Chaos, Solitons and Fractals 142 (2021) 110495

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images

Emtiaz Hussain[a], Mahmudul Hasan[a], Md Anisur Rahman[b,*], Ickjai Lee[c], Tasmi Tamanna[d], Mohammad Zavid Parvez[a]

[a] Department of Computer Science and Engineering, Brac University, Dhaka, Bangladesh
[b] School of Computing & Mathematics, Charles Sturt University, Bathurst, NSW 2795, Australia
[c] Discipline of Information Technology, College of Science & Engineering, James Cook University, Cairns, QLD 4870, Australia
[d] Department of Immunology, Bangladesh University of Health Sciences, Dhaka, Bangladesh

ARTICLE INFO

Article history:
Received 7 October 2020
Accepted 18 November 2020
Available online 23 November 2020

Keywords:
COVID-19
Pneumonia-viral
Pneumonia-bacterial
Deep learning
Convolutional neural network
X-ray
Confusion matrix
Accuracy

ABSTRACT

Background and Objective: The Coronavirus 2019, or shortly COVID-19, is a viral disease that causes serious pneumonia and impacts our different body parts from mild to severe depending on patient's immune system. This infection was first reported in Wuhan city of China in December 2019, and afterward, it became a global pandemic spreading rapidly around the world. As the virus spreads through human to human contact, it has affected our lives in a devastating way, including the vigorous pressure on the public health system, the world economy, education sector, workplaces, and shopping malls. Preventing viral spreading requires early detection of positive cases and to treat infected patients as quickly as possible. The need for COVID-19 testing kits has increased, and many of the developing countries in the world are facing a shortage of testing kits as new cases are increasing day by day. In this situation, the recent research using radiology imaging (such as X-ray and CT scan) techniques can be proven helpful to detect COVID-19 as X-ray and CT scan images provide important information about the disease caused by COVID-19 virus. The latest data mining and machine learning techniques such as Convolutional Neural Network (CNN) can be applied along with X-ray and CT scan images of the lungs for the accurate and rapid detection of the disease, assisting in mitigating the problem of scarcity of testing kits.

Methods: Hence a novel CNN model called CoroDet for automatic detection of COVID-19 by using raw chest X-ray and CT scan images have been proposed in this study. CoroDet is developed to serve as an accurate diagnostics for 2 class classification (COVID and Normal), 3 class classification (COVID, Normal, and non-COVID pneumonia), and 4 class classification (COVID, Normal, non-COVID viral pneumonia, and non-COVID bacterial pneumonia).

Results: The performance of our proposed model was compared with ten existing techniques for COVID detection in terms of accuracy. A classification accuracy of 99.1% for 2 class classification, 94.2% for 3 class classification, and 91.2% for 4 class classification was produced by our proposed model, which is obviously better than the state-of-the-art-methods used for COVID-19 detection to the best of our knowledge. Moreover, the dataset with x-ray images that we prepared for the evaluation of our method is the largest datasets for COVID detection as far as our knowledge goes.

Conclusion: The experimental results of our proposed method CoroDet indicate the superiority of CoroDet over the existing state-of-the-art-methods. CoroDet may assist clinicians in making appropriate decisions for COVID-19 detection and may also mitigate the problem of scarcity of testing kits.

1. Introduction

COVID-19 pandemic started officially as soon as the pneumonia of unknown causes was reported in Wuhan city of Hubei province of China on December 31, 2019 [1–3]. The infectious disease was at first called SARS-CoV-2, and later it was named COVID-19 by World

* Corresponding author.
E-mail addresses: ehrupok@gmail.com (E. Hussain), hasanpiyal@gmail.com (M. Hasan), arahman@csu.edu.au, arahman@csu.edu.au (M.A. Rahman), ickjai.lee@jcu.edu.au (I. Lee), tasmitamanna02@gmail.com (T. Tamanna), zavid.parvez@bracu.ac.bd (M.Z. Parvez).

https://doi.org/10.1016/j.chaos.2020.110495
0960-0779/

Health Organization (WHO). This new virus took only 30 days to spread from Wuhan to other parts of China [4]. At the initial stage, the COVID outbreak was declared to be a Public Health Emergency of International Concern on 30 January, 2020 [5] and later on 11 March, 2020, it was declared as a Pandemic by WHO. There are total 14,892,577 cases of C0VID-19 where active cases are 89,45,979 killing 6,14,208 people worldwide till 21st July, 2020 [6].

The United States of America observed it's first seven cases on January 20, 2020 [7], which is now the most affected country by COVID-19 with total 39,63,376 cases and 1,43,889 deaths reported on 21st July 2020 [8]. Although this virus is commonly found among animals such as bats and cats, because of its zoonotic nature, the virus was transmitted from wild animals to the human body and later spread through human to human contact. Previously, two families of coronaviruses, namely severe acute respiratory syndrome virus (SARS-CoV) and the Middle East respiratory syndrome virus (MERS-CoV) caused severe respiratory diseases and deaths in humans [9]. COVID-19 genome has mutated quite a number of times in its lifetime till now.

A study from University College London (UCL) identified 198 recurring mutations to the virus [10]. A study from the University of Glasgow, which also analyzed mutations, said these changes did not amount to different strains of the virus, but they advised that only one type of the virus is currently circulating [11]. The most common symptoms of COVID-19 are fever, sneezing, cough, sore throat, throat swelling, headache, weakness, malaise, and breathlessness [12].

The mostly used technique for COVID-19 detection is a real-time reverse transcription-polymerase chain reaction (RT-PCR), even though immunological tests are also available in many countries. Radiological images of the chest, such as computed tomography (CT) scanning and X-ray images, have rendered important roles in the rapid diagnosis to early management of this disease [13]. As the RT-PCR has a low sensitivity of 60%–70%, and it is also a time-consuming method, detecting pathological effects of COVID-19 by examining images of lungs of patients can ensure early treatment [14,15]. It is to be noted that CT scanning is a more sensitive technique to detect pneumonia of COVID-19 and can be utilized as a screening tool along with RT-PRC [16]. Pathological changes of lungs CT scan can be observed for a long duration after the first appearance of the symptoms [17].

In a study, it was indicated that the survivors COVID-19 pneumonia, which is the most significant lung disease caused by this infection, started showing changes in their CT scan images ten days after the symptoms started [18]. The beginning of this pandemic, many countries, including China, are fighting insufficiency test kits as well as a scarcity of skilled laboratory technologists, resulting in a production of a high rate of false-negatives. As a result, doctors are encouraged to make decisions only based on X-ray and chest CT results [17,19]. CT and X-ray images are widely used for COVID-19 detection in countries with a shortage of testing kits. Researchers claim that combination of clinical image features and laboratory results may help in early detection of COVID-19 [9,16,20–22].

Some studies have shown that they noticed changes in chest X-ray and CT scan images even before the appearance of clinical features of COVID-19 [23]. Significant discoveries have been made in studies by investigating X-RAY and CT scan results. For example, Kong et al. [9] claimed about opacities at the right infrahilar space found in a COVID-19 patient. Zu et al. [13] and Chung et al. informed that 33% of chest CT scans have reported round lung opacities. Zhao et al. [21] discovered ground-glass opacities (GGO) or mixed GGO in most of the patients with consolidation and vascular dilation in the lesions of some patients. Furthermore, Li and Xia [22] reported GGO and consolidation, thickening of interlobular septa and air bronchogram sign, with or without increased broncho-vascular markings to be the common CT findings in COVID-19 patients. Moreover, Yoon et al. [24] presented that one in three patients under study had a single nodular opacity in the left lower lung region, the other two had four and five irregular opacities in both lungs.

Convolutional Neural Network (CNN) has become one of the most popular methods in Artificial intelligence (AI) in recent years. CNN has been successfully used in medical image analysis such as MRI [25], X-ray [26], CT scans [27], Ultrasonography [28] etc. CNN has also been very successful in natural language processing [29], computer vision [30], audio recognition [31] and speech recognition [32]. Furthermore, a neural network is a series of algorithms that recognize relationships in a set of data through a process that is very similar to human brain operation. This algorithm is very effective for pattern recognition and image processing. It takes images as input and builds a model that processes the images to extract the features from those images [33] and recognizes a pattern [34]. By using the pattern, CNN identifies the similarities of new inputs as accurately as possible. This algorithm is very popular because of its simple structure, adaptability, reduced training parameters, and low complexity of the network model.

