Review

Utilisation of deep learning for COVID-19 diagnosis

S. Aslani*, J. Jacob

Centre for Medical Image Computing and Department of Respiratory Medicine, University College London, London, UK

The COVID-19 pandemic that began in 2019 has resulted in millions of deaths worldwide. Over this period, the economic and healthcare consequences of COVID-19 infection in survivors of acute COVID-19 infection have become apparent. During the course of the pandemic, computer analysis of medical images and data have been widely used by the medical research community. In particular, deep-learning methods, which are artificial intelligence (AI)-based approaches, have been frequently employed. This paper provides a review of deep-learning-based AI techniques for COVID-19 diagnosis using chest radiography and computed tomography. Thirty papers published from February 2020 to March 2022 that used two-dimensional (2D)/three-dimensional (3D) deep convolutional neural networks combined with transfer learning for COVID-19 detection were reviewed. The review describes how deep-learning methods detect COVID-19, and several limitations of the proposed methods are highlighted.
© 2022 The Authors. Published by Elsevier Ltd on behalf of The Royal College of Radiologists. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

The SARS-CoV-2 (COVID-19) virus, detected in December 2019[1] has, as of April 2022, infected over 507 million people across the world and resulted in a global death toll of over 6 million people.[2] Effective and efficient primary screening for COVID-19 infection has been a cornerstone of management.[3] The standard primary screening tool for COVID-19 has been the reverse transcription polymerase chain reaction (RT-PCR) test,[4] in which the ribonucleic acid (RNA) of the COVID-19 virus is identified in sputum samples obtained from the upper respiratory tract; however, early in the pandemic, several studies highlighted a variable sensitivity of RT-PCR tests, which were influenced by the time of collection of the specimen relative to the time of infection.[5,6] During the acute phase of the pandemic, laboratory services were often overwhelmed by the volume of testing required, resulting in diagnostic delays. As a result, an additional diagnostic screening method considered alongside RT-PCR was the examination of chest radiographs (CXR).[7,8] CXRs were favoured as the necessary equipment is easy to access, they are fast to perform and interpret, and portable systems can markedly reduce the chances of virus transmission[9,10]; however, a major challenge in CXR screening during the pandemic was the limited number of expert radiologists available for interpreting imaging data.[11]

In some European centres, computed tomography (CT) imaging was used to screen patients for COVID-19.[6,12,13] CT imaging was shown to be more sensitive to CXRs in diagnosing COVID-19, particularly in cases where the diagnosis

* Guarantor and correspondent: S. Aslani, Department of Respiratory Medicine, University College London, Gower Street, London WC1E 6BT, UK.
   E-mail address: a.shahab@ucl.ac.uk (S. Aslani).

was incidental, for example, in the work-up of elective surgical patients.[14,15] CT has also been valuable in assessing the lungs of patients with worsening respiratory complications and in patients with negative RT-PCR test results where COVID-19 infection remained in the differential. Yet the specificity of CT for diagnosing COVID-19 (as with CXRs) is limited[16,17] making radiological interpretation of imaging for COVID-19 diagnosis challenging.[18–20]

Known workforce shortages of radiologists, combined with the low specificity of chest imaging methods in diagnosing COVID-19, led many research groups to develop AI-based algorithms to support clinicians and radiologists diagnosing COVID-19.[21] The present review provides an overview of previously proposed AI methods using deep-learning (DL) algorithms for diagnosing COVID-19 using CXR and CT.

Application of DL in COVID-19 detection

DL in COVID-19 detection using CXR images

As mentioned above, the convenience and ubiquity of CXRs in the assessment of COVID-19 led to an exponential growth in the acquisition of CXR data. The resulting large datasets were leveraged by AI researchers to develop automated DL algorithms for COVID-19 detection. COVID-19 diagnosis has typically been considered in as a two-group or three-group classification challenge. When considered a two-group problem, computer algorithms were trained to distinguish between COVID-19 imaging and imaging acquired from healthy controls. Classification in this way, however, does not allow inference of how an algorithm would interpret a CXR containing non-COVID-19 pneumonia. A three-group classification task aims to distinguish between COVID-19 pneumonia, non-COVID-19 pneumonia, and normal CXR imaging.

The method proposed by Apostolopoulos et al.[22] is one of the earliest AI-based approaches proposed for COVID-19 detection. In this paper, authors used state-of-the-art convolutional neural networks (CNN) including VGG19,[23] MobileNet v2,[24] Inception,[25] Xception,[26] and Inception ResNet v2.[25] The process of transfer learning was used for COVID-19 diagnosis. In DL analyses, transfer learning is a commonly applied process where computer models previously trained for a specific task (e.g., classification of two image classes such as cats and dogs) reuse the stored knowledge gained in the initial task and apply it to a new but related task (e.g., classification of two image classes such as COVID-19 and non-COVID-19 CXRs). In the study by Apostolopoulos et al.[22] the best accuracies achieved for two-group (normal, COVID-19) and three-group (normal, bacterial pneumonia, COVID-19) classification were 98.75% and 93.48%, respectively. Moreover, they tested the proposed model on a separate dataset with additional viral pneumonia cases, and the reported performance was 96.78% and 94.72% for two-group and three-group classification, respectively.

