

Supervised framework for COVID-19 classification and lesion localization from chest CT

Junyong Zhang^{*1}, Yingna Chu^{*2}, Na Zhao²

Abstract

Background: Quick and precise identification of people suspected of having COVID-19 plays a key function in imposing quarantine at the right time and providing medical treatment, and results not only in societal benefits but also helps in the development of an improved health system. Building a deep-learning framework for automated identification of COVID-19 using chest computed tomography (CT) is beneficial in tackling the epidemic.

Aim: To outline a novel deep-learning model created using 3D CT volumes for COVID-19 classification and localization of swellings.

Methods: In all cases, subjects' chest areas were segmented by means of a pre-trained U-Net; the segmented 3D chest areas were submitted as inputs to a 3D deep neural network to forecast the likelihood of infection with COVID-19; the swellings were restricted by joining the initiation areas within the classification system and the unsupervised linked elements. A total of 499 3D CT scans were utilized for training worldwide and 131 3D CT scans were utilized for verification.

Results: The algorithm took only 1.93 seconds to process the CT amount of a single affected person using a special graphics processing unit (GPU). Interesting results were obtained in terms of the development of societal challenges and better health policy.

Conclusions: The deep-learning model can precisely forecast COVID-19 infectious probabilities and detect swelling areas in chest CT, with no requirement for training swellings. The easy-to-train and high-functioning deep-learning algorithm offers a fast method to classify people affected by COVID-19, which is useful to monitor the SARS-CoV-2 epidemic. [*Ethiop. J. Health Dev.* 2020; 34(4):000-000]

Key words: COVID-19, CT scan, deep learning, neural network, DeCoVNet, RT-PCR, computed tomography

Introduction

Coronavirus disease 2019 (COVID-19) has been prevalent around the world from December 2019 (1,2). It spreads quickly among humans and produces acute respiratory distress and/or organ failure (3). A total of 1,391,890 cases of COVID-19 had been reported globally as of 8 April 2020 and the mortality rate was 82,589, with a significant rise of affected people in Europe and North America. The reverse transcription-polymerase chain reaction (RT-PCR) test (4) is conducted to find the presence of the virus. However, the detection of COVID-19 disease using RT-PCR is time-consuming. High false-negative rates and low sensitivities create obstacles to the early identification and treatment of presumptive affected persons (3,6).

Even though initial studies have shown promising outputs by utilizing chest CT to diagnose COVID-19 and detect affected areas, most of the current techniques depend on the generally utilized supervised learning regime. This necessitates a tremendous amount of research on the hand-operated naming of data; conversely, in such an epidemic scenario, physicians have little time to conduct time-consuming hand-operated marking, and can fail to apply such controlled deep-learning techniques. The 2019-nCoV infection triggers extreme respiratory disease clusters close to serious acute respiratory syndrome coronavirus and was correlated with ICU admission and high mortality. Significant gaps in our understanding of the origin, epidemiology, and length of human infection and continuum of disease need to be resolved in future research.

Additional work is necessary to further grasp the current COVID-19 in order to improve antiviral agents and vaccines (2,3). The condition is normally verified by RT-PCR, and although it is the most promising test for COVID-19, it typically takes up to four hours to get the result. A CT scan however, can reveal the condition of the lungs instantly and more precisely than a chest X-ray. RT-PCR should be done along with a CT scan to confirm positivity (5,6). The demand for RT-PCR kits is very high worldwide, so there is a need for an alternative method, such as chest CT, to establish that it can support incoming patients' needs. It can also address the aim of societal health development goals where RT-PCR is not widely available. CT is a non-invasive imaging method that can identify a few characteristic pulmonary manifestations related to COVID-19 (7,8). CT may also act as an important method of initial testing and treatment of the disease. CT distributes few identical imaging properties among pneumonia type COVID-19 and other categories of pneumonia, and remains complex to distinguish.

