RESEARCH

A novel deep learning-based method for COVID-19 pneumonia detection from CT images

Ju Luo[1], Yuhao Sun[2], Jingshu Chi[1], Xin Liao[2] and Canxia Xu[1*]

Abstract

Background: The sensitivity of RT-PCR in diagnosing COVID-19 is only 60–70%, and chest CT plays an indispensable role in the auxiliary diagnosis of COVID-19 pneumonia, but the results of CT imaging are highly dependent on professional radiologists.

Aims: This study aimed to develop a deep learning model to assist radiologists in detecting COVID-19 pneumonia.

Methods: The total study population was 437. The training dataset contained 26,477, 2468, and 8104 CT images of normal, CAP, and COVID-19, respectively. The validation dataset contained 14,076, 1028, and 3376 CT images of normal, CAP, and COVID-19 patients, respectively. The test set included 51 normal cases, 28 CAP patients, and 51 COVID-19 patients. We designed and trained a deep learning model to recognize normal, CAP, and COVID-19 patients based on U-Net and ResNet-50. Moreover, the diagnoses of the deep learning model were compared with different levels of radiologists.

Results: In the test set, the sensitivity of the deep learning model in diagnosing normal cases, CAP, and COVID-19 patients was 98.03%, 89.28%, and 92.15%, respectively. The diagnostic accuracy of the deep learning model was 93.84%. In the validation set, the accuracy was 92.86%, which was better than that of two novice doctors (86.73% and 87.75%) and almost equal to that of two experts (94.90% and 93.88%). The AI model performed significantly better than all four radiologists in terms of time consumption (35 min vs. 75 min, 93 min, 79 min, and 82 min).

Conclusion: The AI model we obtained had strong decision-making ability, which could potentially assist doctors in detecting COVID-19 pneumonia.

Keywords: Artificial intelligence, Deep learning, COVID-19, Community acquired pneumonia, CT image

Background

Since the beginning of 2020, the highly contagious novel coronavirus disease 2019 (COVID-19) has spread widely all over the world, causing a tremendous impact on the health and epidemic prevention systems of global countries, claiming millions of lives. After the novel coronavirus infects the human body, it mainly corrodes the lung region and causes lung inflammation, acute respiratory distress, or multiple organ failure in severe cases [1–4].

Reverse transcription polymerase chain reaction (RT-PCR) is used by doctors to evaluate whether patients are infected with novel coronavirus pneumonia (NCP). However, its harsh testing environment inevitably affects the rapid screening of suspected cases. As rapid RT-PCR testing becomes more available, challenges remain,

*Correspondence: xucanxia2000@163.com

[1] Third Xiangya Hospital, Central South University, NO.138, Tongzipo Road, Changsha 410013, Hunan, China
Full list of author information is available at the end of the article

Luo *et al. BMC Medical Informatics and Decision Making*   (2022) 22:284

Page 2 of 7

including high false negative rates and the sensitivity sometimes reported as low as 60–70% [5, 6].

As an important supplement to RT-PCR testing, radiographic imaging techniques, such as X-ray examination and computed tomography (CT), similarly play an indispensable role in the auxiliary diagnosis of NCP [7]. CT can detect early COVID-19 in patients with a negative RT-PCR test [8]. During the radiological examination of confirmed cases, researchers found that patients without symptoms, or before patients develop symptoms or after symptoms resolved, chest X-rays and CT images already showed changes associated with pneumonia induced lesions [9–11].

Artificial intelligence (AI) has profoundly transformed the way we live our lives, especially in the field of medical science. Deep learning (DL), a branch of AI, has made tremendous progress in diagnosis assistance and prognosis prediction along with the accumulation of abundant medical data and the improvement of computer algorithms. During the pandemic, due to the rapid increase in the number of new and suspected COVID-19 cases, DL had the potential to aid in the rapid evaluation of CT scans for the differentiation of COVID-19 findings from other clinical problems.

Some studies have already demonstrated the potential for AI-based diagnosis. Harmon et al. utilized a series of deep learning algorithms trained in a diverse multinational cohort of 1280 patients to detect COVID-19 pneumonia and achieved up to 90.8% accuracy [12]. Zhang et al. developed an AI system that can diagnose NCP and differentiate it from other common pneumonia and normal controls [13]. Fang et al. developed an early-warning system with deep learning techniques to predict COVID-19 malignant progression [14]. Yildirim et al. proposed a hybrid approach for diagnosing COVID-19 on chest X-ray images. The accuracy values obtained in the two different datasets were 99.05 and 97.1%, respectively [15]. Another study developed a hybrid model to diagnose COVID-19 from X-ray images and compared it with AlexNet, Resnet50, GoogLeNet, and VGG16. This hybrid model achieved the highest accuracy [16].