COVID-19 detection using CNN has become a well-established research technique after it has become a global pandemic. We have found excellent CNN based research works using both X-ray images and CT scan images to detect and classify COVID-19. Though these CNN techniques have achieved extraordinary results, they are not an alternative to actual testing methods yet. These techniques are very helpful in association with actual testing methods, but there is huge room for research and improvement before commercial use. A large number of researchers and data scientists are putting their efforts into building highly accurate and reliable deep learning-based detection techniques for detecting COVID-19. Researchers are focusing on deep learning techniques to specify features from both X-ray and CT images of COVID-19 patients. Ioannis et al. [35] proposed CNN architecture VGG19 to classify COVID-19 positive, COVID-19 negative, Normal patients, and achieved an accuracy of 93.48% using X-ray images. In another study, Wang and Wong [36] proposed a model known as COVID-Net for similar 3 class classification, which achieved an accuracy of 92.4% in X-ray images. Sethy et al. [37], in their research work, used ResNet50+SVM to classify COVID-19 and Normal and achieved an accuracy of 95.38%. Hemdan et al. [38] proposed a model COVIDX-net for 2 class classification which achieved an accuracy of 90.0%.

In another research, Narin et al. [39] achieved an accuracy of 98% with their Deep CNN ResNet-50 model for 2 class classification. Ying et al. [40] proposed DRE-Net and classified COVID *vs.* non-COVID patients with an accuracy of 86% using CT scan images. Wang et al. [41] proposed a model M-Inception which achieved 82.9% accuracy for 2 class classifications in CT scan images. Zheng et al. [42] proposed a Unet+3D Deep Network, which achieved an accuracy of 90.8%. Their model was also performed on CT images. Xu et al. [43] proposed ResNet+Location Attention for 3 class classification (COVID-19, Pneumonia-viral, and Normal) which achieved an accuracy of 86.7% using CT images. Except for these works, there are some other works which were conducted based on CNN using CT images [44,45]. Ozturk et al. [46] used X-ray images and proposed a model known as DarkCovidNet, which achieved an accuracy of 98.08% for 2 class and 87.02% for 3 class classification. In their study, they proposed a 17 layer CNN model. Most recently, Khan et al. [47] proposed a CNN model Coronet, which is based on Xception, has been used for 2 class, 3 class, and 4 class classification using X-ray images. They achieved an accuracy of 99%, 95%, and 89.6%, respectively.

In this study, a 22-layer CNN model is proposed, which performs better than previously published CNN models in terms of 2 class, 3 class, and 4 class classification using X-ray images. Our pro-

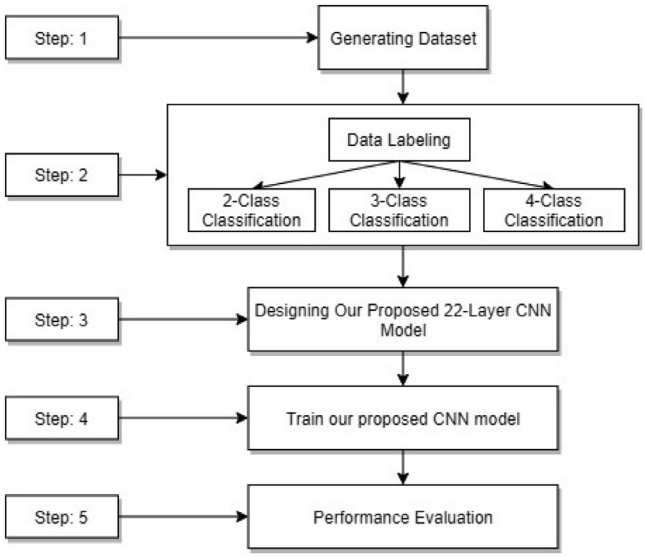

Fig. 1. A flowchart of our proposed CoroDet method.

posed CNN model is based on a 22-layer network, which consists of convolutional, max pooling, dense layer, flatten layer, and three activation functions, namely Sigmoid, ReLU, and Leaky ReLU. We have evaluated the performance of our proposed model in terms of accuracy, precision, recall, F1 score, specificity, sensitivity, and confusion matrix. Though it is not replaceable against the existing testing methods, it can be very helpful to assist in the identification of COVID-19. The main contributions of the paper are as follows:

1. A 22-layer CNN architecture, which has achieved an accuracy of 99.1% for 2 class classification, 94.2% for 3 class classification, and 91.2% for 4 class classification. To the best of our knowledge, the accuracy of our proposed CoroDet method is higher than the state-of-the-art-methods for COVID detection;
2. A demonstration of the same model for both 2 class, 3 class, and 4 class classification;
3. Construction of the largest X-ray image database for COVID-19 classification and experiments with the database.

The rest of the paper is organized as follows. In Section II, we have described our proposed model of work for the whole study. In Section III, we have presented the experimental results and discussion. Finally, we conclude our study with future plans in Section IV.

2. Corodet: our proposed method for COVID detection

This section discusses our proposed method called CoroDet, which consists of a new 22-layer CNN model for COVID detection using chest X-ray and CT images. The model performance is evaluated based on accuracy, precision, recall, F1 score, specificity, sensitivity, and confusion matrix for 2 class classification, 3 class classification, and 4 class classification.

Fig. 1 shows workflow of our proposed CoroDet method to demonstrate the workflow from the beginning to end. The details on every step of CoroDet is given below.

2.1. Generating dataset called COVID-R

In this study, we generated one of the biggest datasets for COVID-19 detection, using both X-ray and CT images. A brief introduction on the dataset is given in Table 1. From Table 1, we can see that the dataset has 2843 COVID-19 images. There are 3108 normal images and 1439 pneumonia (both viral and bacterial) images. Therefore, the total images in the COVID-R dataset are 7390.

Table 1
A brief introduction on the COVID-R dataset.

Class	Number of images
COVID-19	2843
Normal	3108
Pneumonia (Viral + Bacteria)	1439

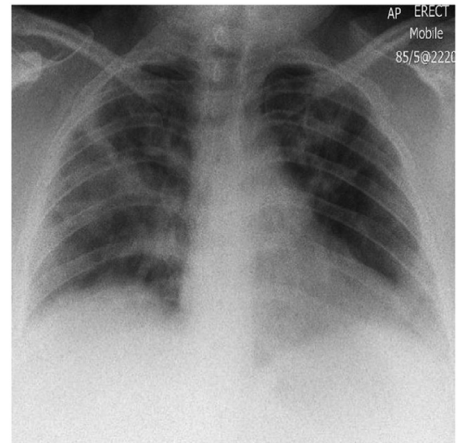

Fig. 2. COVID-19 sample.

Fig. 3. Normal sample.

Our COVID-R dataset is prepared by combining and modifying the eight different publicly available data repositories namely Covid-chestxray-dataset [48], COVID-CT [49], Figure1-COVID-chestxray-dataset [50], Actualmed-COVID-chestxray-dataset [51], SARS-COV-2 Ct-Scan Dataset [52], COVID-19 Xray Dataset (Train & Test Sets) with COVID-19 CNN Pneumonia Detector [53], COVID-19 Radiography Database [54], and COVID-19 Patients Lungs X Ray Images 10,000 [55].

The four sample images from the COVID-R dataset are shown in Fig. 2–5. From each class, we selected one image and added it to the figures. These images are discussed in the experimental result section. Fig. 2 shows the P-A view of the chest radiograph of a COVID patient showing hyperlucent lung fields that indicate lungs hyperinflation due to obstruction of small airways (caused by bronchiolitis or bronchitis). Multiple patchy opacities are seen in both lung fields representing consolidations caused by bacterial co-infection. Blunting of costo-phrenic angles is seen suggesting para-pneumonic pleural effusion. Fig. 3 shows the normal chest

Fig. 4. Pneumonia-viral sample.

