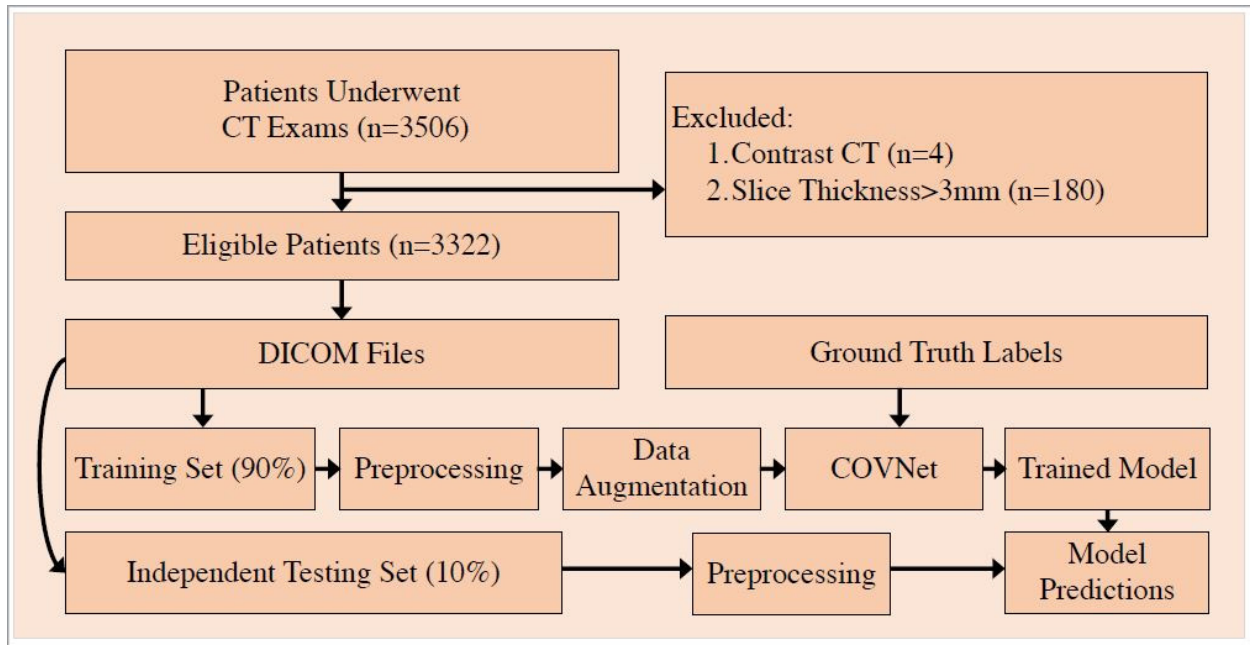


Fig 1. Flow diagram. We collected a dataset of 3506 patients with chest CT exams. After exclusion, 3,322 eligible patients were included for the model development and evaluation in this study. CT exams were extracted from DICOM files. The dataset was split into a training set (to training the model), and the independent testing set at the patient level. A supervised deep learning framework (COVNet) was developed to detect COVID-19 and community acquired pneumonia. The predictive performance of the model was evaluated by using an independent testing set. COVNet = COVID-19 detection neural network.