Artificial intelligence (AI) that utilizes deep-learning techniques has recently shown huge accomplishments in the field of medical imaging because of its ability to extract features (9-11). In particular, deep learning has been used to diagnose and distinguish bacterial and viral pneumonia in pediatric chest radiographs (12,13). Various efforts have been made to identify different imaging characteristics of chest CT. Magnetic resonance imaging (MRI) is more efficient than a CT scan, but MRI has certain limitations. It cannot be used

^{*}Department of Radiology, Yantai Mountain Hospital, No. 91, Jiefang Road, Yantai, Shandong, 264001, China
Department of Radiology, Yantai Yuhuangding Hospital, Laishan Branch, No.1181, East Gangcheng Street, Laishan District, Yantai, Shandong, 264003, China

for those patients who have metallic implants or pacemakers, for example. While a CT scan is less expensive than MRI (14,15), the process is not cost-effective. Even so, in cases of emergency and severe lung infection, it is mandatory to conduct CT scans. So, in other words, CT scanning is an alternative to understanding lung infection due to COVID-19 and can be accessed by the majority of people worldwide.

RT-PCR is regarded as a good method for supporting people affected by COVID-19 (4). Conversely, the RT-PCR assay is not sufficient in many areas that are affected badly, as was the case during the initial spread of the disease. The laboratory test also has the problem of high false-negative rates because of several parameters, such as sample making and monitoring of quality (5). In scientific research, readily available imaging technology, such as chest X-ray and thoracic CT, provides considerable help to physicians (6-11). In China, for example, several people were reported to have been diagnosed with COVID-19 when characteristic manifestations of CT scans were monitored (5). Suspected cases were also hospitalized or quarantined for further laboratory tests, even with no scientific indications. Due to the strong false-positive rate in nucleic acid testing, many affected people needed to be checked many days apart before a definitive diagnosis could be made.

Observations from CT images play a significant role in restricting viral spread and also in combating COVID-19. Recent innovations, such as artificial intelligence (AI), further enhance the capacity of imaging devices, and assist medical specialists. Accordingly, we analyze below the rapid response of the medical imaging community (AI empowered) to COVID-19.

First, AI-powered image acquisition may help to dramatically optimize the scanning phase and reshape the workflow, with limited interaction with people infected with COVID-19, offering the best security for imaging technicians. Second, Diego *et al.* note that AI will boost job performance by correctly delineating pathogens in X-ray and CT pictures, enabling eventual restriction of further advancement of pathogens. In fact, computer-aided systems enable radiologists to create scientific choices, e.g. for cancer detection, follow-up and prognosis.

In this review article, we explore the range of medical imaging and interpretation procedures used with COVID-19, comprising image processing, segmentation, evaluation and follow-up. This model attains state-of-the-art functioning (94.4 per cent area under the curve (AUC)) in 6,716 national lung cancer testing cases, and also functions a separate scientific validation group of 1,139 people. Two reader tests were executed where existing CT imaging was not usable. The model overwhelms all six radiologists with an

absolute reduction of 11 per cent in false positives and 5 per cent in false negatives. Where initial CT imaging was present, the functioning of the model was constant with the same radiologists. This offers the opportunity to minimize the testing procedure by means of computer support and automation. If there remains a number of asymptomatic persons undetected, there is a good chance for deep-learning models to enhance the precision, and uptake of chest cancer testing globally.

Visual examination of histopathology slides is the key technique utilized by pathologists for the evaluation of the level, form and subcategory of chest tumors. The most common subcategories of chest cancer involve a visual examination by a qualified pathologist. A deep convolutionary neural network (inception v3) of full-slide photos from The Cancer Genome Atlas is trained to identify them specifically in regular chest tissue. The efficiency of the approach is equivalent to that of pathologists with a mean area under the curve (AUC) of 0.97. The framework was pre-tested on many databases of frozen tissues. In addition, the network is trained to forecast the popular generally transformed genes in lung adenocarcinomas (LUAD). Such results denote that deep-learning frameworks may help pathologists in the identification of subcategory gene transformations in cancer.

Feng Pan *et al.* composed and examined anonymized data in order to assist good scientific conclusions and COVID-19 therapy. Diagnosis and release requirements as per the WHO draft diagnosis and care guidelines depends on the following factors: (a) epidemiological history; (b) scientific signs; and (c) laboratory identification. Patients with reported COVID-19 have been admitted and separated for treatment. The discharge conditions are that the affected person: (a) has been afebrile for around three days; (b) has significantly improved respiratory indications; (c) shows improvements to radiological abnormalities in chest radiography or chest CT scans; and (d) has two successive negative COVID-19 nucleic acid tests over two days (7).