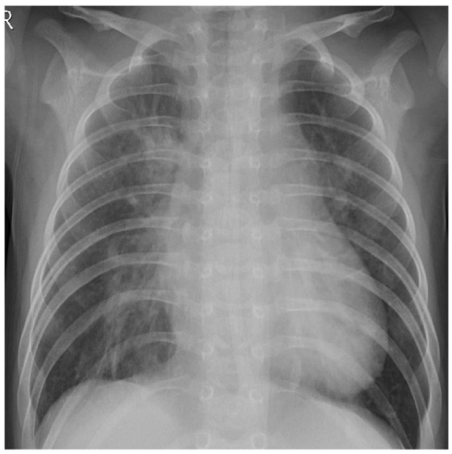

Fig. 5. Pneumonia-bacterial sample.

radiograph or P-A chest film showing the normal size and shape of the chest wall with normal position and appearance of the trachea, heart, and mediastinum. There is no focal opacity, air-fluid level, hyperlucency, or increased bronco-vascular markings. Costo-phrenic and cardio-phrenic angles are free as well. In Fig. 4, an X-ray image of non-COVID viral pneumonia is shown that shows diffuse bilateral Ground Glass Opacities (GGO), indicating diffuse alveolar damage and septal thickenings. Fig. 5 shows an X-ray image of bacterial pneumonia in which a focal opacity is indicating consolidation (without cavitation) on the upper lobe of the right lung is seen while costo-phrenic and cardio-phrenic angles are free with no air-fluid level.

2.2. Data labelling

After the dataset generation process, we prepare three subsets from the COVID-R dataset for 2 class classification, 3 class classification, and 4 class classification that we use to train our proposed CNN model. We considered different sets of images for 2 class classification, 3 class classification, and 4 class classification and then labeled the images accordingly. The tables from Tables 2–4 show the number of images that we used from the COVID-R dataset for 4 class classification, 3 class classification, and 2 class classification, respectively. Note that, due to hardware limitations, we choose a small number of images for training purposes. From the tables, we can notice that for both 2 class, 3 class, and 4 class classification, the number of COVID-19 images and normal images is 500 and 800, respectively. However, for 4 class classification, the number of images for Pneumonia-viral and Pneumonia-bacteria are both 400. For the 3 class classification, the number of Pneumonia-bacteria image is 800.

Table 2
The number of images used to train the model for 4 class classification.

Class	Number of images
COVID-19	500
Normal	800
Pneumonia-viral	400
Pneumonia-bacteria	400

Table 3
The number of images used to train the model for 3 class classification.

Class	Number of images
COVID-19	500
Normal	800
Pneumonia-bacteria	800

Table 4
The number of images used to train the model for 2 class classification.

Class	Number of images
COVID-19	500
Normal	800

2.3. Our proposed 22-layer CNN model

In this section, we discuss our proposed 22-layer CNN model for the detection and classification of COVID-19 using X-ray and CT scan images. Our proposed model is based on five basic components, namely convolutional layer, pooling layer, flatten layer, dense layer, and activation function. The components are used in different layers of our proposed model. A detailed discussion of each basic component is given below.

2.3.1. Convolutional layer

This is the layer where the convolution process happens, and the CNN model learns. This layer performs most of the computations for a CNN model. It is the most important component of a convolutional network. It has some parameters and hyperparameters, namely filters, kernels, and K. Convolution layers extract features using these filters and then learns from them. For this reason, this layer is also known as the feature extraction layer. The input images are compared segment by segment to distinguish the similarities and the differences among them. The segments are called features. The convolution layer extracts one or more features from the input images, and then using the image matrix, it creates a dot product and produces one or more matrix. If we have an image of size 5×5 and the pixel values of the images are either 0 or 1, and the size of the filter matrix is 3×3, then the 3×3 filter matrix will be multiplied by 5×5 image matrix, which is known as a feature map. The filter moves from left to right with a specific stride value until it parses the full width. Then it moves down to the beginning left side of the image with the same stride and repeats the process until the full image is traversed [56]. The main goal of this convolution layer is to extract high-level features like edges. Moreover, various operations are performed in this layer, including blur, sharpen, edge detection, by applying various filters.

If we have a volume of size $W1 \times H1 \times D1$ then we need four hyper parameters namely number of filters K, their spatial extent F, the stride S, and the amount of zero padding P [57]. Then an output of size $W2 \times H2 \times D2$ is produced by convolutional layer.

The equations for W2, H2, and D2 in convolutional layer are given below:

$$W2 = \frac{W1 - F + 2P}{S + 1}, \tag{1}$$

$$H2 = \frac{H1 - F + 2P}{S + 1}, \tag{2}$$

$$D2 = D1 = K. \tag{3}$$

Note that we have used nine conv2D layers in our proposed CoroDet model.

2.3.2. Pooling layer

The pooling layer is used repeatedly in convolution net to reduce the size of the volume when the image size is too large. This layer makes the computation fast, prevents from overfitting, and reduces memory. There are various kinds of pooling layers, such as max pooling, average pooling, and sum pooling. Max pooling takes the largest value from the feature map, average pooling takes the average by calculating for each patch of the feature map, and sum pooling takes the sum of all elements in the feature map. The most popular and common type of pooling layer is max pooling. The two hyperparameters are required for the pooling layer, namely filter (F) and stride (S).

If the volume of an input image is W1 × H1 × D1 then an output of size W2 × H2 × D2 is produced by pooling layer [57]. The equations for W2, H2, and D2 in pooling layer are given below:

$$W2 = \frac{W1 - F}{S + 1}, \tag{4}$$

$$H2 = \frac{H1 - F}{S + 1}, \tag{5}$$

$$D2 = D1. \tag{6}$$

Note that, for every Conv2D layer, we have used one Maxpool2D layer. Therefore, we have used nine MaxPool2D layers in our proposed model.

2.3.3. Flatten layer

In our model, after using the pooling layer, we used a flatten layer to flat the whole network. The flatten layer transforms the entire pooled feature map matrix into a single column. Then this is being passed to the neural network for further processing.

2.3.4. Dense layer

After the flatten layer, we have used two dense layers. The dense layer is also known as a fully connected layer. In this layer, the input from the past layers is flattened from a matrix into a vector. After flattening, the volume of the previous layer is given input to the fully connected layer like a neural network. By looking at the output of the previous layer, this layer decides which features mostly match a particular class. It works on high-level features that have specific weights. For this reason, a fully connected layer provides the correct probabilities for the different classes as it computes the products between the weights and the previous layer. By using the activation function, the outputs are classified.

2.3.5. Activation function

We have used Sigmoid function as shown in Eq. (7) with another dense layer and ReLU function as shown in Eq. (8). We have also used an activation function Leaky ReLU as shown in Eq. (9) which has proved to give best performance with Maxpooling2D. The three activation functions are shown as follows:

$$f(x) = \frac{1}{1 + \exp(-x)}, \tag{7}$$

Table 5
A summary on our proposed 22-layer model.

Layer	Output Shape	Parameter
conv2d_1 (Conv2D)	(64, 37, 37)	25,340
Max_pooling2d_1 (MaxPooling)	(64, 18, 18)	0
conv2d_2 (Conv2D)	(128, 19, 19)	19,696
Max_pooling2d_2 (MaxPooling)	(128, 9, 9)	0
conv2d_3 (Conv2D)	(64, 21, 21)	75,956
Max_pooling2d_3 (MaxPooling)	(64, 10, 10)	0
conv2d_4 (Conv2D)	(128, 21, 21)	296,168
Max_pooling2d_4 (MaxPooling)	(128, 10, 10)	0
conv2d_5 (Conv2D)	(256, 13, 13)	73,728
Max_pooling2d_5(MaxPooling)	(256, 6, 6)	0
conv2d_6 (Conv2D)	(128, 15, 15)	296,168
Max_pooling2d_6 (MaxPooling)	(128, 7, 7)	0
conv2d_7 (Conv2D)	(256, 17, 17)	73,728
Max_pooling2d_7 (MaxPooling)	(256, 8, 8)	0
conv2d_8 (Conv2D)	(128, 15, 15)	32,768
Max_pooling2d_8 (MaxPooling)	(128, 7, 7)	0
conv2d_9 (Conv2D)	(256, 15, 15)	296,168
Max_pooling2d_9 (MaxPooling)	(256, 7, 7)	0
Flatten_1 (Flatten)	4072	0
Dense_1(Dense)	512	1,683,376
Leaky_relu_1 (LeakyReLU)	512	0
Dense_1 (Dense)	1	513

$$ReLU = \max(0, x), \tag{8}$$

$$f(x) = 1(x < 0)(\alpha x) + 1(x >= 0)(x). \tag{9}$$

In the above equations, x is a real-valued number, and α is a small constant. The value of the Sigmoid function ranges from 0 to 1. The Sigmoid function is an S-shaped curve. It is comparatively easy, but it has a vanishing gradient problem. The output is not centered at zero; as a result, the gradient updates go too far in different directions, which makes the optimization harder. On the other hand, ReLU is an efficient and straightforward activation function as it avoids the vanishing gradient problem. Leaky ReLU has been used to solve the "dying ReLU" problem. Instead of the function being 0 where x <0, a Leaky ReLU will have a small negative slope.