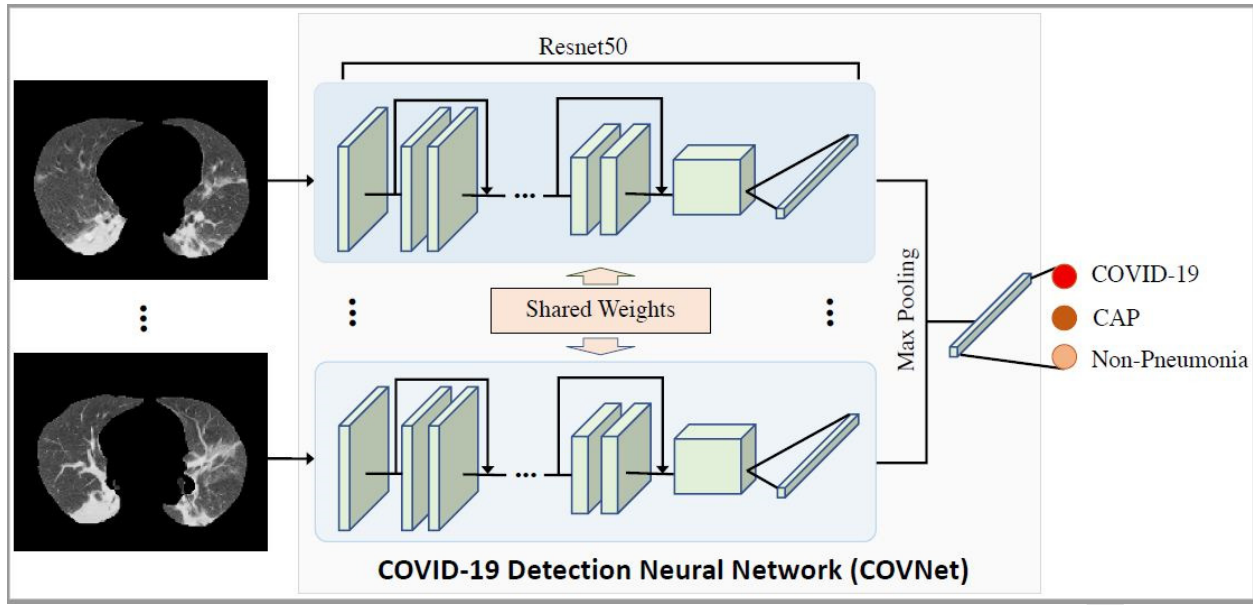


Fig 2. COVID-19 detection neural network (COVNet) architecture. The COVNet is a convolutional neural network (CNN) using ResNet50 as the backbone. It takes as input a series of CT slices and generates a classification prediction of the CT image. The CNN features from each slice of the CT series are combined by a max-pooling operation and the resulting feature map is fed to a fully connected layer to generate a probability score for each class.

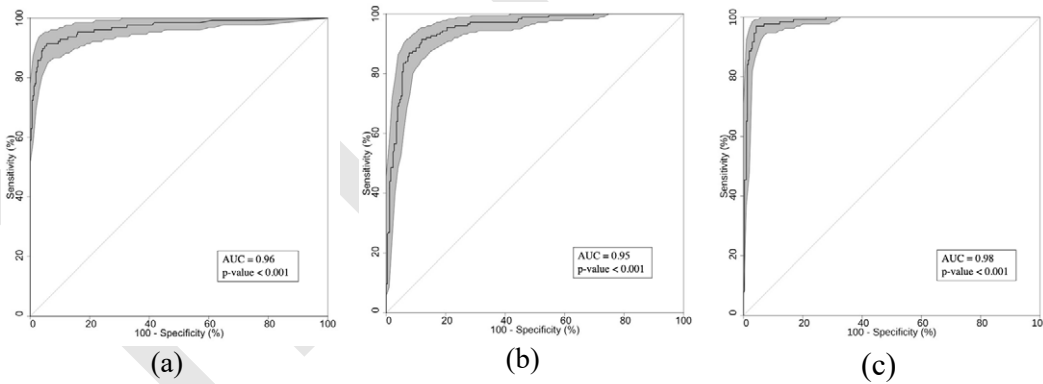


Fig 3. Receiver operating characteristic curves of the model. Each plot illustrates the receiver operating characteristic (ROC) curve of the algorithm (black curve) on the independent testing set for (a) COVID-19 with AUC = 0.96 (p-value < 0.001) (b) CAP with AUC = 0.95 (p-value < 0.001) and (c) Non-Pneumonia with AUC = 0.98 (p-value < 0.001). The gray region indicates the 95% CI. COVID-19 = coronavirus disease 2019, CAP = community acquired pneumonia, CI = confidence interval.