Ozturk et al.,[28] designed a 19-layer CNN known as DarkCOVIDNet, which was trained and tested in two-group and three-group classification tasks. They specifically visualised the heat maps of the proposed model using the Grad-CAM[33] approach. A heat map or saliency map is a graphical two-dimensional representation of CNN information that uses a colour-coding system to represent areas of differing importance within the image. An example of a CXR heat map produced using the Grad-CAM method can be seen in Fig 1. The outputs of the heat maps in the study by Ozturk et al.[28] were assessed qualitatively by an expert radiologist. The optimal performance obtained by the DarkCOVIDNet was 98.08% and 87.02% for two-group and three-group classification tasks, respectively.

Rahimzadeh et al.[29] proposed a twin CNN (TCNN) architecture, using two well-known CNNs, Xception[26] and ResNet,[34] to extract parallel deep features from an image. A deep or latent feature is the consistent model response or output at the last node or layer. The extracted latent features were combined for the final prediction. The accuracy of two-group and three-group classification tasks were 99.05% and 94.40%. The results suggested that a TCNN approach could boost algorithm performance for COVID-19 prediction. Similar to Rahimzadeh et al.,[29] Ouchicha et al.[30] proposed a TCNN with shared layers known as CVDNet. The designed architecture was tested on a three-group classification task producing an accuracy of 96.69%.

A challenging area for computer algorithms when assessing and classifying diseases on CXRs lies in the region of the diaphragm. Diaphragm contain areas of high density that can confuse algorithms and, in the context of COVID-19 classification, result in misdiagnosis. Heidari et al.[27] proposed a pre-processing algorithm that can boost the performance of a CNN by identifying and removing the diaphragmatic area with multi-stage image processing algorithms. The results in this paper suggest that using their proposed processing pipeline can improve model performance in COVID-19 detection to 98.1% and 94.5% for two-group and three-group classification tasks, respectively.

SqueezeNet is a well-designed CNN proposed for natural image classification, which uses fewer parameters than other models.[35] In the study by Ucar et al.,[31] the building blocks of SqueezeNet were used to design a new network called COVIDiagnosis-Net. To perform tuning of model hyperparameters, they used a Bayesian optimisation[36] algorithm. Moreover, multi-scale offline image augmentation was performed to overcome imbalances in data classes. The output accuracy of COVIDiagnosis-Net for the three-group classification task was reported as 98.30%.

