Computers in Biology and Medicine 121 (2020) 103795

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks

Ali Abbasian Ardakani [a], Alireza Rajabzadeh Kanafi [b], U. Rajendra Acharya [c,d,e,f], Nazanin Khadem [g], Afshin Mohammadi [g,*]

[a] Medical Physics Department, School of Medicine, Iran University of Medical Sciences (IUMS), Tehran, Iran
[b] Department of Radiology, Razi Hospital, Guilan University of Medical Sciences, Rasht, Iran
[c] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[d] Department of Biomedical Engineering, School of Science and Technology, Singapore University of Social Sciences, Singapore
[e] School of Medicine, Faculty of Health and Medical Sciences, Taylor's University, 47500, Subang Jaya, Malaysia
[f] Department of Biomedical Informatics and Medical Engineering, Asia University, Taiwan
[g] Department of Radiology, Faculty of Medicine, Urmia University of Medical Science, Urmia, Iran

ARTICLE INFO

Keywords:
Computed tomography
Coronavirus infections
COVID-19
Deep learning
Lung diseases
Pneumonia
Machine learning

ABSTRACT

Fast diagnostic methods can control and prevent the spread of pandemic diseases like coronavirus disease 2019 (COVID-19) and assist physicians to better manage patients in high workload conditions. Although a laboratory test is the current routine diagnostic tool, it is time-consuming, imposing a high cost and requiring a well-equipped laboratory for analysis. Computed tomography (CT) has thus far become a fast method to diagnose patients with COVID-19. However, the performance of radiologists in diagnosis of COVID-19 was moderate. Accordingly, additional investigations are needed to improve the performance in diagnosing COVID-19. In this study is suggested a rapid and valid method for COVID-19 diagnosis using an artificial intelligence technique based. 1020 CT slices from 108 patients with laboratory proven COVID-19 (the COVID-19 group) and 86 patients with other atypical and viral pneumonia diseases (the non-COVID-19 group) were included. Ten well-known convolutional neural networks were used to distinguish infection of COVID-19 from non-COVID-19 groups: AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, and Xception. Among all networks, the best performance was achieved by ResNet-101 and Xception. ResNet-101 could distinguish COVID-19 from non-COVID-19 cases with an AUC of 0.994 (sensitivity, 100%; specificity, 99.02%; accuracy, 99.51%). Xception achieved an AUC of 0.994 (sensitivity, 98.04%; specificity, 100%; accuracy, 99.02%). However, the performance of the radiologist was moderate with an AUC of 0.873 (sensitivity, 89.21%; specificity, 83.33%; accuracy, 86.27%). ResNet-101 can be considered as a high sensitivity model to characterize and diagnose COVID-19 infections, and can be used as an adjuvant tool in radiology departments.

1. Introduction

In December 2019, a new coronavirus disease, termed coronavirus disease 2019 (COVID-19), was reported in Wuhan, China [1]. The most common symptoms of COVID-19 are fever, dyspnea, cough, myalgia, and headache [2]. At the time of the study, no specific drug or treatment was available.

According to the WHO, all COVID-19 diagnoses must be confirmed by molecular assays, such as the reverse-transcription polymerase chain reaction (RT-PCR) [3]. Besides RT-PCR and, medical imaging, computed tomography (CT) has become a vital method to assist in the diagnosis and management of patients with COVID-19. The prominent chest CT findings of COVID-19 are ground-glass opacity (GGO), multifocal patchy consolidation, and a 'crazy-paving' pattern with a peripheral distribution [4–6].

The critical role of radiologists is to provide early diagnosis and

* Corresponding author.
E-mail addresses: A.ardekani@live.com (A.A. Ardakani), Alireza_r245@yahoo.com (A.R. Kanafi), aru@np.edu.sg (U.R. Acharya), Nazanin.khadem74@gmail.com (N. Khadem), Afshin.mohdi@gmail.com (A. Mohammadi).

https://doi.org/10.1016/j.compbiomed.2020.103795
Received 26 March 2020; Received in revised form 10 April 2020; Accepted 27 April 2020
Available online 30 April 2020

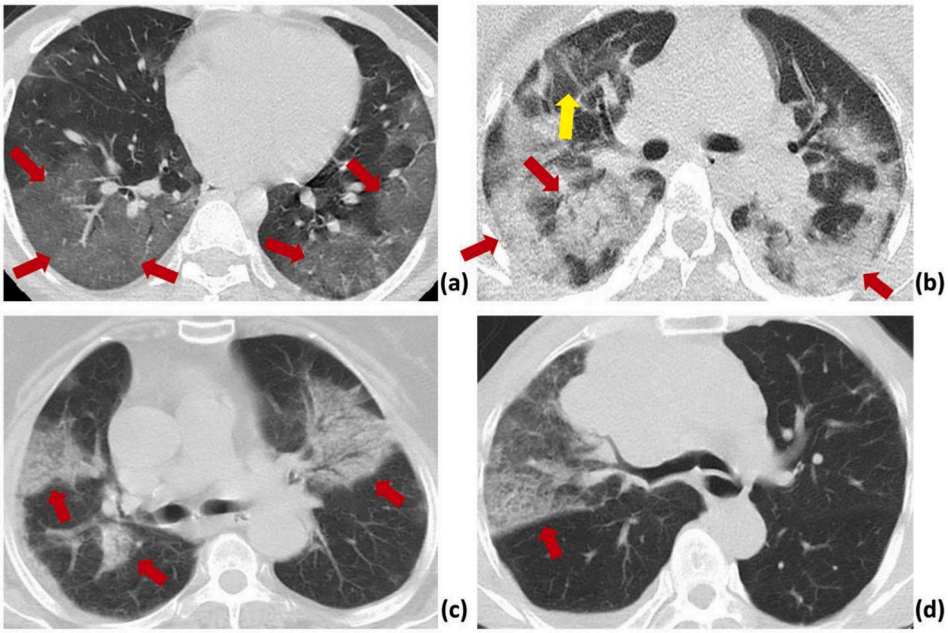

Fig. 1. CT sample images of patients with pneumonia.
(a): A 28-year-old male with confirmed COVID-19 pneumonia. The red arrows indicate ground-glass opacity in the right and left lower lobes.
(b): A 33-year-old female patient with confirmed COVID-19 pneumonia. The red and yellow arrows indicate typical mixed ground-glass opacity-consolidation patterns in both lobes, and a ground-glass opacity pattern in the peripheral right upper lobe, respectively.
(c): An 81-year-old female patient with H1N1 influenza. The red arrows indicate infection with mixed ground-glass opacity and consolidation pattern in both lobes
(d): A 72-year-old male patient with atypical pneumonia. The red arrows indicate ground-glass opacity in the right middle lobe.