Methods

Image acquisition: We obtained CT scan images from local hospitals or specialized clinics where COVID-19 patients were being treated. All of the scans were of subjects who were lying down horizontally in the supine position, with both arms stretched above their shoulders.

The scan images were analyzed by specialized software for further understanding. Certain limitations were in place when collecting the images for this research. Specifically, no images were collected or analyzed from any patients who did not show any detectable symptoms for COVID-19. This group of people was not considered due to their lack of symptoms and their absences of lesions.

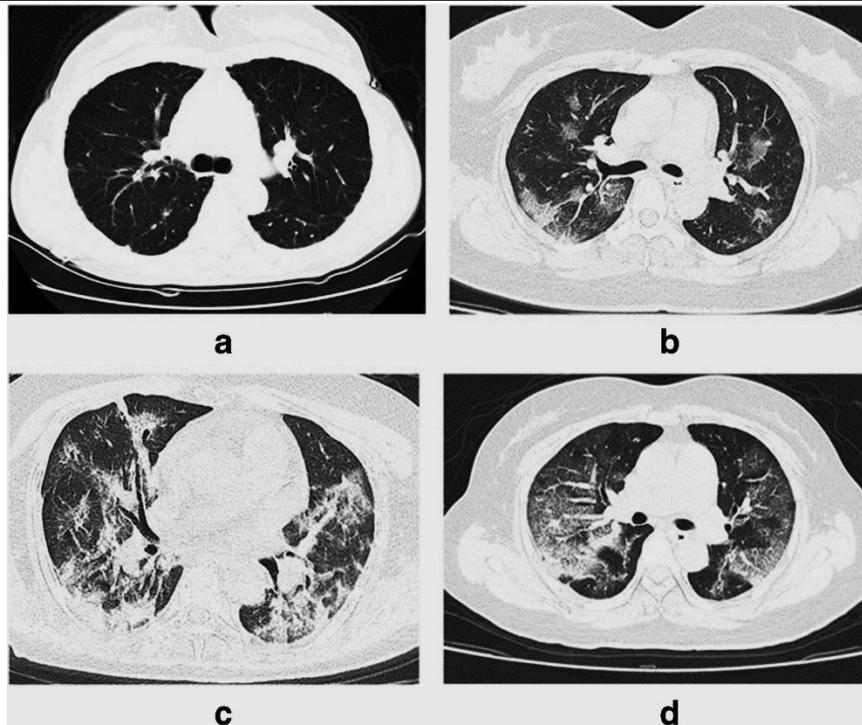


Figure 1a-1d: COVID-19 patient chest CT images, adapted from Li et al., 2017

Ground-truth label: Early multifocal stage with minor patchy reflecting and interstitial anomalies, especially in the secondary sector of the two-sided chest. In the development stage, the swellings may grow in size and quantity; they may increase to several ground-glass opacity (GGOs) with more penetration into the two sides of the chest. Extreme cases of pulmonary disperse

aggregation can happen and pleural effusion is seldom demonstrated.(17-21)

The final diagnosis of COVID-19 requires a mix of epidemiological properties, medical signs and indications, chest CT, lab outputs and RT-PCR. The corresponding covers on the picture are shown in Figure 2.

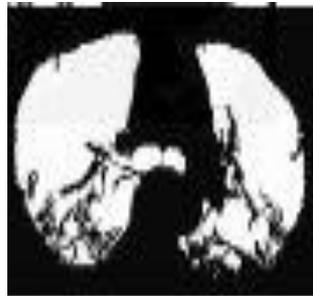


Figure 2: Mask of the corresponding cover CT image

The proposed CoVNet: We suggested a deep convolutional neural system to identify COVID-19. As seen in Figure 3, the proposed system took the quantity of the CT as the input and its 3D chest cap(22). The 3D pulmonary cover was developed by an earlier trained U-

Net (24). The proposed neural network is divided into three phases. The network stem stage is the initial stage and has a 3D vanilla convolution; the size of the kernel is 5 7 7.