One of the most successful strategies used to improve classification performance is Ensemble Learning (EL).[37] EL combines the output of several independent deep-learning models, each of which may have individual and distinct strengths. The expectation is that the various models will show complementary performance when combined, making the ensemble greater than the sum of its parts and more robust to unseen data. Rajaraman et al.[38] combined the predictions of nine individual DL-based models for two- and three-group classification tasks using several well-known ensemble strategies including max voting, averaging, weighted averaging, and stacking. Moreover,
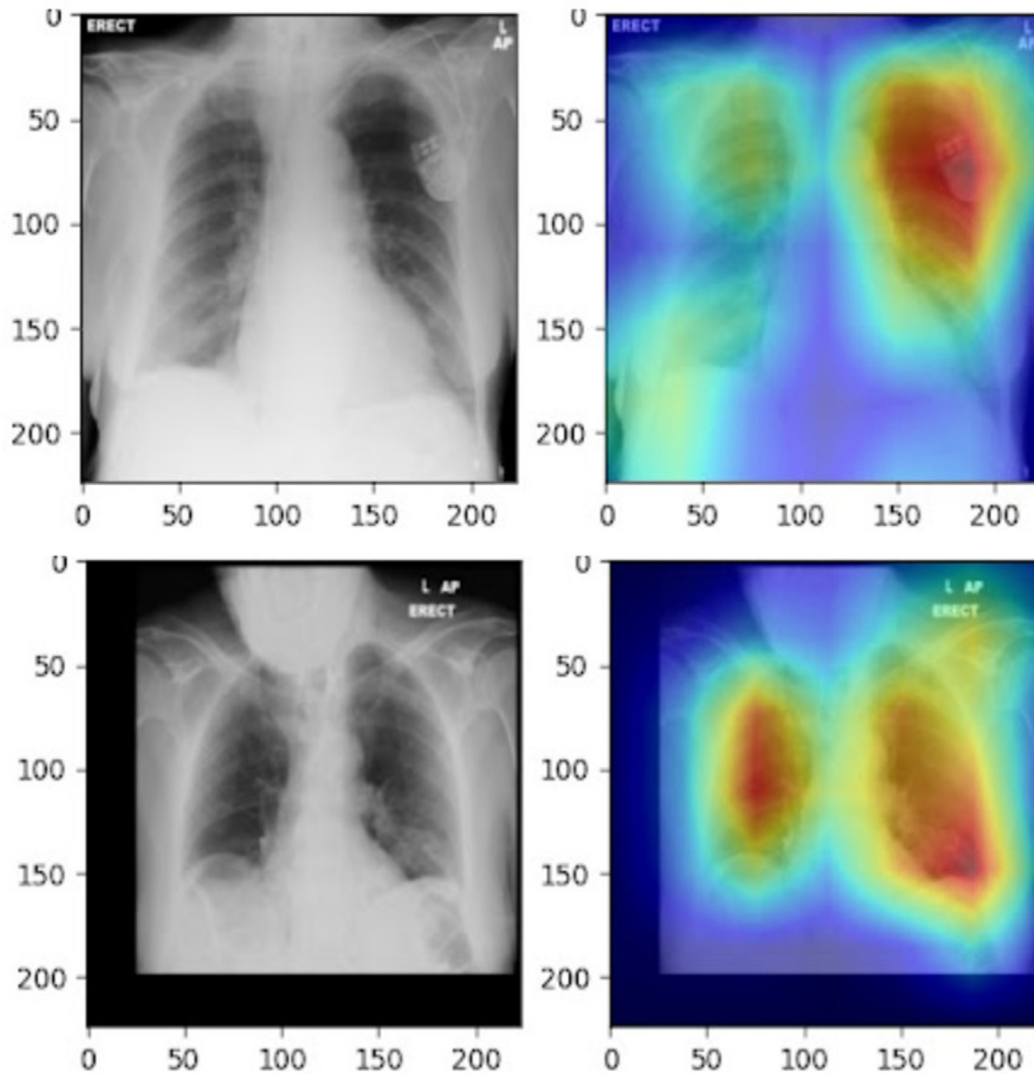
Figure 1 Representation of CNN outputs on a heat map using Grad-CAM.[33] The first column shows the original CXR images from NCCID dataset[72] and the second column shows the heat-map representation of the model output for the corresponding images.[73] Warmer colours show strong signals (higher values) and colder colours show weak signals (lower values).

they also proposed an iterative pruning strategy to fine-tune the model. The iterative pruning strategy was able to identify the optimal number of layers for a given network and, in so doing, decreased the complexity of the model without compromising model performance. The accuracy reported for the three-group classification task was 98% without pruning and 99% when the pruning strategy was employed.

Other CNN based models for COVID-19 detection include multi-dilation CNN[39]; CoroNet,[21] which uses the Xception[26] architecture; COVID-CAPS[40] and Convolutional capsnet,[41] which use a capsule network-based framework[42]; and algorithms based on ResNet[43] and MobileNet[44] architectures.

Training DL models requires a considerable amount of data; however, the main challenge faced by the models mentioned previously in this review lay in the small sample size of the datasets available for model training. This was essentially a consequence of difficulties in data sharing in the early stages of the pandemic and the focus on acute clinical care in the emergency setting. Data limitations can result in early models demonstrating poor generalisability for unseen or out-of-distribution data. To overcome the handicap of limited data, most of the recently proposed algorithms based on DL approaches utilise transfer learning to boost performance; however, the main disadvantage of transfer learning is that of negative transfer, which occurs as the initial and target tasks are not similar enough to allow satisfactory training of the model. To avoid this, the authors in[32] took the novel approach of generating synthetic CXR images (using an algorithm called COVIDGAN) using generative adversarial networks (GANs).[45] Specifically, they used an auxiliary classifier generative adversarial network (ACGAN)[46] architecture to enhance COVID-19 detection. The results suggested that the synthetic images produced by COVIDGAN can increase COVID-19 detection accuracy by up to 10%.

Table 1
Detailed information about the proposed artificial intelligence deep-learning-based models for COVID-19 detection using radiographic images.

