



Automated detection of COVID-19 cases using deep neural networks with X-ray images

Tulin Ozturk^a, Muhammed Talo^b, Eylul Azra Yildirim^c, Ulas Baran Baloglu^d, Ozal Yildirim^{e,*}, U. Rajendra Acharya^{f,g,h}

^a Department of Radiology, Medikal Park Hospital, Elazığ, Turkey

^b Department of Software Engineering, Firat University, Elazığ, Turkey

^c Computer Engineer, Ministry of Health, Ankara, Turkey

^d Department of Computer Engineering, University of Bristol, Bristol, UK

^e Department of Computer Engineering, Munzur University, Tunceli, Turkey

^f Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

^g Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan

^h International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto, Japan

ARTICLE INFO

Keywords:

Coronavirus (COVID-19)

Deep learning

Chest X-ray images

Radiology images

ABSTRACT

The novel coronavirus 2019 (COVID-2019), which first appeared in Wuhan city of China in December 2019, spread rapidly around the world and became a pandemic. It has caused a devastating effect on both daily lives, public health, and the global economy. It is critical to detect the positive cases as early as possible so as to prevent the further spread of this epidemic and to quickly treat affected patients. The need for auxiliary diagnostic tools has increased as there are no accurate automated toolkits available. Recent findings obtained using radiology imaging techniques suggest that such images contain salient information about the COVID-19 virus. Application of advanced artificial intelligence (AI) techniques coupled with radiological imaging can be helpful for the accurate detection of this disease, and can also be assistive to overcome the problem of a lack of specialized physicians in remote villages. In this study, a new model for automatic COVID-19 detection using raw chest X-ray images is presented. The proposed model is developed to provide accurate diagnostics for binary classification (COVID vs. No-Findings) and multi-class classification (COVID vs. No-Findings vs. Pneumonia). Our model produced a classification accuracy of 98.08% for binary classes and 87.02% for multi-class cases. The DarkNet model was used in our study as a classifier for the you only look once (YOLO) real time object detection system. We implemented 17 convolutional layers and introduced different filtering on each layer. Our model (available at (<https://github.com/muhammedtalo/COVID-19>)) can be employed to assist radiologists in validating their initial screening, and can also be employed via cloud to immediately screen patients.

1. Introduction

COVID-19 presentation, which began with the reporting of unknown causes of pneumonia in Wuhan, Hubei province of China on December 31, 2019, has rapidly become a pandemic [1–3]. The disease is named COVID-19 and the virus is termed SARS-CoV-2. This new virus spread from Wuhan to much of China in 30 days [4]. The United States of America [5], where the first seven cases were reported on January 20, 2020, reached over 300,000 by the 5th of April 2020. Most coronaviruses affect animals, but they can also be transmitted to humans because of their zoonotic nature. Severe acute respiratory syndrome Coronavirus

(SARS-CoV) and the Middle East respiratory syndrome Coronavirus (MERS-CoV) have caused severe respiratory disease and death in humans [6]. The typical clinical features of COVID-19 include fever, cough, sore throat, headache, fatigue, muscle pain, and shortness of breath [7].

The most common test technique currently used for COVID-19 diagnosis is a real-time reverse transcription-polymerase chain reaction (RT-PCR). Chest radiological imaging such as computed tomography (CT) and X-ray have vital roles in early diagnosis and treatment of this disease [8]. Due to the low RT-PCR sensitivity of 60%–70%, even if negative results are obtained, symptoms can be detected by examining

* Corresponding author.

E-mail address: oyildirim@munzur.edu.tr (O. Yildirim).

<https://doi.org/10.1016/j.complbiomed.2020.103792>

Received 5 April 2020; Received in revised form 21 April 2020; Accepted 26 April 2020

Available online 28 April 2020

0010-4825/© 2020 Elsevier Ltd. All rights reserved.

radiological images of patients [9,10]. It is stated that CT is a sensitive method to detect COVID-19 pneumonia, and can be considered as a screening tool with RT-PCR [11]. CT findings are observed over a long interval after the onset of symptoms, and patients usually have a normal CT in the first 0–2 days [12]. In a study on lung CT of patients who survived COVID-19 pneumonia, the most significant lung disease is observed ten days after the onset of symptoms [13].

At the beginning of the pandemic, Chinese clinical centers had insufficient test kits, which are also producing a high rate of false-negative results, so doctors are encouraged to make a diagnosis only based on clinical and chest CT results [12,14]. CT is widely used for COVID-19 detection in countries such as Turkey, where a low number of test kits at onset of the pandemic were available. Researchers state that combining clinical image features with laboratory results may help in early detection of COVID-19 [6,11,15–17]. Radiologic images obtained from COVID-19 cases contain useful information for diagnostics. Some studies have encountered changes in chest X-ray and CT images before the beginning of COVID-19 symptoms [18]. Significant discoveries have been realized by investigators in imaging studies of COVID-19. Kong et al. [6] observed right infrahilar airspace opacities in a COVID-19 patient. Yoon et al. [19] reported that one in three patients studied had a single nodular opacity in the left lower lung region. In contrast, the other two had four and five irregular opacities in both lungs. Zhao et al. [16] not only found ground-glass opacities (GGO) or mixed GGO in most of the patients, but they also observed a consolidation, and vascular dilation in the lesion. Li and Xia [17] reported GGO and consolidation, interlobular septal thickening and air bronchogram sign, with or without vascular expansion, as common CT features of COVID-19 patients. Peripheral focal or multifocal GGO affecting both lungs in 50%–75% of patients is another observation [9]. Similarly, Zu et al. [8] and Chung et al. [6] discovered that 33% of chest CTs can have rounded lung opacities. In Fig. 1, chest X-ray images taken at days 1, 4, 5 and 7 for a 50-year-old COVID-19 patient with pneumonia are given, and explanations of these images are also provided [20].

Application of machine learning methods for automatic diagnosis in the medical field have recently gained popularity by becoming an adjunct tool for clinicians [21–25]. Deep learning, which is a popular research area of artificial intelligence (AI), enables the creation of end-to-end models to achieve promised results using input data, without the need for manual feature extraction [26,27]. Deep learning techniques have been successfully applied in many problems such as arrhythmia detection [28–30], skin cancer classification [31,32], breast cancer detection [33,34], brain disease classification [35], pneumonia detection from chest X-ray images [36], fundus image segmentation [37], and lung segmentation [38,39]. The COVID-19 epidemic's rapid rise has necessitated the need for expertise in this field. This has increased interest in developing the automated detection systems based on AI techniques. It is a challenging task to provide expert clinicians to

every hospital due to the limited number of radiologists. Therefore, simple, accurate, and fast AI models may be helpful to overcome this problem and provide timely assistance to patients. Although radiologists play a key role due to their vast experience in this field, the AI technologies in radiology can be assistive to obtain accurate diagnosis [40]. Additionally, AI approaches can be useful in eliminating disadvantages such as insufficient number of available RT-PCR test kits, test costs, and waiting time of test results.

Recently, many radiology images have been widely used for COVID-19 detection. Hemdan et al. [41] used deep learning models to diagnose COVID-19 in X-ray images and proposed a COVIDX-Net model comprising seven CNN models. Wang and Wong [42] proposed a deep model for COVID19 detection (COVID-Net), which obtained 92.4% accuracy in classifying normal, non-COVID pneumonia, and COVID-19 classes. Ioannis et al. [43] developed the deep learning model using 224 confirmed COVID-19 images. Their model achieved 98.75% and 93.48% success rates for two and three classes, respectively. Narin et al. [44] achieved a 98% COVID-19 detection accuracy using chest X-ray images coupled with the ResNet50 model. Sethy and Behera [45] classified the features obtained from various convolutional neural network (CNN) models with support vector machine (SVM) classifier using X-ray images. Their study states that the ResNet50 model with SVM classifier provided the best performance. Finally, there are also several recent studies on COVID-19 detection that employed various deep learning models with CT images [46–51].

In this study, a deep learning model is proposed for the automatic diagnosis of COVID-19. The proposed model has an end-to-end architecture without using any feature extraction methods, and it requires raw chest X-ray images to return the diagnosis. This model is trained with 125 chest X-ray images, which are not in a regular form and were obtained hastily. Diagnostic tests performed after 5–13 days are found to be positive in recovered patients [52]. This crucial finding shows us that recovered patients may continue to spread the virus. Therefore, more accurate methods for the diagnosis is needed. One of the most important disadvantages of chest radiography analyses is an inability to detect the early stages of COVID-19, as they do not have sufficient sensitivity in GGO detection [8]. However, well-trained deep learning models can focus on points that are not noticeable to the human eye, and may serve to reverse this perception.