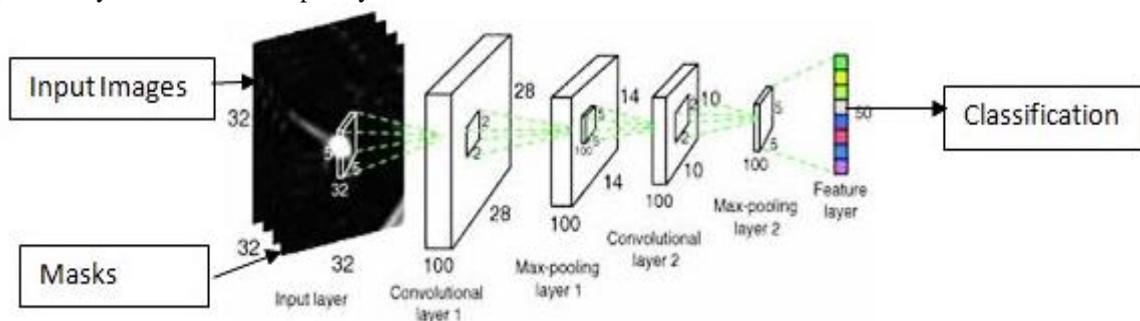


Figure 3: Proposed architecture of convolutional neural network

The second stage consists of two residual 3D units. Throughout the ResBlock, a 3D function map was

transferred to both 3D convolution with a batch norm layer and a link consisting of a 3D convolution which

was absent throughout for the aspect orientation. The resulting feature maps were used in an individual fashion. The third stage comprises a progressive classifier with three 3D convolution layers and a completely linked layer with the soft max initiation feature. ProClf slowly extracts knowledge in CT

$$G[m, n] = (f * h)[m, n] = \sum_j \sum_k h[j, k] f[m - j, n - k]$$

The 3D chest cover of the CT input chest quantity supports to minimize context details and to properly identify COVID-19. Detection of a 3D chest cover was an established problem. A basic 2D U-Net utilizing CT images is trained in the training package. To achieve ground-reality chest covers, chests areas are segmented by means of an unsupervised learning technique (25), and the effects of rest segmentation were used as ground-level reality covers. The 3D chest cover of increasing CT quantity was attained by measuring the qualified 2D U-Net frame-by-frame without the use of any time details. The whole preparation and research protocols for the U-Net and DeCoVNet COVID-19 categorization are shown in Figure 2.

Lesion localization: The swelling localization was to join the initiation areas created by the deep classification system with the unsupervised chest segmentation

quantities through 3D max-pooling and eventually explicitly outputs the likelihood of being COVID-positive or COVID-negative(23). The indexes of rows and columns of the result matrix are marked with m and n correspondingly.

technique. The technique is shown in Figure 3. The right section has certain swelling areas from DeCoVNet by implementing the CAM technique suggested by Simonyan & Zisserman (25). In the left hand of Figure 3, we isolated possible COVID-19 swelling areas from unsupervised chest segmentation tests. After implementing the 3D connected element technique (24) to the CT scan, we noticed that the swelling areas were prone to the 3DCC algorithm and can be used for the localization of swellings. To obtain the reaction map, variance was estimated in a 7-7 window for the entire pixel as a 3DCC initiation. The 3DCC initiation area with the maximum size was then chosen and referred to as R3dcc. Finally, the CAM initiation area with the greatest overlap with R3dcc was chosen as the last product of the COVID-19 swelling localization (see Figure 4).

