



Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Deep learning model for distinguishing novel coronavirus from other chest related infections in X-ray images

Fareed Ahmad^{a,b,*}, Muhammad Usman Ghani Khan^a, Kashif Javed^c

^a Department of Computer Science, University of Engineering and Technology, Lahore, Pakistan

^b Quality Operations Laboratory, Institute of Microbiology, University of Veterinary and Animal Sciences, Lahore, Pakistan

^c Department of Electrical Engineering, University of Engineering and Technology, Lahore, Pakistan

ARTICLE INFO

Keywords:

Ensemble learning
Data augmentation
Feature fusion
Deep learning models
Transfer learning
Novel coronavirus classification

ABSTRACT

Novel Coronavirus is deadly for humans and animals. The ease of its dispersion, coupled with its tremendous capability for ailment and death in infected people, makes it a risk to society. The chest X-ray is conventional but hard to interpret radiographic test for initial diagnosis of coronavirus from other related infections. It bears a considerable amount of information on physiological and anatomical features. To extract relevant information from it can occasionally become challenging even for a professional radiologist. In this regard, deep-learning models can help in swift, accurate and reliable outcomes. Existing datasets are small and suffer from the balance issue. In this paper, we prepare a relatively larger and well-balanced dataset as compared to the available datasets. Furthermore, we analyze deep learning models, namely, AlexNet, SqueezeNet, DenseNet201, MobileNetV2 and InceptionV3 with numerous variations such as training the models from scratch, fine-tuning without pre-trained weights, fine-tuning along with updating pre-trained weights of all layers, and fine-tuning with pre-trained weights along with applying augmentation. Our results show that fine-tuning with augmentation generates best results in pre-trained models. Finally, we have made architectural adjustments in MobileNetV2 and InceptionV3 models to learn more intricate features, which are then merged in our proposed ensemble model. The performance of our model is statistically analyzed against other models using four different performance metrics with paired two-sided *t*-test on 5 different splits of training and test sets of our dataset. We find that it is statistically better than its competing methods for the four metrics. Thus, the computer-aided classification based on the proposed model can assist radiologists in identifying coronavirus from other related infections in chest X-rays with higher accuracy. This can help in a reliable and speedy diagnosis, thereby saving valuable lives and mitigating the adverse impact on the socioeconomics of our community.

1. Introduction

Zoonotic diseases are caused by microbes that are transmitted from animals to humans [1]. The scope, scale and worldwide effect of zoonoses pose a risk not only to animals and humans but also to global health safety [2]. Approximately 1500 pathogens are recognized to induce infections in individuals [3]. Among these infections, 61% of the recognized and 75% of the emerging infectious diseases in humans are zoonotic in origin [2,4]. It is approximated that infectious illnesses cause nearly 16% of all the deaths and 44% of the deaths in low-income countries [2]. Annually, zoonotic illnesses cause 2.7 million mortalities and 2.5 billion illnesses in individuals [5]. Emerging zoonotic infections are responsible for various notable and disastrous epidemics [6].

Coronavirus is of zoonotic origin [7]. A study proposes that the virus is transferred to humans either by means of pangolins or bats [8]. No direct relationship between humans and other species is reported. However, the virus is an extremely mutated microorganism that can comfortably cross the species barrier [9,10]. It can affect the cells of human airways. As a result, it can induce pneumonia and critical respiratory diseases, kidney malfunction and can even cause death [11]. The pathogen can persist in the air and on various surfaces from several hours to multiple days [12]. Health executives have observed that the pathogen spreads through coughed and sneezed droplets of 5–10 μm and are around 30 times tinier than a human hair [12,13]. Many scientists also believe that virus can persist in aerosol form [14,15]. World health organization (WHO) is currently assessing the role of

* Corresponding author. Department of Computer Science, University of Engineering and Technology, Lahore, Pakistan.

E-mail addresses: fareed.ahmad@uvas.edu.pk (F. Ahmad), usman.ghani@uet.edu.pk (M.U. Ghani Khan), kashif.javed@uet.edu.pk (K. Javed).

<https://doi.org/10.1016/j.combiomed.2021.104401>

Received 6 December 2020; Received in revised form 6 April 2021; Accepted 11 April 2021

Available online 21 April 2021

0010-4825/© 2021 Elsevier Ltd. All rights reserved.

aerosol transmission [16]. Recently, WHO said that it cannot be ruled out in closed, crowded and poor ventilated settings [17]. Thrilling analysis foretells that sneezing can cause the particles of the virus to move up to 27 feet [18]. Its ease of dispersion, tremendous capability for ailment and death rate make it a likely candidate for a bio-warfare [19–21].

In USA, UK and Europe, the epidemic is extensively prevalent and millions of individuals are infected and thousands have died because of the infection. As of February 18, 2021, approximately 27.82 million cases are recorded and 490,717 people lost their lives in the United States [22]. The country's topmost contagious disease specialist predicts that without proper safety measures, the lethal pathogen might kill up to 2.4 million individuals [23]. The UK and the European states such as France, Italy, Germany and Spain recorded that about 15.88 million individuals became affected and 0.34 million lost their lives [22]. The British health executives believe that the virus can infect 80% of the population and 0.5 million individuals can lose their lives [24]. All over the globe, billions of people are either living their lives under lockdowns or in self-quarantine. If the virus is allowed to continue its route, the healthcare system will be overburdened, economics will crumble and millions of valuable lives will be lost [23,25].

The most commonly used techniques for the identification of coronavirus are enzyme-linked immunosorbent assay (ELISA) and Reverse transcription-polymerase chain reaction (RT-PCR) [26]. The primary screening method for recognizing COVID-19 is RT-PCR, which identifies the pathogen's RNA from lower respiratory tract specimens. These specimens are collected in several ways, such as nasopharyngeal or oropharyngeal swabs. While RT-PCR is regarded as a golden standard [27,28], it is a time consuming, complicated and sensitive manual approach [29]. The test results can be affected by low sensitivity, sampling errors and low viral load [30,31]. In the case of inadequate viral load, the results of the test can be falsely negative [28,32]. An alternative method used for COVID-19 screening employs chest radiographic images (i.e. X-rays or CT-Scans). Radiologists analyze these images to assess the apparent symptoms related to COVID-19. Initial studies revealed that patients show variations in chest radiographic images indicative of those concerned with SARS-CoV-2 viral infection [28,33]. Other studies suggest that radiographic analyses should be used as a principal mechanism for the screening of the virus in afflicted regions [28]. The radiography-based tests can be performed quickly and are readily available in our healthcare system. It gives them a real compliment to PCR testing (in some cases, even showing higher sensitivity) [28,32]. A study based on 1014 cases recorded a 97% sensitivity of chest radiographic images for the diagnosis of COVID-19, while the average time interval within initial negative and positive RT-PCR was roughly five days [28]. The major issue with radiographic analysis is that they require an expert radiologist to assess resulting medical images as the visual indicators can be difficult to interpret [34,35]. Though, computer-aided designs can help radiologists to instantly and accurately interpret radiographic images to recognize disorders that cause COVID-19 and other chest-related infections. In current times, deep learning models can facilitate in early diagnosis, quick prevention and remedy of the infections caused by the novel virus [36].

A combination of convolutional neural networks (CNN) (e.g. InceptionV3 [37] and AlexNet [38]), large-scale databases (e.g. ImageNet [39]) and an effective overfitting checking method (e.g. dropout [40]) have shown an improved prediction accuracy, performance and outstanding generalization ability to solve medical [41,42], biological [36,43] and complex computer vision tasks [44]. The advantage of deep networks for image classification is that a model is trained end-to-end but suffers from overfitting on modest datasets. It is a challenging task for deep networks to obtain the same good performance on small datasets as they can obtain on large datasets [45]. A solution to this problem is implementing pre-trained deep designs on a small dataset [46]. This technique is known as transfer learning which replaces the last few layers of the pre-trained model and fine-tunes it on a given

dataset. The technique can be very useful if augmentation along with proper hyper-parameters are set and efficient fine-tuning strategies are applied. Another technique that is also being effectively applied on image classification tasks is ensemble learning. It fuses the features from various deep models into a classifier of high-grade quality thus, achieving better and reliable predictive performance. Image processing and deep learning models in bio-medical image processing and analysis have yielded promising results particularly in the field of chest radiology. The techniques are generally used to conduct lung nodule classification [47], pulmonary tuberculosis identification [48] and especially for novel coronavirus classification in radiological images [35,49–52]. Next, a summary of pneumonia and COVID-19 related works from the current literature is presented., ,

Ayan and Ünver [53] proposed a quick diagnostic technique for pneumonia based on chest X-ray images using Xception and VGG-16. Their work shows that Xception generates better results in identifying pneumonia and VGG-16 performs good for healthy cases. Varshni et al. [54] apply six deep models (Xception, VGG-16, ResNet 50, DenseNet-169 and DenseNet-121) for feature retrieval on X-ray images and then use numerous machine learning classifiers (SVM, K-nearest neighbors, Naive Bayes, and Random Forest) for classification. The results are further improved by optimizing the hyper-parameters. Stephen et al. [55] proposed a CNN model trained from scratch to classify and detect pneumonia in chest X-rays images [56]. The design retrieves relevant features from the images and uses them to classify the infection. The model achieved an accuracy of 93.73% on a meager dataset with the help of augmentation, hyper-parameter tuning and fine-tuning. In Ref. [57], the authors propose a compressed sensing based deep model for automatic classification of pneumonia using chest radiology images to help the medical professionals in early diagnosis of the disease. The dataset consists of 5863 X-ray images from Kaggle. Wide-ranging simulation outcomes have shown that the proposed technique predicts the presence of pneumonia with 97.34% accuracy. Chouhan et al. [58] devised an ensemble method that combines the outputs from different models to classify pneumonia in chest X-ray images using transfer learning. The ensemble model attained an accuracy of 96.4% with a recall of 99.62% on the test data. The outcomes of the convolution models on different abnormalities based on the freely available OpenI dataset [59] show that the same deep CNN model does not give acceptable results for all types of abnormalities [60]. Yet, ensemble model considerably increases the classification accuracy as compared to a particular deep learning model.