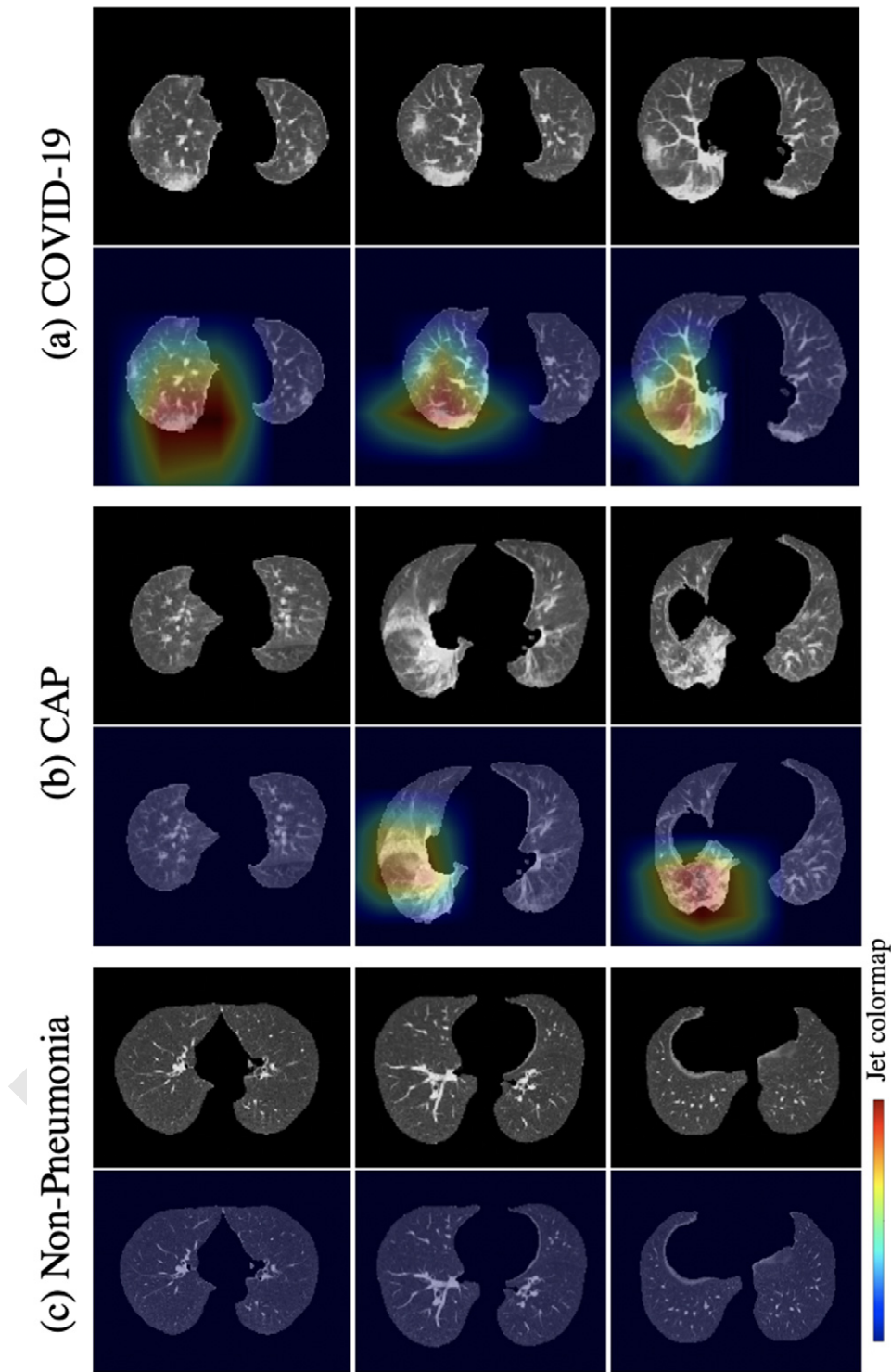


Fig 4. Representative examples of the attention heatmaps generated using Grad-CAM method for (a) COVID-19, (b) CAP, and (c) Non-Pneumonia. The heatmaps are standard Jet colormap and overlapped on the original image, the red color highlights the activation region associated with the predicted class. COVID-19 = coronavirus disease 2019, CAP = community acquired pneumonia.