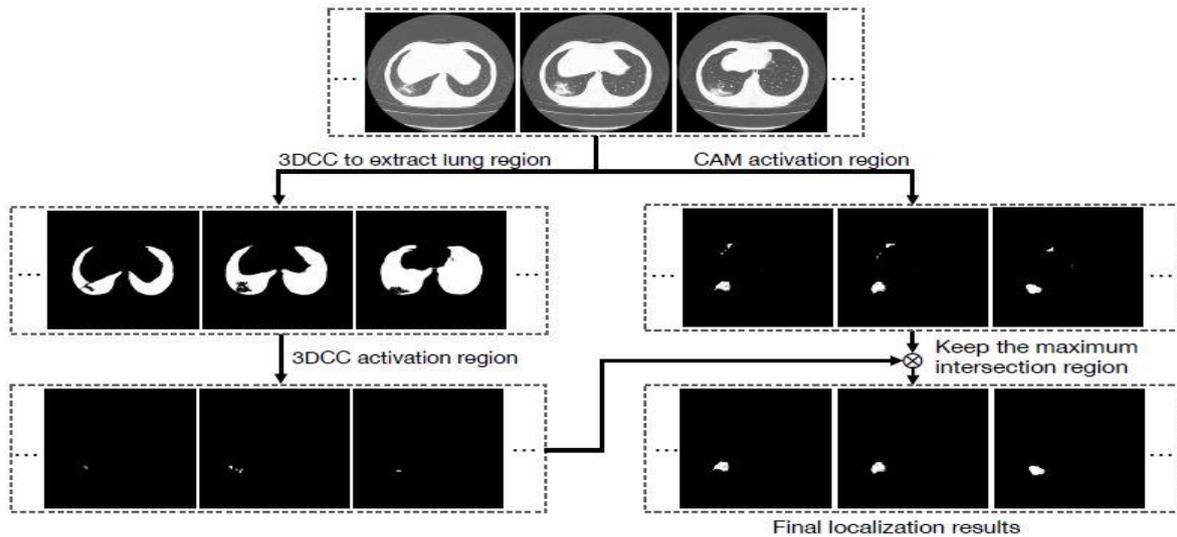


Figure 4: **Lesion localization**

Data preprocessing and data augmentation: Both CT volumes were pre-processed in a standardized way prior to preparing the 2D U-Net to segment the chests. Initially, the measurement unit was transferred to the Hounsfield unit (HU) scale, and the value was linearly transformed from 16-bit to 8-bit after the HU window threshold was calculated. After that, all CT volumes were re-sampled to the same spatial resolution, by which the CT volumes were modified with no impact of the cylindrical scanning limits of the CT scanner. This stage was also applied to the ground-truth chest covers that were obtained.

Pre-processing of DeCoVNet: For all of the CT scans, the chest covers created by the trained U-Net created the quantity of the cover. Next, the quantity of the CT was joined to the quantity of the cover to get the quantity of the CT mask. Finally, the CT mask quantity was re-sampled to a set spatial resolution without adjusting the amount of DeCoVNet training and testing parts(26). The overall quantity of items in the database was 14,116, varying from 73 to 250 pixels. The summary of different segmentation is given in Table 1.

Table 1: Summary of image segmentation methods in COVID-19 applications

Modality	Method	Target ROI	Application	Highlights
CT	U-Net	Lung	Diagnosis	Weakly-supervised method by pseudo labels
CT	U-Net	Lung Lesion	Quantification	
CT	U-Net	Lung Lung lobes lesion	Quantification	
CT	U-Net	Lung lobes Lesion	Quantification	
CT	U-Net/ Commercial Software	Lung Leson	Diagnosis	Combination of 2D and 3D methods
CT	U-Net	Lesion	Diagnosis	
CT	UNet++	Lesion	Diagnosis	
CT	UNet++	Lung Lesion	Diagnosis	Joint segmentation and classification
CT	VB-Net	Lung Lung lobes Lung segments Lesion	Quantification	Human-in-the-loop
CT	Commercial Software	Lung Lesion Trachea Bronchus	Quantification	

Data augmentation: During preparation, multiple random on-the-fly data increase techniques were used, including: (1) crop square patches at the center of input frames with a scale factor randomly chosen between 0.7 and 1, and crop sizes resized to 224×224; (2) rotation with an angle randomly selected between $\Delta = -25$ to 25; (3) random horizontal reflection, i.e. flipped the picture in the left-right direction, with a probability $p = 0.5$; and (4) adjust contrast by randomly darkening or brightening, with a factor ranging from 0.5 to 1.5.

Training and testing procedures: The DeCoVNet software was built in accordance with the PyTorch (open source machine learning) library (21). It is a suggested 3D deep convolutional neural network which

is assigned to detect COVID-19 from CT volumes. It's not just a typical CT scan. It is more effective and less time-consuming in detecting, and more precise in finding the actual condition of, the lungs (27). The proposed DeCoVNet was trained on a complete structure basis, which implies that CT volumes were received as inputs and that only the last production was tracked with no hand-operated interference applying automated sampling. The network has been trained over 100 iterations through the Adam optimizer (algorithm) (15) at a stable learning speed of $1e-5$ (1×10^5). Since the duration of the CT quantity of all the affected persons was not known, the batch dimension was made to 1.

Data-raising strategies were not applied throughout the testing process. The trained DeCoVNet took the pre-processed CT mask quantity of the entire affected persons and established both a positive COVID-19 probability and a negative COVID-19 probability. Predicted odds for all affected persons and their related ground-truth marks were then gathered for analytical review.