Ref.	Methods	Data groups	Tasks	Accuracy per-task (%)
Heidari et al.[27]	VGG16	445 COVID-19 5,179 Non-COVID-19 pneumonia 2,880 Normal	Two groups (normal, COVID-19) Three groups (normal, pneumonia, COVID-19)	98.1 94.5
Apostolopoulos et al.[22]	VGG19 Mobile Net Inception ResNets	224 COVID-19 700 Bacterial pneumonia 504 normal	Two groups (normal, COVID-19) Three groups (normal, pneumonia, COVID-19)	98.7 93.5
Ozturk et al.[28]	Custom CNN	130 COVID-19 500 pneumonia 500 normal	Two groups (normal, COVID-19) Three groups (normal, pneumonia, COVID-19)	98.08 87.20
Ranimzaden et al.[29]	TCNN	180 COVID-19 6,054 Pneumonia 8,851 Normal	Two groups (normal, COVID-19, Three groups (normal, pneumonia, COVID-19)	99.05 94.40
Khan et al.[21]	Xception	290 COVID-19 330 Bacterial pneumonia 330 Viral pneumonia 310 Normal	Two groups (normal, COVID-19) Three groups (normal, pneumonia, COVID-19) Four groups (normal, pneumonia bacterial, pneumonia viral, COVID-19)	98.80 94.52 89.60
Ouchicha et al.[30]	TCNN	219 COVID-19 1,345 Viral pneumonia 1,341 Normal	Three groups (normal, pneumonia, COVID-19)	96.69
Ucar et al.[31]	Deep-SqueezeNet	76 COVID-19 4,290 Non-COVID-19 pneumonia 1,583 Normal	Three groups (normal, pneumonia, COVID-19)	98.30
Waheed et al.[32]	COVIDGAN	403 COVID-19 721 Normal	Two groups (normal, COVID-19)	95.00

A summary of the proposed methods using CXR images for COVID-19 detection can be seen in Table 1.

DL in COVID-19 detection using CT

In several countries, such as The Netherlands,[47,48] CT was the primary imaging methodology used to assess the lung to identify COVID-19 infection. The improved spatial resolution of CT over CXR can enhance the sensitivity of COVID-19 detection. This is particularly valuable in detecting subtle early disease or identifying ground-glass densities hidden by the heart or hemidiaphragm on frontal CXRs[6,49]; however, workforce limitations in radiology departments during the pandemic meant that reading vast numbers of CT studies to detect COVID-19 became a challenging task.[50] When coupled with reader intra- and interobserver variabilities, AI experts began to propose automated algorithms to diagnose COVID-19 on CT.

An early method proposed by Chen et al.[51] was based on a 2D Unet++[52] pre-trained on ImageNet.[53] The model was initially trained to segment regions of interest within the lungs on CT and then predict suspicious lesions within these regions. The best accuracy reported by this approach was 92.59%. Similarly, Gunraj et al.[54] proposed a COVID-19 detection algorithm based on a pre-trained 2D CNN combined with optimised neural network architecture, informed by human prerequisites for model sensitivity and positive predictive value for COVID-19 detection. Specifically, they benefited from a machine-driven design exploration strategy proposed in[55] to design the network architecture automatically and identify the designed architecture patterns/blocks. The proposed model, called COVIDNet-CT, had a reported accuracy for a three-group (normal, non-COVID-19 pneumonia, and COVID-19) classification task of 99.10%.

Another study used the COVNet TCNN model,[56] which is based on two parallel ResNet50s (CNNs that are 50 layers deep)[34] with weight sharing. Similar to the previously mentioned models, this method first identified the lungs using a 2D Unet[57] and then trained the TCNN to detect features of COVID-19 within the lungs. The authors reported an area under the receiver operating characteristic curve (AUROC) value of 96.50%. Utilising the same concept, Wang et al.[58] designed a TCNN architecture based on 3D networks. They used a 3D U-Net[59] for lung segmentation and then combined two 3D-ResNets for COVID-19 detection. They utilised a prior-attention mechanism and proposed a new prior-attention residual learning block to boost the model's performance. The reported accuracy and AUROC were 93.30% and 97.30%, respectively. Song et al.[60] proposed an architecture based on three parallel 2D ResNets with weight sharing to extract different levels of lung CT features, including global features, detailed local features, and rational features. They directly extracted global features from the whole lung area as an initial processing step using ResNet50. A feature pyramid network algorithm was utilised to segment the lung area and generate sub-region/images at different scales. Then, based on the defined sub-

regions, a ResNet50 was used to extract local features at each region and relational features between regions. They later combined these features for the final COVID-19 prediction. The proposed method had an accuracy of 93% and an AUROC of 95%.

Acar et al.[61] proposed a pipeline based on GAN[45] algorithms to increase the performance of the previously developed CNNs for COVID-19 detection. They initially extracted the lungs using a deep network called BDCU-Net.[62] They then used a GAN model to synthesise new lung CT images, thereby increasing the number of samples available for training. For their final step, they utilised several well-known pre-trained 2D CNNs to diagnose COVID-19 and demonstrated that using GANs in the baseline models could increase COVID-19 diagnostic accuracy up to 9%.