2. Material and methods

2.1. X-ray image DataSet

In this study, X-ray images obtained from two different sources were used for the diagnosis of COVID-19. A COVID-19 X-ray image database was developed by Cohen JP [53] using images from various open access sources. This database is constantly updated with images shared by

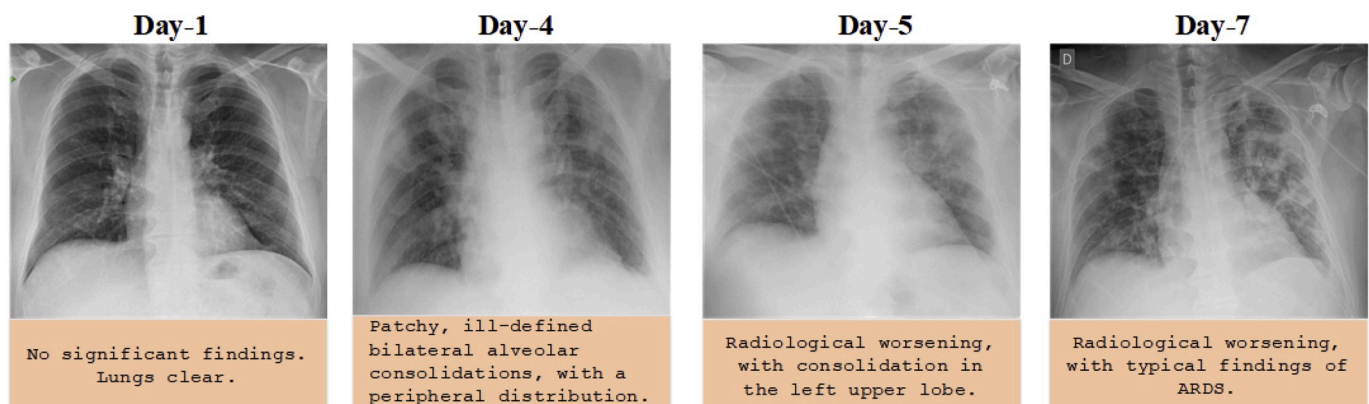


Fig. 1. Chest X-ray images of a 50-year-old COVID-19 patient with pneumonia over a week [20].

researchers from different regions. At present, there are 127 X-ray images diagnosed with COVID-19 in the database. Fig. 2 shows a few COVID-19 cases obtained from the database and the findings of the experts.

There are 43 female and 82 male cases in the database that were found to be positive. In this dataset, a complete metadata is not provided for all patients. The age information of 26 COVID-19 positive subjects is given, and the average age of these subjects is approximately 55 years. Also, the ChestX-ray8 database provided by Wang et al. [58] was used for normal and pneumonia images. In order to avoid the unbalanced data problem, we used 500 no-findings and 500 pneumonia class frontal chest X-ray images randomly from this database.

2.2. The proposed DarkCovidNet model

The advent of deep learning technology has revolutionized artificial intelligence [26,27]. The word deep refers to the increase in the size of this network with the number of layers. The structure is named after convolution, a mathematical operator. A typical CNN structure has a convolution layer that extracts features from the input with the filters it applies, a pooling layer to reduce the size for computational performance, and a fully connected layer, which is a neural network. By combining one or more such layers, a CNN model is created, and its internal parameters are adjusted to accomplish a particular task, such as classification or object recognition.

Instead of initiating a deep model development from scratch, a more rational approach is to construct a model using already proven models. Therefore, while designing the deep model used in this study, the Darknet-19 model [59] is chosen as the starting point. Darknet-19 is the classifier model that forms the basis of a real-time object detection system named YOLO (You only look once) [59]. This system has the state-of-the-art architecture designed for object detection. The DarkNet classifier is used on the basis of this successful architecture. We designed

the DarkCovidNet architecture (available at <https://github.com/muhammedtalo/COVID-19>), inspired by the DarkNet architecture that has proven itself in deep learning, instead of building a model from scratch. We have used fewer layers and filters as compared to the original DarkNet architectures. We gradually increased the number of filters such as to 8, 16, 32. To better understand this new model, it is helpful to understand the basics of the Darknet-19, which consists of 19 convolutional layers and five pooling layers, using Maxpool. These layers are typical CNN layers with different filter numbers, sizes, and stride values. Let the letter C denote a convolutional layer, and M denote a Maxpool layer. As C_1 is taken as the input layer, Darknet-19 has a layer layout as follows.

$$C_1-M_1-C_2-M_2-C_3-C_4-C_5-M_3-C_6-C_7-C_8-M_4-C_9-C_{10}-C_{11}-C_{12}-C_{13}-M_5-C_{14}-C_{15}-C_{16}-C_{17}-C_{18}-C_{19}$$

For input signal X (image) and kernel K , the two-dimensional convolution operation can be defined as follows.

$$(X * K)(i, j) = \sum_m \sum_n K(m, n) X(i - m, j - n) \quad (1)$$

where $*$ represents the discrete convolution operation. The K matrix slides over the input matrix with stride parameter. The Leaky rectified linear unit (Leaky ReLu) is used as an activation function in the DarkNet architecture. The calculation of the leaky ReLu function is given in equation (2).

$$f(x) = \begin{cases} 0.01x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (2)$$

A schematic presentation for the flow of input data from convolution layer (C) and Max-pooling (M) layer, respectively, is given in Fig. 3.

This model ends with Avgpool and Softmax layers that produce the outputs. Eventually, a deep model with large number of layers is essential for the feature extraction of a real-time object detection system.

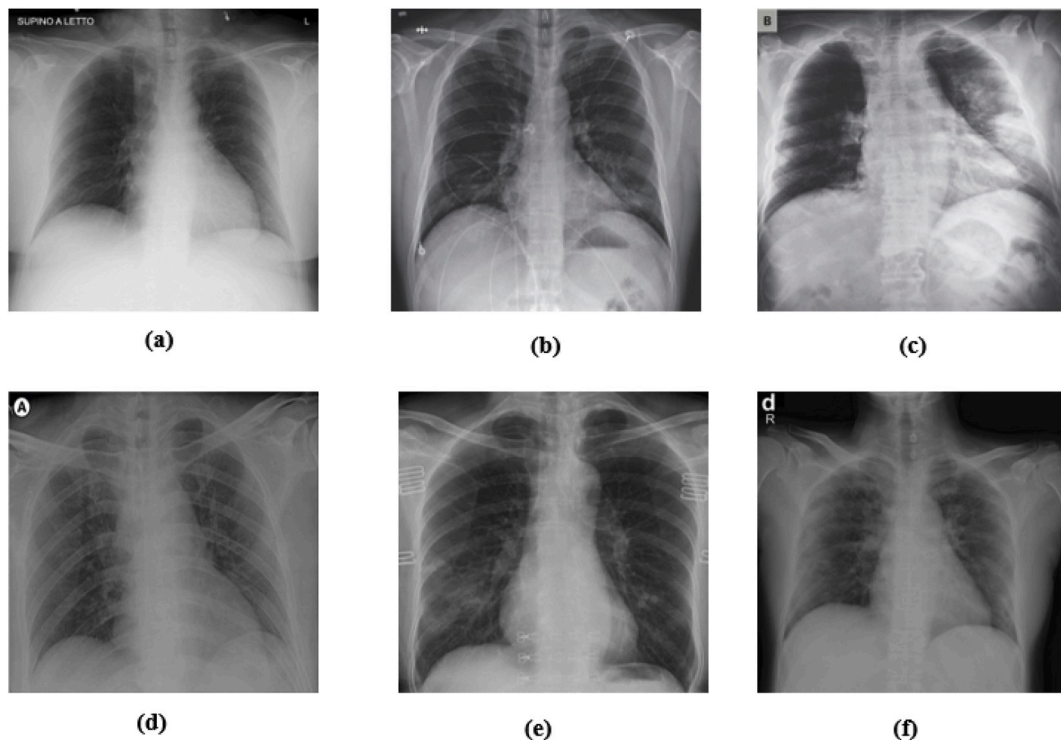


Fig. 2. A few COVID-19 cases and findings by dataset: (a) Cardio-vascular shadow within the limits [54], (b) Increasing left basilar opacity is visible, arousing concern about pneumonia [5], (c) Progressive infiltrate and consolidation [55], (d) Small consolidation in right upper lobe and ground-glass opacities in both lower lobes [56], (e) Infection demonstrates right infrahilar airspace opacities [6], and (f) Progression of prominent bilateral perihilar infiltration and ill-defined patchy opacities at bilateral lungs [57].