Gozes et al. [50] presented a swift AI development cycle using a CNN-based analysis of CT-scan images. Xu et al. [51] proposed a three-class model that can distinguish among normal, viral and COVID-19 cases. The segmentation-based approach attains an accuracy of 86.7% on a dataset with 618 CT-scan images. Wang et al. [52] suggested a deep design to retrieve graphical features from CT-scan images for COVID-19 classification. The design processes a dataset consisting of 1065 CT-scan images (325 COVID-19 images and 740 viral pneumonia images) of patients. The model obtains a test accuracy of 79.3%. Ioannis et al. [61] presented a transfer learning technique for classifying various chest-related infections on a dataset with 504 normal, 700 bacterial pneumonia and only 224 COVID-19 X-ray images. Their technique's specificity, sensitivity and accuracy were 96.46%, 98.66% and 96.78% respectively. Although various models were compared but results were based on a small number of COVID-19 instances. In Ref. [35], the authors provide a combined public dataset and proposed a deep model called COVID-Net for the detection of COVID-19. The architecture relies on a tailored CNN model, which uses the chest X-rays as inputs. The dataset consists of 358 COVID-19, 5538 pneumonia, 8066 normal images. COVID-Net achieves an accuracy of 93.3% but COVID-19 class images were limited. Ucar and Korkmaz [62] train a pre-trained deep network SqueezeNet with the help of Bayesian optimization technique to identify COVID-19 re-lated infections in chest X-ray images. The dataset consists of 66 COVID-19, 1349 normal and 3895 pneumonia

X-ray images. Although the model shows promising results on the small dataset but its performance on a large COVID-19 related dataset was not checked. Abbas et al. [63] exhibit a CNN model that performs dimensionality reduction to transform a high dimensional feature space into a lower one. The dataset comprises 80 normal, 11 SARS and 105 COVID-19 X-ray images. The model attains an accuracy of 95.12%. Khan et al. [64] employ the Xception model to automatically detect coronavirus infections in chest X-ray images. The dataset consists of 310 normal, 330 bacterial pneumonia, 327 viral pneumonia and 284 COVID-19 images. On this small dataset, the model achieves an accuracy of 89.6% while achieves an accuracy of 95% by combining the two pneumonia classes. Ashfar et al. [65] proposed a capsule network known as COVID-CAPS. The network achieves an accuracy of 95.7% on 94,323 X-ray images related to common thorax disorders. Gian-chandani et al. [66] apply two deep ensemble models, one comprising of VGG16 and ResNet152V2 for binary (COVID-19, Normal) classification and another consisting of VGG16, DenseNet201 for three-class (COVID-19, bacterial, and viral) classification. For the binary model 1525 images are used for each class and for three-class models only 401 images are used for each class. Although results are good but may not reproduce for a larger dataset with many classes. A simple train, validation, test split is used, although cross-validation is mostly recommended for limited datasets. Singh et al. [67] applies genetic and particle swarm optimization based CNN's for screening of COVID patients. These genetic models apply a simple binary classification approach, they are slow to converge but still deliver good results. However, training, validation, and test curves are widely separated, which reflect overfitting. Hyperparameter tuning is also applied to further improve the model's performance. The author [68] also applies ensemble learning using three CNN models (VGG16, DenseNet201, and ResNet152V2) on CT-Scan images to generate excellent results on a larger four class (COVID-19, pneumonia, tuberculosis, and normal) dataset. A simple train test split is applied instead of applying cross-validation.

Most of the above mentioned approaches for the identification of pneumonia apply machine and deep learning methods along with fine-tuning, augmentation, and ensemble learning. These approaches have produced good results and are being used to recognize coronavirus related infections in chest radiology. Most of the above mentioned COVID-19 classification techniques were tested on small datasets to produce promising results, but there is no guarantee that these models would generate similar results on large datasets. Furthermore, most of those datasets were very unbalanced and most of the models did not employ hyper-parameter tuning, cross-validation, and statistical analysis. Table 1 provides a summary of the comparison of the research works related to CNN models.

In this work, we prepared a relatively larger and well-balanced

dataset consisting of X-ray images of patients with bacterial, viral and novel coronavirus infections. We also included X-ray images of healthy individuals. Fig. 1 shows some of the samples. These images are analyzed by applying transfer learning along with fine-tuning on pre-trained models to explain whether these pre-trained networks can provide better results when data is scarce. Apart from implementing fine-tuning and transfer learning, we also applied traditional data augmentation methods like rotation and reflection [69], which resolves the scarcity issue of training data by enriching it with transformed original examples. Finally, we propose a deep ensemble learning model composed of MobileNetV2 and InceptionV3 models as shown in Fig. 2. It produces far better results than currently available fine-tuned pre-trained architectures. The devised model obtains state-of-the-art performance on image dataset associated to various chest X-ray infections.

The contribution of our work can be summarized as follows.

- (i) We prepared a large and well-balanced dataset with four classes (viral, bacterial, COVID-19 and normal). Each class consists of 1000 images.
- (ii) Our work integrates fine-tuning, augmentation, transfer learning and hyper-parameter tuning into one model.
- (iii) We propose a deep design by merging the features of MobileNetV2 and InceptionV3 models using an ensemble approach. We have made architectural adjustments to deep models by adding three dense layers at the end of each model to learn more intricate features before merging their features through the addition layer. We also added a dropout layer after the addition layer to handle overfitting.
- (iv) We compare our ensemble model with state-of-the-art deep models in terms of four performance metrics namely specificity, recall, F-score and accuracy.
- (v) We also compared our ensemble model with another ensemble approach presented in the literature, to show that our model is far superior to the previous model.
- (vi) To statistically validate the performance of the proposed design, we employ 5-fold cross-validation with the paired two-sided *t*-test.

2. Materials and methods

A deep learning framework for feature fusion of deep models using ensemble learning is presented in Fig. 3. The proposed framework consists of seven main steps: (i) Image resizing of large-scale images (ii) Dataset splitting (iii) Data augmentation (iv) Fine-tuning of deep models (v) Merging best performing models using ensemble learning (vi) Adding additional dense and dropout layers and finally performing

Table 1
A comparison of different CNN models for COVID-19 Classification.

Approach	No. Of Classes	Dataset Details	Aug-mentation?	Dataset Balanced?	Transfer Learning?	5-fold CV?	Ensemble Model?	Statistical Comparison?	Tuning Hyper parameters?
[30]	5	N 191, B 54, T 57, V 20, C 180		×	×	×	×	×	
[35]	3	N 8851, P 6012, C 180		×	×	×	×	×	×
[62]	3	N 1349, P 3895 C 66		×	×	×	×	×	
[65]	5	Not specified	×	×	×	×	×	×	×
[70]	3	3 N 8066, P 5521 C 183		×	×	×	×	×	×
[71]	3	N 8851, P9579C99		×	×	×	×	×	×
[72]	4	N 7595, B 2780 C 313, UP 6012	×	×	×	×	×	×	
[73]	2	N 500, C 184		×	×	×	×	×	×
[74]	4	N 1203, B 931 V 660, C 68		×	×	×	×	×	×
[75]	3	N 1579, V 1485 C 423		×	×		×	×	×
[66]	3	N 401, V 401 C 401		×		×		×	

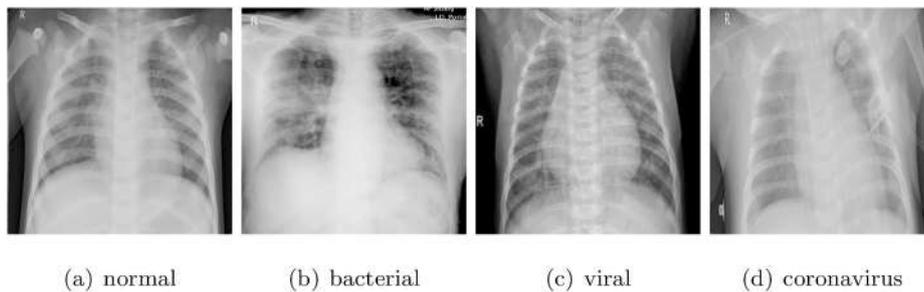


Fig. 1. Example posteroanterior chest radiograph images of our dataset.