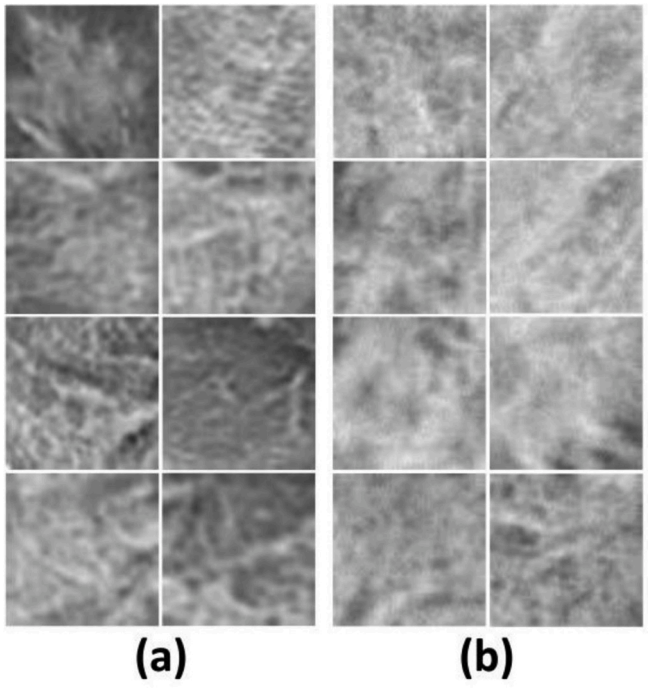

Fig. 2. CT sample patch images of patients with COVID-19 **(a)** and other atypical and viral pneumonia diseases **(b)**.

treatment to distinguish the COVID-19 infection from other conditions, which may have similar findings at CT [4,5]. For example, GGO is a common finding among other atypical and viral pneumonia diseases such as influenza, severe acute respiratory syndrome (SARS), and Middle East respiratory syndrome (MERS) [5,7]. However, the specificity and sensitivity of radiologists in diagnosis of COVID-19 were high and moderate, respectively [8]. Hence, further investigations are needed to help and improve the radiologist's performance. In addition, congestion of patients, as well as a high workload of radiologists, which can increase fatigue, affect diagnostic performance [9,10].

Although for the first time, Lodwick used the term computer-aided diagnosis (CAD) in the literature in 1966 [11], serious attempts on CAD began in the 1980s [12]. Recently, CAD systems have been an inseparable part of a routine clinical practice that assist radiologists in the process of diagnosis. CAD systems have advantages over radiologists. They are reproducible and detect the subtle changes that cannot be observed by the visual inspection [13]. The CAD systems have been widely used to assist radiologists in detecting lung abnormalities. We have included a brief literature review and report the highlights of computerized methods in lung disease management. The system proposed by Than et al. using Riesz and Gabor transforms obtained an accuracy of 98.73% in detecting lung disease. The accuracy increased to 99.53% with the combination of Riesz, Gabor, fractal dimension, grey level co-occurrence matrix, and grey level run-length matrix based features [14]. Gu et al. indicated that a deep convolutional neural network (CNN) could detect lung nodules with a competition performance metric of 0.7967. They proposed a density-based spatial clustering of applications with noise (DBSCAN) algorithm to improve the detection sensitivity of the network [15]. Also, the CAD system based on a 3D skeletonization feature proposed by Zhang et al. can assist radiologists to differentiate lung nodules from interferential vessels [16]. Transforming an image to the time–frequency domain by the optimal fractional S-transform (OFrST) method can provide meaningful information and features, which can be used to diagnose diseases. This method was applied on lung CT images by Sun et al. to detect and differentiate nodules from the vessels. They employed the Teager–Kaiser energy (TKE) in the time–frequency domain to obtain the energy distribution and characterize lung nodules with 97.87% sensitivity [17]. Sun et al. concluded that the CNN represented higher performance than deep belief network (DBN) and stacked denoising autoencoder (SDAE) in diagnosing malignant lung nodules with an area under the curve (AUC) of 0.899 [18]. Furthermore, CAD systems are a useful tool to monitor lung function after transplantation. In this regard, Barbosa et al. compared the quantitative CT (qCT), pulmonary function testing (PFT), and semi-quantitative image scores (SQS) metrics and found that pulmonary function testing and qCT metrics demonstrated the highest accuracy for monitoring bi- and unilateral lung transplantation with an AUC of 0.771 and 0.817, respectively [19]. The potential of CAD systems can go even further and help radiologists to detect lung abnormalities by computerized analysis of pulmonary sounds [20] and flow and gas concentration of breath [21].

In this study, we propose a CAD system based on deep learning to classify COVID-19 infection versus other atypical and viral pneumonia