2.3.6. Discussion on 22-layer CNN model

In Table 5 and Fig. 6, we present a summary on our proposed 22-layer CNN model. From Table 5 and Fig. 6, we can see that our proposed 22-layer CoroDet model has nine Conv2d layers and nine maxpool2d layers. We have used batch normalization (BatchNorm) to increase the stability of the model by using standardized input size. We have also used the Adam optimizer for weight updates, cross-entropy loss function, and selected learning rate. The most challenging part for us to build the model was to put all the layers, activation functions, and optimizer values together to achieve the best outcome from our proposed model in terms of accuracy and other evaluation criteria. We build the model through a rigorous testing process by adding layers, changing activation functions, and changing the value of the optimizers. We aim to find out what gives us the best possible outcome of the model. For example, at the beginning, we started with a simple 12-layer CNN model, which gave us an accuracy of 81% for 4-class classification.

Note that we used ReLu as an activation function with the pooling layer instead of LeakyReLu at the beginning. Later, we observed that LeakyRelu might give us a better output; therefore, we replaced ReLu with LeakyReLu in the pooling layer. After identifying which activation functions and optimizer values work the best, then we focused on adding more Conv2D as a convolutional layer to make the model deeper and to obtain a better outcome. We kept adding Conv2D layer and Maxpool2D layer with our basic 12-

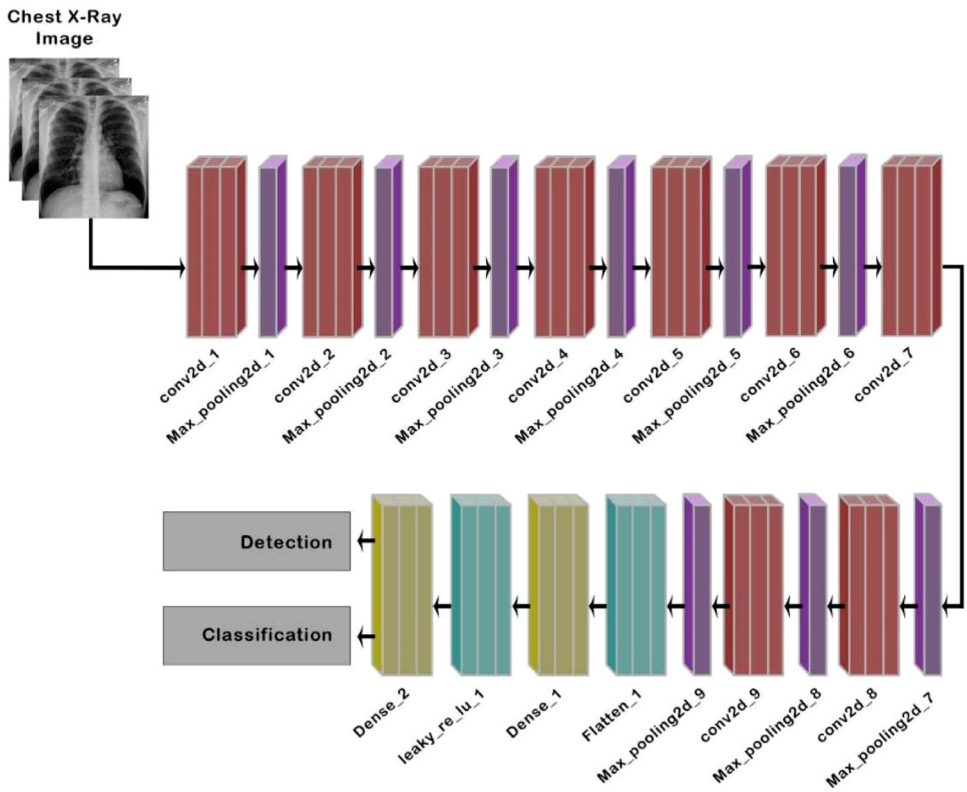

Fig. 6. Architecture of proposed 22 layer CNN model.

layer model and noticed the improvement of accuracy for 4-class, 3-class, and 2-class classification. An empirical justification for our 22-layer model is given in the experimental result section.

The combination of the layers, activation functions, and optimizer values made our model a unique model. We built the model by trying every possible combination of the layers, activation functions, and optimizer values and determine which combination works better and then finalized the model that can capture such small changes in features of the images. Our model is shown better than the previously published and predefined models that we discussed in the experimental result section, although most of the existing methods have deeper layers and more parameters than our proposed model.

2.4. Train our proposed CNN model

We trained our proposed model using the labeled data that we prepared in Step 2. A model on 4 class classification is built to identify chest X-ray images of four categories, namely COVID-19, Pneumonia-bacterial, Pneumonia-viral, and Normal. We then build two other models for 3 class (COVID-19, Pneumonia, and Normal) classification, and 2 class (COVID-19 and Normal) classification.

We implemented the proposed model using Keras and Tensorflow 2.0. ImageNet dataset was used to pre-train the model. The experiment was carried out on the Google Colaboratory Ubuntu server that has a Tesla K80 graphics card. We retrain the model using our prepared dataset and Adam optimizer. The value for the learning rate is 0.0001, the size of the batch is 10, and the value for the epoch is 50. Data shuffling was enabled for training. Note that data shuffling was performed before every epoch. The 5-fold cross-validation approach is used to assess the performance of 4 class, 3 class, and 2 class classification. We randomly divide the training set into 5 subsets, where each subset has the same number of records. 4 out of 5 subsets were considered to train the CNN model, and the other subset was considered for validation. This approach of validation is repeated 5 times by shifting the validation and training sets. The average result is calculated based on the result of each individual fold.

2.5. Performance evaluation

In order to analyze the performance of our proposed CNN model, we calculated precision, recall, F1 score, accuracy, sensitivity, specificity, and confusion matrix for each fold. At first, we computed these for 4 class 5 folds, and then the process is repeated for 2 and 3 class classification as well. The definitions of accuracy, F1 score, precision, and recall can be obtained in [58].

Tables 6 –8 show the performance of our proposed CNN model in terms of precision, recall, F1 score, accuracy, sensitivity, specificity, and confusion matrix. At first, the performance for each fold has been demonstrated, and then we computed the average of the 5 folds. The process is repeated for 4 class classification, 3 class classification, and 2 class classification.

In order to obtain the overall performance for each classification, we have also computed average class wise precision, recall, and F1 score as shown in Tables 9–11. All the results represent a very good performance of our proposed model for all the classes. In Fig. 7–9 we have plotted the confusion matrix for each fold to evaluate the model for each shuffle. For almost all the folds, the model has shown a very good performance for 4 class classification, 3 class classification, and 2 class classification.

In order to demonstrate the better performance of our proposed 22-layer CNN model, we have also plotted model accuracy in Fig. 10 and model loss in Fig. 11 for our 50 epochs for 4 class classification. These figures show that our proposed 22-layer CNN model is trained perfectly, and the model is not either under-fit or over-fit. This explains why we achieved better results in terms of accuracy, precision, recall, F1 score, specificity, sensitivity, and con-

Table 6
Evaluation based on 4 class classification for 5 fold.

	Sensitivity	Specificity	Precision	Recall	F1 score	Accuracy
Fold 1	94.7	95.8	96.5	93.6	92.4	92.4
Fold 2	90.1	91.4	92.1	93.2	91.2	91.6
Fold 3	93.2	95.3	93.2	92.2	90.1	93.5
Fold 4	89.6	91.6	91.2	90.4	89.2	90.3
Fold 5	91.2	93.3	87.2	90.1	87.3	88.2
Average	91.76	93.48	92.04	91.9	90.04	91.2

Table 7
Evaluation based on 3 class classification for 5 fold.