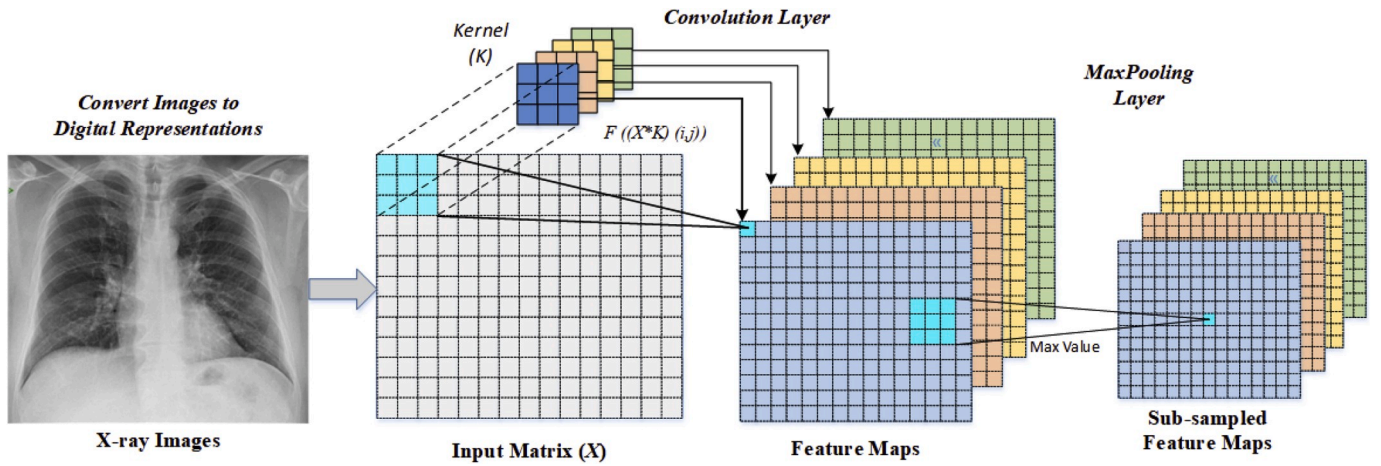


Fig. 3. A schematic presentation of convolution and Max-pooling layer operations.

In this study, we encountered a problem in classifying the images with subtle details. The model performing such classification should have a structure which can capture and learn small differences rather than being very deep, like ResNets or ResNext [60] models. An illustration of the proposed model used in this study is shown in Fig. 4.

The proposed model has 17 convolution layers. In Fig. 4, each DN (DarkNet) layer has one convolutional layer followed by BatchNorm, and LeakyReLU operations, while each $3 \times \text{Conv}$ layer has the same setup *three* times in successive form. The batch normalization operation is used to standardize the inputs, and this operation has other benefits, such as reducing training time and increasing stability of the model. LeakyReLU is a variation of the ReLU operation used to prevent dying neurons. Unlike ReLU or sigmoid activation functions, which have zero value in the negative part of their derivatives, LeakyReLU has a small epsilon value to overcome the dying neuron problem. Similar to the Darknet-19 model, the Maxpool method is used in all the pooling operations. Maxpool downsizes an input by taking the maximum of a region determined by its filter. When working with two classes, the proposed model performs the COVID-19 detection task. If three different classes of images are used in the input, the same model performs the classification task to determine the labels of the input chest X-ray images as COVID-19, Pneumonia, or No-Findings. Finally, the layer details and layer parameters of the model are given in Table 1. The developed deep

Table 1

The layers and layer parameters of the proposed model (for the binary classification task).

Number of Layer	Layer Type	Output Shape	Number of Trainable Parameters
1	Conv2d	[8, 256, 256]	216
2	Conv2d	[16, 128, 128]	1152
3	Conv2d	[32, 64, 64]	4608
4	Conv2d	[16, 66, 66]	512
5	Conv2d	[32, 66, 66]	4608
6	Conv2d	[64, 33, 33]	18,432
7	Conv2d	[32, 35, 35]	2048
8	Conv2d	[64, 35, 35]	18,432
9	Conv2d	[128, 17, 17]	73,728
10	Conv2d	[64, 19, 19]	8192
11	Conv2d	[128, 19, 19]	73,728
12	Conv2d	[256, 9, 9]	294,912
13	Conv2d	[128, 11, 11]	32,768
14	Conv2d	[256, 11, 11]	294,912
15	Conv2d	[128, 13, 13]	256
16	Conv2d	[256, 13, 13]	294,912
17	Conv2d	[2, 13, 13]	4608
18	Flatten	[338]	0
19	Linear	[2]	678

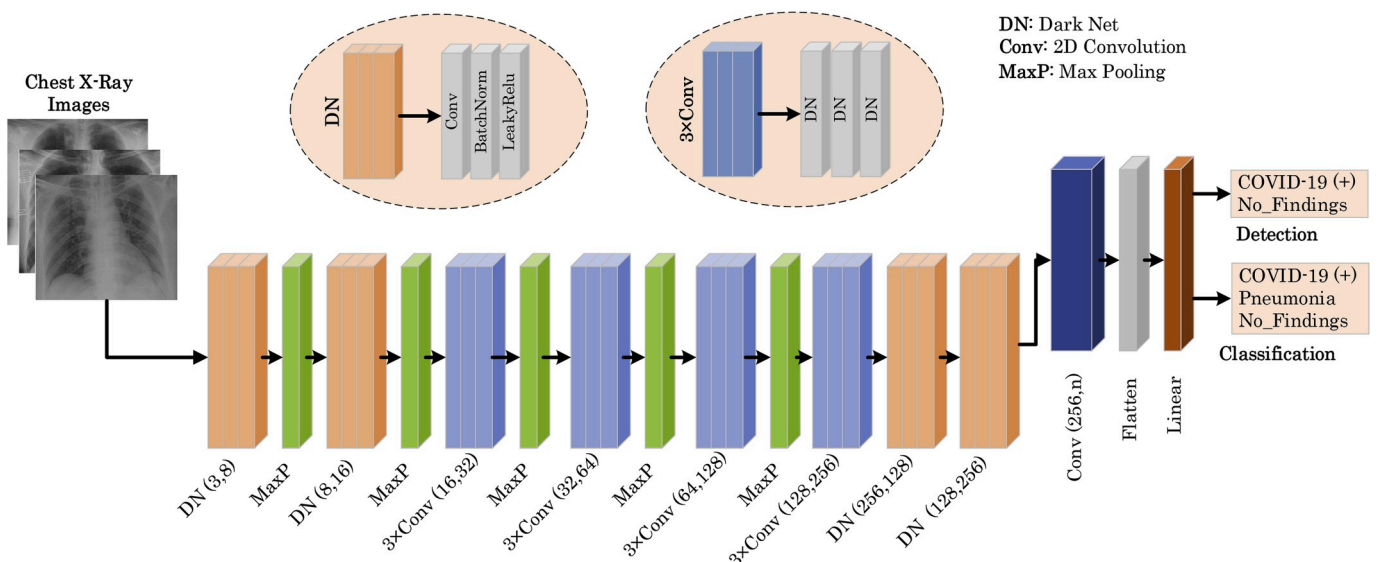


Fig. 4. The architecture of the proposed model (DarkCovidNet).

learning model consists of 1,164,434 parameters. We have used the Adam optimizer for weight updates, cross entropy loss function and selected learning rate as $3e-3$.

3. Experimental results

We performed experiments to detect and classify COVID-19 using X-ray images in two different scenarios. First, we have trained the DarkCovidNet deep learning model to classify X-ray images into three categories: COVID-19, No-Findings, Pneumonia. Secondly, the DarkCovidNet model is trained to detect two classes: COVID-19 and *No-Findings* categories. The performance of the proposed model is evaluated using the 5-fold cross-validation procedure for both the binary and triple classification problem. Eighty percent of X-ray images are used for training and 20% for validation. The experiments are repeated five times as shown in Fig. 5. All of the split k pieces are wrapped in folds to use in the validation stage. We have trained DarkCovidNet for 100 epochs. In Fig. 6 The training and validation loss graphs of the multi-class classification and validation accuracy graphs are shown for the Fold-1.

It can be noted from Fig. 6 that there is a significant increase in loss values in the beginning of the training, which decrease substantially in the later stage of the training. The main reason for this sharp increase and decrease is attributed to the number of data in the COVID-19 class, which is far less than the other two (Pneumonia and No-Findings) classes. However, when the deep model examines all X-ray images over and over again for each epoch during the training, these rapid ups and downs are slowly reduced in the later part of the training.