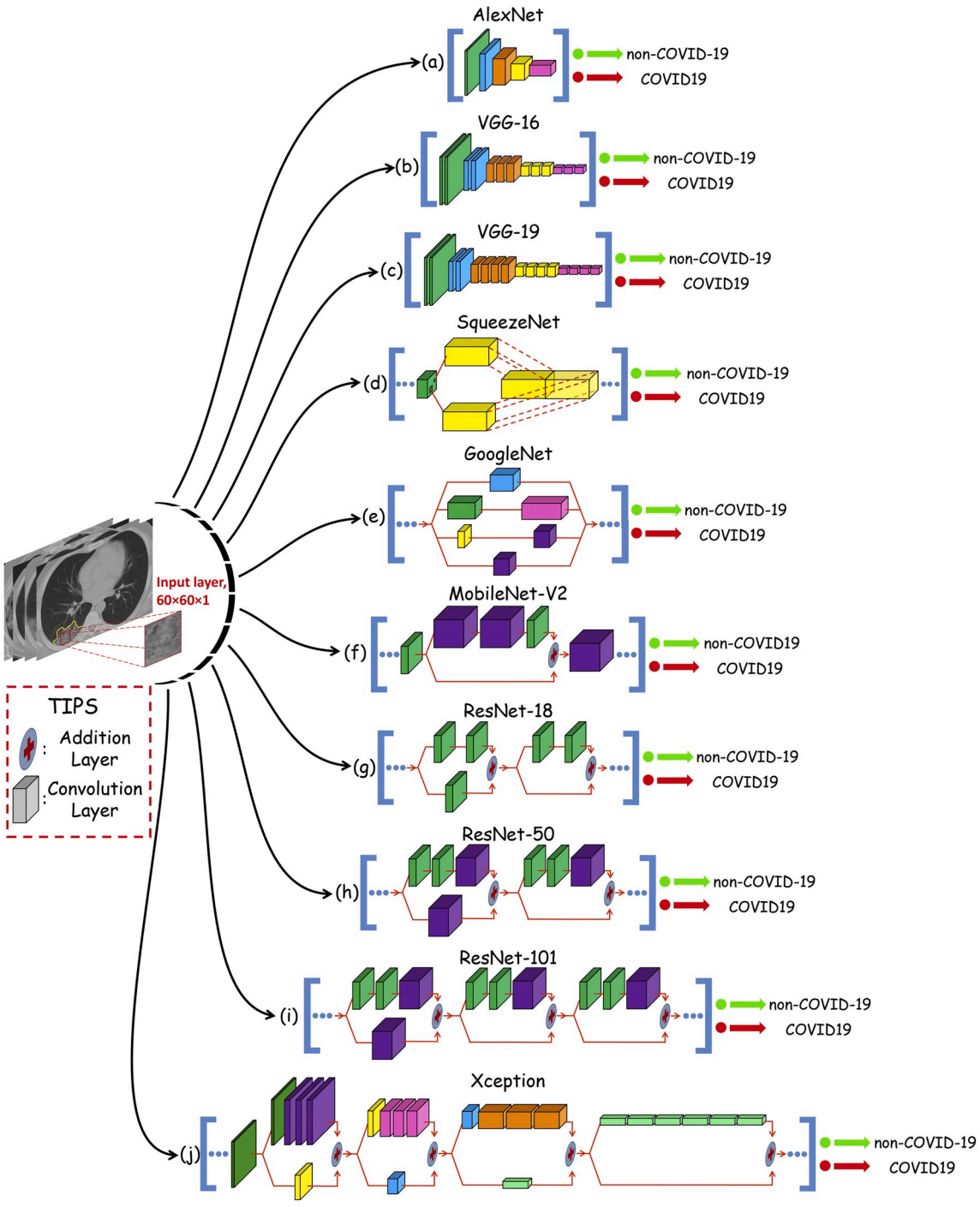

Fig. 3. An overview of the ten pre-trained networks architecture used in this study. In each network, the convolution layers with the same color are in the same size. The straight arrows represent the direction of flow and computation. All convolution layers of the AlexNet, VGG-16 and VGG-19 networks are depicted. However, for the other networks, convolution layers of the most repetitive (core) part are depicted.

diseases. We hypothesized that the deep learning method would help radiologists in diagnosing infection related to COVID-19. We used ten well-known pre-trained convolutional neural networks (CNN) to diagnose infections related to the COVID-19.

2. Patients and methods

2.1. Patients

During the COVID-19 pandemic, the entirety of patients representing flu-like symptoms with an initial diagnosis of the novel coronavirus, regardless of demographic values such as age and sex, were included in the study. Prior to enrolment, chest high-Resolution CT (HRCT) examination was obtained from all patients during the acute phase of the disease.

The inclusion criterion was the confirmation of diagnosis of COVID-19 through RT-PCR performed on nasopharyngeal swabs samples. Patients with concurrent pulmonary infections, as confirmed by laboratory tests and negative RT-PCR, were excluded. In addition, patients with CT imaging suggestive of chronic lung diseases and subsequent pulmonary involvement were excluded. Imaging studies were performed between 3 and 6 days from the onset of flu-like symptoms. We retrospectively

Table 1
Characteristics of ten convolutional neural networks used in this study.

Network	Depth	Parameters (10^6)	Input Layer Size	Output Layer Size
AlexNet	8	61.0	60-by-60	2-by-1
VGG-16	16	138	60-by-60	2-by-1
VGG-19	19	144	60-by-60	2-by-1
SqueezeNet	18	1.24	60-by-60	2-by-1
GoogleNet	22	7	60-by-60	2-by-1
MobileNet-V2	53	3.5	60-by-60	2-by-1
ResNet-18	18	11.7	60-by-60	2-by-1
ResNet-50	50	25.6	60-by-60	2-by-1
ResNet-101	101	44.6	60-by-60	2-by-1
Xception	71	22.9	60-by-60	2-by-1

Table 2
Demographic characteristics data of patients with COVID-19 and other atypical and viral pneumonia (Non-COVID-19 group) diseases.

Characteristics	COVID-19	Non-COVID-19	P-value*
	Mean ± SD		
Age (year)	50.22 ± 10.85	61.45 ± 15.04	<0.001
Gender	**Number (Percentage)**		**P-value****
Female	48 (44.44)	35 (40.70)	0.600
Male	60 (55.56)	51 (59.30)	

p-values were obtained using *t*-test (*) and chi-square test (**).

analyzed the HRCT images of patients from September 2019 to December 2019 with other causes of atypical and viral pneumonia as adenoviral or H1N1 flu from the PACS of our university hospital.

2.2. Image acquisition

HRCT images of all subjects were acquired from a 16-MDCT scanner (Alexion, Toshiba Medical System, Japan) using the high-resolution protocol as follows: patient in the supine position with the arms above the head; 1- to 2-mm slice thickness in increments of up to 10 mm from the lung apices through the hemidiaphragm, at deep inspiration; tube voltage, 120 kVp; tube current–time, 50–100 mAs; pitch, 0.8–1.5, and matrix size, 512 × 512 pixel. For further analysis, parenchymal window settings were set to a window level and a window width of −600 HU and 1650 HU, respectively. All CT examinations were performed without the use of contrast agent, and images were reconstructed in the transverse plane using a high spatial resolution algorithm.

2.3. CT image pre-processing and analysis

All CT images of patients were converted to the grey-scale, and then reviewed by a radiologist with more than 15 years of experience in thoracic imaging. Fig. 1 indicates the CT sample images of patients with COVID-19 and other atypical and viral pneumonia diseases. The radiologist defined the rectangular region of infections on CT images. Then, the regions were cropped and resized to 60 × 60 pixels (Fig. 2). One region of infection was defined per slice for further analysis. The largest region was selected if there are more than one regions of infections.

To evaluate the performance of a radiologist in diagnosing COVID-19 infection versus other atypical and viral pneumonia diseases (non-COVID-19 group), another radiologist who did not define the region of infection and was blinded to the results of the laboratory test reviewed both patches and CT slices. The performances were reported based on visual inspection of patches and slices separately.