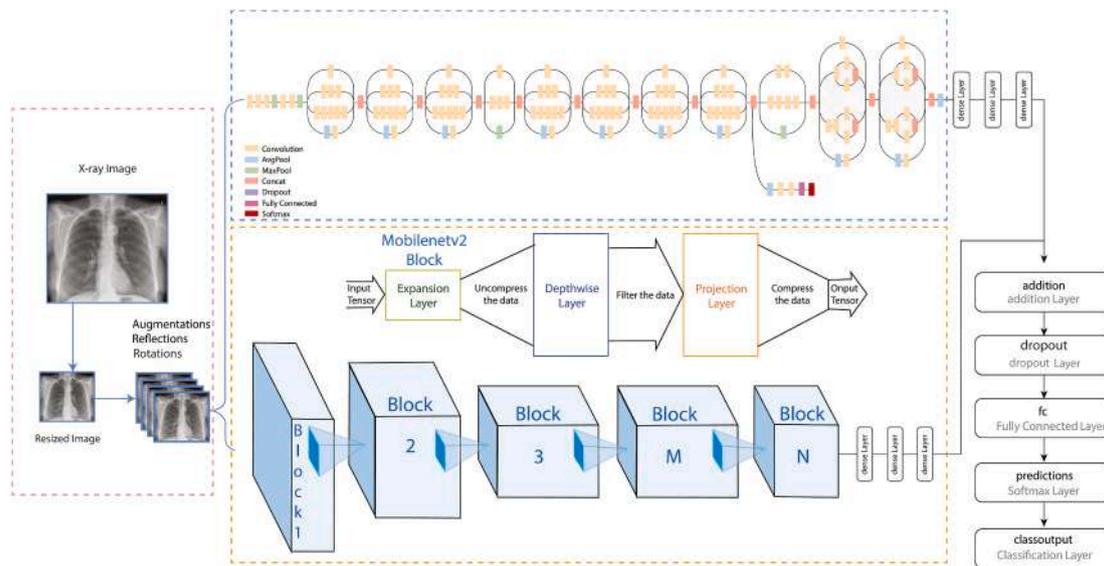


Fig. 2. Various phases of our proposed method.

classification on the proposed ensemble model.

The classes column gives the number of classes in a dataset. In the dataset details column, N, B, V, T, P, UP, C represent the number of Normal, Bacterial, Viral, Tuberculosis, Pneumonia, Pneumonia of Unknown type, Covid-19 instances respectively. Augmentation? column describes whether the research work uses data augmentation. Dataset Balanced? shows whether the dataset is balanced. Transfer Learning? means whether transfer learning was employed. Similarly the 5- fold CV? column indicates whether cross validation was used. The Ensemble Model? column shows whether the approach uses ensemble learning. The Statistical Comparison? column shows whether the approach performs statistical comparison with other deep models. The last column Tuning Hyper parameters? shows whether the approach tunes hyper parameters during model training.

Algorithm 1 describes the pseudo code of our proposed deep ensemble method. Initially, X-ray images, which are provided as input are resized according to the required image scale. Then, these rescaled images are split into training and test sets. We apply augmentation to these images to generate an augmented dataset. A list of different deep learning models is retrieved. These deep learning models are loaded, trained, fine-tuned and evaluated one by one. Each fine-tuned model is appended to Model-List one by one. Finally, the best two models are merged and some additional dense and dropout layers are added to

further improve the ensemble model. We perform classification to obtain the class labels of the bacterial, viral, COVID-19 images as output.

2.1. Data description and augmentation

Our dataset consists of X-ray images belonging to four categories namely bacterial, viral, COVID-19 and normal. There are 4000 images in the dataset and each category has 1000 images. X-ray images of bacterial, viral and normal categories were collected from the Kaggle dataset [76]. As far as the images of COVID-19 are concerned, 900 images were gathered from Mendeley dataset [77] and the remaining hundred images were collected from two open-source repositories namely (i) Italian Society of Medical and Interventional Radiology (SIRM) [78] and (ii) Radiopaedia [79].

Augmentation is employed to the proposed model to enhance the size of the training data, overcome over-fitting and devise a more generalized model. Different augmentation techniques like random rotation, random vertical/horizontal reflection and random vertical/horizontal shear are employed to generate an augmented dataset. During the experimentation, 80% of the dataset is used for training a model and the remaining 20% of the data is used for testing purposes.

Algorithm 1: Our proposed deep ensemble model

Data: Images of Chest X-rays

Result: Classified X-ray images of Bacterial, Viral, COVID-19 infections

Image \leftarrow *Read* – *image*

Resized – *images* \leftarrow *Resized* – *image*(*images*, *image* – *scale*)

train, *test* \leftarrow *Split*(*Resized* – *images*)

Augmented – *images* \leftarrow *Augmentation*(*train*)

Models \leftarrow *DeepLearningModels*[]

for *each Model* in *Models* **do**

Load – *model* \leftarrow *load*(*Model*)

Train – *model* \leftarrow *train* – *model*(*Loaded* – *model*, *train*, *test*)

Tuned – *model* \leftarrow *parameter* – *tuning*(*Trained* – *model*)

Model – *List.append*(*Resultant* – *model*)

model1, *model2* \leftarrow *max*(*Model* – *List*)

model1, *model2* \leftarrow *Remove* – *Last* – *Layers*(*model1*, *model2*)

model1 \leftarrow *Add* – *DenseLayers*(*model1*)

model2 \leftarrow *Add* – *DenseLayers*(*model2*)

Merged – *model* \leftarrow *Merge*(*model1*, *model2*)

Merged – *model* \leftarrow *Add* – *DropoutLayer*(*Merged* – *model*)

Classified – *images* \leftarrow *Classification*(*Merged* – *model*, *labels*)

return Class of Bacterial, Viral, COVID-19 Image;

2.2. Pre-processing

The chest X-ray images in the dataset are massive in size, with few images with a dimensions of 1007×1024 pixels. These images are resized to a standard size 299×299 pixels. But we had to further tailor these resized images to meet the requirements of deep learning models. Pre-trained models like MobileNetV2 and InceptionV3 have an input size of 224-by-224 and 299-by-299 respectively. The images were finally resized and provided to the given deep model. Some of the chest X-ray images were cropped, which helped in eliminating the noise produced by the extended black background.

2.3. Hyper-parameters tuning

In our work, we tried to estimate the impact of various hyper-parameters on the performance of the selected deep models. These hyper-parameters include the number of epochs, learning rate and batch size. We have investigated five deep models namely AlexNet, SqueezeNet, DenseNet201, MobileNetV2 and InceptionV3. For each of these models, we have fine tuned the parameters. The different values of the learning rate that we tried are $\{9e-4, 9e-5, 1e-5, 9e-6, 7e-6, 5e-6, 3e-6, 1e-6, 8e-7, 6e-7, 4e-7, 1e-7, 1e-8, 5e-10\}$. For the batch size, the values are $\{24, 36, 32, 48\}$. The values of number of epochs are $\{10, 30, 40, 50, 85, 100, 135, 150, 200, 300, 400, 500, 600, 700, 800, 900, 950\}$.

We found that mostly the deep models displayed the best results with the batch size values of $\{48, 32\}$. For the initial epochs, learning rate of $\{1e-8, 1e-7, 9e-6\}$ were very fruitful and gradually we increased the learning rate. The details of various hyper-parameters used during the training of different deep models are presented in Table 2. To maintain a stable utilization of the GPU resource, we have identified the value twenty-four as the smallest batch size.

2.4. Overview of convolutional neural networks (CNNs)

The CNNs rely on the foundation of conventional neural networks and usually consist of a convolutional layer, a pooling layer and a fully connected layer. After the convolution operation, we generally perform pooling to reduce the dimensionality. This enables us to decrease the number of parameters that decreases the training time. It also helps in overcoming the overfitting problem. The pooling operation down samples each feature map individually, reducing the height and width while keeping the significant features intact. After these layers, various fully connected layers try to acquire mid-level feature maps. Executing full connection in these layers needs a large number of weight parameters. For further details, we refer the readers to Ref. [80].