The multi-class classification performance of the DarkCovidNet model has been evaluated for each fold, and the average classification performance of the model is calculated. The overlapped as well as each separate confusion matrix (CM) are shown in Fig. 7. The overlapped CM is created using the sum of CMs obtained in all folds. Thus, it is aimed to obtain an idea about the general perforations of the model. The DarkCovidNet model achieved an average classification accuracy of 87.02% to classify: no findings, COVID-19, and Pneumonia categories. Sensitivity, specificity, precision, F1-score, and accuracy values are shown in Table 2 for the detail analysis of the model for the 3-class problem.

It can be noted from the overlapped confusion matrix of the multi-class classification task that the deep learning model classified COVID-19 better than the classes of pneumonia and no findings. The obtained sensitivity, specificity, and F1-score values are 85.35%, 92.18%, and 87.37%, respectively.

Secondly, the result of confusion matrixes for the binary classification problem in detecting COVID-19 positive are shown in Fig. 8. In addition, sensitivity, specificity precision, F1-score, and accuracy results for the binary classification task are given in Table 3.

It can be noted from Table 3 that the proposed model has achieved an average accuracy of 98.08% in detecting COVID-19 and the obtained

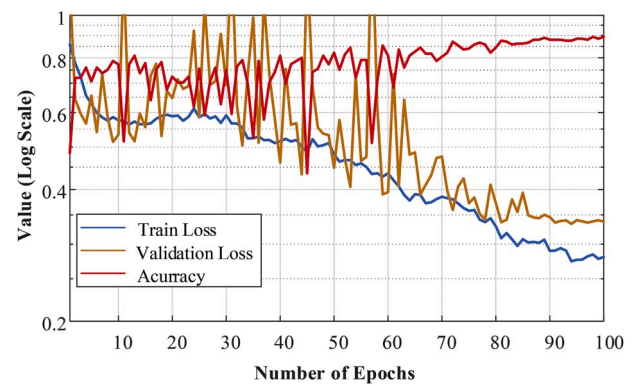


Fig. 6. Validation, training loss and validation accuracy curves obtained for DarkCovidNet model in fold-1.

average sensitivity, specificity, and F1-score values of 95.13%, 95.30%, and 96.51%, respectively.

4. Evaluation of the model outputs by the radiologist

This section includes the interpretation of results of the DarkCovidNet model by an expert radiologist. The DarkCovidNet model is designed for the automatic detection of COVID-19 using X-ray images, without requiring any handcrafted feature extraction techniques. The developed model helps to provide a second opinion to expert radiologists in health centers. It may significantly reduce the workload of clinicians and assist them to make an accurate diagnosis in their daily routine work. The proposed model can save time (the diagnostic process is fast); hence specialists can focus on more critical cases. In this work, we have shared the outputs of the model with expert radiologists to confirm model robustness. We shared the top prediction errors of the model and the actual labels of the X-ray dataset with radiologists. Additionally, we also used the Grad-CAM [61] heat map approach to visually depict decisions made by the deep model. The heatmap highlights important areas that the model emphasizes on the X-ray. In this way, we have ensured that the outcome of the model is approved by a radiologist. In the clinical setting, an illustration of the model which can provide the second opinion to radiologists is shown in Fig. 9.

The radiologist comments on the output of DarkCovidNet model are as follows:

- The model performed outstandingly in detecting COVID-19 cases for the binary class task.
- The DarkCovidNet model is successful in detecting COVID-19 findings.
- Clinically, pneumonia images are also later included in the study. Therefore, the model evaluated patients with COVID-19 as pneumonia. Since COVID-19 pneumonia is a subset of pneumonia diseases evaluated by the model, the diagnosis is correct, although the interpretation seems to be incorrect. Therefore, patients identified as COVID-19 are evaluated as pneumonia (see Fig. 10 (a)). For this reason, the success rate of the model in the multi-class classification problem is relatively low as compared to the binary class.
- The model is sensitive in detecting pneumonia disease. Although the model can predict pneumonia positively and marked as no findings in the dataset, this patient has a mass (see Fig. 10 (b)).
- The model made the incorrect predictions in poor quality X-ray imagery and in patients with acute respiratory distress syndrome (ARDS), in which the lung image is diffuse and much lung ventilation is lost (see Fig. 10 (c)).
- The model is useful to detect COVID-19 with presence of a heat map in normal subjects. Its effectiveness diminishes in pneumonia and ARDS cases. The heat map showed a greater concentration area in

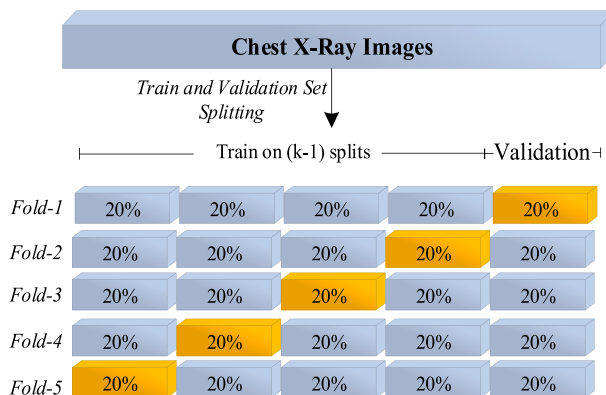


Fig. 5. Schematic representation of training and validation scheme employed in the 5-fold cross-validation procedure.

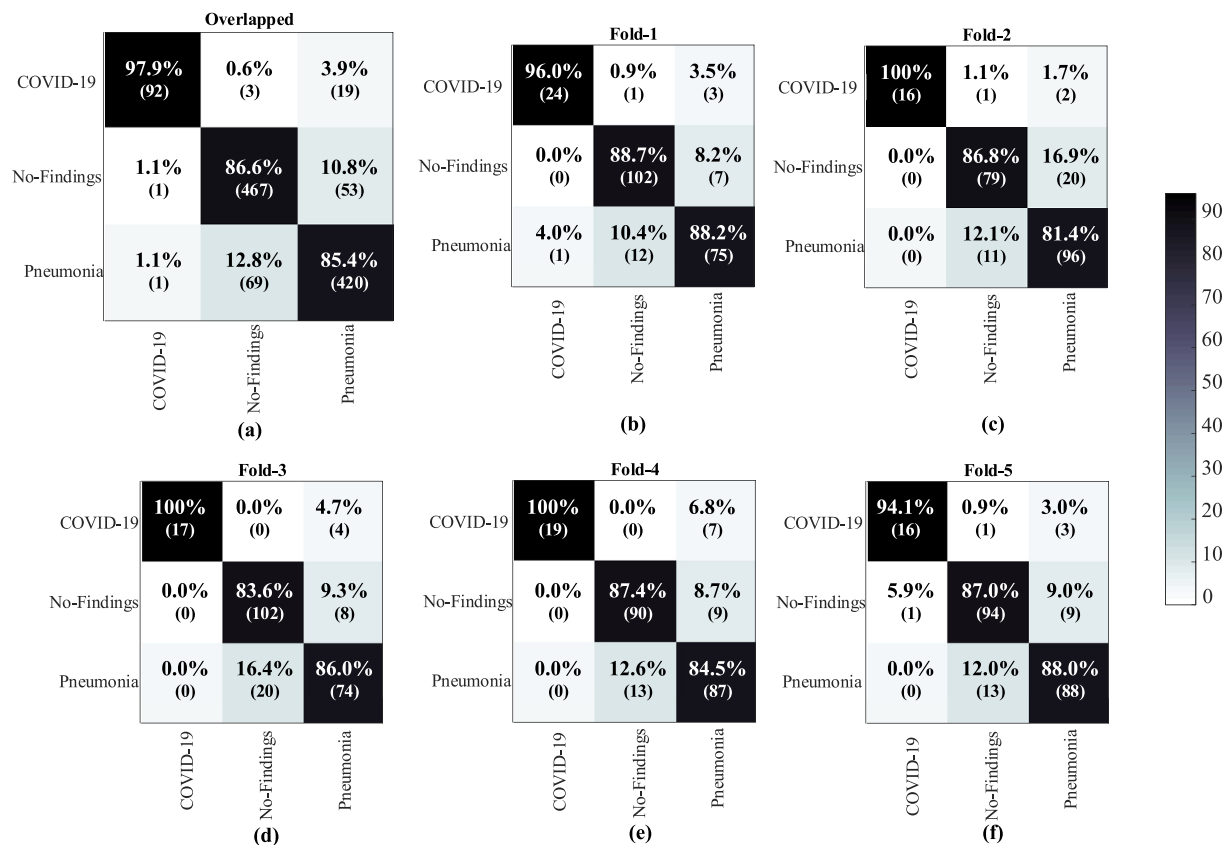


Fig. 7. The overlapped and 5-fold confusion matrix results of the multi-class classification task: (a) overlapped confusion matrix, (b) Fold-1 CM, (c) Fold-2 CM, (d) Fold-3 CM, (e) Fold-4 CM, and (f) Fold-5 CM.