One of the main challenges in training AI models for COVID-19 detection using CT is having good quality and abundant manual labels of regions of COVID-19-affected lung parenchyma. A strategy that has been used to mitigate against limited labels has been to use weakly supervised DL algorithms[63] that can efficiently work on datasets without labels but can maintain sufficient accuracy for COVID-19 detection. Inspired by the work of Yang et al.,[64] Hu et al.[63] initially proposed a 2D Unet multi-view segmentation model with additional attention layers to extract the lungs from the CT. Later they modified the VGG[23] network adding a weakly supervised multi-scale learning algorithm to increase COVID-19 detection performance. The accuracy and AUROC for this model were 87.4% and 89.60%. Similarly, Wang et al.[65] proposed a weakly supervised framework based on a 2D Unet and a 3D CNN. As a first stage, they used a 2D Unet model to segment the lungs on CT images. Then they used the segmented mask to extract the lung area and combined the slices to create a 3D lung mask. They then trained a 3D CNN encoder called DeCoVNet by concatenating the original 3D scans with the lung masks and combined this with a weakly supervised COVID-19 lesion localisation algorithm to classify the scans. The accuracy and AUROC for this approach were 90.10% and 95.90%, respectively.

Other similar approaches based on CNNs for COVID-19 detection include methods based on Unet architectures,[66,67] methods that utilised attention-based networks,[68] and other CNN architectures focused on segmentation and classification.[20,50,69,70] A summary of the various CNN methods that used CT can be seen in Table 2.

Discussion

The present review described the application of DL-based AI models to detect and diagnose COVID-19 on CXR and CT. In total, 30 studies performed within the time frame of the review were assessed. Sixteen studies with pioneering approaches to COVID-19 diagnosis are described, detailing the model architectures and their reported performance. (The remaining 14 studies used similar approaches, and their model architectures are summarised

briefly.) Yet, comparing performance between models is an essentially futile exercise as none of the models described examined a common standardised dataset, thereby making an unbiased definitive comparison impossible.

Based on the present review, it is clear that there are several challenges related to the application of AI for COVID-19 detection. Invariably when image analysis of CXR or CT images for COVID-19 detection is performed, the ground truth for diagnosis is the RT-PCT test. Yet, solely using a molecular test to confirm the presence of lung disease is flawed. A patient may have COVID-19 infection with no lung manifestations. The CXR and CT may be clear, but the AI model will be forced to classify the lungs as COVID positive. A consequence is that the model will be trained on faulty ground truth data. A comparable scenario is when in times of endemic infection, a patient with non-COVID-19 pneumonia may acquire nosocomial COVID-19 that does not affect the lungs. The model will again be forced to learn that the lung abnormalities are those of COVID-19 even though the underlying aetiology might be a different infectious agent. False-negative RT-PCR results have also not been reported infrequently. These cases can result in spurious training data where the AI model will be forced to learn that basic COVID-19 features on CXR or CT should be classified as not being COVID-19. In all these examples, the flawed assumption that the RT-PCT test is a surrogate for lung infection is a major constraint to AI model performance.

Despite the good accuracy and performance quoted by the various AI models described in this review, there has been almost no evaluation as to how these models perform on real-world CT and CXR examinations. A recent study[71] highlighted quite dramatically that when some of these algorithms were re-implemented, the regions of the image that were the key determinant of how the image was classified came from features outside of the lungs. On CXR images, non-pathological features in the images, such as laterality markers (identifying the left or right side of the patient on the image), image edges, the diaphragm, and the cardiac silhouette strongly influenced predictions of COVID-19 status. These features instead of emphasising COVID-19-related damage, are more likely to reflect differences across training datasets where the CXR acquisition anterior–posterior (AP) versus posterior–anterior (PA) or patient position differed between centres contributing data to the imaging database.[71] These observations underscore the need for interpretability and transparency of AI model outputs so that human readers can be confident in the logic by which AI models come to their conclusions. These requirements are essential for safety and trust in AI systems.

It is also worth considering that the high accuracy reported for several of the AI methods can result from undesired bias within the datasets used, such as training and testing the models on the same dataset. For example, the AI model proposed by Acar et al.[61] reported the highest accuracy among all the proposed AI methods that analysed CT scans; however, when their model was tested on an external dataset, the accuracy decreased by 8.5%, suggesting that the model was overfitted to the training dataset.

Table 2
Detailed information about the proposed artificial intelligence deep-learning-based models for COVID-19 detection using computed tomography images.