To assist physicians and radiologists in improving the accuracy of diagnosis and relieving the fatigue of checking large CT images. This research focused on the identification of COVID-19 infections in lung CT images in open source multi-institute and multidisease datasets. And propose a novel deep learning-based method for COVID-19 detection from CT images that achieves state-of-the-art accuracy.

Methods

Study population

This retrospective study included 222 patients positive for COVID-19, 88 community-acquired pneumonia (CAP), and 127 normal cases. Each patient contained 150–200 CT images provided with the Digital Imaging and Communications in Medicine (DICOM) format. All patients and CT images were obtained from the 2021 IEEE ICASSP Signal Processing Grand Challenge. All CT scans in training and validation were obtained by a SIEMENS SOMATOM Scope scanner with a normal radiation dose and a slice thickness of 2 mm. The test data-set was obtained from a different medical center with various slice thicknesses and radiation doses. COVID-19 cases were collected from February 2020 to April 2020, whereas CAP cases and normal cases were collected from April 2018 to December 2019 and January 2019 to May 2020, respectively. The diagnosis of COVID-19 infection was based on positive RT-PCR test results, clinical parameters, and CT scan manifestations identified by three experienced thoracic radiologists. The diagnosis of CAP was based on clinical parameters and CT scan manifestations identified by three experienced thoracic radiologists.

CT images datasets

The study population was divided into training, validation, and test datasets. Patient labels provided by three radiologists showed a high degree of agreement (more than 90%). Based on the patient-level labels, we built a slice-level labeled CT image data-set. Every case was analyzed by the radiologist to identify and label slices with evidence of infection. The labeled CT-image data-set contained 14,976 slices demonstrating infection and 40,553 slices without infection. The data-set was then divided into training, validation, and test sets, which are described in Table 1.

Table 1 The information of all datasets

	Training datasets (patients/slices/infection slices)	Validation datasets (patients/slices/infection slices)	Test datasets (patients)
Normal	52/9321/0	24/4525/0	51
CAP	41/7520/2468	19/3451/1028	28
COVID-19	116/20,208/8104	55/10,504/3376	51

Luo *et al. BMC Medical Informatics and Decision Making*   (2022) 22:284

Page 3 of 7

Data preprocessing

Data preprocessing mainly included lung region extraction and format conversion of slice images. U-Net with batch normalization after each layer as an automatic lung segmentation model for image preprocessing. The U-Net model performs precise segmentation on individual slices and extracts the right-left lung separately, including air pockets, tumors, and effusions. Therefore, helping the classification network focus more on the association between ground glass shadows and slice lesions. In the data-set, slices were displayed in DICOM format. To better adapt to the network, this study carried out a format conversion. The converted PNG format images were sorted according to the slice location value in the DICOM format to ensure that slices and labels corresponded to each other. The conversion threshold was adjusted according to the lung window range so that the lung tissue in the image was displayed more clearly and the lung texture highlighted.

Classification model

Classification and discrimination algorithms mainly included backbone network training based on ResNet-50 and custom discrimination rules based on clinical diagnosis experience. In order to prevent overfitting, the Batch-Normalization layer, decay learning rate method and early stop strategy work together to inhibit over-fitting. In the model training phase, we focused on the processing of input data. Through the observation of the images after lung segmentation, it was found that the first 10% and the last 14% part of the patient slice sequence rarely included image semantic information that was associated with distinguishing features. Experiments show that the classification accuracy of the network trained after truncating the patient slice set can be significantly improved. After the slice sequence was truncated at a certain ratio at the head and tail, the patient-level symptoms of the current sequence were classified according to the relative number of remaining slice prediction categories. When the current sequence was judged to be COVID-19 or CAP, the system directly output the prediction result. When the sequence was classified as normal, the system again compared the quantity of COVID-19 and CAP slices in the sequence. If their number exceeded 20% of the sequence length, the sequence was further judged as COVID-19 or CAP according to the relative quantity of corresponding slices. Figure 1 shows the architecture of proposed method. Relevant codes and models can be freely accessed at https://github.com/philiplaw1984/COVID-19/.