The training of CNNs initiates in a feed-forward manner as it begins from the initial layer to the final output layer. Then, the error propagates in reverse order as it starts from the final layer to the convolution layer. Let p be a neuron in layer a , which accepts input from a neuron q of layer $a - 1$ in the forward pass. The input [81] is calculated as below.

$$In_p^a = \sum_{q=1}^n W_{pq}^a x_q + b_p \quad (1)$$

where b_p and W_{pq}^a are the bias term and weight vector of the a th layer respectively. A nonlinear function such as rectified linear activation function (ReLU) [81] can be used to compute the output, which is given below.

$$Out_p^a = \text{Max}(0, In_p^a) \quad (2)$$

All the nodes in the convolution and fully connected layers use Equations (1) and (2) to compute the input and output. The pooling layer utilizes a $K \times K$ square window sliding on the $N \times N$ features map, and takes the average or maximum value of the features inside the window. It therefore, reduces the spatial dimension of the feature map

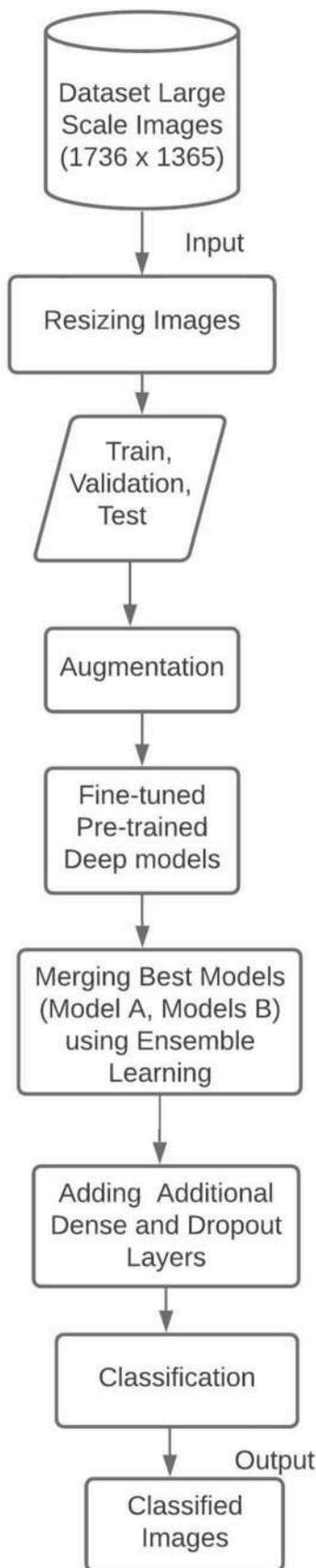


Fig. 3. Flowchart of various steps of our proposed method.

from $N \times N$ to $N/K \times N/K$ as it generates a single value for the $K \times K$ region. We use the SoftMax function [82] to calculate the classification probability of every pathogen in the final layer as given below.

$$Out_p^i = \frac{e^{In_p^i}}{\sum_k e^{Out_k^i}} \quad (3)$$

Back propagation procedure is used to trained the CNN's. It can be used to minimize the cost function [82] related to the unknown weight W , which is given below.

$$C = -\frac{1}{m} \sum_{n=1}^m \ln(p(y^n|X^n)) \quad (4)$$

where m represents the number of instances in a training set, X^n is the n th instance in the training set and y^n is its corresponding label and its true classification probability is $p(y^n|X^n)$. Using stochastic gradient descent (SGD) over the mini-batches of magnitude N , the cost function is reduced and training costs are estimated by the mini batch cost. If W^s denotes the weights at iteration s for the convolution layer a and C is the cost of mini-batch, we will then update the weights [82] at the subsequent iteration as given below.

$$\begin{cases} \gamma^s = \gamma \left[\frac{sN}{m} \right] \\ V_a^{s+1} = \mu V_a^s + \gamma^s \alpha_a \frac{\partial C}{\partial W_a} \\ W_a^{s+1} = W_a^s + V_a^{s+1} \end{cases} \quad (5)$$

where α_a represents the learning rate of layer a , γ is the rate of scheduling, which diminishes the original learning rate α at the end of the specific number of epochs and μ denotes the momentum, which portrays the impact of previously updated weights in the recent iteration.

2.5. Transfer learning and fine-tuning deep learning models

During the training process, the weights of various layers of pre-trained models are updated after every iteration, as given by Equation (5). There are 314 layers in total and 25 million training parameters in the InceptionV3 with a depth of 48 layers. On the other hand, in MobileNetV2, there are 155 layers and 3.5 million parameters with a depth of 53 layers. Different deep models have different number of training parameters and depths as shown in Table 3.

For the training and optimization of these deep models, an abundant amount of data is a pre-requisite, as it serves as a fuel for these models. However, for a modest dataset, it is pretty challenging to acquire the appropriate local minimum for the cost function and the model will undergo overfitting. Originally pre-trained weights are applied for the MobileNetV2 and InceptionV3 models. These models are fine-tuned on our dataset by applying different values of the learning rate, batch size and the number of epochs. The initial layers in these models possess general features and the succeeding layers contain domain specific features. To preserve the features from the initial layers and decrease the pace of learning in the remaining transferred layers, the primary learning rate is fixed to a small value. However, to increase the pace of learning in the newly appended layers than in the transferred layers, the rate is set to a higher value in the fully connected layer. The final fully connected layer of the model consists of 1000 neurons that correspond to classes in the ImageNet. To gain the domain specific features of novel coronavirus and other infections in X-ray images, this layer is set to 4 neurons according to the number of classes in our dataset. The pre-trained models and ensemble learning are explained in the following sub-sections.

2.5.1. MobileNetV2

MobileNetV2 is a deep convolutional network with 53 layers. The

model is an extension of MobileNetV1, which proposes a depth wise separable con-volution layer, which miraculously lessens the size and complexity cost of the design. The current design introduces a more beneficial module with an inverted residual structure, where the input and output of the residual block are narrow bottleneck layers. The intermediate layer is an expanded representation that uses lightweight depth wise convolutions to filter features. Moreover, non-linearities are removed in the thinner layers to maintain the representational ability [83]. MobileNetV2 can generate state-of-the-art performances in image classification.

2.5.2. InceptionV3

InceptionV3 is a pre-trained deep architecture that consists of 48 layers. The architecture is trained on the ImageNet repository that consists of more than a million images [37]. The architecture can categorize images into 1000 classes. As a result, the architecture retrieves valuable features for a broad category of images. InceptionV3 is an enhanced variant of InceptionV2 that achieves enormous proficiency in image classification tasks by factorizing 5×5 convolution layers into two simplistic 3×3 layers. The representational bottleneck is eliminated by adding a regularization part to the loss function. The novel InceptionV3 model narrows down overfitting and achieves label smoothing to a large amount. The design also factorizes a 7×7 convolution layer and joins many discrete layers with batch normalization routine, providing even higher accuracy with less computational complexity. Fig. 4 shows the complete design of the InceptionV3 module.

2.5.3. Ensemble learning for classification

These kinds of networks possess nonlinear architectures that acquire complex associations from the input data by stochastic optimization and back-propagation. Thus, making them greatly sensitive to random weight initializations and the noise present in the training data. These concerns can be alleviated by employing an ensemble model by training numerous deep models and fusing their predictions, where a particular model's shortcomings are compensated by the predictions of the other deep model. Combined predictions are demonstrated to generate better results than a specific deep model [72].

There are many ensembles learning strategies stated in the literature like stacking, max voting, weighted averaging, boosting, simple and blending and several others that minimize the error and improve performance and generalization of deep models. Various research works [47,78,84] depict that ensemble models generate better results for the classification of tuberculosis in X-ray images. An averaging ensemble method utilized to pre-trained models assisted researchers in improving cardiomegaly classification using chest X-ray images [60].

Initially, we conduct feature extraction from InceptionV3 and MobileNetV2. But before extracting features, three dense layers are appended to these deep models, which help the models in learning

Table 2
The range of hyper-parameters used for different pre-trained models.

Deep Model	Batch Size	Learning Rate
AlexNet	48, 32	5e-10, 1e-8, 8e-7, 6e-7, 1e-7, 1e-6, 1e-5
SqueezeNet	48, 32, 24	1e-8, 8e-7, 1e-7, 1e-6, 9e-5
DenseNet201	48, 32	1e-7, 1e-6, 1e-5
MobileNetV2	48, 32, 24	1e-8, 6e-7, 4e-7, 1e-7, 9e-6, 5e-6, 3e-6, 1e-6, 8e-5
InceptionV3	36, 32, 24	1e-8, 1e-6, 8e-5, 2e-5, 1e-5
Our ensemble model (InceptionV3 + MobileNetV2)	48	9e-8, 9e-7, 7e-7, 4e-7, 1e-7, 7e-6, 1e-6

intricate features. The succeeding addition layer helps in merging features from these models. Then, a 0.5 dropout layer is added before classification, which helps in overcoming the hurdles of overfitting and long training time [85]. Finally, the ensemble learning model performs classification, which produces far better results than any particular deep learning design.

3. Results

This section presents the experimental results of the five deep learning models. Initially, we select the best two models and retrieve features from these pre-trained models by employing the ensemble method. Finally, for classification, a fully-connected, SoftMax layer and classification layer are appended after the dropout layer to the ensemble model. We evaluate the performance of these deep models under four different perspectives. When these deep models are (i) trained from the start (ii) fine-tuned without pre-learned weights (iii) fine-tuned with all layers unfrozen and (iv) augmentation and fine-tuned with all layers unfrozen. For each above strategy, specificity, recall, F-score and accuracy are determined for deep models on the COVID-19 dataset, as shown in Table 4. We can find the following observations.