Table 2

Sensitivity, specificity, precision, F1-score, and accuracy values obtained for each fold of the proposed model.

Folds	Performance Metrics (%)				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
Fold-1	88.17	93.66	90.97	89.44	89.33
Fold-2	84.57	90.61	89.38	86.63	84.89
Fold-3	84.13	91.14	89.88	86.54	85.78
Fold-4	83.66	92.29	90.61	86.42	87.11
Fold-5	85.83	92.75	89.71	87.57	88.00
Average	85.35	92.18	89.96	87.37	87.02

the X-rays of patients with COVID-19 than the area in which the disease is not seen (see Fig. 11).

- The model may be more useful evaluate the efficacy of treatment based on the heatmap. It can also assist experts in terms of diagnosis, follow-up, treatment, and isolation of patients.

Fig. 12 shows the difference between a few COVID and pneumonia case images. The following primary findings are frequently observed in the chest X-rays of COVID-19 patients [15].

- Ground-glass opacities (GGO) (bilateral, multifocal, subpleural, peripheral, posterior, medial and basal).
- A crazy paving appearance (GGOs and inter-/intra-lobular septal thickening).
- Air space consolidation.
- Bronchovascular thickening (in the lesion).
- Traction bronchiectasis.

Similarly, chest X-ray findings of pneumonia patients are observed as

follows [62].

- Ground-glass opacities (GGO) central distribution, unilateral
- Reticular opacity
- Vascular thickening
- Distribution more along the bronchovascular bundle
- Bronchial wall thickening

In COVID-19, isolated lobar or segmental consolidation without GGO, multiple tiny pulmonary nodules, tree-in-bud, pneumothorax, cavitation, and hilar lymphadenopathy smoother interlobular septal thickening with pleural effusion are rare, while these findings can often be seen in pneumonia [9].

In the COVID-19 epidemic, radiological imaging plays an important role in addition to the diagnostic tests performed for the early diagnosis, treatment, and isolation stages of the disease. Chest radiography can detect a few characteristic findings in the lung associated with COVID-19. Deep learning models are sensitive in detecting COVID-19 lung involvement and hence the diagnostic accuracy rate is high. During the evaluation of the model, X-ray radiographs of COVID-19 patients confirmed positive by the PCR Test are used. The model can easily detect GGO, consolidation areas, and nodular opacities, which are the pathognomic findings of patients for COVID-19 on X-ray radiography. In COVID-19, bilateral, lower lobe, and peripheral involvement is observed, and the proposed model can detect localization of the lesion. These models are particularly important in identifying early stages of COVID-19 patients. Early diagnosis of the disease is important to provide immediate treatment and to prevent disease transmission. The models can also play an indispensable role in patients lacking early symptoms. There is a margin of error in patients with diffuse late lung parenchyma and in patients with significantly reduced lung ventilation due to poor quality X-ray images. X-rays that are not of optimal quality are difficult

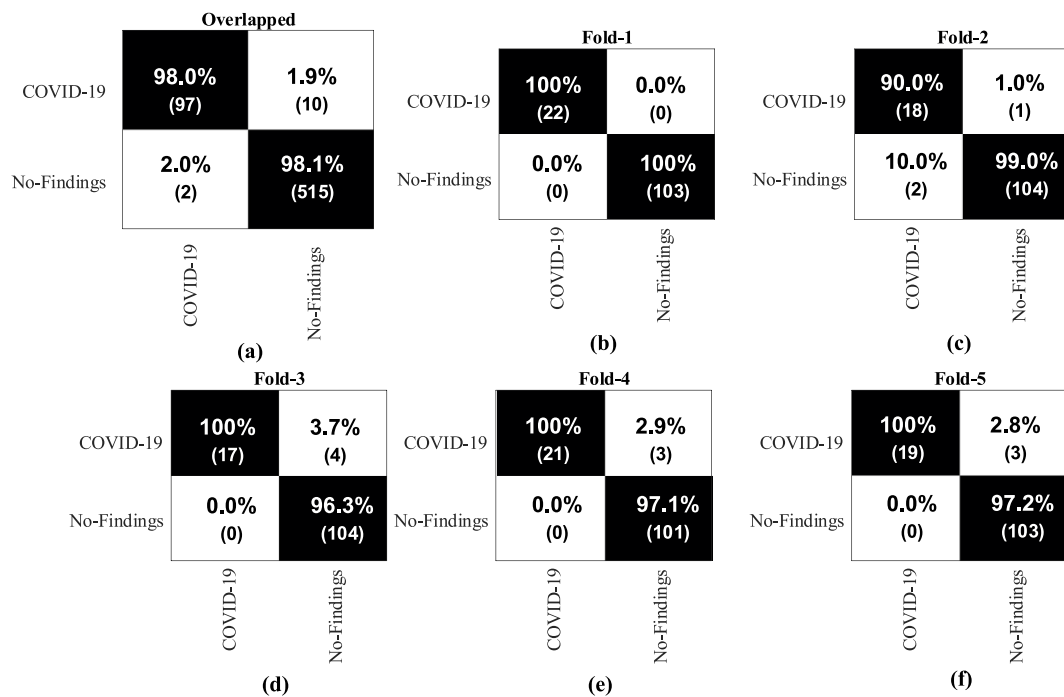


Fig. 8. The overlapped and 5-fold confusion matrix results for the binary classification task: (a) Overlapped confusion matrix, (b) Fold-1 CM, (c) Fold-2 CM, (d) Fold-3 CM, (e) Fold-4 CM, and (f) Fold-5 CM.

Table 3

Sensitivity, specificity, precision, F1-score, and accuracy values for No findings and COVID-19 classes of the proposed model.

Folds	Performance Metrics (%)				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
Fold-1	100	100	100	100	100
Fold-2	96.42	96.42	94.52	95.52	97.60
Fold-3	90.47	90.47	98.14	93.79	96.80
Fold-4	93.75	93.75	98.57	95.93	97.60
Fold-5	93.18	93.18	98.58	95.62	97.60
Overlapped					
COVID-19	90.65	99.61	97.97	94.17	98.07
No-Findings	99.61	90.65	98.09	98.84	98.07
Average	95.13	95.3	98.03	96.51	98.08

to evaluate by radiologists. The clinical and radiological images of later-stage patients are well established and it is easier to detect the findings by experts. The role of deep learning models is more prominent in screening and diagnosis when the infection is at its early stages.

The models can be readily used in healthcare centers. There is no need to wait long hours for the radiologists to screen the images. Thus, healthcare workers and patient relatives can focus on isolation of suspicious cases so that treatment can begin. Hence, the spread of the disease can be significantly reduced. The patients can seek a second opinion if they are diagnosed as positive by our system. Hence, waiting time can be significantly reduced, and it will alleviate clinician workload.

5. Discussion

Dr. Cohen has compiled the radiology images from various sources of COVID-19 cases for researches to develop an accurate model to diagnose

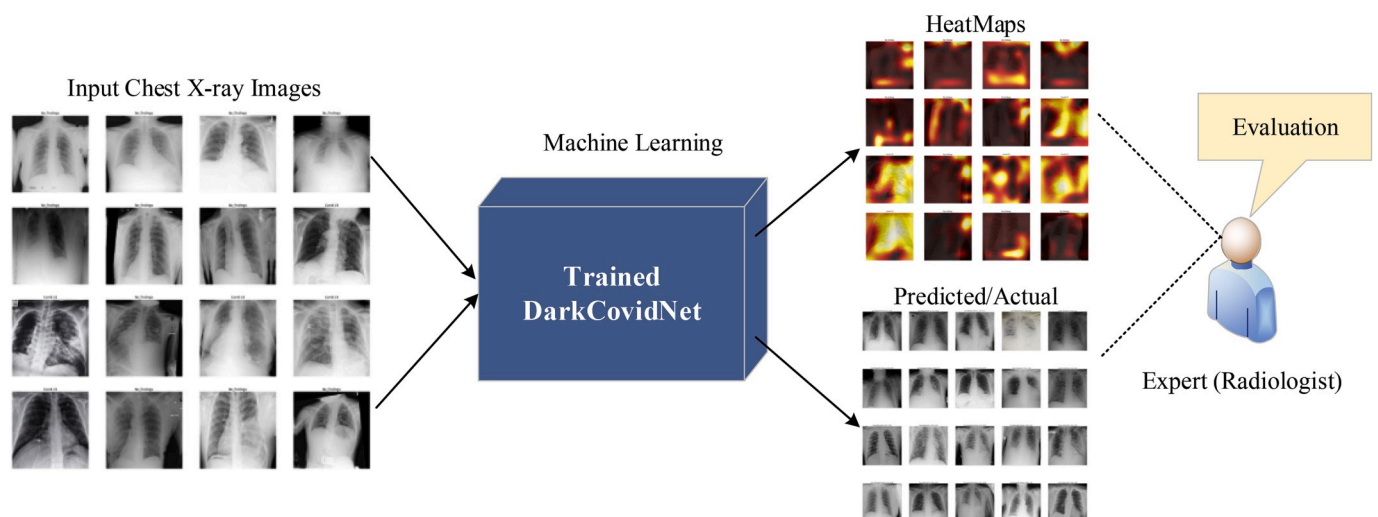


Fig. 9. An illustration of performance evaluation of the model outputs by an expert.