Statistical analysis: COVID-19 classification findings were documented and evaluated using receiver operational characteristics (ROCs) and precision recall curves (PRCs). The area under the curve and the area under the accuracy recall curve were determined. In addition, several operating points on the ROC curve were chosen. In order to quantitatively test the efficacy of the regulated swelling localization algorithm, the calculation parameter outlined by Nair & Hinton (21) was used to compute the swelling impact rate as follows. For any of the CT scans forecasted as successful by DeCoVNet, promising 3D swelling cover was forecasted by the suggested swelling localization algorithm; if the core of the forecasted 3D swelling cover was placed in either of the labelled boxes, it was a

strong hits; else, it could not be achieved. Finally, the impact probability was estimated by subtracting the number of hits by the number of misses.

Results

The detailed configuration of the 2DCIfNet was compared, as set out in Table 2. As a consequence of multi-scale testing, Figure 5 shows several instances of COVID-19 class initiation maps (CAMs) received at various stages of functionality, i.e. Conv3, Conv4 and Conv5. Hot areas denote where contagion occurs. The hotter an area is, the more likely it is to be contaminated. The suggested model learns to catch the spreading of swellings of various scales. Notably, we found mid-level layers, i.e. Conv3 and Conv4, to learn how to identify small swellings (GGOs) most often, particularly those shared externally and sub-plurally. However, they are not capable of capturing bigger patchy-like swellings, and this may be due to the limited mid-layer receptive field. On the other hand, the high-level sheet, i.e. Conv5, which has a relatively broad receptive recorded, learns well to identify huge piece-like swellings covering and expansion, in the middle of and peribronchially.

Table 2: Detailed structure of the compared 2DCIfNet

Stages	Layers	Output size
2D Stem	Conv2d(2, 64)@7×7+BN+ReLU	64×56×84
2D ResBlocks	ResBlock(64, 128)@3×3	128×56×84
	MaxPool2d	128×28×42
	ResBlock(128, 128)@3×3	128×28×42
2D ConvLayers	Conv2d(128, 128)@3×3+ReLU	128×28×42
	Conv2d(128, 128)@3×3+ReLU	128×28×42
	GlobalMaxPool2d	128×1×1
Classifier	StackSlices	T×128×1×1
	SliceMaxPool	128×1
	FullyConnected(128, 2)	2

For each CT scan, DeCoVNet forecast the possibility of COVID-19. Through applying them to their arbitrary

names, the ROC and PR) curves were plotted (see Figure 6). The AUC value of the ROC was 0.959.