2.4. Deep learning study

In this study, ten well-known pre-trained CNN were used to distinguish infection of COVID-19 from non-COVID-19 group: 1- AlexNet, 2- VGG-16, 3- VGG-19, 4- SqueezeNet, 5- GoogleNet, 6- MobileNet-V2, 7- ResNet-18, 8- ResNet-50, 9- ResNet-101, and 10- Xception (Fig. 3). AlexNet is a type of feedforward CNN with 8-layer deep. It contains five convolution layers (conv1 to conv5) and three fully-connected layers (fc6 to fc8) [22]. It was proposed by Alex Krizhevsky and trained on 1 million images to classify images into 1000 different classes (Fig. 3a). VGG-16 is a combination of five convolutional blocks (13 convolutional layers) and tree fully-connected layers (fc6 to fc8) [23]. It was trained on a million images of 1000 classes (Fig. 3b). VGG-19 uses 19 layers, including five convolutional blocks (16 convolutional layers) and tree fully-connected layers (fc6-8) [23]. Compared to VGG-16, the VGG-19 is a deeper CNN architecture with more layers (Fig. 3c). SqueezeNet is a compact CNN with up to 18 learnable layers deep to classify images into 1000 different classes [24]. The network starts with a stand-alone convolution layer followed by eight fire modules and ends with a final convolution layer. A scheme of one sample module is illustrated in Fig. 3d. The GoogleNet is a deep model, which is trained on either the ImageNet or Places365 datasets. In this study, we used the GoogleNet trained on the ImageNet dataset [25]. The network is 22-layers deep, starts with three convolution layers, followed by 9 inception blocks, and ends with a fully-connected layer. The inception block as a core part of the GoogleNet is shown in Fig. 3e. MobileNet-V2 is a light-weight CNN with 53 layers deep (52 convolution and one fully connected layers) [26]. The primary part architecture of the network is based on inverted residual and linear bottlenecks. The network starts with three convolution layers, followed by 16 inverted residual and linear bottleneck blocks and ends with one convolution layer and one fully-connected layer. The scheme of the inverted residual block, which is the main part of the network and is illustrated in Fig. 3f. ResNet is a type of deep network based on residual learning. This kind of learning can facilitate the training of networks by considering the layer inputs as a reference [27]. All types of ResNet, ResNet-18, ResNet-50, and ResNet-101, are versions of ResNet that have their specific residual block. ResNet-18 with 22-layers are deep, starts with a convolution layer followed by 8 residual blocks and ends with a fully-connected layer. The main components of the residual blocks of ResNet-18 are shown in Fig. 3g. ResNet-50 is similar to ResNet-18 but has different residual blocks scheme and different number (16) of residual blocks that contain in the network (Fig. 3h). The ResNet-50 contains 50 layers and same is true for ResNet-101. Therefore, it is 101 layers deep with 33 residual blocks (Fig. 3i). The last network, Xception, is CNN based on depthwise separable convolution layers [28]. It starts with two convolution layers, followed by depthwise separable convolution layers, four convolution layers, and a fully-connected layer. The main component of depthwise separable convolution layers are shown in Fig. 3j.

We used a transfer-learning to optimize the CNNs to the datasets. In this regard, the input layer of the CNNs were replaced with a new one, which is consistent with the size of the infection patches (i.e., 60 × 60 × 1). In addition, the dimension of the last fully connected layer of all networks was set to the number of the classes, i.e., two (Table 1). All networks were trained as follows: optimizer, SGDM; initial learning rate, 0.01; validation frequency, 5. The dataset was shuffled at every epoch, and the training process stopped if the training process did not change significantly. For all networks, the dataset was divided into 80% and 20% for training and validating sets, respectively. The same training and validation datasets were selected for all networks to facilitate the performance comparison of networks.

2.5. Statistical analysis

The Kolmogorov-Smirnov test was used to check the normality of all quantitative data. In addition, the age and sex distributions among COVID-19 and non-COVID-19 groups was evaluate by the two-tailed independent samples *t*-test and chi-square test, respectively. A P-value smaller than 0.05 was considered as statistically significant.

Table 3
Diagnostic Performance of ten networks and the radiologist in diagnosis of COVID-19 pneumonia disease in training and validating datasets.

Network	Group	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	AUC	Confusion Matrix (True Class: C / NC; Predict Class)
AlexNet	Training	93.38	71.81	82.60	76.81	91.56	0.913	381 / 115 (C); 27 / 293 (NC)
	Validation	89.21	68.63	78.92	73.98	86.42	0.894	91 / 32 (C); 11 / 70 (NC)
VGG-16	Training	80.63	86.76	83.70	85.90	81.76	0.928	329 / 54 (C); 79 / 354 (NC)
	Validation	80.39	86.27	83.33	85.42	81.48	0.926	82 / 14 (C); 20 / 88 (NC)
VGG-19	Training	94.85	79.41	87.13	82.16	93.91	0.955	387 / 84 (C); 21 / 324 (NC)
	Validation	92.16	78.43	85.29	81.03	90.91	0.943	94 / 22 (C); 8 / 80 (NC)
SqueezeNet	Training	80.15	85.29	82.72	84.50	81.12	0.905	327 / 60 (C); 81 / 348 (NC)
	Validation	78.43	87.25	82.84	86.02	80.18	0.899	80 / 13 (C); 22 / 89 (NC)
GoogleNet	Training	84.31	85.29	84.80	85.15	84.47	0.936	344 / 60 (C); 64 / 348 (NC)
	Validation	81.37	90.20	85.29	88.30	82.73	0.927	83 / 11 (C); 19 / 91 (NC)
MobileNet-V2	Training	97.30	89.95	93.63	90.64	97.09	0.990	397 / 41 (C); 11 / 367 (NC)
	Validation	97.06	87.25	92.16	88.39	96.74	0.982	99 / 13 (C); 3 / 89 (NC)
ResNet-18	Training	93.87	89.95	91.91	90.33	93.62	0.975	383 / 41 (C); 25 / 367 (NC)
	Validation	95.10	88.23	91.67	88.99	94.73	0.970	97 / 12 (C); 5 / 90 (NC)
ResNet-50	Training	96.56	100	98.28	100	96.68	0.993	394 / 0 (C); 14 / 408 (NC)
	Validation	90.20	100	94.12	100	91.07	0.990	92 / 0 (C); 10 / 102 (NC)
ResNet-101	Training	100	99.26	99.63	99.27	100	0.999	408 / 3 (C); 0 / 405 (NC)
	Validation	100	99.02	99.51	99.03	100	0.994	102 / 1 (C); 0 / 101 (NC)
Xception	Training	98.77	100	99.38	100	98.79	0.997	403 / 0 (C); 5 / 408 (NC)
	Validation	98.04	100	99.02	100	98.08	0.994	100 / 0 (C); 2 / 102 (NC)
Radiologist (Patch Based)	Validation	61.76	58.82	60.29	60.00	60.61	0.603	63 / 42 (C); 39 / 60 (NC)
Radiologist (Slice Based)	Validation	89.21	83.33	86.27	84.25	88.54	0.873	91 / 17 (C); 11 / 85 (NC)

C, COVID-19 and NC, non-COVID-19 cases.

2.5.1. Performance evaluation of networks

In order to compare performance of the CNNs and the radiologist, five performance indices were calculated as follows:

$$Accuracy = \frac{N_{TN} + N_{TP}}{N_{TN} + N_{FN} + N_{TP} + N_{FP}} \quad (1)$$

$$sensitivity = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (2)$$

$$Specificity = \frac{N_{TN}}{N_{TN} + N_{FP}} \quad (3)$$

$$Positive\ \ Predictive\ \ Value(PPV) = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (4)$$

$$Negative\ \ Predictive\ \ Value(NPV) = \frac{N_{TN}}{N_{TN} + N_{FN}} \quad (5)$$

In this study, the positive and negative cases were assigned to COVID-19 and non-COVID-19 infections, respectively. Hence, N_{TP} and N_{TN} represent the number of correctly diagnosed COVID-19 and non-COVID-19 infections, respectively. N_{FP} and N_{FN} indicate the number of incorrectly diagnosed COVID-19 and non-COVID-19 infections, respectively. In addition, the area under the ROC curve (AUC) was calculated to indicate the overall performance of the CNNs and the radiologist [29]. Statistical analysis was performed the SPSS software (version 24, IBM SPSS Statistics 19, IBM Corporation, USA).