	Sensitivity	Specificity	Precision	Recall	F1 score	Accuracy
Fold 1	94.7	96.7	97.5	94.6	92.6	95.7
Fold 2	91.1	92.1	91.1	91.2	93.4	93.3
Fold 3	92.2	94.2	93.2	92.2	90.1	93.4
Fold 4	91.6	93.6	94.2	91.4	89.2	94.4
Fold 5	94.2	96.2	94.2	93.1	91.3	94.2
Average	92.76	94.56	94.04	92.5	91.32	94.2

Table 8
Evaluation based on 2 class classification for 5 fold.

	Sensitivity	Specificity	Precision	Recall	F1 score	Accuracy
Fold 1	96.7	98.7	97.5	96.6	97.4	99.1
Fold 2	95.1	97.1	98.1	97.2	95.4	99.5
Fold 3	97.2	99.2	97.2	93.2	97.1	98.8
Fold 4	96.6	99.6	96.2	94.4	96.2	98.6
Fold 5	91.2	92.2	99.2	95.1	98.3	99.6
Average	95.36	97.36	97.64	95.3	96.88	99.12

Table 9
Average class wise precision, recall and F1 score for 4 class classification.

Class	Precision	Recall	F1 score
COVID-19	94.27	96.17	97.51
Normal	96.25	95.15	95.35
Pneumonia Viral	87.85	86.85	87.85
Pneumonia Bacteria	86.15	84.27	87.95

Table 10
Average class wise precision, recall and F1 score for 3 class classification.

Class	Precision	Recall	F1 score
COVID-19	95.37	97.47	98.62
Normal	97.25	96.15	96.45
Pneumonia	88.85	87.85	88.95

Table 11
Average class wise precision, recall and F1 score for 2 class classification.

Class	Precision	Recall	F1-score
COVID-19	99.27	98.17	98.51
Normal	98.25	97.15	97.35

fusion matrix. These figures explain how accurately the model has been trained. We also obtained a similar accuracy and loss graph results for 3 class classification and 2 class classification.

3. Experimental results and discussion

This section includes the experimental results and discussion on the performance of our proposed method. We compare the accuracy of our proposed CoroDet method with previously published techniques for 4 class classification, 3 class classification, and 2 class classification for COVID-19 detection. We also provide an empirical justification for our proposed 22-layer model. Moreover, the outcomes of the CoroDet method is evaluated by a physician in order to demonstrate the performance of CoroDet more precisely.

Table 12
Accuracy comparison our proposed model *vs.* existing models.

Model	4 class	3 class	2 class	Number of Images
Ioannis et al. [35]	N/A	93.48%	N/A	224, 700, 504
Wang and Wong [36]	N/A	92.4%	N/A	53, 5526, 8066
Sethy and Behr [37]	N/A	N/A	95.38%	25, 25
Hemdan et al [38]	N/A	N/A	90%	25, 25
Narin et al [39]	N/A	N/A	98%	50, 50
Wang et al [41]	N/A	N/A	82.9%	195, 258
Zheng et al [42]	N/A	N/A	90.8%	313, 229
Xu et al [43]	N/A	86.7%	N/A	219, 224, 175
Ozturk et al [46]	N/A	87.02%	98.08%	125, 500, 500
Khan et al. [47]	89.6%	95%	99%	284, 327, 330, 310
CoroDet	91.2%	94.2%	99.1%	500, 400, 400, 800

From the discussion in Introduction, we can see that researchers from all over the world proposed a number of techniques for the detection of COVID-19. Some of the proposed models are for multi class classification, some of them are for 2 class classification, and some of them for both multi class and 2 class classification. Table 12 shows the accuracy comparison between our proposed CoroDet method and ten existing techniques for 2 class classification, 3 class classification, and 4 class classification. Note that, the models in Table 12 were trained using different number of images from different data sources. However, our proposed CoroDet model was trained by using the combined images from all the data sources.

The number of images used to train the models can be seen from the last column of Table 12. In that column, the first number indicates the number of COVID-19 images used to train a model. If there are two values in the number of images column then the first value indicates the number of COVID-19 images, and the second value indicates the number of normal images which are used

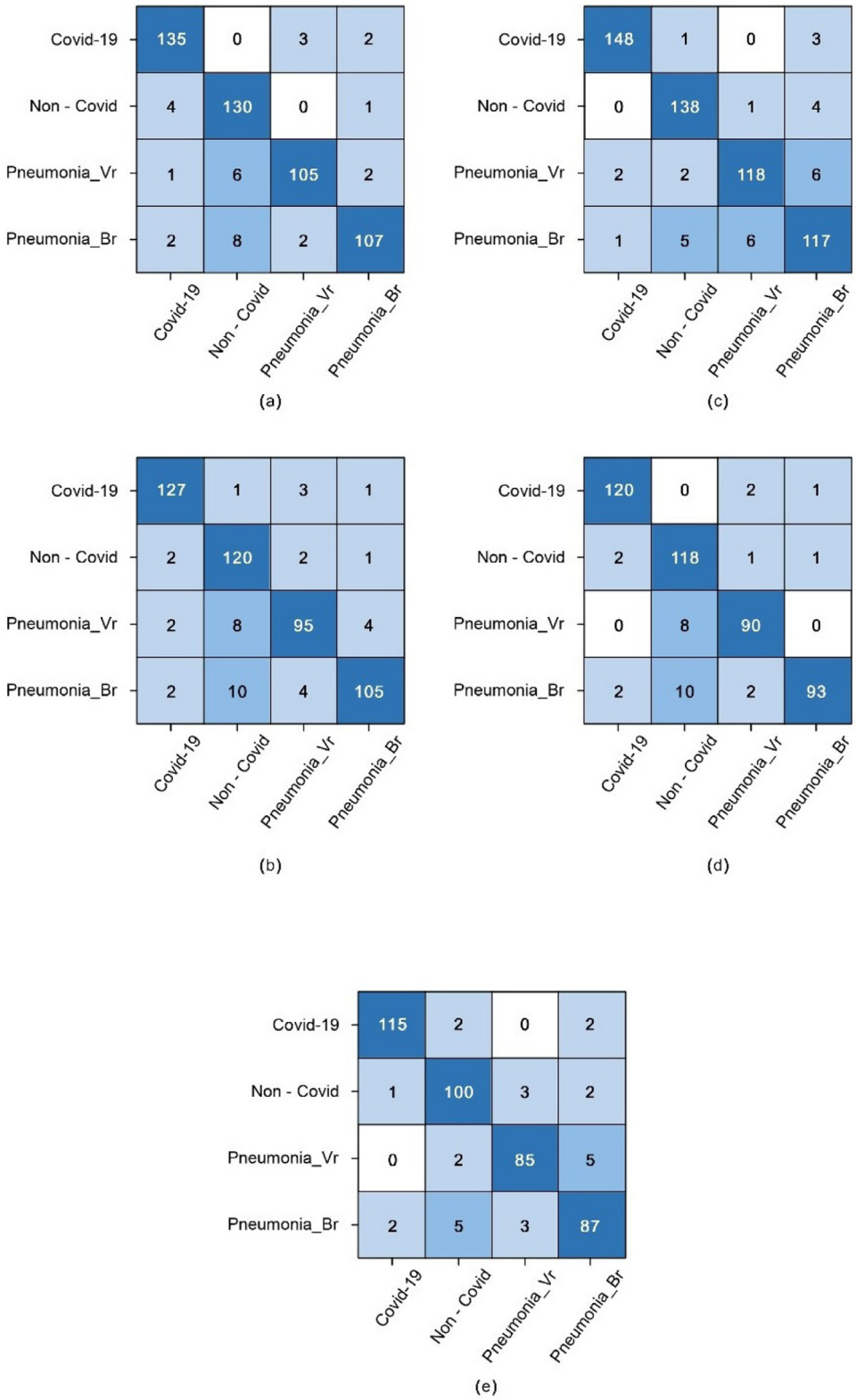

Fig. 7. Confusion matrix for each fold in 4 class classification.