- (i) When the dataset is limited, conventional shallow CNN models produce better results as compared to deeper models. This can be seen from the performance metrics of these CNN models trained from start. The CNN models produce comparatively low specificity, recall, F-score and accuracy because they are not thoroughly trained, due to an immense number of parameters and inadequate training instances.
- (ii) Fine-tuning can also assist in improving performance metrics like F-score and accuracy etc. of these CNN models, even if trained from the start. The results show that the CNN models only fine-tuned on the primary COVID-19 dataset can considerably improve F-score and accuracy, even if no pre-trained weights are used during the training process.
- (iii) Pre-trained weights are used for deep models in the last two approaches. Results show that these pre-trained weights help in substantially increasing the performance of models.
- (iv) Augmentation is also very effective in enhancing a deep model's performance mainly when the training data is limited. The deep models along with standard augmentation techniques can make these models accomplish improved performance. As shown by the results, the strategy where augmentation is applied models show an increase up to 1.77% in F-score and up to 0.84% in accuracy.
- (v) Our proposed deep ensemble model produces a specificity, recall, F-score and accuracy of 98.97%, 96.89%, 96.89%, 96.90%, and 98.45% respectively. The model generates results better as compared to any of the selected CNN models. The proposed design employing augmentation and fine-tuning procedures to accomplished an increase of up to 4.24% in F-score, and up to 2.02% in accuracy over all the selected models.
- (vi) We compare our results with a previous ensemble approach [66] that uses VGG16 and DenseNet201 models, to show that our model is superior in all respects. The model was trained using the same hyper-parameters as suggested in the study, only the batch

Table 3
The number of training parameters and depth of Pre-trained Models.

Pre-trained Model	Depth	Weight Parameters
AlexNet	8	61 million
SqueezeNet	18	1.24 million
DenseNet201	201	20.0 million
MobileNetV2	53	3.5 million
InceptionV3	48	25 million

size was set to 24 to speed up the training process and attain high accuracy. However, the model only achieved specificity, recall, FScore, and accuracy of 91.35%, 74.05%, 65.34%, and 87.04%, respectively. This clearly depicts that deep CNN models need large training data and hundreds of epochs to generalize well for unseen examples.

- (vii) The confusion matrices further elaborate the results of the proposed model for the COVID-19 dataset as shown in Fig. 5. The results of all the ma-trices of five-fold cross validation show that there are ten misclassifications in the Normal class and 1 misclassification in the COVID-19 class. How-ever, mis-classifications are relatively high in the other two classes (bac-terial 32 and viral 81).
- (viii) Fig. 6 shows ROC curves for 5-fold cross-validation to elaborate the performance of the presented model. The ROC curve indicates that the ensemble model performs very good on the COVID-19 dataset.

Over-fitting can be a significant hurdle, particularly with a small number of training instances. A model may achieve significant accuracy on the training data but it may not generalize well for new examples. So, a vital issue to review is that whether there is over-fitting or the sug-gested design has generalized well for given examples. To make this comparison, we evaluate the performance of the model by evaluating the gap within the training and validation curves against the number of epochs. The more widespread is the space between the curves, the greater is over-fitting.

Fig. 7 shows the change in accuracy and loss between training and valida-tion curves for the proposed ensemble model as the number of epochs increases. The curve is related to fold-2 of the model, which was further trained to increase the performance of the model. The curve further shows a slight increase in ac-curacy and a decrease in loss after 10 epochs. The figure also demonstrates that the validation and train curves move side by side without a gap, which shows that there is no over-fitting and the proposed model has generalized correctly for the training examples.

3.1. Statistical analysis procedure

To statistically estimate the performance of our proposed model against its competing models, we repeated our evaluation procedure 5 times. Then, we follow the steps given below.

- (a) We consider the 5 accuracy values of our model and those of a competing model and apply a paired two-sided *t*-test at the 5% significance level.
- (b) Our null hypothesis is that the difference between accuracy values of our model and those of a competing model comes from a normal distribution with mean equal to zero and unknown

variance. The alternative hypothesis is that the mean is not equal to zero.

The same statistical analysis procedure has been applied on F scores, recall and specificity. Our results in Table 4 show that our model is statistically significant than the other models.

3.2. A comparison against previous works on coronavirus classification

A comparison of our work with previously applied approaches for the classification of COVID-19 in chest X-ray images is presented in this section. This elaborates that only limited COVID-19 images are used in most of these approaches, secondly, datasets are mostly unbalanced, thirdly, 5- Folds cross-validation is applied by only by few approaches, and finally, statistical comparison of deep models and hyper-parameter tuning is not applied by any of the approaches. While considering the performance metrics in Table 5, the proposed ensemble model out-performs the considered state-of-the-art approaches, achieving an overall classification accuracy of 98.45%.

4. Discussions

The COVID-19 pandemic has adversely affected the health of people and worldwide economies. Speedy diagnosis has often been substandard and sero-logical tests have not been freely available. The likelihood to use chest X-ray images along with ensemble learning as part of the diagnostic technique could help us in the battle against COVID-19 infections and different respiratory pathogens that might emerge in the days to come [72]. Machine and deep learning techniques are producing excellent outcomes in different fields [73,88] but they earned vital attention because of the new outbreak of diseases like COVID-19 [74,89] and Ebola [90]. These methods can help in the quick and precise diagnosis of a disease. Deep CNNs in particular are highly effective in the image classification tasks and various pre-trained deep models exist for upcoming tasks while applying transfer learning, fine-tuning, and augmen-tation techniques. These techniques are a benchmark in several domains such as object classification [91], image identification [46], biomedical image analysis [92], bacterial classification [31,80,93], and especially disease classification using chest X-rays [94]. A specific type of these deep models known as the ensemble learning model performs image classification, which generates far better results than any individual deep learning model [72]. Our study reveals that transfer learning from the non-biological tasks can substantially enhance current methods of classification for radiology images while also separately yielding feasible performances for a small dataset as shown in Table 4. Some notable investigations [78,79,84] by researchers produce confirmation that transfer learning can generate outstanding outcomes, especially in the case of small datasets [78,95]. Fine-tuning is utilized on multiple pre-trained deep models using Chest X-ray images, to help deep models converge swiftly and acquire features related to a specific domain. The current study unveils that fine-tuning a deep model is essential for its reusability [96]. With such a design, we maintain the initial architectural model, and the pre-learned weights, are appropriated to initialize the pre-trained deep model. The learned weights are therefore renewed through-out the procedure of fine-tuning, helping the model to discover features specific to the related problem. Recently, several types of research have confirmed that fine-tuning is effective and effective for different types of classification issues in the biological field [97].

In the proposed ensemble model, two best performing deep models (MobileNetV2 and InceptionV3) selected for ensemble learning. The ensemble model, possessing the characteristics of both models, was fine-tuned to attain domain-specific features, which are necessary for the identification of novel coro-navirus and other chest related infections in X-ray images.

MobileNetV2 depends on a streamlined architectural design that

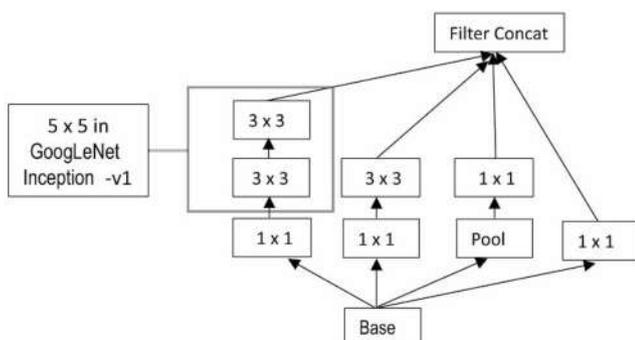


Fig. 4. InceptionV3 module.