3. Results

3.1. Demographic characteristics

In this study, 108 patients with laboratory-proven COVID-19 pneumonia (COVID-19 group) including 60 males and 48 females with mean age of 50.22 ± 10.85 (mean age ± SD) were studied. Eighty-six patients with atypical pneumonia diseases including 51 male and 35 female with mean age of 61.45 ± 15.04 were enrolled as controls (nonCOVID-19 group). No significant differences was seen between COVID-19 and non-COVID-19 groups in the term of sex distribution ($P > 0.05$). However, the mean age of COVID-19 group was significantly lower than the non-COVID-19 group ($P < 0.001$, Table 2).

1020 image patches including 510 COVID-19 and 510 non-COVID-19 were extracted from CT slices by a specialist radiologist. The dataset was divided into 816 (with 50%–50% distribution), and 102 (with 50%–50% distribution) for training and validating process.

3.2. Diagnosis of COVID-19 by deep learning and radiologist

Table 3 indicates the diagnostic performances of ten networks and

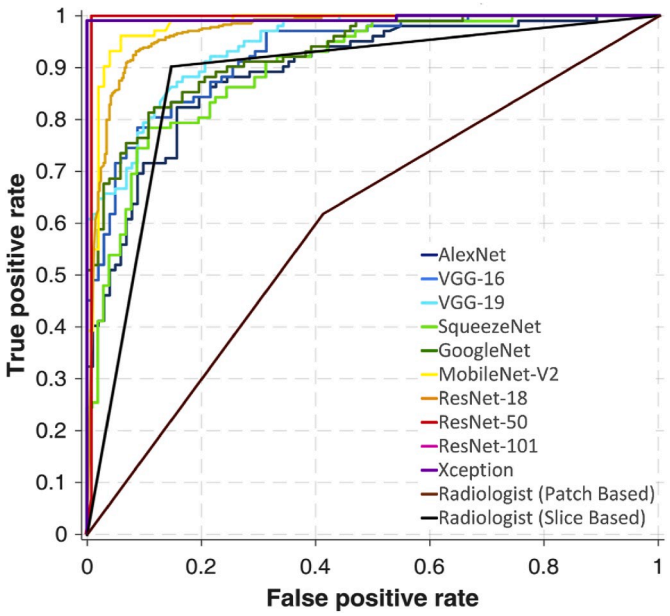

Fig. 4a. (a)ROC curves of individual ten networks and the radiologist on validating dataset.

the radiologist in detail with confusion matrices. Networks could distinguish COVID-19 from non-COVID-19 groups with AUC in the range of 0.894–0.994. The best performance was achieved by ResNet-101 and Xception networks. In this regard, ResNet-101 network achieved AUC of 0.999 (sensitivity, 100%; specificity, 99.26%; accuracy, 99.63%; PPV, 99.27%; and NPV, 100%) and 0.994 (sensitivity, 100%; specificity, 99.02%; accuracy, 99.51%; PPV, 99.03%; and NPV, 100%) in training and validating datasets, respectively. Likewise, Xception networks could diagnose COVID-19 infection with AUC of 0.997 (sensitivity, 98.77%; specificity, 100%; accuracy, 99.38%; PPV, 100%; and NPV, 98.79%) and 0.994 (sensitivity, 98.04%; specificity, 100%; accuracy, 99.02%; PPV, 100%; and NPV, 98.08%) in training and validating datasets, respectively.

The radiologist classified COVID-19 and non-COVID-19 groups based on CT slices with AUC of 0.873 (sensitivity, 89.21%; specificity, 83.33%; accuracy, 86.27%; PPV, 84.25%; and NPV, 88.54%) in validating dataset. Moreover based on patches analysis, the radiologist reached a lower performance with AUC of 0.603 (sensitivity, 61.76%; specificity, 58.82%; accuracy, 60.29%; PPV, 60.00%; and NPV, 60.61%).

Fig. 4a and b represent ROC curves and radar plot of diagnostic indices of the radiologist and all networks in validating phase, respectively. The training and validating process of the CNNs are indicated in Fig. 5.

4. Discussion

In this study, ten well-known CNNs were used to provide a comprehensive view of the role of artificial intelligence in diagnosing COVID-19 infection. The results showed that deep learning could distinguish COVID-19 from other atypical and viral pneumonia diseases with high accuracy. The best results were found for the ResNet-101 and Xception networks. Although the Xception network implied the highest performances in classifying COVID-19 and non-COVID-19 infections, it did not have the best sensitivity. In contrast, ResNet-101 could diagnose COVID-19 infection with the highest sensitivity and implied lower specificity compare to the Xception network. It is desired to introduce a method with the highest sensitivity to diagnose all patients with COVID-19. In this regard, the ResNet-101 has an advantage over other networks, especially Xception, due to its highest sensitivity and AUC. ResNet-101 is trained based on residual learning. This kind of learning can facilitate the training of networks by considering the layer inputs as a reference. In addition, residual networks are optimized easier, and the accuracy can be improved as the depth increases [27]. This specific residual learning can lead to better training and provide robust model. Hence, the best performance can be obtained using ResNet-101. The radiologist classified two groups based on CT slices and patches. The results show that the

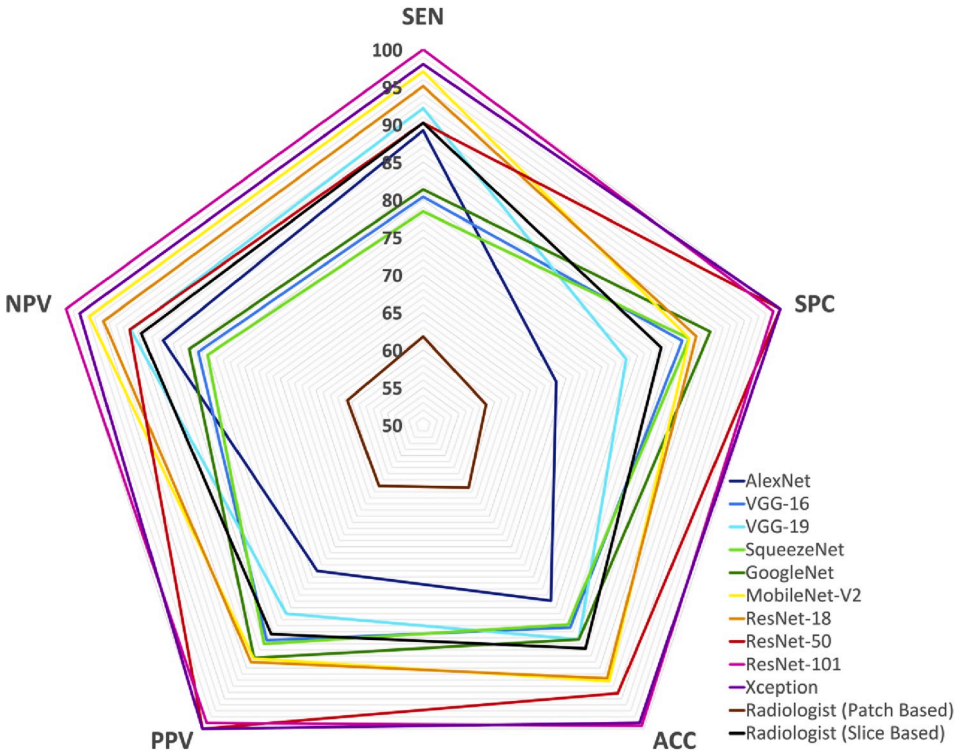

Fig. 4b. (b)Radar plot of individual ten networks and the radiologist on validating dataset.