AI vs. doctors

A radiologist with different qualifications, including two experts and two novices, was defined as a radiologist who had clinical experience of two years or less, and an expert was defined as a radiologist who had clinical experience of five years or more. All four radiologists were asked to classify the data of the same validation set, and their results were compared with the AI model. The accuracy in classifying normal, CAP, and COVID-19 patients were calculated. The sensitivity of identifying COVID-19 patients and time consumption of diagnosis for the entire validation set were also calculated.

Results

Ablation comparison results in the validation dataset

For the model evaluation, 98 patient's slice sequences were divided into validation dataset, including 55 COVID-19 cases, 19 CAP cases, and 24 normal cases. Experiments showed that the classification accuracy of the network trained after truncating the patient slice set or introduce customized rules can be significantly improved. The results are shown in Table 2.

Classification model performance in the validation dataset

For the three-category classification task, Fig. 2 shows the confusion matrix. In the validation dataset, the accuracy of was 92.86%. In addition, The sensitivity for detecting COVID-19 were calculated. The sensitivity was 94.55%. The total diagnosis time of all 98 cases in the validation set was 35 min. The other results are shown in Table 3.

Radiologist performance in the validation dataset

For the three-category classification task, Fig. 3 shows the confusion matrix. In the validation dataset, the accuracy of the two novice radiologists was 86.73% and 87.75%. The accuracy of the two expert radiologists was 94.90% and 93.88%. The sensitivity of the two novice radiologists for identifying COVID-19 was 81.82% and 89.09%. The sensitivity of both expert radiologists for identifying COVID-19 was 94.55%. The total diagnosis time of the two novice radiologists was 75 min and 93 min, and the total diagnosis time of the two expert radiologists was 79 min and 82 min.

AI vs. doctors

In the validation set, the diagnostic accuracy of AI model for all three classifications was 92.86%. The sensitivity of the model for COVID-19 was 94.55%. The results indicated that AI model performed significantly better than all four radiologists in terms of time consumption. The
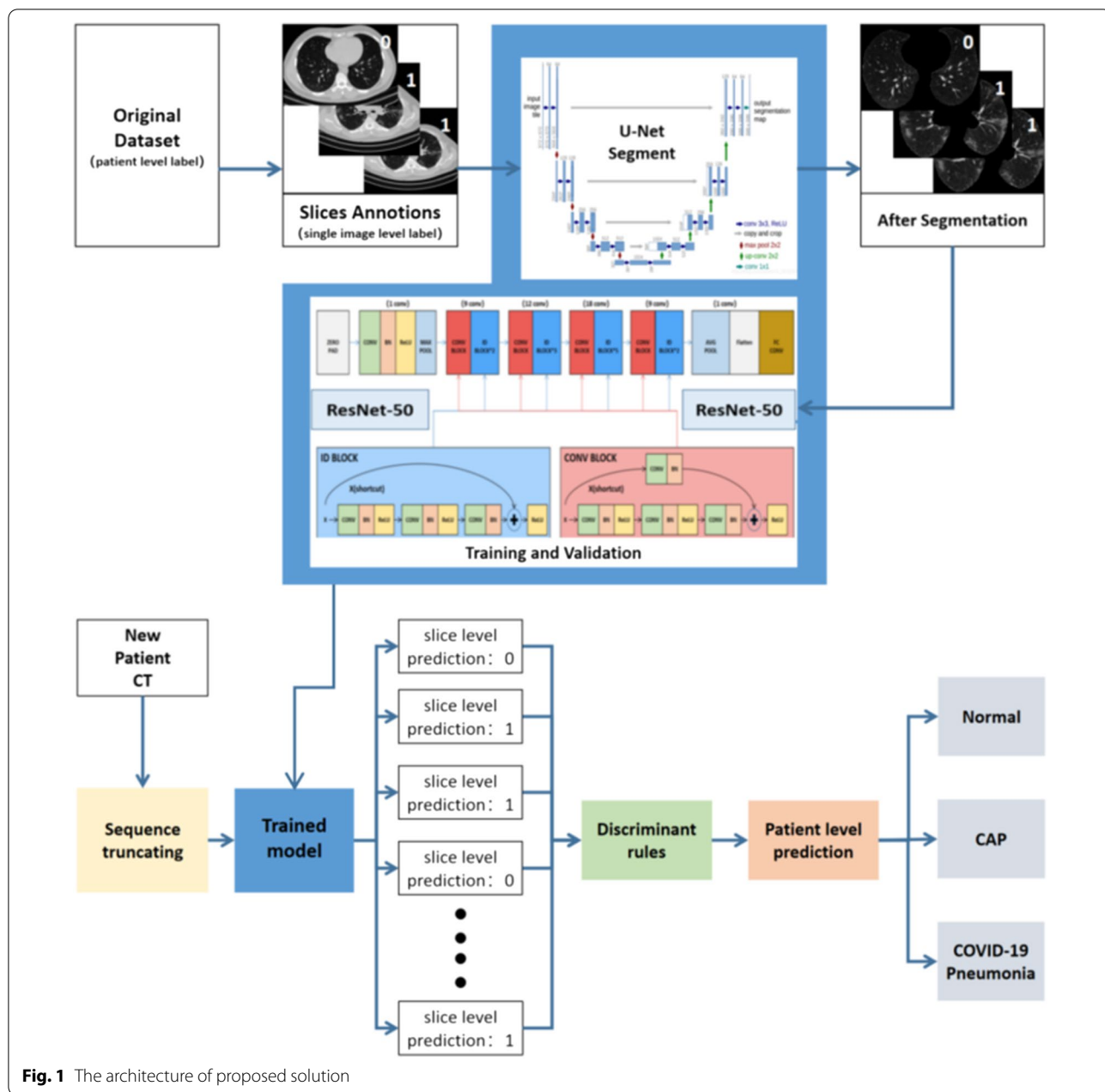
Luo *et al. BMC Medical Informatics and Decision Making* (2022) 22:284