to train the model. If there are three values in the number of images column then the first value indicates the number of COVID-19 images, the second value indicates the number of Pneumonia images, and the third value indicates the number of normal images which are used to train the model. If there are four values in the number of images column then the first value indicates the number of COVID-19 images, the second value indicates the number of Pneumonia-viral images, the third value indicates the number of Pneumonia-bacteria images, and the fourth value indicates the number of normal images which are used to train the model. For 2 class classification of Ozturk et al [46], 125 COVID-19 images and 500 normal images are used to train the model. For 2 class clas-

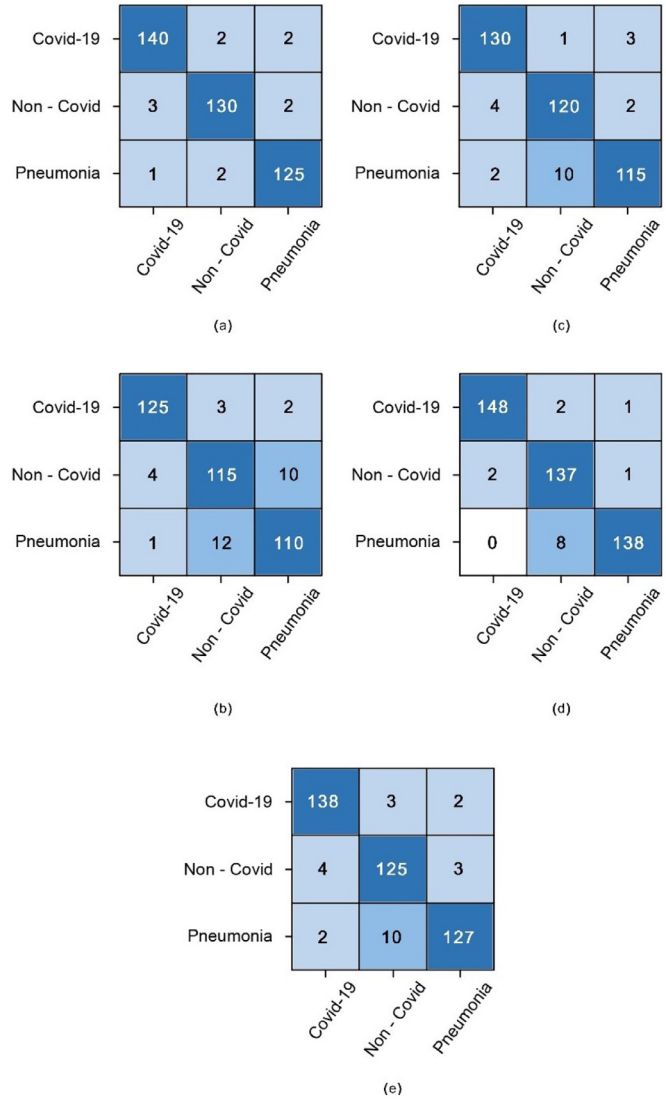

Fig. 8. Confusion matrix for each fold in 3 class classification.

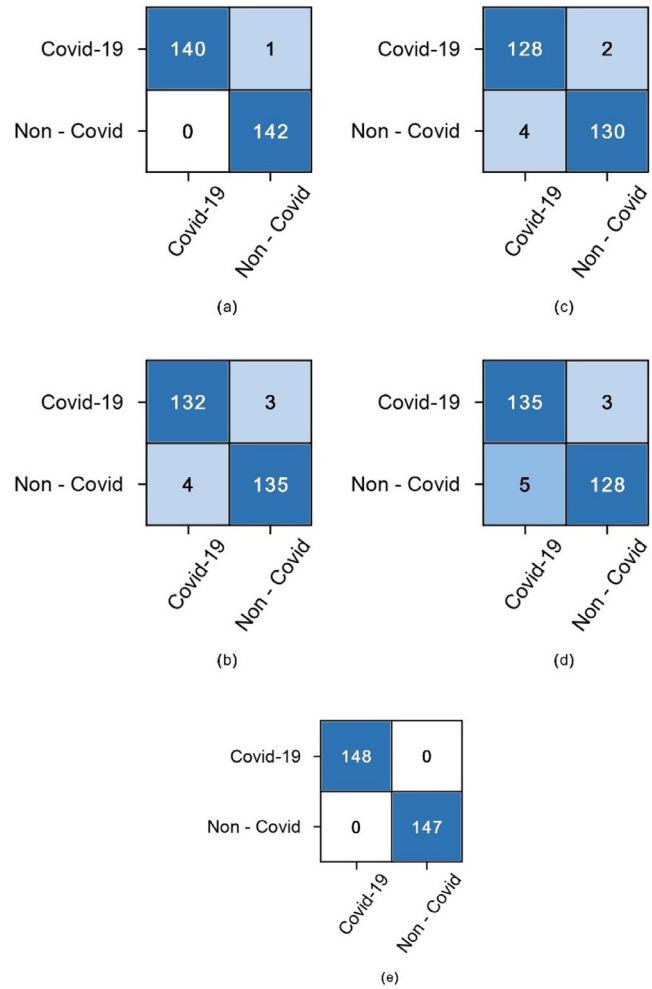

Fig. 9. Confusion matrix for each fold in 2 class classification.

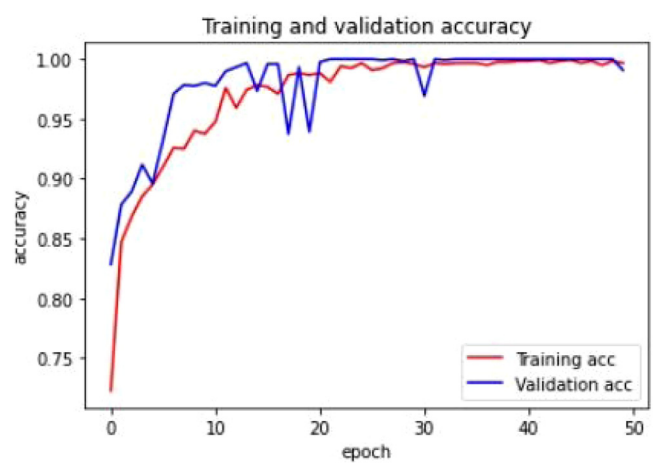

Fig. 10. Model accuracy graph for 4 class classification.

sification of Khan et al. [47], 284 COVID-19 images and 310 normal images are used to train the model. For 3 class classification of Khan et al. [47], 284 COVID-19 images, 657 Pneumonia images, and 310 normal images are used to train the model.

It is worth to mention that we have used the highest number (i.e. 500) of COVID-19 images to train our CoroDet model compared to the existing models. For 4 class classification, we used 500 COVID-19, 400 Pneumonia-viral, 400-Pneumonia-bactria, and 800 normal images to train our CoroDet model. For 3 class classification, we used 500 COVID-19, 800 Pneumonia, and 800 normal images to train our CoroDet model. For 2 class classification, we used 500 COVID-19, and 800 normal images to train our model. Table 12 illustrates that our proposed CoroDet method shows better accuracy than all the existing techniques for 2 class classification and 4 class classification. Khan et al. [47] achieved slightly higher accuracy than CoroDet for 3 class classification. However, for 3 class classification, the accuracy of CoroDet is better than the rest of the existing techniques. For 2 class classification, CoroDet achieved 99.1% accuracy which is the highest accuracy for 2 class classification to the best of our knowledge.

3.1. Empirical justification on our proposed 22-layer model

In this section, we have provided some empirical analysis on our 22-layer proposed model where the number of layers in the model is not more than or less than 22. We prepare 12-layer, 16-layer, 18-layer, 20-layer, and 22-layer models for 2 class, 3 class, and 4 class classifications. Note that, for every model, we use the same parameters and factors such as activation function and op-

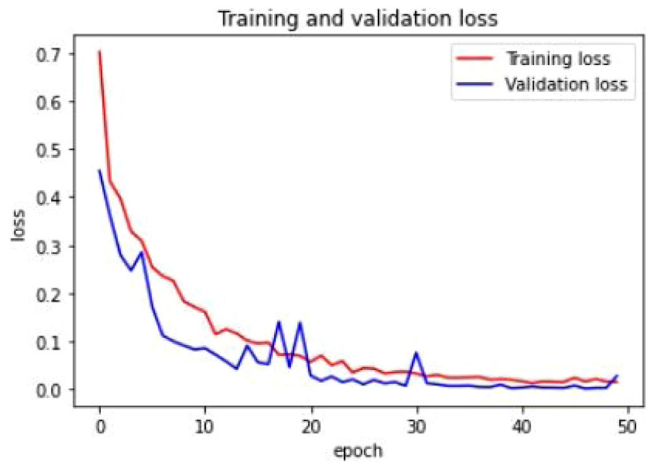

Fig. 11. Model loss graph for 4 class classification.