radiologist failed to diagnose COVID-19 based on patches. The main reason is that a radiologist needs the whole lung lobe to evaluate the intensity of lesions and compare them with the mean density of adjacent organs and structures. However, since computerized image analysis can recognize the pattern quantitatively with more precision, the CNNs could classify two groups with improved performance. The results of the present study indicate that the discriminative power of the proposed CNN is improved compared with the radiologist (AUC of ResNet-101 vs. radiologist: 0.994 vs. 0.873). Also, the deep learning technique can accurately aid radiologists to diagnose infections related to COVID-19.

Only one study was found in the literature which employed deep learning for COVID-19 diagnosis [30]. The investigators proposed COVNet, which is based on ResNet-50, to diagnose COVID-19, with a resulting 90% sensitivity and 96% specificity. They examined and modified only one CNN, and in contrast, herein we compared the ten most widely used networks to provide comprehensive results, and help the other researchers to prevent redundant analysis and select the best network for their works. Our results indicate that ResNet-101 has an advantage over ResNet-50, the network they used, due to its higher AUC. Unlike us, they fed the whole lung region to their network to differentiate CAP form COVID-19 patients. In their approach, some redundant data such as interferential vessels can be misdiagnosed as pathology [16, 31]. Hence, we used ROI based analysis to increase the performance of the network. In addition, they differentiated COVID-19 from community-acquired pneumonia (CAP) diseases. Differentiating CAP from other atypical and viral pneumonia diseases like COVID-19 is not

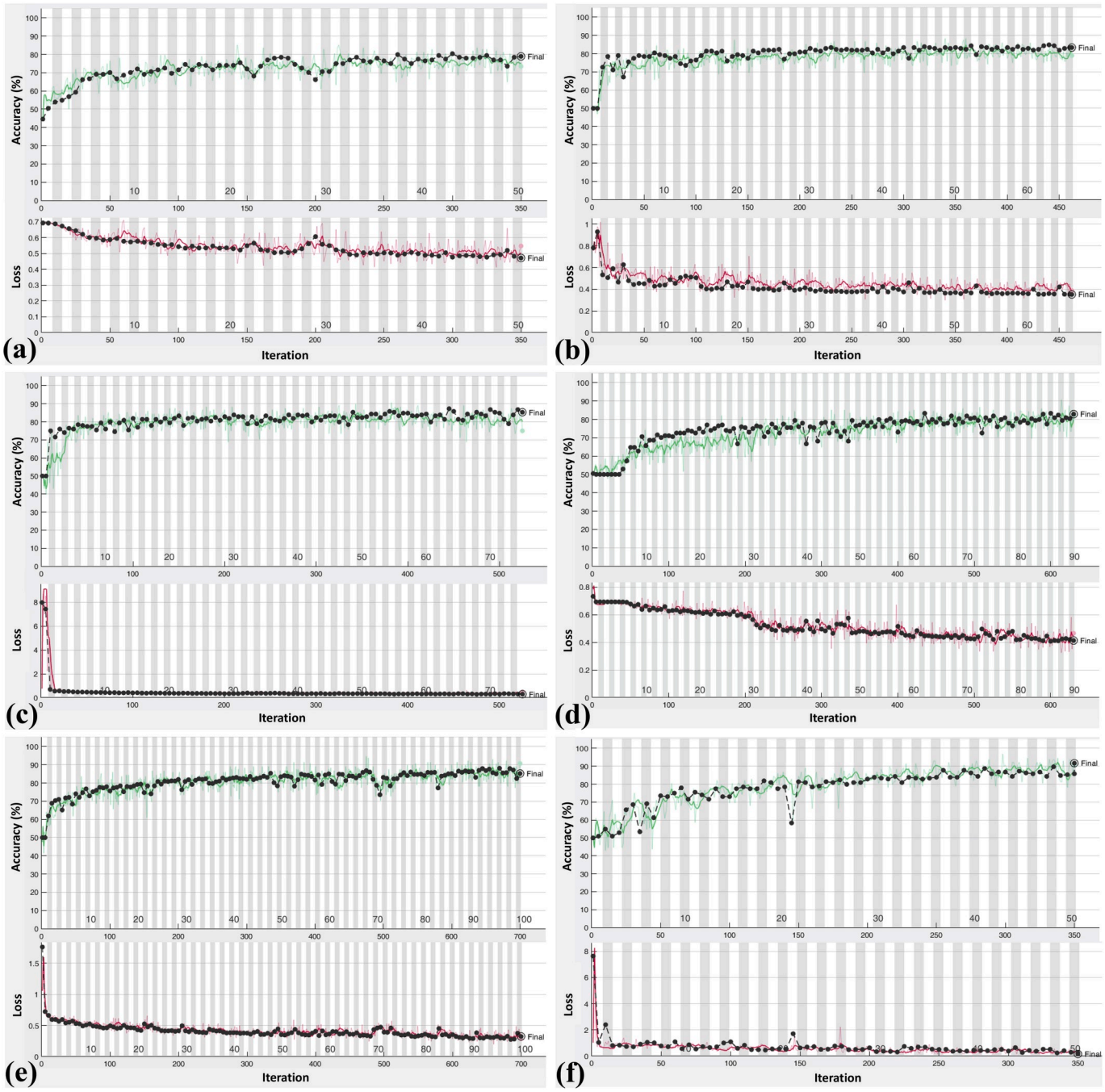

Fig. 5. Accuracy (green line) and loss (red line) plot of ten convolutional neural network for training and validating datasets: **(a)**, AlexNet; **(b)**, VGG-16; **(c)**, VGG-19; **(d)**, SqueezeNet; **(e)**, GoogleNet; **(f)**, MobileNet-V2; **(g)**, ResNet-18; **(h)**, ResNet-50; **(i)**, ResNet-101; **(j)**, Xception.