Page 4 of 7

Fig. 1 The architecture of proposed solution

Table 2 Comparison of ablation experiment results in three classification tasks

Ablation comparison	Sens (C1)	Sens (C2)	Sens (C3)	Acc
Without sequence truncating	0	0.7895	0.8727	0.6429
Without customized rules	1	0.3158	0.4909	0.5816
Proposed method	1	0.7895	0.9455	0.9286

"C1" represents "Normal", "C2" represents "CAP", and "C3" represents "COVID-19"

Fig. 2 The confusion matrix for the three-category classification task in the validation dataset

Table 3 Evaluation metrics of the model

	Acc	Sens	Spe	Fpr	Fnr	Fdr	F1
Normal	0.9898	1	0.9865	0.0135	0	0.04	0.9796
CAP	0.9388	0.7895	0.9747	0.0253	0.2105	0.1176	0.8334
COVID-19	0.9286	0.9636	0.907	0.093	0.0364	0.0714	0.9458

Fig. 3 The confusion matrix for the three-category classification task in the validation set of radiologists
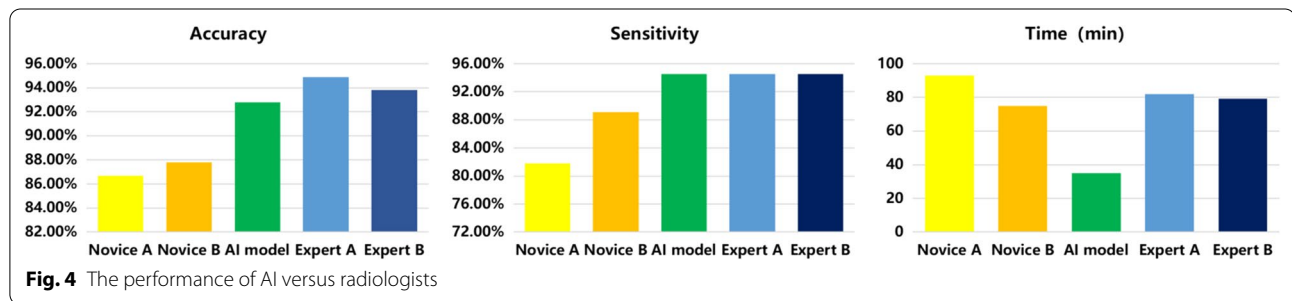
Fig. 4 The performance of AI versus radiologists

Fig. 5 The confusion matrix for the three-category classification task in the test set

AI model performed better than novice radiologists and better than expert radiologists in terms of the accuracy and sensitivity of diagnosis, as shown in Fig. 4.

Classification model performance in the test set
For the three-category classification task, Fig. 5 shows the confusion matrix. In the test dataset, the accuracy was 93.84%. In addition, we calculated the sensitivity of the three classes. The sensitivity (C1) was 0.9803, the

sensitivity (C2) was 0.8928, and the sensitivity (C3) was 0.9215.