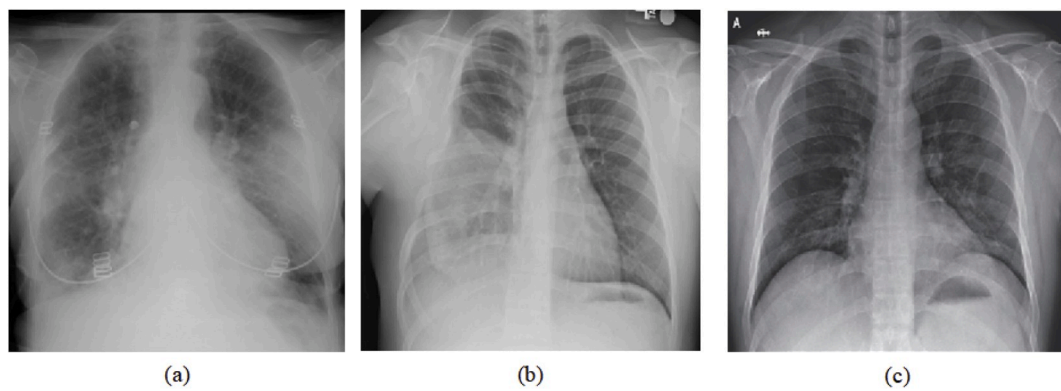


Fig. 10. Images evaluated by the radiologist and DarkCovidNet model: (a) Predicted as Pneumonia by model but actual class is COVID-19, (b) Predicted as Pneumonia by model but actual class is No-Findings, (c) Model is correctly detected as multifocal GGO.

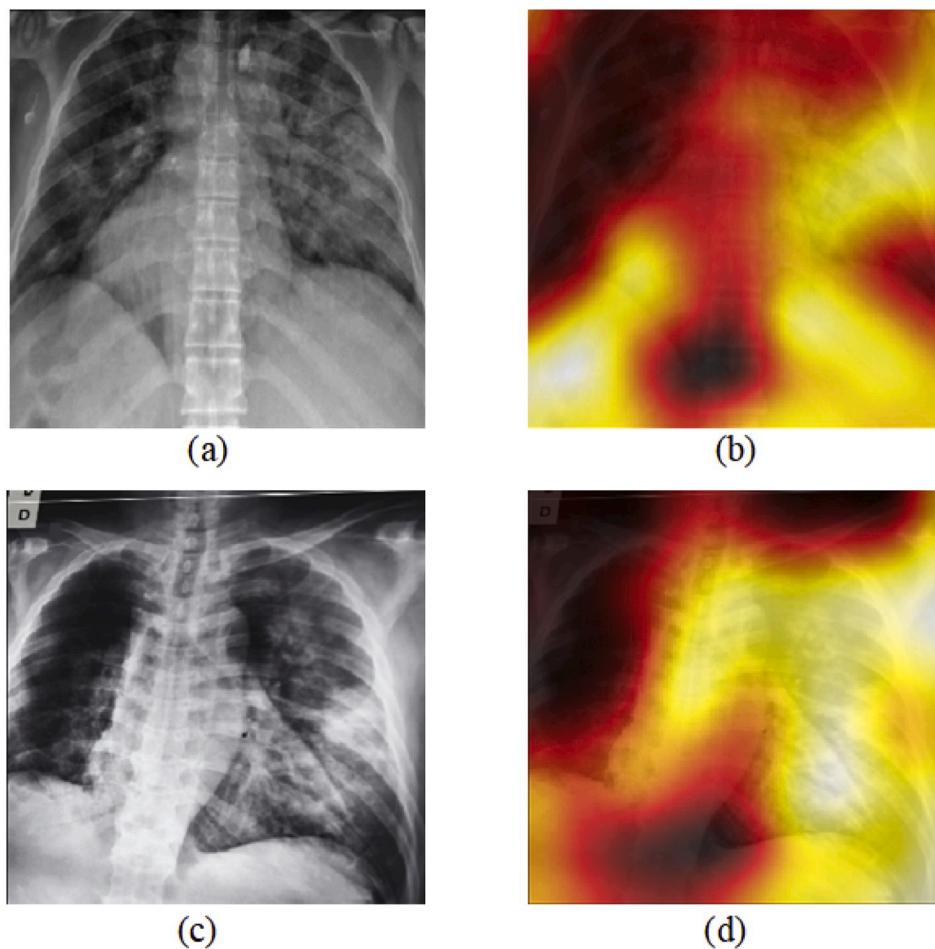


Fig. 11. X-ray images and the corresponding heat maps: (a) first X-ray image, (b) heat map of (a), (c) second X-ray image, and (d) heat map of (c).

this disease accurately [53]. Most of the studies mentioned in this part of the article used the COVID-19 image data from this source, and for other cases, such as Non-COVID-19, images are obtained from various publicly available sources. Hemdan et al. [41] proposed COVIDX-Net to diagnose COVID-19 in X-ray images. They have obtained 90% accuracy rate using 25 COVID-19 positive and 25 normal images. Wang and Wong [42] designed COVID-Net, a deep learning-based model, for COVID19 detection. COVID-Net achieved a 92.4% success rate using a total of 16,756 radiography images obtained from different open access data. Transfer learning, which has gained increasing attention, is used by Ioannis et al. [43] for the same purpose as COVID-Net. In this study, 224

approved COVID-19, 700 pneumonias, and 504 normal radiology images are used. They achieved a 98.75% performance for the 2-class and 93.48% performance for the 3-class problem. Narin et al. [44] employed three different CNN models (ResNet50, InceptionV3, and Inception-ResNetV2) using 50 open access COVID-19 chest X-ray images [53] shared by Joseph Cohen and 50 normal images taken from the Kaggle repository [63]. However, in this study, non-COVID images are images belonging to children between the ages of 1–5 years.

The researchers also attempted different models and combinations of various methodologies. Sethy and Behera [45] used CNN models to extract image features and then classified them using the SVM classifier.