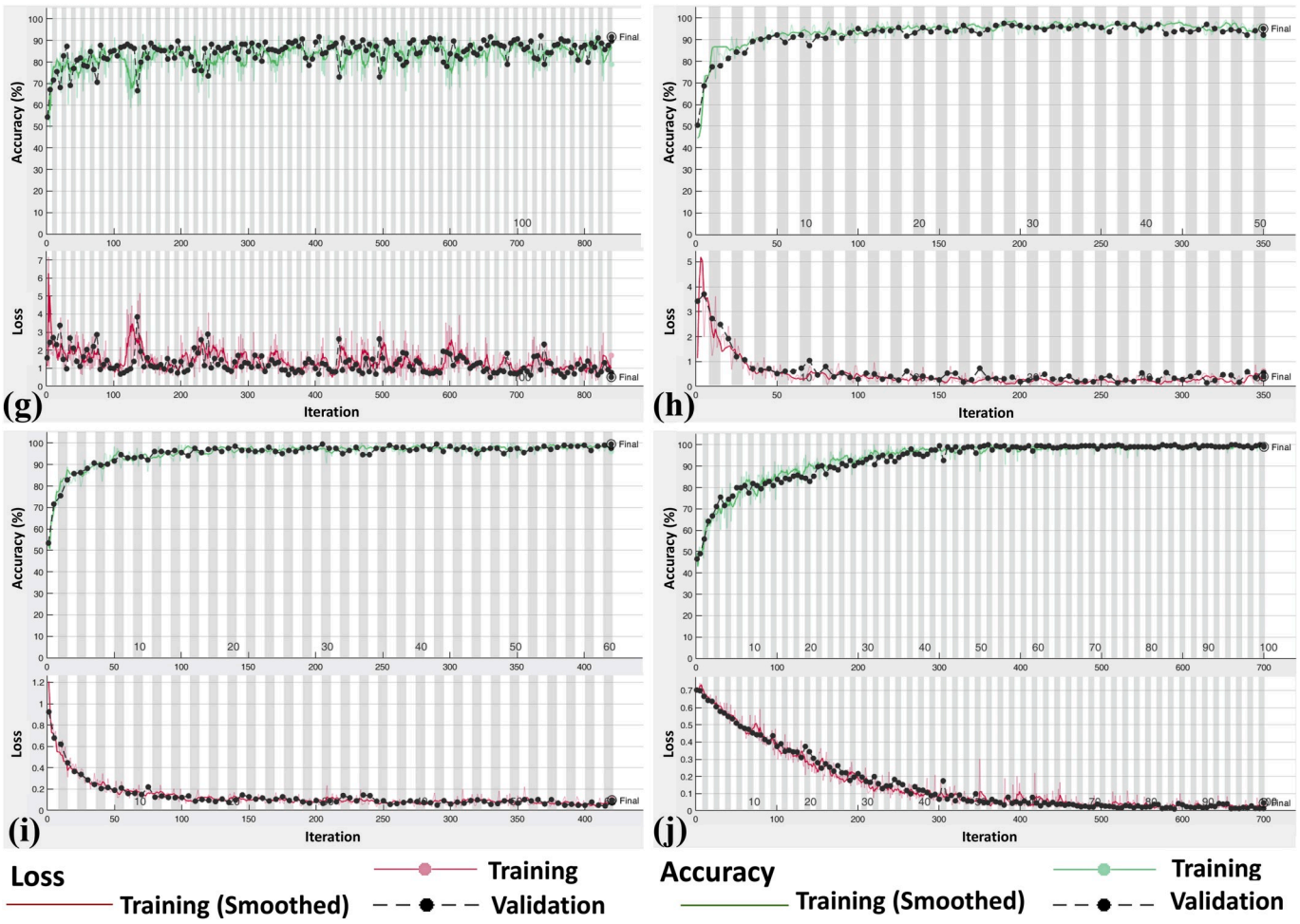

Fig. 5. (*continued*).

critical in clinics and CT images. The most critical and important issue for radiologists is differentiating COVID-19 and other atypical and viral pneumonia diseases, which are the same in CT imagery and have similar symptoms [32]. Contrary to their study, we include two groups with similar indications, COVID-19, and other atypical and viral pneumonia diseases, to obviate this problem and help radiologists in diagnostic tasks. Few patients with COVID-19 may have initial negative RT-PCR results [33]. Hence, to eliminate any confounding factors, the non-COVID-19 patients who referred to the department before the COVID-19 outbreak were included to ensure that no patients with COVID-19 exist in the non-COVID-19 group. In contrast, they included both patients with COVID-19 and CAP during the same period, which could introduce bias and affect reliability.

Moreover, a clinical method with highest sensitivity and earliest diagnostic ability is not yet confirmed. Currently, two primary modalities, RT-PCR and CT, are used to diagnose COVID-19. In this regard, many studies have utilized CT and RT-PCR to evaluate their diagnostic performances; however, the findings must yet be validated. Bernheim et al. [34] and Yang et al. [35] reported that RT-PCR diagnosed COVID-19 earlier than CT, while in some cases, the CT findings could be normal even ten days after symptoms onset. In contrast, Xie et al. [36], Ai et al. [37] and Huang et al. [38] found some cases with positive CT findings and an initial negative RT-PCR screening result, which later became positive upon repeating swab tests. On the other hand, in the pandemic phase of COVID-19, requests for CT Exam has substantially increased. The increased workload can affect the diagnostic performance of radiologists [9,10]. The results of our study indicate that deep learning can characterize the pattern of abnormality effectively, and may help physicians in managing workflow and diagnosing infection related to the COVID-19.

The limitations of this study are given below.

First, performance of the proposed CAD system was not compared with radiologists. So, future studies should plan to compare the CAD system with radiologists.

Second, few patients with COVID-19 may have negative RT-PCR results [39]. Therefore, these patients may have been incorrectly excluded from the present study.

RT-PCR is a current standard method for diagnosing of COVID-19. This molecular assay has limitations including time consumption, high cost, shortage of the kit, and requires well-equipped laboratories. These limitations are overwhelming, especially in middle and low-income countries, and may not be effective in reducing the risk of spreading the disease. Furthermore, since health systems are often weaker in lower and middle income countries, they require technical and financial support to better manage and diminish the risk of COVID-19 efficiently [40]. CT as a first-line imaging modality is fast, widely available, and does not waste materials. In this study, we proposed a CAD system based on CT images to improve diagnostic performance, which can implement in any radiology department and also in deprived areas via telecommunication and analyzing the images remotely. Accordingly, the proposed system could reduce radiologist workload burden and could play a role as an auxiliary tool with regard to decision-making as to whether an infection is COVID-19 related.

5. Conclusion

In conclusion, a CAD approach based on CT images with promising potential was proposed to distinguish infection of COVID-19 from other atypical and viral pneumonia diseases. This study showed that the ResNet-101 can be considered as a promising model to characterize and diagnose COVID-19 infections. This model does not impose substantial cost, and can be used as an adjuvant method during CT imaging in radiology departments.

Declaration of competing interest

None.

References

[1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. 382 (2020) 727–733.

[2] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (2020) 497–506.

[3] World Health Organization (WHO), Novel Coronavirus. (2019-nCoV) Technical Guidance: Laboratory Guidance. Geneva: WHO, Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance.

[4] J.P. Kanne, B.P. Little, J.H. Chung, B.M. Elicker, L.H. Ketai, Essentials for radiologists on COVID-19: an update—radiology scientific expert panel, Radiology (2020), https://doi.org/10.1148/radiol.2020200527, 200527.