Table 13
A justification on our proposed 22-layer model.

Layer	2 class	3 class	4 class
12-layer	81.5%	85.6%	87.5%
16-layer	83.8%	88.8%	91.3%
18-layer	86.5%	91.5%	94.1%
20-layer	89.4%	93.2%	96.5%
22-layer	91.2%	94.2%	99.1%

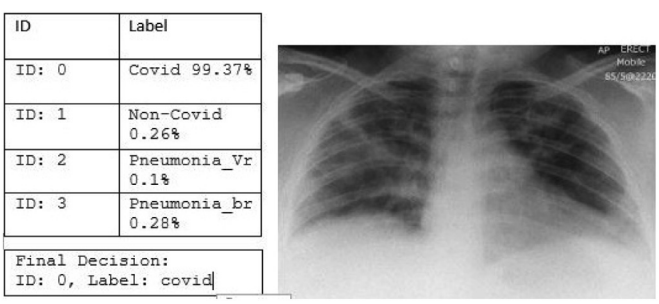

Fig. 12. An example on COVID detection using a randomly selected test image.

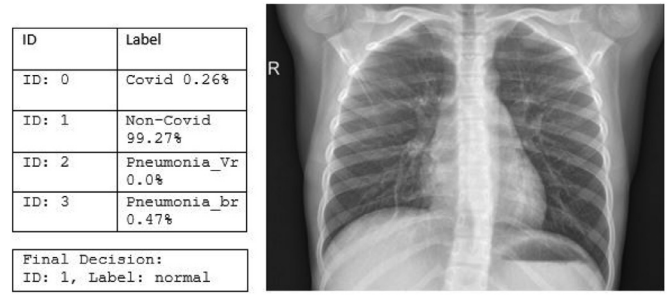

Fig. 13. An example of Non-COVID detection using a randomly selected test image.

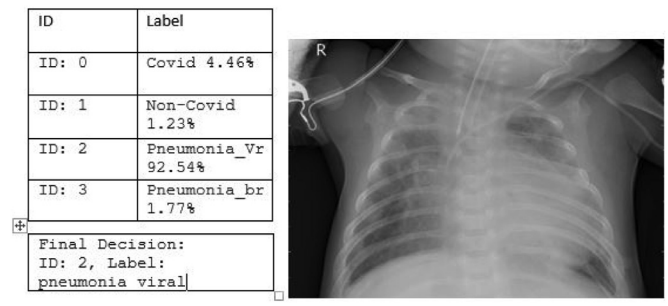

Fig. 14. An example of pneumonia-viral detection using a randomly selected test image.

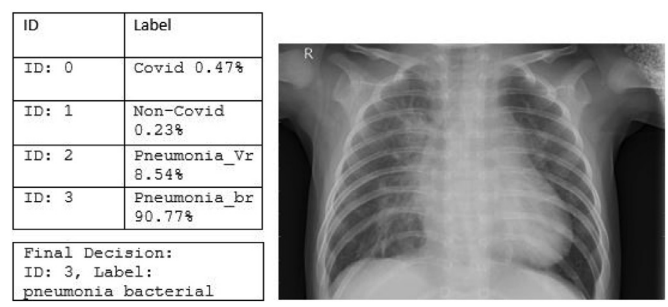

Fig. 15. An example on pneumonia-bacterial detection using a randomly selected test image.

timizer values. Table 13 shows accuracy for 12-layer, 16-layer, 18-layer, 20-layer, and 22-layer for 2 class, 3 class, and 4 class classifications. From Table 13, we can see that accuracy for the 22-layer model is better than others.

Note that, after 22-layer, we stopped adding more Conv2D layer and Maxpool2D layers as the model was not giving us better accuracy. The accuracy became constant after our 22-layer model for all the classes. It is worth mentioning that, if we reduce the layers from our 22-layer model, then the accuracy of the model decreases through all the other factors of the model are the same. That is why we decided to use 22 layers in our proposed model.

3.2. Clinical evaluation of CoroDet

Our proposed CoroDet was prepared with the impression that it may help clinicians and assist them in making appropriate decisions for diagnosis. To demonstrate the performance of CoroDet to the clinicians, we randomly selected some X-ray images from the COVID-R dataset and gave them input to the model and obtained output on the images as shown in Figs. 12–15.

The images are explained by a clinician. Fig. 12 shows lung hyperinflation, which is evident by increased transparency of the lungs field and reduced broncho-vascular markings than the normal lungs. It is caused by obstructed air in the lungs, which may be due to small airway obstruction in the case of bronchitis and/or bronchiolitis associated with COVID-19 infection. Hyperinflated lungs are larger than the normal lungs, while the gap between the ribs is also raised. Fig. 13 shows normal translucency with normal broncho-vascular markings and cardiothoracic ratio with free cardio-phrenic and costo-phrenic angles. The trachea is centrally placed, and stomach gas is visible under the left diaphragm. So this figure represents a normal chest radiogram.

Fig. 14 shows increased opacity distributed bilaterally, giving ground glass appearance, which is the feature commonly found in viral pneumonia. Fig. 15 shows a round space area lower segment of the right upper lobe indicating consolidation with increased broncho-vascular markings on the same side. This image directs the diagnosis of bacterial pneumonia. Note that the images from Fig. 12 to Fig. 15 are correctly detected by the CoroDet model that shows the usability of the model. The clinician makes the following comments on the outcome of the model:

- Opacities in the images became more obvious and clearly demarcated after performing this model especially consolidation of bacterial pneumonia and scattered opacities of COVID-19 be-

came easier to notice. However, no significant change was observed in the case of GGO of viral pneumonia and normal chest radiogram. This is probably because of haziness throughout the lung fields in case of GGO has nothing to show more markedly, which is also true for a normal chest X-ray;
- The model can detect COVID-19, Normal, Pneumonia-Viral, and Pneumonia-Bacterial separately and put a label of which one it is based on the percentage of the detection;
- Model performed poorly for the low-quality X-ray images.

Though we achieved extraordinary results compared to other existing techniques, in the opinion of the physician, it is not ready yet to replace the traditional testing techniques for COVID-19 detection. However, it can certainly be very helpful to assist the clinicians to take proper decisions on COVID-19 detection. In the future, we intend our model to be more effective and accurate. We hope to train our model with a large number of images. We also want to try different feature extraction techniques to evaluate for which feature extraction our model gives the most accurate result.