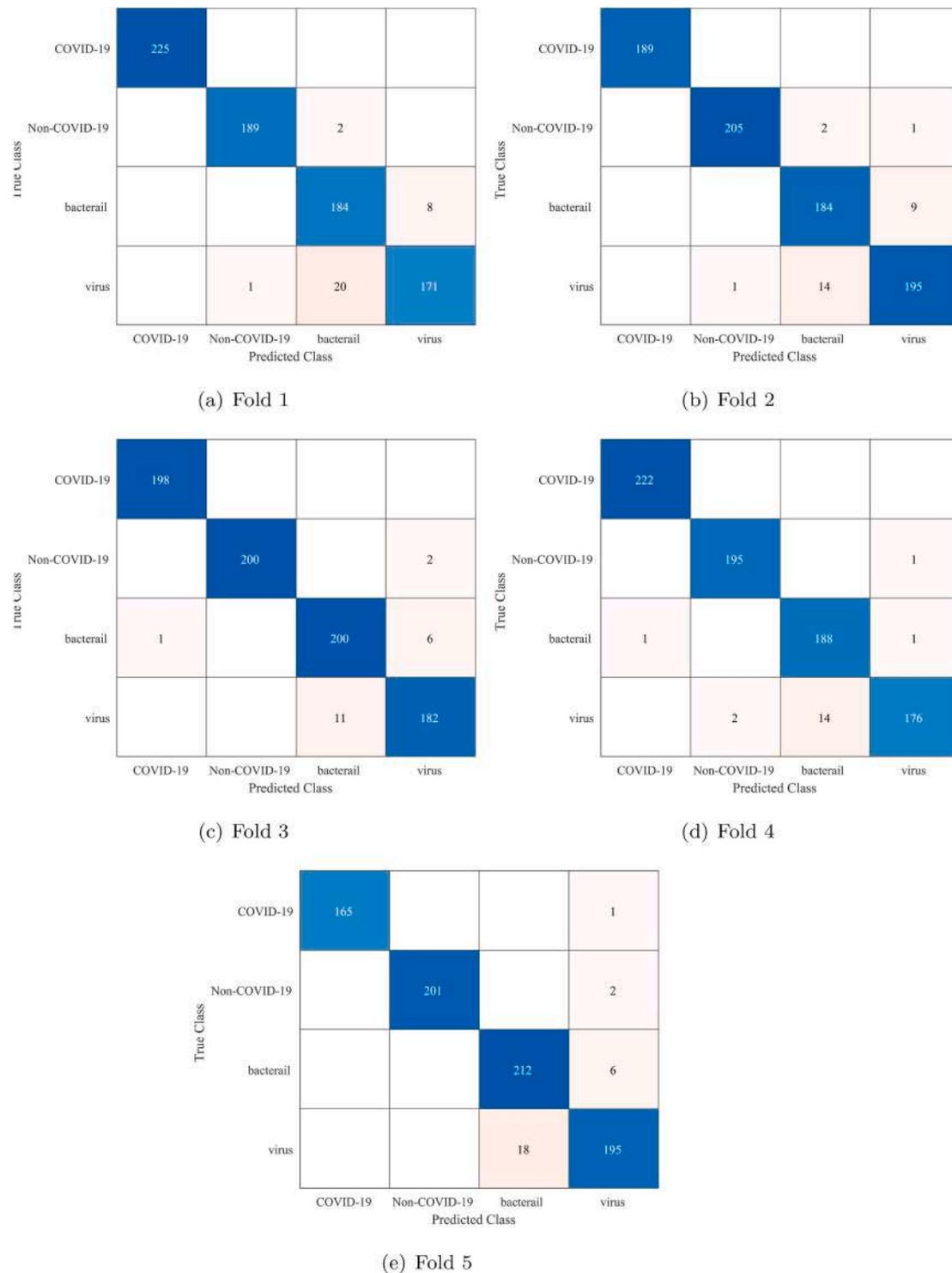


Fig. 5. Confusion matrices for 5-fold cross-validation.

applies depth-wise separable convolutions with different layers for filtering and merging. The factorization has the impact of substantially diminishing computational cost and design dimensions [98]. Such type of network possesses a lesser number of parameters to adjust, as compared to standard Convolution networks, which reduces overfitting.

A recent investigation manifested that pre-trained InceptionV3 model and fine-tuned using chest X-ray films relating to the examination of pulmonary nodules, accomplished fantastic results for the diagnosis of thoracic disease, similar to the conclusion of expert radiologists [99]. Another research also utilizes the InceptionV3 model and transfer

learning using chest X-rays for the classification of pneumonia [58]. The InceptionV3 architecture utilizes factorized inception blocks, facilitating the interface to pick appropriate kernel-sizes for the convolution layers. That allows the design to gain both high and low-level features with larger and smaller convolution layers [83].

The research also depicts that augmentation can significantly help in reducing overfitting and improving the performance of deep models. All pre-trained models applied augmentation technique produced an increase of 1.54–4.24% increase in F-score, and 0.74–2.02% increase in accuracy over prior models, as in Table 4. Investigations also maintain our opinion

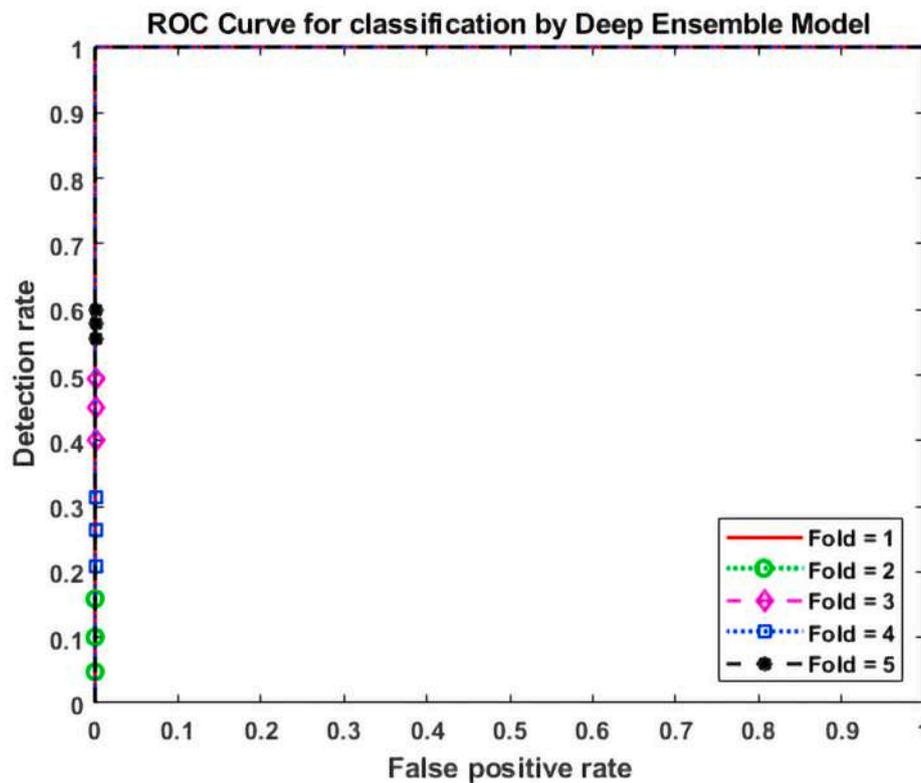


Fig. 6. ROC curve of the deep ensemble model for 5-fold cross-validation.

Table 4

A comparison of performances of our ensemble model and various deep learning models.

Model	Methods	Specificity	Recall	F-score	Accuracy
Our ensemble method (InceptionV3+MobileNetV2)		98.97 ± 0.22	96.89 ± 0.65	96.90 ± 0.65	98.45 ± 0.32
Ensemble method [66] (Vgg16+Densenet201)		● 91.35 ± 0.42	● 74.05 ± 0.96	● 65.34 ± 1.10	● 87.04 ± 0.57
AlexNet	Trained on ODS-NPTW	76.15 ± 0.37	28.48 ± 1.46	24.18 ± 1.82	64.36 ± 0.64
	FT on ODS-NPTW	92.69 ± 0.70	78.11 ± 2.20	77.77 ± 2.17	89.03 ± 1.03
	FT on ODS-PTW	97.41 ± 0.43	92.21 ± 1.25	92.20 ± 1.33	96.13 ± 0.62
	FT on ADS-ALUF	● 97.67 ± 0.46	● 93.05 ± 1.08	● 92.96 ± 1.13	● 96.50 ± 0.64
SqueezeNet	Trained on ODS-NPTW	75.28 ± 0.31	25.85 ± 0.95	14.00 ± 1.32	62.91 ± 0.69
	FT on ODS-NPTW	92.46 ± 0.49	77.30 ± 1.13	77.35 ± 1.15	88.69 ± 0.65
	FT on ODS-PTW	97.24 ± 0.38	91.77 ± 0.87	91.78 ± 0.95	95.87 ± 0.54
	FT on ADS-ALUF	● 97.73 ± 0.25	● 93.21 ± 0.47	● 93.19 ± 0.47	● 96.60 ± 0.33
Densenet201	Trained on ODS-NPTW	74.64 ± 1.43	23.73 ± 4.03	22.20 ± 4.68	61.99 ± 2.14
	FT on ODS-NPTW	87.46 ± 0.74	62.36 ± 2.49	62.55 ± 2.33	81.21 ± 1.01
	FT on ODS-PTW	97.36 ± 0.43	92.08 ± 1.35	92.04 ± 1.41	96.05 ± 0.65
	FT on ADS-ALUF	● 97.91 ± 0.38	● 93.67 ± 1.35	● 93.67 ± 1.33	● 96.86 ± 0.32
MobileNetV2	Trained on ODS-NPTW	74.48 ± 0.57	23.46 ± 1.73	14.29 ± 0.85	61.74 ± 1.06
	FT on ODS-NPTW	92.33 ± 0.54	77.04 ± 1.06	76.68 ± 1.13	88.50 ± 0.74
	FT on ODS-PTW	97.91 ± 0.21	93.76 ± 0.63	93.71 ± 0.66	96.86 ± 0.32
	FT on ADS-ALUF	● 98.24 ± 0.27	● 94.71 ± 1.01	● 94.68 ± 1.01	● 97.36 ± 0.42
InceptionV3	Trained on ODS-NPTW	72.16 ± 0.58	16.57 ± 1.33	16.31 ± 1.00	58.24 ± 0.59
	FT on ODS-NPTW	85.01 ± 0.41	55.02 ± 1.09	55.29 ± 1.17	77.53 ± 0.52
	FT on ODS-PTW	98.08 ± 0.45	94.19 ± 1.63	94.21 ± 1.60	97.11 ± 0.70
	FT on ADS-ALUF	● 98.49 ± 0.38	● 95.42 ± 1.31	● 95.43 ± 1.28	● 97.74 ± 0.58