Ref.	Methods	Data groups	Tasks	Accuracy (%)	AUROC (%)
Chen et al.[51]	2D Unet++	51 COVID-19 55 Other disease	Two groups (non-COVID-19, COVID-19)	92.59	-
Gunraj et al.[54]	2D CNN	-	Three groups (normal, pneumonia, COVID-19)	99.10	-
Li et al.[56]	2D Unet + 2D TCNN	1,296 COVID-19 1,735 Pneumonia 1,325 Non-pneumonia	Three groups (non-pneumonia, pneumonia, COVID-19)	-	96.5
Wang et al.[58]	3D Unet + 3D TCNN	1,315 COVID-19 2,406 ILD 936 Normal	Three groups (normal, ILD, COVID-19)	93.30	97.30
Song et al.[60]	Shared 2D CNNs	88 COVID-19 100 Bacterial pneumonia 86 Normal	Three groups (normal, pneumonia, COVID-19)	93.00	95.00
Acar et al.[61]	BDCUnet + GAN + 2D CNN	1,607 COVID-19 1,667 Normal	Two groups (normal, COVID-19)	99.51	-
Hu et al.[63]	2D Unet + 2D WS-CNN	80 COVID-19 78 Pneumonia 72 Normal	Three groups (normal, pneumonia, COVID-19)	87.4	89.60
Wang et al.[65]	2D Unet + 3D WS-CNN	313 COVID-19 299 Non-COVID-19	Two groups (non-COVID-19, COVID-19)	90.10	95.90

AUROC, area under the receiver operating characteristic curve; ILD, interstitial lung disease.

Of the studies reviewed in Tables 1 and 2, performance was better for methods that used CT images; however, models that use CT and 3D processing algorithms utilise more complex algorithmic pipelines and need more complex computational resources. They also have an increased likelihood of overfitting the model to their training data as the number of parameters used in the 3D DL networks is higher. Furthermore, as mentioned previously, creating manual labels of imaging features on CT images is an expensive process requiring the input of experts that are typically in short supply. Models that use CXR images are intrinsically less complex as they are based on 2D DL networks with single end-to-end pipelines. As the size of CXR datasets is typically exponentially larger than CT datasets, the training data have a better chance of being more representative of the various disease manifestations in the lungs; however, as previously described, relying on RT-PCR results as the ground truth for lung damage is a major constraint to the models.

According to the literature, most of the proposed algorithms analysing COVID-19 focus on distinguishing COVID-19 from non-COVID pneumonia; however, clinically, there are other differentials of COVID-19 on CXR, including other viral or bacterial infections, aspiration pneumonia, etc. The main limitation of the proposed methods is that they ignore other types of lung damage/disease, limiting the ability of an AI algorithm to differentiate COVID-19 from other kinds of illness. Moreover, a patient may simultaneously have COVID-19 and a bacterial or viral pneumonia.

There are limitations to this review. The focus was on the image analysis aspects of these studies, and studies where clinical data were also built into AI models were not explored. The content of this review describes a rapidly evolving research field. Papers published until March 2022 were reviewed, but since this time, several new AI models have been released and were beyond the remit of this review.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Joseph Jacob reports financial support was provided by Wellcome Trust.

Acknowledgements

References

1. Kong WH, Li Y, Peng MW, *et al*. SARS-CoV-2 detection in patients with influenza-like illness. *Nat Microb* 2020;**5**(5):675–8.
2. WHO. Coronavirus (COVID-19) dashboard. https://COVID19.who.int/. [Accessed 26 April 2022].
3. Emanuel EJ, Persad G, Upshur R, *et al*. Fair allocation of scarce medical resources in the time of COVID-19. *Mass Medical Soc* 2020;**47**(1):3–6.
4. Wang W, Xu Y, Gao R, *et al*. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA* 2020;**323**(18):1843–4.
5. Yang Y, Yang M, Shen C, *et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections*. MedRxiv; 2020 https://doi.org/10.1101/2020.02.11.20021493.
6. Ai T, Yang Z, Hou H, *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**(2):E32–40.
7. Guan Wj, Ni Zy, Hu Y, *et al*. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;**382**(18):1708–20.
8. Huang C, Wang Y, Li X, *et al*. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;**395**(10223):497–506.

9. Jacobi A, Chung M, Bernheim A, *et al.* Portable chest X-ray in coronavirus disease-19 (COVID-19): a pictorial review. *Clin Imaging* 2020;**64**:35–42.

10. Rubin GD, Ryerson CJ, Haramati LB, *et al.* The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. *Radiology* 2020;**296**(1):172–80.

11. Chuna AG, Pavlova M, Gunraj H, *et al.* COVID-Net MLSys: designing COVID-Net for the clinical workflow. In: *2021 IEEE Canadian conference on electrical and computer engineering (CCECE)*. Piscataway, NJ: IEEE; 2021. p. 1–5.

12. Fang Y, Zhang H, Xie J, *et al.* Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 2020;**296**(2):E115–7.

13. Xie X, Zhong Z, ZhaoW, *et al.* Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. *Radiology* 2020;**296**(2):E41–5.