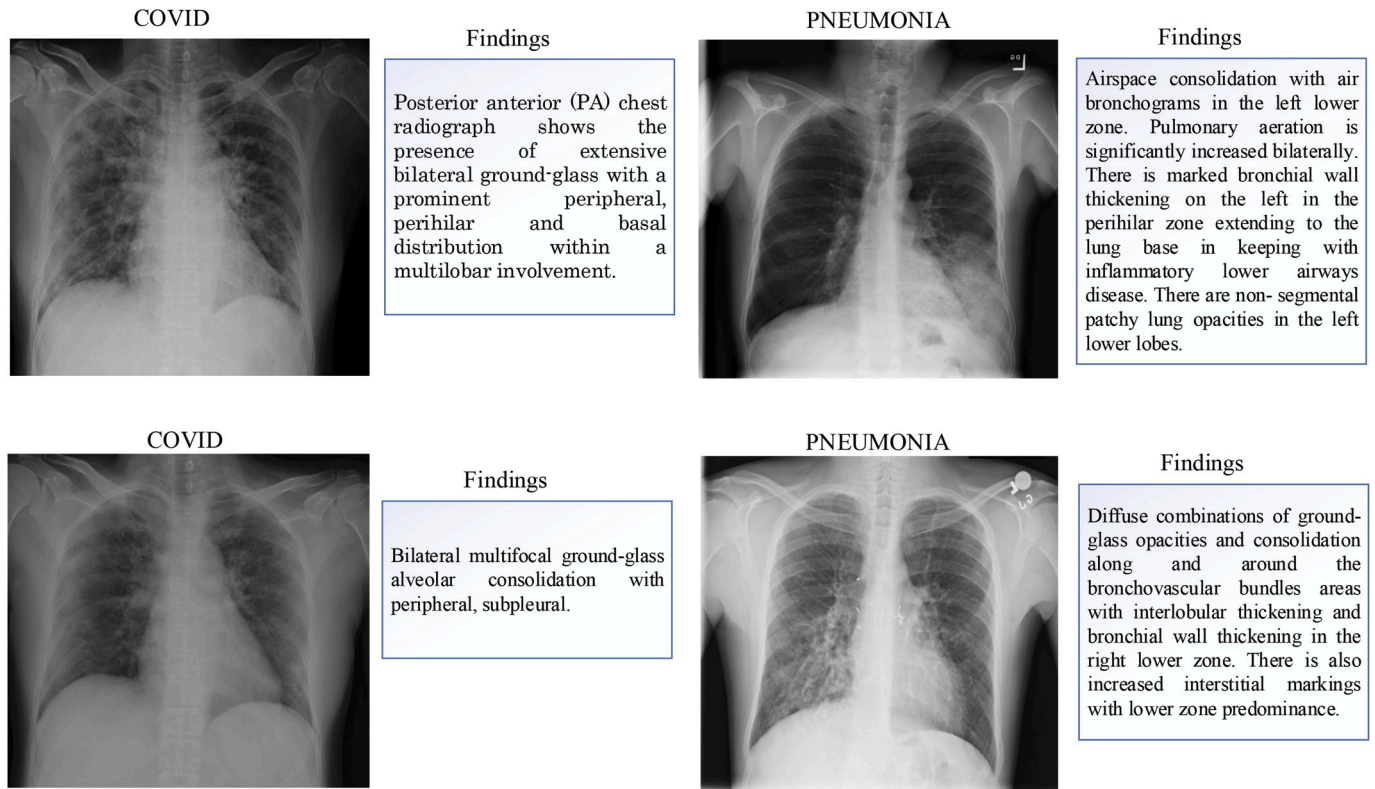


Fig. 12. Differences observed by the radiologist between some COVID and pneumonia case images.

In this study, a 95.38% accuracy was achieved using ResNet50 and SVM in combination with 50 images. Ying et al. achieved an 86% success rate using CT images with a deep model built on the pre-trained ResNet50, termed DRE-Net [46]. Wang et al. [47] achieved a classification accuracy of 82.9% using the modified Inception (M-Inception) deep model developed using CT images. Zheng et al. [48] proposed a three-dimensional deep CNN model to detect COVID-19 from CT imagery and reported a 90.8% accuracy. Xu et al. [49] achieved an 86.7% performance rate in detecting COVID-19 using ResNet coupled with CT images. Most of these earlier studies have few data to develop the model.

In this study, a deep learning model called DarkCovidNet was used for the detection of COVID-19. We have used a total of 1125 images (125 COVID-19(+), 500 Pneumonia and 500 No-Findings) to develop our model. We obtained an accuracy of 98.08% and 87.02% for binary and three classes, respectively. We obtained superior performance compared to other studies in the literature (Table 4).

The COVID-19 public image data has limited data, which is a limitation of this study. Public sources from X-ray images are constantly updated. Few studies conducted prior to this study have used 25 and 50 images in each class [41,42,44,45]. In this study, a total of 125 COVID-19 positive data are used. In previous studies, common techniques such as VGG and ResNet are generally used. However, the proposed new model is based on the DarkNet method. In addition, an effective screening process is required to eliminate images that are not useful present in the database. Since COVID-19 is a new epidemic, fewer X-ray images are available for developing the automated diagnostic system. The main advantages of the model are as follows:

- The model classified chest X-ray images without using a feature extraction technique.
- It is an effective approach that can assist experts for diagnosis.
- The heatmaps produced by the model are evaluated by an expert radiologist. The model focuses on localizing effective regions on chest X-ray images.

The proposed model can be used for the diagnosis of COVID-19 using X-ray radiographs. X-ray radiographs are preferred because they are readily accessible for disease diagnosis. They are widely used in health centers worldwide during the pandemic. The model has the ability to diagnose COVID-19 within seconds. CT is a costly process and not readily accessible as they are usually only located in larger health centers. In addition, when CT is compared to X-ray, the amount of radiation received by the patient is more. Hence, it is recommended to use a deep learning model with X-ray imagery, as it is more accessible with lower radiation dose as compared to CT. Patients diagnosed as COVID positive by the model can be directed to advanced centers for confirmation, followed by treatment without delay. In addition, patients who have been diagnosed negatively by the model can be prevented from undergoing PCR tests and occupying health centers unnecessarily.

In the future, we intend to validate our model by incorporating more images. This developed model can be placed in a cloud to provide diagnosis instantly and to help rehabilitate affected patients immediately. This should reduce clinician workload significantly. We will address CT images for COVID-19 detection and compare the obtained results with the proposed model trained using X-ray images. Also, we will try to collect local radiology images for COVID-19 cases and evaluate them with our model from sites in Turkey. After the necessary tests are done, we aim to deploy the developed model in local hospitals for screening.

6. Conclusion

In this study, we have proposed a deep learning based model to detect and classify COVID-19 cases from X-ray images. Our model is fully automated with an end-to-end structure without the need for manual feature extraction. Our developed system is able to perform binary and multi-class tasks with an accuracy of 98.08% and 87.02%, respectively. The performance of the developed model is assessed by expert radiologists and is ready to be tested with a larger database. This

Table 4

Comparison of the proposed COVID-19 diagnostic method with other deep learning methods developed using radiology images.

Study	Type of Images	Number of Cases	Method Used	Accuracy (%)
Ioannis et al. [43]	Chest X-ray	224 COVID-19 (+) 700 Pneumonia 504 Healthy	VGG-19	93.48
Wang and Wong [42]	Chest X-ray	53 COVID-19 (+) 5526 COVID-19 (−) 8066 Healthy	COVID-Net	92.4
Sethy and Behra [45]	Chest X-ray	25 COVID-19 (+) 25 COVID-19 (−)	ResNet50+ SVM	95.38
Hemdan et al. [41]	Chest X-ray	25 COVID-19 (+) 25 Normal	COVIDX-Net	90.0
Narin et al. [44]	Chest X-ray	50 COVID-19 (+) 50 COVID-19 (−)	Deep CNN ResNet-50	98
Ying et al. [46]	Chest CT	777 COVID-19 (+) 708 Healthy	DRE-Net	86
Wang et al. [47]	Chest CT	195 COVID-19 (+) 258 COVID-19 (−)	M-Inception	82.9
Zheng et al. [48]	Chest CT	313 COVID-19 (+) 229 COVID-19 (−)	UNet+3D Deep Network	90.8
Xu et al. [49]	Chest CT	219 COVID-19 (+) 224 Viral pneumonia 175 Healthy	ResNet + Location Attention	86.7
Proposed Study	Chest X-ray	125 COVID-19 (+) 500 No-Findings	DarkCovidNet	98.08
		125 COVID-19 (+) 500 Pneumonia 500 No-Findings		87.02

system can be used in remote places in countries affected by COVID-19 to overcome a shortage of radiologists. Also, such models can be used to diagnose other chest-related diseases including tuberculosis and pneumonia. A limitation of the study is the use of a limited number of COVID-19 X-ray images. We intend to make our model more robust and accurate by using more such images from our local hospitals.

Declaration of competing interest

The authors declare no conflicts of interest.