The experiments demonstrated that the proposed solution was effectively utilized to solve the classification problem of COVID-19, CAP, and normal cases based on lung CT images. All the datasets in this study were obtained from the 2021 IEEE ICASSP Signal Processing Grand Challenge. Our proposed method was the first place winner of this challenge.

Discussion
In this study, an AI system for the diagnosis of COVID-19 pneumonia based on ResNet-50 and U-Net were developed. U-Net is an automatic lung segmentation method based on the semantic segmentation architecture, which will help the network focus on the lung region and extract distinguishing features [17]. Lung segmentation is able to extract all the lung regions, thereby helping the classification network focus more on the association between ground glass shadows and solid lesions. Then,

the ResNet-50 was used for every CT slice classification. ResNet-50, which is often used as the backbone of the classification network, is also widely used in COVID-19 detection tasks [18]. Çınar et al. developed a model based on the layers of ResNet50. With this new model, the diagnosis of pneumonia can be made early and accurately [19]. Narin et al. [20] used ResNet-50 to solve the detection of novel coronavirus infection in chest X-ray images and obtained satisfactory accuracy. Aram et al. [21] used ResNet-50 as the feature extractor in the training phase, and in the COVNet proposed by Li et al. [22], parallel ResNet-50 with shared weights constituted their model.

However, the AI system is often hampered by problems with low universality due to the uniformity of data. Therefore, our study was specifically designed to maximize the potential for generalizability. All the CT images come from an open source dataset collected by multiple institutes in Canada. The training and validation datasets have a normal radiation dose and a slice thickness of 2 mm. The CT images from the test dataset were obtained from different medical centers with various exposure doses, slice thicknesses, and ranges of values in the Hounsfield Unit. Surprisingly, AI system's performance in the test dataset was equally matched with the validation dataset. Thus it can be seen that our AI system has excellent universality and robustness.

Several factors set this study apart from similar prior efforts: (1) to expand the training images, we performed additional slice-level annotations on all patient slice sequences marked as "Training" the dataset. (2) The U-Net-based semantic segmentation architecture was used to segment each lung slice. Only the lung tissue, not the original CT images, was input into the network for training. (3) We chose the ResNet-50 model, which is commonly used in COVID-19 detection, as the backbone of the classification network. However, we defined the discriminant rules based on the clinical diagnosis and treatment experience of medical imaging. Recently, most AI and COVID-19 studies have focused on big data and complex computer algorithms. Few studies have introduced discriminant rules into the AI model.

RT-PCR from respiratory samples is the standard for diagnosis. However, the sensitivity of testing varies from 33 to 80% [23–25]. Chest CT imaging findings are nonspecific and overlap with other infections, especially CAP, so the diagnostic value of CT imaging for COVID-19 is limited [26]. However, with the help of the AI model, doctors could distinguish CAP and COVID-19 more easily. Moreover, under some circumstances, patients admitted to the hospital with abnormal chest CT imaging findings compatible with RT-PCR were negative. If the model was highly accurate for COVID-19, doctors could increase the frequency of RT-PCR tests and transfer and treat patients earlier. In that way, it would be of great help to the patient and public health.

Our study carried out an AI versus radiologist experiment. The results showed that the AI model detected COVID-19 with a high accuracy of 93.84% and sensitivity of 92.15%, which was better than novice radiologists and equivalent to expert radiologists. This indicates that the AI model could significantly improve the ability of novice radiologists. In addition to diagnostic accuracy and sensitivity, this study also found that the AI method showed a significant saving of time compared to all radiologists. When the COVID-19 outbreak occurred, many patients rushed to the hospital, and a massive number of suspected patients needed to be excluded. COVID-19 has the potential to overwhelm the local health care system. However, with the assistance of the AI model, COVID-19 could be diagnosed quickly and accurately. The AI model will provide strong support to the medical system, especially in areas with weak medical infrastructure.

There are several limitations to this study. First, the sample size of our research was smaller than that of other multinational studies. All the patients and CT images were not collected by our own but from an open source dataset released by the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing. Second, all COVID-19 cases in this study were based on RT-PCR positivity and CT positivity. However, CT images are often negative despite a positive RT-PCR test [27]. The AI model may not have the ability to identify CT-negative cases. However, more suitable for the situation when the RT-PCR is negative but patient is highly suspected to be infected. If the AI model was diagnosed with COVID-19, repetitive RT-PCR is indispensable. Finally, the AI algorithm aimed to classify normal, CAP, and COVID-19 pneumonia cases. However, many other pneumonias, such as fungal infection and influenza pneumonia, also need to be distinguished. The situation in real clinical practice may be more complicated, so the AI model also needs to be improved.