14. Tian S, Hu W, Niu L, *et al.* Pulmonary pathology of early-phase 2019 novel coronavirus (COVID-19) pneumonia in two patients with lung cancer. *J Thorac Oncol* 2020;**15**(5):700–4.

15. Shatri J, Tafilaj L, Turkaj A, *et al.* The role of chest computed tomography in asymptomatic patients of positive coronavirus disease 2019: a case and literature review. *J Clin Imaging Sci* 2020;**10**:35.

16. 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;**323**(11):1061–9.

17. Chung M, Bernheim A, Mei X, *et al.* CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology* 2020;**295**(1):202–7.

18. Bai HX, Hsieh B, Xiong Z, *et al.* Performance of radiologists in differentiating COVID-19 from non-COVID-19 viral pneumonia at chest CT. *Radiology* 2020;**296**(2):E46–54.

19. Mei X, Lee HC, Diao Ky, *et al.* Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat Med* 2020;**26**(8):1224–8.

20. Gunraj H, Sabri A, Koff D, et al. COVID-Net CT-2: enhanced deep neural networks for detection of COVID-19 from Chest CT images through bigger, more diverse learning. arXiv 2021; 210107433.

21. Khan AI, Shah JL, Bhat MM. CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. *Comput Methods Programs Biomed* 2020;**196**:105581.

22. 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**(2):635–40.

23. Simonyan K, Zisserman A. *Very deep convolutional networks for large-scale image recognition*. 2014:14091556. arXiv.

24. Howard AG, Zhu M, Chen B, *et al. Mobilenets: efficient convolutional neural networks for mobile vision applications*. 2017:170404861. arXiv.

25. Szegedy C, Ioffe S, Vanhoucke V, *et al.* Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-first AAAI conference on artificial intelligence, 4–9 february, san francisco, California*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence; 2017.

26. Chollet F. Xception: deep learning with depthwise separable convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway, NJ: IEEE; 2017. p. 1251–8.

27. Heidari M, Mirniaharikandehei S, Khuzani AZ, *et al.* Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. *Int J Med Inform* 2020;**144**:104284.

28. 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.

29. Rahimzadeh M, Attar A. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Inform Med Unlocked* 2020;**19**:100360.

30. Ouchicha C, Ammor O, Meknassi M. CVDNet: a novel deep learning architecture for detection of coronavirus (COVID-19) from chest X-ray images. *Chaos Solitons Fractals* 2020;**140**:110245.

31. Ucar F, Korkmaz D. COVIDiagnosis-Net: deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Med Hypoth* 2020;**140**:109761.

32. Waheed A, Goyal M, Gupta D, *et al.* COVIDgan: data augmentation using auxiliary classifier GAN for improved COVID-19 detection. *IEEE Access* 2020;**8**:91916–23.

33. Selvaraju RR, Cogswell M, Das A, *et al.* Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision*. Piscataway, NJ: IEEE; 2017. p. 618–26.

34. He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway, NJ: IEEE; 2016. p. 770–8.

35. Iandola FN, Han S, Moskewicz MW, *et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and¡ 0.5 MB model size. arXiv.* 2016:160207360.

36. Wu J, Chen XY, Zhang H, *et al.* Hyperparameter optimization for machine learning models based on Bayesian optimization. *J Elect Sci Technol* 2019;**17**(1):26–40.

37. Ganaie M, Hu M, Malik A, *et al. Ensemble deep learning: a review. arXiv.* 2021:210402395.

38. Rajaraman S, Siegelman J, Alderson PO, *et al.* Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays. *IEEE Access* 2020;**8**:115041–50.

39. Mahmud T, Rahman MA, Fattah SA. CovXNet: a multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput Biol Med* 2020;**122**:103869.

40. Afshar P, Heidarian S, Naderkhani F, *et al.* COVID-caps: a capsule network-based framework for identification of COVID-19 cases from X-ray images. *Pattern Recognit Lett* 2020;**138**:638–43.

41. Toraman S, Alakus TB, Turkoglu I. Convolutional capsnet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solitons Fractals* 2020;**140**:110122.

42. Hinton GE, Sabour S, Frosst N. Matrix capsules with EM routing. In: *International conference on learning representations*; 2018. p. 115. https://openreview.net/pdf?id=HJWLfGWRb.

43. Oh Y, Park S, Ye JC. Deep learning COVID-19 features on CXR using limited training data sets. *IEEE Trans Med Imaging* 2020;**39**(8):2688–700.

44. Apostolopoulos ID, Aznaouridis SI, Tzani MA. Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. *J Med Biol Eng* 2020;**40**(3):462–9.

45. Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. *Adv Neural Inf Process Syst* 2014;**27**:2672–80 https://papers.nips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html.

46. Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier gans. *ICML'17: Proc 34th Int Conf Machine Learn* 2017;**70**(August):2642–51 https://dl.acm.org/doi/10.5555/3305890.3305954.