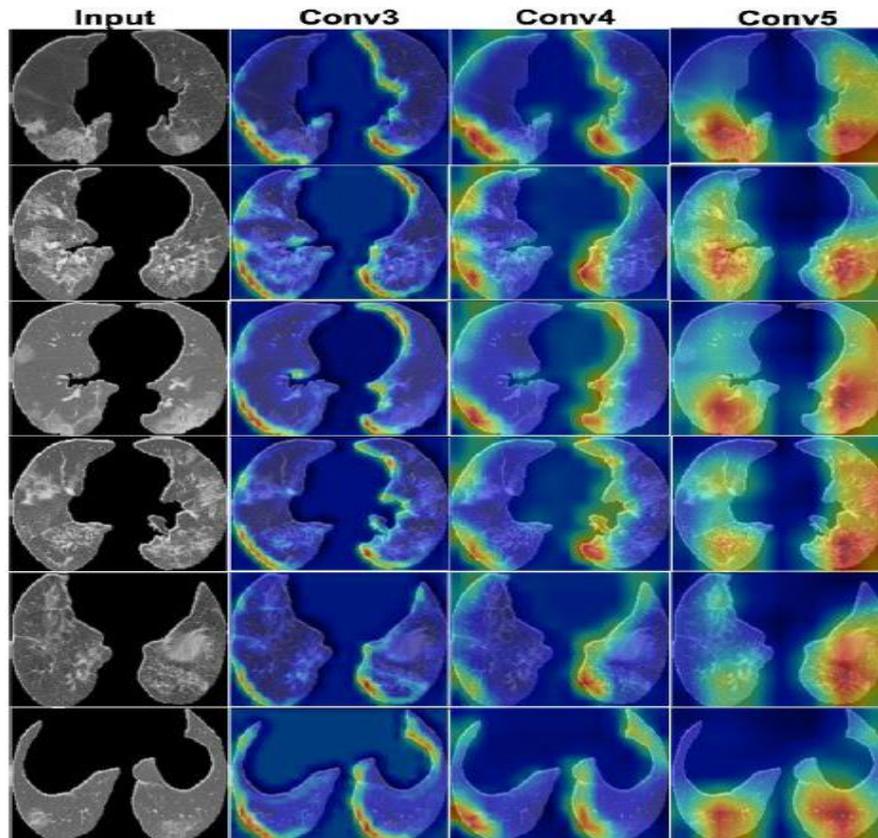


Figure 5: Class initiation maps of chest

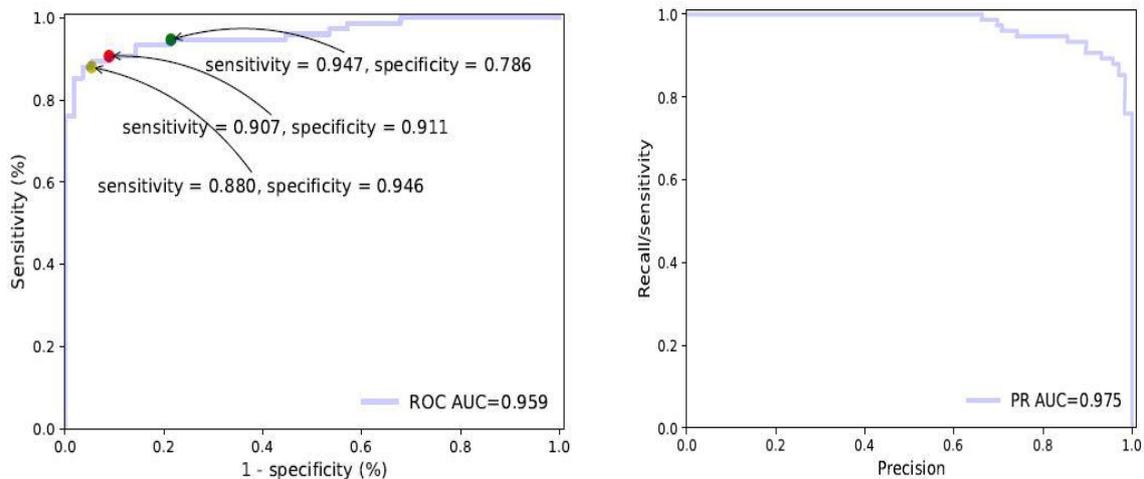


Figure 6a-b: COVID-19 classification outputs estimated through the receiver operating characteristic curve and precision-recall curve

Discussion

The suggested model learns to catch the spreading of swellings of various scales. Notably, we found mid-level layers, i.e. Conv3 and Conv4, to learn how to identify small swellings (GGOs) most often, particularly those shared externally and sub-pleurally. However, they are not capable of capturing bigger patchy-like swellings, and this may be due to the limited mid-layer receptive field. In the other side, the high-level sheet, i.e. Conv5, which has a relatively broad receptive recorded, learns well to identify huge piece-like swellings, as like insane covering and expansion, which are mostly scattered in middle and peribronchially.

Conclusions

Without the necessity to annotate COVID-19 swellings in CT volumes for processing, the proposed deep-

learning method has achieved a strong COVID-19 rating efficiency and a fair swelling localization check. As a consequence, our algorithm has a tremendous potential to be used in scientific applications for precise and accelerated detection of COVID-19, which is of immense benefit to frontline medical personnel and is therefore critical to managing this pandemic globally. The signature characteristics of CT imaging in the chest and Wuhan sensitivity or near touch background are strongly predictive of COVID-19 pneumonia, while RT-PCR stays as the recommendation norm. Standard CT properties of COVID-19 pneumonia consist of multifocal two-sided GGOs with irregular consolidation, well-known peripheral sub-pleural sharing, and favorite posterior or lower lobe predilection. Thin-sliced chest CT can aid early detection, direct scientific decision-making and track

disease development, and play a vital function in initial avoid and management of COVID-19. Albeit, it is not cost-effective to do so, it is crucial to know the exact condition of the lungs of suspected COVID patients. More care should be taken on radiologists' role in the war against this new contagious disease.