ODS, ADS, NPTW, PTW, ALUF, FT stands for original dataset, augmented dataset, no pre-trained weights, pre-trained weights, all layers un-frozen, fine-tuned. A ● denotes that our deep learning ensemble model is statistically better than its competing model.

that augmentation can help in increasing the performance and producing a more generalized prototype without the menace of overfitting [88,95]. The current research also depicts that merging features from deep models using ensemble learning can help in increasing the accuracy and overcoming deficiencies of individual models. As a result, F-score and accuracy of the model increased over all the applied deep models by 1.54–4.24% and 0.74–2.02%, respectively.

Various researches propose that contrary to the conventional CNN models, ensemble learning models by merging deep CNN's, acquire

more useful features from images in the training data. These ensemble models have accomplished outstanding results in image classification tasks in various domains [89,100,101], along with pneumonia classification [58], cardiovascular tissues identification [102] and especially in the area of radiology images [90,103].

5. Conclusions and future works

In the current research, a new coronavirus classification procedure is

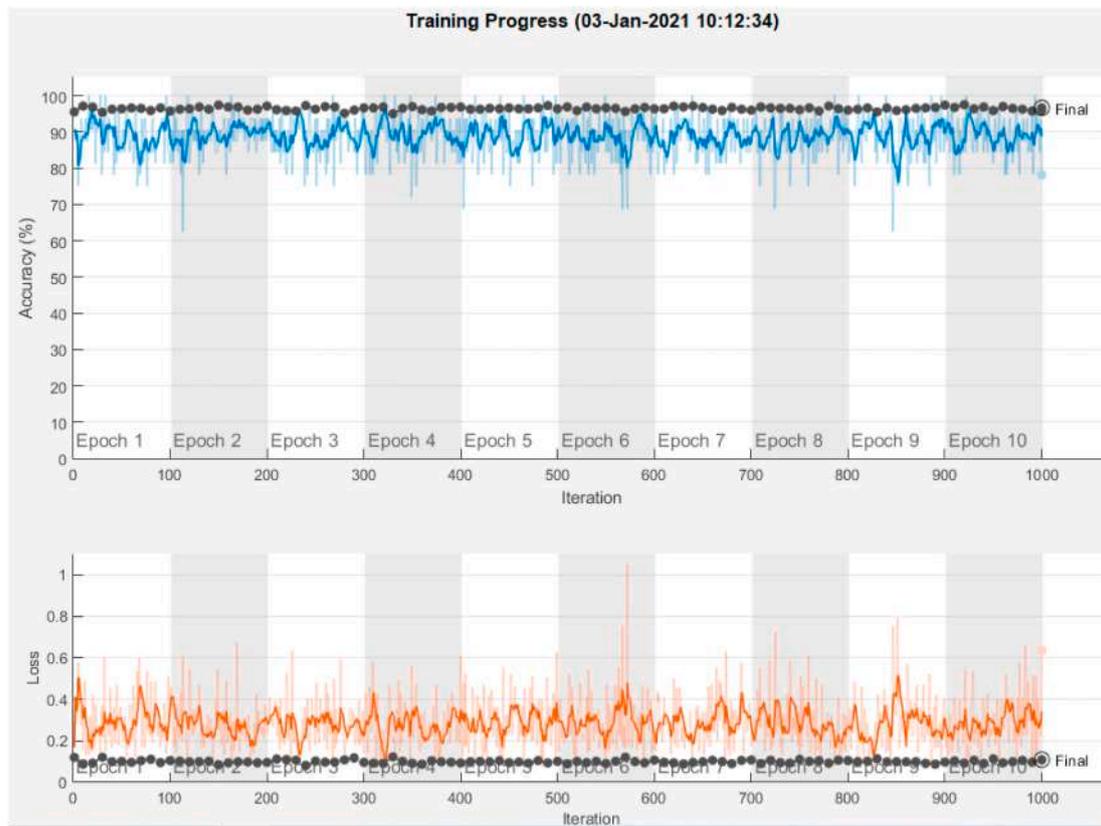


Fig. 7. Learning curves for training and validation accuracy (blue, black dotted lines) and training and validation loss (orange, black dotted lines) of fold-2 of fine-tuned, pre-trained, ensemble model for various infections in X-rays.

devised, which takes advantage of ensemble learning, transfer learning, fine-tuning, and data augmentation techniques for differentiating between various chest infections. Ensemble learning assists in blending the qualities of separate models while overcoming deficiencies of specific models. Transfer learning addresses the need for an abundant quantity of training data. Fine-tuning aids the model converge quickly and attain domain-related features. Data augmentation makes datasets more diverse, which enhances the generalization capacity of the model and thus assists in handling overfitting. The suggested design comprises of MobileNetV2 and InceptionV3 designs, which produces far better classification performance than any of the selected deep models. We also compared our model with another ensemble approach presented in the literature, to show that our model is far superior to the presented approach. The model attained a specificity, recall, F-score and accuracy of 98.97%, 96.89%, 96.90%, and 98.45%, which can significantly be beneficial for radiologists and diagnostic staff in the accurate classification of the novel coronavirus and other bacterial and viral infections in chest X-rays. As a result, the proposed model can assist in the accurate

and swift diagnosis of chest related infections and thereby limiting the social and economic impact on the community.

Although the results are good, the dataset is still limited and noisy. For viable deep learning solutions that can be accepted as a standard for classifying COVID-19 in X-rays, there is a need to develop a significantly larger dataset without any background noise. As future work, a deep lung segmentation model like U-Net can be used to extract the Lung counter from the noisy X-ray image to develop a significantly larger dataset for COVID-19, which can then be used as input for CNN models. This can be a leap towards a highly accurate and practical deep learning solution for the community.

Consent to participate

The authors declares their consent for participation.

Table 5

A comparison of Our model with previous approaches for COVID-19 Classification.

(c) In case the null hypothesis is rejected, the performance of our model is statistically different from the other model. We consider it to be a win for our model if the mean accuracy value of our model is greater than that of the competing model. We denote it by a •. Otherwise, it's a loss for our model. We denote it by a ◦.

(d) If *t*-test does not reject the null hypothesis, the performance of our model is not statistically different from the other model and we consider it to be a tie. A tie will be represented by no symbol.

Approach	Classes	COVID-19 Images	Balanced Dataset?	5-folds CV?	Statistical Comparison?	Tuning Hyper parameters?	Accuracy
Our Model	4	1000					98.45%
Covid-Net [35]	3	180	×	×	×	×	93.3%
CoroNet [64]	4	284	×	×	×	×	89.6%
XGB [86]	4	130	×		×	×	79.52%
DELT [87]	4	305	×		×	×	90.13%
COVID-ResNet [74]	4	68	×	×	×	×	96.23%

Consent for publication

The authors declares their consent for publication.

Code availability

Code will be shared on request.

Authors' contributions

- (i) We prepared a large and well-balanced dataset with four classes (viral, bacterial, COVID-19 and normal). Each class consists of 1000 images.
- (ii) Our work integrates fine-tuning, augmentation, transfer learning and hyper-parameter tuning into one model.
- (iii) We propose a deep design by merging the features of Mobile-NetV2 and InceptionV3 models using an ensemble approach. We have made architectural adjustments to deep models by adding three dense layers at the end of each model to learn more intricate features before merging their features through the addition layer. We also added a dropout layer after the addition layer to handle overfitting.
- (iv) We compare our ensemble model with state-of-the-art deep models in terms of four performance metrics namely specificity, recall, F-score and accuracy.
- (v) We also compared our ensemble model with another ensemble approach presented in the literature, to show that our model is far superior to the previous model.
- (vi) To statistically validate the performance of the proposed design, we employ 5-fold cross-validation with the paired two-sided t-test.

Availability of data and material

COVID-19 dataset will be shared on request.

Funding

No funds, grants, or other support was received.

Declaration of competing interest

The authors declares no conflicts of interest, financial or otherwise.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compbiomed.2021.104401>.