References

- [1] F. Wu, S. Zhao, B. Yu, et al., A new coronavirus associated with human respiratory disease in China, *Nature* 579 (7798) (2020) 265–269.
- [2] C. Huang, Y. Wang, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet* 395 (10223) (2020) 497–506.
- [3] World Health Organization, Pneumonia of Unknown Cause—China. Emergencies Preparedness, Response, Disease Outbreak News, World Health Organization (WHO), 2020.
- [4] Z. Wu, J.M. McGoogan, Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention, *Jama* 323 (13) (2020) 1239–1242.
- [5] M.L. Holshue, C. DeBolt, et al., First case of 2019 novel coronavirus in the United States, *N. Engl. J. Med.* 328 (2020) 929–936.
- [6] W. Kong, P.P. Agarwal, Chest imaging appearance of COVID-19 infection, *Radiology: Cardiothoracic Imaging* 2 (1) (2020), e200028.
- [7] T. Singhal, A review of coronavirus disease-2019 (COVID-19), *Indian J. Pediatr.* 87 (2020) 281–286.
- [8] 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* (2020), <https://doi.org/10.1148/radiol.2020200490>. In press.
- [9] J.P. Kanne, B.P. Little, J.H. Chung, B.M. Elicker, L.H. Ketani, Essentials for radiologists on COVID-19: an update—radiology scientific expert panel, *Radiology* (2020), <https://doi.org/10.1148/radiol.2020200527>. In press.
- [10] 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* (2020), <https://doi.org/10.1148/radiol.2020200343>. In press.
- [11] E.Y. Lee, M.Y. Ng, P.L. Khong, COVID-19 pneumonia: what has CT taught us? *Lancet Infect. Dis.* 20 (4) (2020) 384–385.
- [12] A. Bernheim, X. Mei, et al., Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection, *Radiology* (2020), <https://doi.org/10.1148/radiol.2020200463>. In press.
- [13] F. Pan, T. Ye, et al., Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia, *Radiology* (2020), <https://doi.org/10.1148/radiol.2020200370>. In press.
- [14] C. Long, H. Xu, et al., Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? *Eur. J. Radiol.* 126 (2020) 108961.
- [15] H. Shi, X. Han, et al., Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study, *Lancet Infect. Dis.* 24 (4) (2020) 425–434.
- [16] W. Zhao, Z. Zhong, X. Xie, Q. Yu, J. Liu, Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study, *Am. J. Roentgenol.* 214 (5) (2020) 1072–1077.
- [17] Y. Li, L. Xia, Coronavirus Disease 2019 (COVID-19): role of chest CT in diagnosis and management, *Am. J. Roentgenol.* (2020) 1–7.
- [18] J.F.W. Chan, S. Yuan, 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* 395 (10223) (2020) 514–523.
- [19] S.H. Yoon, K.H. Lee, et al., Chest radiographic and CT findings of the 2019 novel coronavirus disease (COVID-19): analysis of nine patients treated in Korea, *Korean J. Radiol.* 21 (4) (2020) 494–500.
- [20] Edgar Lorente, COVID-19 pneumonia - evolution over a week. <https://radiopaedia.org/cases/COVID-19-pneumonia-evolution-over-a-week-1?lang=us>.
- [21] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, C. I. Sánchez, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
- [22] J. Ker, L. Wang, J. Rao, T. Lim, Deep learning applications in medical image analysis, *Ieee Access* 6 (2017) 9375–9389.
- [23] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, *Annu. Rev. Biomed. Eng.* 19 (2017) 221–248.
- [24] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, *Comput. Methods Progr. Biomed.* 161 (2018) 1–13.
- [25] F. Murat, O. Yildirim, M. Talo, U.B. Baloglu, Y. Demir, U.R. Acharya, Application of deep learning techniques for heartbeats detection using ECG signals-Analysis and Review, *Comput. Biol. Med.* 120 (2020) 103726.
- [26] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444.
- [27] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: *Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.
- [28] Ö. Yildirim, P. Plawiak, R.S. Tan, U.R. Acharya, Arrhythmia detection using deep convolutional neural network with long duration ECG signals, *Comput. Biol. Med.* 102 (2018) 411–420.
- [29] A.Y. Hannun, P. Rajpurkar, M. Haghighpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, *Nat. Med.* 25 (1) (2019) 65.
- [30] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, A. Gertych, R. San Tan, A deep convolutional neural network model to classify heartbeats, *Comput. Biol. Med.* 89 (2017) 389–396.
- [31] A. Esteva, B. Kuprel, R.A. Novoa, et al., Dermatologist-level classification of skin cancer with deep neural networks, *Nature* 542 (7639) (2017) 115–118, <https://doi.org/10.1038/nature21056>.
- [32] N.C. Codella, Q.B. Nguyen, S. Pankanti, D.A. Gutman, B. Helba, A.C. Halpern, J. R. Smith, Deep learning ensembles for melanoma recognition in dermoscopy images, *IBM J. Res. Dev.* 61 (4/5) (2017), 5–1.
- [33] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, U.R. Acharya, Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images, *Pattern Recogn. Lett.* 133 (2020) 232–239.
- [34] A. Cruz-Roa, A. Basavanahally, et al., March). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks, in: *Medical Imaging 2014: Digital Pathology*, vol. 9041, International Society for Optics and Photonics, 2014, p. 904103.
- [35] M. Talo, O. Yildirim, U.B. Baloglu, G. Aydin, U.R. Acharya, Convolutional neural networks for multi-class brain disease detection using MRI images, *Comput. Med. Imag. Graph.* 78 (2019) 101673.

- [36] P. Rajpurkar, J. Irvin, et al., Chexnet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, 2017 arXiv preprint arXiv:1711.05225.
- [37] J.H. Tan, H. Fujita, S. Sivaprasad, S.V. Bhandary, A.K. Rao, K.C. Chua, U. R. Acharya, Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network, *Inf. Sci.* 420 (2017) 66–76.
- [38] G. Gaál, B. Maga, A. Lukács, Attention U-Net Based Adversarial Architectures for Chest X-Ray Lung Segmentation, 2020 arXiv preprint arXiv:2003.10304.
- [39] J.C. Souza, J.O.B. Diniz, J.L. Ferreira, G.L.F. da Silva, A.C. Silva, A.C. de Paiva, An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks, *Comput. Methods Progr. Biomed.* 177 (2019) 285–296.
- [40] F. Caobelli, Artificial intelligence in medical imaging: game over for radiologists? *Eur. J. Radiol.* 126 (2020) 108940.
- [41] E.E.D. Hemdan, M.A. Shouman, M.E. Karar, COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images, 2020 arXiv preprint arXiv:2003.11055.
- [42] L. Wang, A. Wong, COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images, 2020 arXiv preprint arXiv:2003.09871.
- [43] Ioannis D. Apostolopoulos¹, Tzani Bessiana, COVID-19: Automatic Detection from X-Ray Images Utilizing Transfer Learning with Convolutional Neural Networks, arXiv:2003.11617.
- [44] A. Narin, C. Kaya, Z. Pamuk, Automatic Detection of Coronavirus Disease (COVID-19) Using X-Ray Images and Deep Convolutional Neural Networks, 2020 arXiv preprint arXiv:2003.10849.
- [45] P.K. Sethy, S.K. Behera, Detection of Coronavirus Disease (COVID-19) Based on Deep Features, 2020.
- [46] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, Y. Chong, Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, *medRxiv* (2020).
- [47] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, B. Xu, A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19), *medRxiv* (2020).
- [48] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, X. Wang, Deep learning-based detection for COVID-19 from chest CT using weak label, *medRxiv* (2020), <https://doi.org/10.1101/2020.03.12.20027185>.
- [49] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, et al., Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia, 2020 arXiv preprint arXiv:200209334.
- [50] M. Barstugan, U. Ozkaya, S. Ozturk, Coronavirus (COVID-19) Classification Using CT Images by Machine Learning Methods, 2020 arXiv preprint arXiv:2003.09424.
- [51] X. Chen, L. Yao, Y. Zhang, Residual Attention U-Net for Automated Multi-Class Segmentation of COVID-19 Chest CT Images, 2020 arXiv preprint arXiv:2004.05645.
- [52] L. Lan, D. Xu, G. Ye, C. Xia, S. Wang, Y. Li, H. Xu, Positive RT-PCR test results in patients recovered from COVID-19, *Jama* 323 (15) (2020) 1502–1503.
- [53] J.P. Cohen, COVID-19 Image Data Collection, 2020. <https://github.com/ieee8023/COVID-chestxray-dataset>.
- [54] Gallarato Gabriele, Demaria Paolo, Negri Alberto, Baralis Ilaria, Cerutti Andrea, Priotto Roberto, Violino Paolo, COVID-19: caso 56. <https://www.sirm.org/2020/03/21/COVID-19-caso-56/>.
- [55] L.T. Phan, T.V. Nguyen, Q.C. Luong, T.V. Nguyen, H.T. Nguyen, H.Q. Le, Q. D. Pham, Importation and human-to-human transmission of a novel coronavirus in Vietnam, *N. Engl. J. Med.* 382 (9) (2020) 872–874.
- [56] J. Lim, S. Jeon, H.Y. Shin, M.J. Kim, Y.M. Seong, W.J. Lee, S.J. Park, Case of the index patient who caused tertiary transmission of COVID-19 infection in Korea: the application of lopinavir/ritonavir for the treatment of COVID-19 infected pneumonia monitored by quantitative RT-PCR, *J. Kor. Med. Sci.* 35 (6) (2020).
- [57] S.C. Cheng, Y.C. Chang, Y.L.F. Chiang, Y.C. Chien, M. Cheng, C.H. Yang, Y.N. Hsu, First case of coronavirus disease 2019 (COVID-19) pneumonia in taiwan, *J. Formos. Med. Assoc.* (2020).
- [58] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2097–2106.
- [59] J. Redmon, A. Farhadi, Yolo9000: Better, Faster, Stronger, 2017 arXiv preprint.
- [60] 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.
- [61] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: visual explanations from deep networks via gradient-based localization, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 618–626.
- [62] H.X. Bai, B. Hsieh, et al., Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT, *Radiology* (2020), <https://doi.org/10.1148/radiol.2020200823>. In press.
- [63] Chest X-ray images (pneumonia). <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.