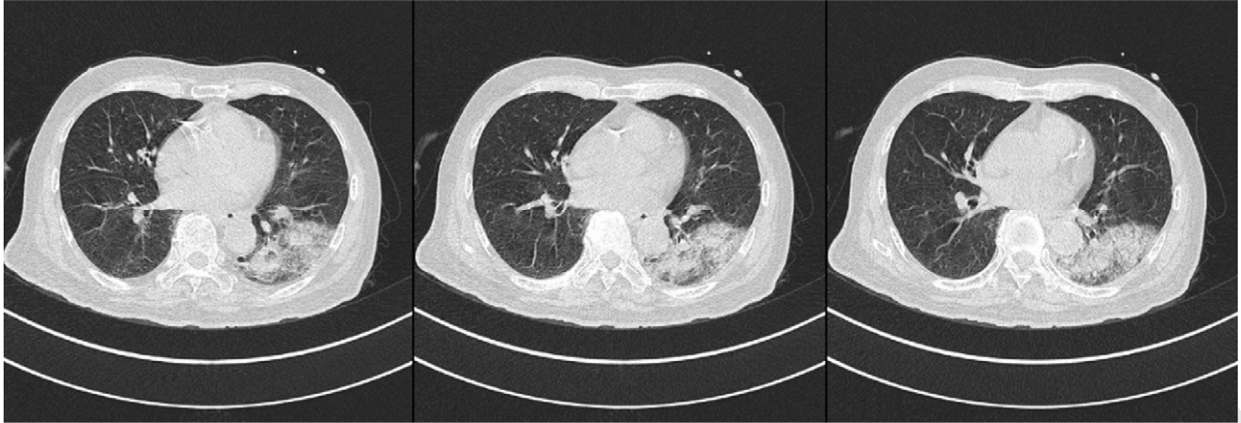


Fig 5. A representative example of community acquired pneumonia case that is misclassified as COVID-19. The consecutive slices around the abnormality are shown from left to right. COVID-19 = coronavirus disease 2019, CAP = community acquired pneumonia.

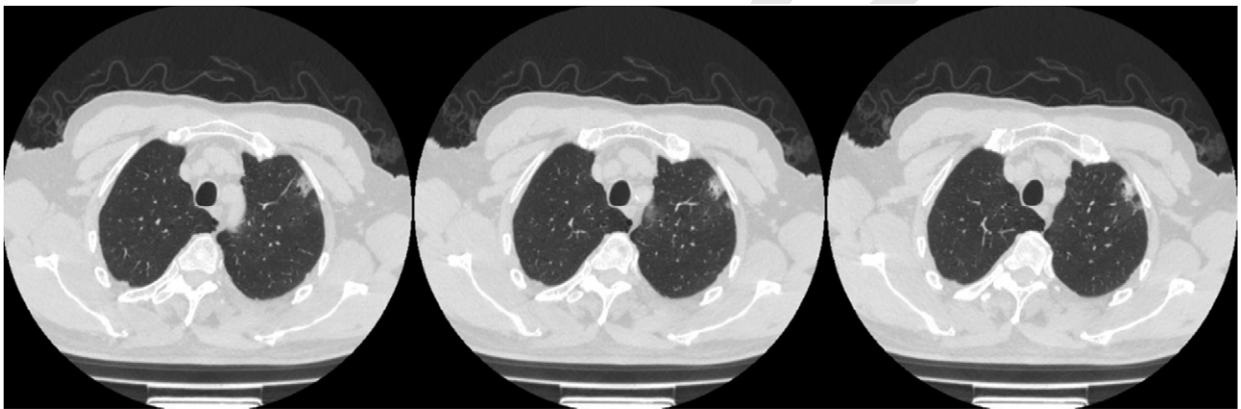


Fig 6. A representative example of COVID-19 case that is misclassified as community acquired pneumonia. The consecutive slices around the abnormality are shown from left to right. COVID-19 = coronavirus disease 2019