Authors' contributions

JZ designed and directed the study. YC and NZ performed radiological CT analysis and image processing. JZ also designed and directed the theoretical study. JZ, YC, NZ equally contributed to draft the paper.

References

1. Li T, Qian R, Dong C, Liu S, Yan Q, Zhu W, *et al.* BeautyGAN: Instance-level facial makeup transfer with deep generative adversarial network. *MM'18: Proceedings of the 26th ACM international conference on multimedia*.
2. Nanni L, Lumini A, Brahnam S. Ensemble of different local descriptors, codebook generation techniques and subwindow configurations for building a reliable computer vision system. *Journal of King Saud University – Science*. 2014;26(2):89-100.
3. Chen C, Dantcheva A, Swearingen T, Ross A. Spoofing faces using makeup: An investigative study. *Identity, Security and Behavior Analysis (ISBA)*, 2017 IEEE International Conference on IEEE: 1-8.
4. Choi Y, Choi M, Kim M, Ha J-W, Kim S, Choo J. StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation. *arXiv:1711.09020*, 2017.
5. Dolhansky B, Canton Ferrer C. Eye in-painting with exemplar generative adversarial networks. *arXiv:1712.03999*, 2017.
6. Gatys LA, Ecker AS, Bethge M, Hertzmann A, Shechtman E. Controlling perceptual factors in neural style transfer. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017: 3730-8.
7. Kim T, Cha M, Kim H, Lee J, Kim J. Learning to discover cross-domain relations with generative adversarial networks. *arXiv:1703.05192*, 2017.
8. Kim T, Kim B, Cha M, Kim J. Unsupervised visual attribute transfer with reconfigurable generative adversarial networks. *arXiv:1707.09798*, 2017.
9. Li Y, Song L, Wu X, He R, Tan T. Anti-makeup: Learning a bi-level adversarial network for makeup-invariant face verification. *arXiv:1709.03654*, 2017.
10. Liao J, Yao Y, Yuan L, Hua G, Kang SB. Visual attribute transfer through deep image analogy. *ACM Transactions on Graphics*. 2017;36(4):120.
11. Liu M-Y, Tuzel O. Coupled generative adversarial networks. *Advances in Neural Information Processing Systems*. 2016:469-77.
12. Liu S, Ou X, Qian R, Wang W, Cao X. Makeup like a superstar: Deep localized makeup transfer network. *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*: 2568-75.
13. Miyato T, Kataoka T, Koyama M, Yoshida Y. Spectral normalization for generative adversarial networks. *arXiv:1802.05957*, 2018.
14. Ulyanov D, Vedaldi A, Lempitsky VS. 2016. Instance normalization: The missing ingredient for fast stylization. *arXiv:1607.08022*, 2016.
15. Wei Z, Sun Y, Wang J, Lai H, Liu S. 2017. Learning adaptive receptive fields for deep image parsing network. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017: 3947-55.
16. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017: 6230-9.
17. Zhu J-Y, Park T, Isola P, Aefros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv:1703.10593*, 2017.
18. Lee CY, Xie S, Gallagher P, Zhang Z, Tu Z. Deeply supervised nets. *arXiv:1409.5185*, 2014.
19. Lin M, Chen Q, Yan S. Network in network. *arXiv:1312.4400*, 2013.
20. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. *ICML'13: Proceedings of the 30th International Conference on International Conference on Machine Learning*, 2013.
21. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. *ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning*, 2010.
22. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, *et al.* Imagenet large scale visual recognition challenge. *arXiv:1409.0575*, 2014.
23. Saxe AM, McClelland JL, Ganguli S. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. *arXiv:1312.6120*, 2013.
24. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. OverFeat: Integrated recognition, localization and detection using convolutional networks. *arXiv:1312.6229*, 2013.
25. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*, 2014.
26. Li X, Zeng X, Liu B, Yu Y. COVID-19 infection presenting with CT halo sign. *Radiology: Cardiothoracic Imaging*. 2020;2(1). <https://doi.org/10.1148/ryct.2020200026>
27. Zheng C, Deng X, Fu Q, Zhou Q, Feng H, Ma H. Deep learning-based detection for COVID-19 from chest CT using weak label. *IEEE Transactions on Medical Imaging*. 2020; <https://doi.org/10.1101/2020.03.12.20027185>.