47. Gietema HA, Zelis N, Nobel JM, *et al.* CT in relation to RT-PCR in diagnosing COVID-19 in The Netherlands: a prospective study. *PloS ONE* 2020;**15**:e0235844.

48. Schalekamp S, Bleeker-Rovers CP, Beenen LF, *et al.* Chest CT in the emergency department for diagnosis of COVID-19 pneumonia: Dutch experience. *Radiology* 2021;**298**(2):e98–106.

49. Ye Z, Zhang Y, Wang Y, *et al.* Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. *Eur Radiol* 2020;**30**(8):4381–9.

50. Fan DP, Zhou T, Ji GP, *et al.* INF-NET: automatic COVID-19 lung infection segmentation from CT images. *IEEE Trans Med Imaging* 2020;**39**(8):2626–37.

51. Chen J, Wu L, Zhang J, *et al.* Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. *Sci Rep* 2020;**10**(1):1–11.

52. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, *et al.* Unet++: a nested u-net architecture for medical image segmentation. In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Cham: Springer; 2018. p. 3–11.

53. Deng J, Dong W, Socher R, *et al.* Imagenet: a large-scale hierarchical image database. In: *2009 IEEE conference on computer vision and pattern recognition*. Piscataway, NJ: IEEE; 2009. p. 248–55.

54. Gunraj H, Wang L, Wong A. COVIDnet-CT: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest CT images. *Front Med* 2020;**7**:608525.
55. Wang L, Lin ZQ, Wong A. COVID-net: a tailored deep convolutional 454 neural network design for detection of COVID-19 cases from chest X-ray images. *Sci Rep* 2020;**10**(1):1—12.
56. Li L, Qin L, Xu Z, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology* 2020;**296**(2):E65—71.
57. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Cham: Springer; 2015. p. 234—41.
58. Wang J, Bao Y, Wen Y, et al. Prior-attention residual learning for more discriminative COVID-19 screening in CT images. *IEEE Trans Med Imaging* 2020;**39**(8):2572—83.
59. Cicek O, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: *International conference on medical image computing and computer-assisted intervention*. Cham: Springer; 2016. p. 424—32.
60. Song Y, Zheng S, Li L, et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Trans Comput Biol Bioinform* 2021;**18**(6):2775—80.
61. Acar E, Şahin E, Yılmaz 1. Improving effectiveness of different deep learning-based models for detecting COVID-19 from computed tomography (CT) images. *Neural Comput Appl* 2021;**33**(24):17589—609.
62. Azad R, Asadi-Aghbolaghi M, Fathy M, et al. Bi-directional ConvLSTM U-Net with densley connected convolutions. In: *Proceedings of the IEEE/CVF international conference on computer vision workshop*. ICCVW); 2019. p. 406—15. https://doi.org/10.1109/ICCVW.2019.00052.
63. Hu S, Gao Y, Niu Z, et al. Weakly supervised deep learning for COVID-19 infection detection and classification from CT images wang 2020 weakly. *IEEE Access* 2020;**8**:118869—83.
64. Yang G, Chen J, Gao Z, et al. Simultaneous left atrium anatomy and scar segmentations via deep learning in multiview information with attention. *Future Gen Comp Syst* 2020;**107**:215—28.
65. Wang X, Deng X, Fu Q, et al. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. *IEEE Trans Med Imaging* 2020;**39**(8):2615—25.
66. Wang B, Jin S, Yan Q, et al. AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical AI system. *Appl Soft Comput* 2021;**98**:106897.
67. Wang G, Liu X, Li C, et al. A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images. *IEEE Trans Med Imaging* 2020;**39**(8):2653—63.
68. Han Z, Wei B, Hong Y, et al. Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning. *IEEE Trans Med Imaging* 2020;**39**(8):2584—94.
69. Ardakani AA, Kanafi AR, Acharya UR, et al. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks. *Comput Biol Med* 2020;**121**:103795.
70. Wang Z, Liu Q, Dou Q. Contrastive cross-site learning with redesigned net for COVID-19 CT classification. *IEEE J Biomed Health Inform* 2020;**24**(10):2806—13.
71. DeGrave AJ, Janizek JD, Lee SI. AI for radiographic COVID-19 detection selects shortcuts over signal. *Nat Mach Intell* 2021;**3**(7):610—9.
72. Jacob J, Alexander D, Baillie JK, et al. Using imaging to combat a pandemic: rationale for developing the UK National COVID-19 Chest Imaging Database. *Eur Resp J* 2020;**56**:2001809.
73. Aslani S, Lilaonitkul W, Gnananthan V, et al. *Optimising chest X-rays for image analysis by identifying and removing confounding factors. arXiv.* 2022:220810320.