[5] Z.Y. Zu, M.D. Jiang, P.P. Xu, W. Chen, Q.Q. Ni, G.M. Lu, L.J. Zhang, Coronavirus disease 2019 (COVID-19): a perspective from China, Radiology, 0 200490. doi: 10.1148/radiol.2020200490.

[6] M. Chung, A. Bernheim, X. Mei, N. Zhang, M. Huang, X. Zeng, J. Cui, W. Xu, Y. Yang, Z.A. Fayad, A. Jacobi, K. Li, S. Li, H. Shan, CT imaging features of 2019 novel coronavirus (2019-nCoV), Radiology 295 (2020) 202–207.

[7] H.J. Koo, S. Lim, J. Choe, S.-H. Choi, H. Sung, K.-H. Do, Radiographic and CT features of viral pneumonia, Radiographics 38 (2018) 719–739.

[8] H.X. Bai, B. Hsieh, Z. Xiong, K. Halsey, J.W. Choi, T.M.L. Tran, I. Pan, L.-B. Shi, D.-C. Wang, J. Mei, X.-L. Jiang, Q.-H. Zeng, T.K. Egglin, P.-F. Hu, S. Agarwal, F. Xie, S. Li, T. Healey, M.K. Atalay, W.-H. Liao, Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT, Radiology (2020), https://doi.org/10.1148/radiol.2020200823, 200823.

[9] T. Nihashi, T. Ishigaki, H. Satake, S. Ito, O. Kaii, Y. Mori, K. Shimamoto, H. Fukushima, K. Suzuki, H. Umakoshi, M. Ohashi, F. Kawaguchi, S. Naganawa, Monitoring of fatigue in radiologists during prolonged image interpretation using fNIRS, Jpn. J. Radiol. 37 (2019) 437–448.

[10] S. Taylor-Phillips, C. Stinton, Fatigue in radiology: a fertile area for future research, Br. J. Radiol. 92 (2019), 20190043.

[11] G. Lodwick, Computer-aided diagnosis in radiology. A research plan, Invest. Radiol. 1 (1965) 72–80.

[12] K. Doi, Computer-aided diagnosis in medical imaging: historical review, current status and future potential, Comput. Med. Imag. Graph. 31 (2007) 198–211.

[13] G. Castellano, L. Bonilha, L.M. Li, F. Cendes, Texture analysis of medical images, Clin. Radiol. 59 (2004) 1061–1069.

[14] J.C.M. Than, L. Saba, N.M. Noor, O.M. Rijal, R.M. Kassim, A. Yunus, H.S. Suri, M. Porcu, J.S. Suri, Lung disease stratification using amalgamation of Riesz and Gabor transforms in machine learning framework, Comput. Biol. Med. 89 (2017) 197–211.

[15] Y. Gu, X. Lu, L. Yang, B. Zhang, D. Yu, Y. Zhao, L. Gao, L. Wu, T. Zhou, Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs, Comput. Biol. Med. 103 (2018) 220–231.

[16] W. Zhang, X. Wang, X. Li, J. Chen, 3D skeletonization feature based computer-aided detection system for pulmonary nodules in CT datasets, Comput. Biol. Med. 92 (2018) 64–72.

[17] L. Sun, Z. Wang, H. Pu, G. Yuan, L. Guo, T. Pu, Z. Peng, Spectral analysis for pulmonary nodule detection using the optimal fractional S-Transform, Comput. Biol. Med. 119 (2020) 103675.

[18] W. Sun, B. Zheng, W. Qian, Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis, Comput. Biol. Med. 89 (2017) 530–539.

[19] E.M. Barbosa, S. Simpson, J.C. Lee, N. Tustison, J. Gee, H. Shou, Multivariate modeling using quantitative CT metrics may improve accuracy of diagnosis of bronchiolitis obliterans syndrome after lung transplantation, Comput. Biol. Med. 89 (2017) 275–281.

[20] F. Pancaldi, M. Sebastiani, G. Cassone, F. Luppi, S. Cerri, G. Della Casa, A. Manfredi, Analysis of pulmonary sounds for the diagnosis of interstitial lung diseases secondary to rheumatoid arthritis, Comput. Biol. Med. 96 (2018) 91–97.

[21] J. Horáček, V. Koucký, M. Hladík, Novel approach to computerized breath detection in lung function diagnostics, Comput. Biol. Med. 101 (2018) 1–6.

[22] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. (2012) 1097–1105.

[23] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 arXiv preprint arXiv:1409.1556.

[24] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters And< 0.5 MB Model Size, 2016 arXiv preprint arXiv:1602.07360.

[25] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, Mobilenetv2: inverted residuals and linear bottlenecks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.

[27] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[28] F. Chollet, Xception: deep learning with depthwise separable convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1251–1258.

[29] A.R. Van Erkel, P.M.T. Pattynama, Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology, Eur. J. Radiol. 27 (1998) 88–94.

[30] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, J. Xia, Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT, Radiology, 0 200905. doi: 10.1148/radiol.2020200905.

[31] G. Zhang, S. Jiang, Z. Yang, L. Gong, X. Ma, Z. Zhou, C. Bao, Q. Liu, Automatic nodule detection for lung cancer in CT images: a review, Comput. Biol. Med. 103 (2018) 287–300.

[32] B.A. Cunha, The atypical pneumonias: clinical diagnosis and importance, Clin. Microbiol. Infect. 12 (2006) 12–24.

[33] C. Long, H. Xu, Q. Shen, X. Zhang, B. Fan, C. Wang, B. Zeng, Z. Li, X. Li, H. Li, Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? Eur. J. Radiol. 126 (2020) 108961.

[34] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z.A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, S. Li, H. Shan, A. Jacobi, M. Chung, Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection, Radiology, 0 200463. doi: 10.1148/radiol.2020200463.

[35] W. Yang, Q. Cao, L. Qin, X. Wang, Z. Cheng, A. Pan, J. Dai, Q. Sun, F. Zhao, J. Qu, F. Yan, Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19):A multi-center study in Wenzhou city, Zhejiang, China, J. Infect. 80 (2020) 388–393.

[36] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, J. Liu, Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing, Radiology, 0 200343. doi: 10.1148/radiol.2020200343.

[37] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases, Radiology, 0 200642. doi: 10.1148/radiol.2020200642.

[38] P. Huang, T. Liu, L. Huang, H. Liu, M. Lei, W. Xu, X. Hu, J. Chen, B. Liu, Use of chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion, Radiology 295 (2020) 22–23.

[39] F.D.A, Fact sheet for healthcare providers: CDC - 2019-nCoV real-time RT-PCR diagnostic panel. https://www.fda.gov/media/134920/download. (Accessed 4 February 2020).

[40] J. Bedford, D. Enria, J. Giesecke, D.L. Heymann, C. Ihekweazu, G. Kobinger, H. C. Lane, Z. Memish, M.-d. Oh, A. Schuchat, COVID-19: towards controlling of a pandemic, Lancet 395 (10229) (2020) 1015–1018, https://doi.org/10.1016/S0140-6736(20)30673-5.