4. Conclusion

COVID-19 has caused a serious negative impact on our day-to-day lives, including public health systems to the global economy. In this paper, we presented a new CNN model to detect COVID-19 using chest X-ray and CT images. The proposed model called CoroDet is capable of providing accurate diagnostics for 2 class classification (COVID and Normal), 3 class classification (COVID, Normal, and Pneumonia), and 4 class classification (COVID, Normal, non-COVID viral pneumonia, and non-COVID bacterial pneumonia). The classification accuracy of the proposed model is 99.1% for 2 classes, 94.2% for 3 class cases, and 91.2% for 4 classes classification, which is the highest achieved accuracy to the best of our knowledge on the datasets used in the experiments. Another contribution in the paper is the largest dataset preparation for the evaluation of classification algorithms. The performance of CoroDet is shown to be better than ten existing techniques in terms of accuracy. The performance of our proposed method indicates the superiority of this method over the existing state-of-the-art-methods. An empirical justification for our 22-layer model is presented. The outcome of the model was explained and accepted by a clinician. Our future goal is to overcome hardware limitations allowing us to use larger image sets to train our proposed CoroDet method and compare the performance of CoroDet with a greater number of existing methods. It is assumed that the use of further images in the training stage may help us to improve the performance of the model in the future.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. Nature 2020;579(7798):265–9.
[2] Huang C, Wang Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan. China Lancet 2020;395(10223):497–506.
[3] World Health Organization. Pneumonia of unknown cause–china. emergencies preparedness. Response, Disease Outbreak News, World Health Organization (WHO); 2020.
[4] Wu Z, Googan JMM. 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 2020;323(13):1239–42.
[5] WHO updates on COVID-19. Apr 3; 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen
[6] COVID-19 CORONAVIRUS PANDEMIC. [Online]. Available: https://www.worldometers.info/coronavirus/.
[7] Holshue ML, De Bolt C, et al. First case of 2019 novel coronavirus in the united states. N Engl J Med vol 2020;328:929–36.
[8] Countries where COVID-19 has spread. [Online]. Available:https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/.
[9] Kong W, Agarwal PP. Chest imaging appearance of COVID-19 infection. Radiology 2020;2(1).
[10] Dorpa L, Acmana M, Richard D, et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. Infect Genet Evol 2020;83:104351.
[11] Lean OM, Orton R, Singer J, Robertson D. No evidence for distinct types in the evolution of SARS-CoV-2. Virus Evol 2020;6(1).
[12] Singhal T. A review of coronavirus disease-2019 (COVID-19). Indian J Pediatr 2020;87:281–6.
[13] Zu ZY, Jiang MD, Xu PP, Chen W, Ni QQ, Lu GM, Zhang LJ. Coronavirus disease 2019 (COVID-19): a perspective from China. Radiology 2020.
[14] Kanne JP, Little BP, Chung JH, Elicker BM, Ketai LH. Essentials for physicians on COVID-19: an update radiology scientific expert panel. Radiology 2020.
[15] Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. Radiology 2020.
[16] Lee EY, Ng MY, Khong PL. COVID-19 Pneumonia: what has CT taught us? Lancet Infect Dis 2020;20(4):384–5.
[17] Bernheim A, Mei X, et al. Chest CT findings in coronavirus disease-19 (COVID19): relationship to duration of infection. Radiology 2020.
[18] Pan F, Ye T, et al. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. Radiology 2020.
[19] Long C, Xu H, et al. Diagnosis of the coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol 2020;126:108961.
[20] Shi H, Han X, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. Lancet Infect Dis 2020;24(4):425–34.
[21] Zhao W, Zhong Z, Xie X, Yu Q, Liu J. Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. Am J Roentgenol 2020;214(5):1072–7.
[22] Li X, Xia L. Coronavirus disease 2019 (COVID-19): role of chest CT in diagnosis and management. Am J Roentgenol 2020;214(6):1–7.
[23] Chan JFW, Yuan S, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 2020;395(1022):514–23.
[24] Yoon SH, Lee KH, Radiol J, et al. Chest radiographic and CT findings of the 2019 novel coronavirus disease (COVID-19): analysis of nine patients treated in Korea. Korean 2020;21(4):494–500.
[25] Zou L, Zheng J, Miao C, Mckeown MJ, Wang ZJ. 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. IEEE Access 2017;5:23626–36.
[26] Liu C, et al. TX-CNN: Detecting tuberculosis in chest x-ray images using convolutional neural network. In: IEEE International conference on image processing (ICIP), Beijing; 2017. p. 2314–18.
[27] Zhao X, Liu L, Qi S, Teng Y, Li J, Qian W. Agile convolutional neural network for pulmonary nodule classification using CT images. Int J CARS 2018;13:585–95.
[28] Liu, Li W, Zhao N, et al. Integrate domain knowledge in training CNN for ultrasonography breast cancer diagnosis Medical image computing and computer assisted intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science, vol 11071. Cham: Springer; 2018.
[29] Alayba AM, Palade V, England M, Iqbal R. A combined CNN and LSTM model for arabic sentiment analysis. International cross-domain conference for machine learning and knowledge extraction. CD-MAKE Lecture Notes in Computer Science, vol 11015. Cham: Springer; 2018.
[30] Peng X, Schmid C. Multi-region two-stream r-CNN for action detection. Computer vision – ECCV 2016. ECCV, Lecture Notes in Computer Science, vol 9908. Cham: Springer; 2016.
[31] Zhang S, Zhang S, Huang T, Gao W, Tian Q. Learning affective features with a hybrid deep model for audio–visual emotion recognition. IEEE Trans Circuits SystVideo Technol 2018;28(10):3030–43.
[32] Abdel-Hamid O, Mohamed A, Jiang H, Penn G. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. In: 2012 IEEE International conference on acoustics. In: Speech and Signal Processing (ICASSP), Kyoto; 2012. p. 4277–80.
[33] Chen Y, Jiang H, Li C, Jia X, Ghamisi P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans. Geosci. Remote Sens. 2016;54(10):6232–51.
[34] Chen L, Wang S, Fan W, Sun J, Naoi S. Beyond human recognition: a CNN-based framework for handwritten character recognition. In: in 2015 3rd IAPR Asian conference on pattern recognition (ACPR), Kuala Lumpur; 2015. p. 695–9.
[35] Apostolopoulos I.D., Mpesiana T.A.. COVID-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. arXiv: 2003116172020.
[36] Wang L., Wong A.. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images. arXiv preprint arXiv:2003098712020;.
[37] Sethy PK, Behera SK. Detection of coronavirus disease (COVID-19) based on deep features. Preprints; 2020. doi:1020944/preprints2020030300v1. 2020030300
[38] Hemdan E.E.D., Shouman M.A., Karar M.E.. COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in x-ray images. arXiv preprint arXiv: 2003110552020;.

[39] Narin A., Kaya C., Pamuk Z.. Automatic detection of coronavirus disease (COVID19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003108492020;.
[40] Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang Z, Chong Y. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. medRxiv 2020.
[41] Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, Xu B. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). medRxiv 2020.
[42] Zheng C, Deng X, Fu Q, Zhou Q, Feng J, Ma H, Wang X. Deep learning-based detection for COVID-19 from chest CT using weak label. medRxiv 2020.
[43] Xu X., Jiang X., Ma C., Du P., Li X., Lv S., et al. Deep learning system to screen coronavirus disease 2019 pneumonia. arXiv preprint arXiv:2002093342020;.
[44] Barstugan M., Ozkaya U., Ozturk S.. Coronavirus (COVID-19) classification using CT images by machine learning methods. arXiv preprint arXiv:2003094242020;.
[45] Chen X., Yao L., Zhang Y.. Residual attention U-Net for automated multi-class segmentation of COVID-19 chest CT images. arXiv preprint arXiv:2004056452020;.
[46] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR. Automated detection of COVID-19 cases using deep neural networks with x-ray images. Comput Biol Med 2020;121:103792.
[47] Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs-Biomed 2020;196:105581.
[48] Cohen JP. Covid-chestxray-dataset. Apr 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset
[49] UCSD-AI4H. COVID-CT. Apr 2020. [Online]. Available: https://github.com/UCSD-AI4H/COVID-CT
[50] Agchung. Figure1-COVID-chestxray-dataset. Apr 2020. [Online]. Available: https://github.com/agchung/Figure1-COVID-chestxray-dataset
[51] Agchung. Actualmed-COVID-chestxray-dataset. May 2020. [Online]. Available: https://github.com/agchung/Actualmed-COVID-chestxray-dataset
[52] Eduardo P. SARS-COV-2 CT-scan datase. June 2020. [Online]. Available: https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset
[53] Khoong WH. COVID-19 x-ray dataset (train & test sets) with COVID-19CNN pneumonia detector. Apr 2020. [Online]. Available: https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets
[54] Rahman T. COVID-19 Radiography database. May 2020. [Online]. Available: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
[55] Sajid N. COVID-19 Patients lungs x ray images 10000. May 2020. [Online]. Available: https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images
[56] Sharma S. Activation functions in neural networks. Sep 2017. [Online]. Available: https://towardsdatascience.com/activation-functions-neuralnetworks-1cbd9f8d91d6
[57] Convolutional neural networks for visual recognition. [Online]. Available: http://cs231n.github.io/convolutional-networks/.
[58] Nicholson C. Evaluation metrics for machine learning - accuracy, precision, recall, and F1 defined; 2019. [Online]. Available: https://pathmind.com/wiki/accuracy-precision-recall-f1