In conclusion, we have seen broad prospects for the combination of deep neural networks and the field of medical imaging. Relying on its powerful feature extraction capability, deep neural networks can obtain excellent discrimination accuracy with the support of sufficient data and appropriate models. Moreover, the customized discriminant rules based on clinical diagnosis experience can impose constraints on the model and improve the interpretability. Combining the above two points, the model this study obtained could have a stronger decision-making ability, which could potentially assist doctors in detecting COVID-19 pneumonia.

Luo *et al. BMC Medical Informatics and Decision Making*   (2022) 22:284

Page 7 of 7

Declarations

Ethics approval and consent to participate
This study conformed to the ethical guidelines of the Declaration of Helsinki and was approved by the ethics committee of the Third Xiangya Hospital of Central South University. All patient and image information was erased by the IEEE Signal Processing Society. According to national legislation and institutional requirements, informed consent was waived by the Ethics Committee of the Third Xiangya Hospital of Central South University due to the public datasets and retrospective nature of this study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Third Xiangya Hospital, Central South University, NO.138, Tongzipo Road, Changsha 410013, Hunan, China. [2]College of Computer Science and Electronic Engineering, Hunan University, Changsha, China.

References
1. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395:507–13.
2. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA. 2020;323:1061–9.
3. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020;382:1199–207.
4. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First case of 2019 novel coronavirus in the United States. N Engl J Med. 2020;382:929–36.
5. Yang Y, Minghui Y, Chenguang S, Fuxiang W, Jing Y, Jinxiu L et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis
6. and monitoring the viral shedding of 2019-nCoV infections. 2020. https://www.medrxiv.org/content/10.1101/2020.02.11.20021493v2.
6. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology. 2020;296:E115–7.
7. Fan DP, Zhou T, Ji GP, Zhou Y, Chen G, Fu H, et al. Inf-Net: automatic COVID-19 lung infection segmentation from CT images. IEEE Trans Med Imaging. 2020;39:2626–37.
8. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. Radiology. 2020;296:E41–5.
9. Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395:514–23.
10. Inui S, Fujikawa A, Jitsu M, Kunishima N, Watanabe S, Suzuki Y, et al. Erratum: chest CT findings in cases from the cruise ship "Diamond Princess" with coronavirus disease 2019 (COVID-19). Radiol Cardiothorac Imaging. 2020;2:e204002.
11. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology. 2020;296:E32-40.
12. Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun. 2020;11:4080.
13. Zhang K, Liu X, Shen J, Li Z, Sang Y, Wu X, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. Cell. 2020;182:1360.
14. Fang C, Bai S, Chen Q, Zhou Y, Xia L, Qin L, et al. Deep learning for predicting COVID-19 malignant progression. Med Image Anal. 2021;72:102096.
15. Yildirim M, Eroğlu O, Eroğlu Y, Çinar A, Cengil E. COVID-19 detection on chest X-ray images with the proposed model using artificial intelligence and classifiers. New Gener Comput. 2022. https://doi.org/10.1007/s00354-022-00172-4.
16. Yildirim M, Cinar A. A deep learning based hybrid approach for COVID-19 disease detections. Traitement du Signal. 2020;37:461–8.
17. Hofmanninger J, Prayer F, Pan J, Rohrich S, Prosch H, Langs G. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. Eur Radiol Exp. 2020;4:50.
18. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra AU. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med. 2020;121:103792.
19. Çinar A, Yıldırım M, Eroğlu Y. Classification of pneumonia cell images using improved ResNet50 model. Traitement du Signal. 2021;38:165–73.
20. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. Pattern Anal Appl. 2021;24:1207–20.
21. Ter-Sarkisov A. COVID-CT-Mask-Net: prediction of COVID-19 from CT Scans using regional features. Res Sq. 2020. https://doi.org/10.21203/rs.3.rs-104621/v1.
22. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. Radiology. 2020;296:E65-71.
23. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA. 2020;323:1843–4.
24. Sethuraman N, Jeremiah SS, Ryo A. Interpreting diagnostic tests for SARS-CoV-2. JAMA. 2020;323:2249–51.
25. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. Ann Intern Med. 2020;173:262–7.
26. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. JAMA. 2020;324:782–93.
27. Yang W, Yan F. Patients with RT-PCR-confirmed COVID-19 and normal chest CT. Radiology. 2020;295:E3.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.