References

- [1] J. Lederberg, M.A. Hamburg, M.S. Smolinski, et al., *Microbial Threats to Health: Emergence, Detection, and Response*, National Academies Press, 2003.
- [2] S.J. Salyer, R. Silver, K. Simone, C.B. Behravesh, Prioritizing zoonoses for global health capacity building—themes from one health zoonotic disease workshops in 7 countries, *Emerg. Infect. Dis.* 23 (2017) S55, 2014–2016.
- [3] R. Franconi, E. Illiano, F. Paoletti, S. Massa, A. Venuti, O.C. Demurtas, Rapid and low-cost tools derived from plants to face emerging/re-emerging infectious diseases and bioterrorism agents, in: *Defence against Bioterrorism*, Springer, 2018, pp. 123–139.
- [4] H. Wang, M. Naghavi, C. Allen, R.M. Barber, Z.A. Bhutta, A. Carter, D.C. Casey, F. J. Charlson, A.Z. Chen, M.M. Coates, et al., Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, in: 1980–2015: a Systematic Analysis for the Global Burden of Disease Study 2015, vol. 388, *The Lancet*, 2016, pp. 1459–1544.
- [5] W.A. Gebreyes, J. Dupouy-Camet, M.J. Newport, C.J. Oliveira, L.S. Schlesinger, Y.M. Saif, S. Kariuki, L.J. Saif, W. Saville, T. Wittum, et al., The global one health paradigm: challenges and opportunities for tackling infectious diseases at the human, animal, and environment interface in low-resource settings, *PLoS Neglected Trop. Dis.* 8 (2014).
- [6] D. Nabarro, C. Wannous, The potential contribution of livestock to food and nutrition security: the application of the one health approach in live-stock policy and practice, *Rev. Sci. Tech. (Int. Off. Epizoot.)* 33 (2014) 475–485.
- [7] P.-I. Lee, P.-R. Hsueh, Emerging threats from zoonotic coronaviruses— from sars and mers to 2019-ncov, *J. Microbiol. Immunol. Infect.* 53 (6) (2020) 365–367.
- [8] D. Cyranoski, Mystery deepens over animal source of coronavirus, *Nature* 579 (2020) 18–19.
- [9] A. Petherick, Developing antibody tests for sars-cov-2, *Lancet* 395 (2020) 1101–1102.
- [10] A. King, The hunt for the next potential coronavirus animal host, Online, <https://www.nationalgeographic.com/animals/2020/03/coronavirus-animal-reservoir-research/>, 2020. (Accessed 3 April 2020).
- [11] W.H. Organization, Zoonotic disease: emerging public health threats in the Region, Online, <http://www.emro.who.int/fr/about-who/rc61/zoonotic-diseases.html/>, 2008. (Accessed 17 July 2011).
- [12] N. Van Doremalen, T. Bushmaker, D.H. Morris, M.G. Holbrook, A. Gamble, B. N. Williamson, A. Tamin, J.L. Harcourt, N.J. Thornburg, S.I. Gerber, et al., Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1, *N. Engl. J. Med.* 382 (2020) 1564–1567.
- [13] W.H. Organization, et al., Modes of Transmission of Virus Causing COVID-19: Implications for IPC Precaution Recommendations: Scientific Brief, World Health Organization, 2020, 27 March 2020, Technical Report.
- [14] D. Lewis, Is the coronavirus airborne? experts can't agree, *Nature* 580 (2020) 175.
- [15] Y. Liu, Z. Ning, Y. Chen, M. Guo, Y. Liu, N.K. Gali, L. Sun, Y. Duan, J. Cai, D. Westerdahl, et al., Aerodynamic Characteristics and Rna Concentration of Sars-Cov-2 Aerosol in Wuhan Hospitals during Covid-19 Outbreak, *BioRxiv*, 2020.
- [16] V.C. Cheng, S.-C. Wong, J.H. Chen, C.C. Yip, V.W. Chuang, O.T. Tsang, S. Sridhar, J.F. Chan, P.-L. Ho, K.-Y. Yuen, Escalating infection control response to the rapidly evolving epidemiology of the coronavirus disease 2019 (covid-19) due to sars-cov-2 in Hong Kong, *Infect. Contr. Hosp. Epidemiol.* 41 (2020) 493–498.
- [17] E. Watts, A. Leck, V. Hu, Personal protective equipment for covid-19 in eye care, *Eye Health* 33 (2020) 22.
- [18] L. Bourouiba, Turbulent gas clouds and respiratory pathogen emissions: potential implications for reducing transmission of covid-19, *Jama* 323 (2020) 1837–1838.
- [19] C. Covid, Global Cases by the Center for Systems Science and Engineering (Csse) at Johns Hopkins University (Jhu). johns hopkins coronavirus resource center, vol. 19.
- [20] A. Dehghani, G. Masoumi, Could sars-cov-2 or covid-19 be a biological weapon? *Iran. J. Public Health* 49 (2020) 143–144.
- [21] S.J. Khurshid, Novel coronavirus (nCoV-2019): is it a bioweapon?, Online; accessed 11-February-2020, <https://www.omicsonline.org/open-access-pdfs/novel-coronavirus-ncov2019-is-it-a-bioweapon.pdf/>, 2020.
- [22] C. for Systems Science, Engineering, Covid-19 dashboard by the center for systems science and engineering (csse) at johns hopkins university (jhu), Online; accessed 18-February-2021, <https://coronavirus.jhu.edu/map.html/>, 2021.
- [23] M.C. Michael D. Shear, J. Glanz, Coronavirus may kill 100,000 to 240,000 in u.s. despite actions, officials say [Online; accessed 31-March-2020], <https://www.nytimes.com/2020/03/31/us/politics/coronavirus-death-toll-united-states.html/>, 2020.
- [24] H. Correspondent, Coronavirus could kill half a million britons and infect 80 indicate, Online, <https://www.independent.co.uk/news/health/>, 2020. (Accessed 26 February 2020), <https://www.independent.co.uk/news/health/coronavirus-news-latest-deaths-uk-infection-flu-a9360271.html/>.
- [25] N. Sadegh, P. Sadegh, Test-based Stabilizing Control of Covid-19 Transmission, 2020.
- [26] cdc gov, Food Safety, 2019.
- [27] V.D. Menachery, R.L. Graham, R.S. Baric, Jumping species—a mechanism for coronavirus persistence and survival, *Curr. Opin. Virol.* 23 (2017) 1–7.
- [28] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in China: a report of 1014 cases, *Radiology* (2020) 200642.
- [29] G. Kolifarhood, M. Aghaali, H.M. Saadati, N. Taherpour, S. Rahimi, N. Izadi, S.S. H. Nazari, Epidemiological and Clinical Aspects of Covid-19; a Narrative Review, vol. 8, *Archives of academic emergency medicine*, 2020.
- [30] Y. Oh, S. Park, J.C. Ye, Deep learning covid-19 features on cxr using limited training data sets, *IEEE Trans. Med. Imag.* 39 (8) (2020) 2688–2700.
- [31] 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) 200343, 200343.
- [32] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, W. Ji, Sensitivity of chest ct for covid-19: comparison to rt-pcr, *Radiology* (2020) 200432.
- [33] M.-Y. Ng, E.-Y. Lee, J. Yang, F. Yang, X. Li, H. Wang, M.M.-s. Lui, C.S.-Y. Lo, B. Leung, P.-L. Khong, et al., Imaging profile of the covid-19 infection: radiologic findings and literature review, *Radiology: Cardio-Thorac. Imag.* 2 (2020), e200034.
- [34] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, et al., Clinical features of patients infected with 2019 novel coronavirus in wuhan, *China, Lancet* 395 (2020) 497–506.
- [35] L. Wang, A. Wong, Covid-net: A Tailored Deep Convolutional Neural Network Design for Detection of Covid-19 Cases from Chest Radiography Images, arXiv, 2020. arXiv–2003.

- [36] B. Zieliński, A. Plichta, K. Misztal, P. Spurek, M. Brzyczczy-W/loch, D. Ochońska, Deep learning approach to bacterial colony classification, *PLoS One* 12 (2017), e0184554.
- [37] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826.
- [38] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: *Advances in Neural Information Processing Systems*, pp. 1097–1105.
- [39] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, in: *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2009, pp. 248–255.
- [40] P. Sermanet, A. Frome, E. Real, Attention for Fine-Grained Categorization, 2014 arXiv preprint arXiv:1412.7054.
- [41] R.H. Abiyev, M.K.S. Ma'aitah, Deep convolutional neural networks for chest diseases detection, *J. Healthc. Eng.* 2018 (2018) 1–11, 4168538.
- [42] A.M. Dawud, K. Yurtkan, H. Oztoprak, Application of deep learning in neuroradiology: brain haemorrhage classification using transfer learning, *Comput. Intell. Neurosci.* 2019 (2019).
- [43] A. Helwan, G. El-Fakhri, H. Sasani, D. Uzun Ozsahin, Deep networks in identifying ct brain hemorrhage, *J. Intell. Fuzzy Syst.* 35 (2018) 2215–2228.
- [44] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, *Int. J. Comput. Vis.* 115 (2015) 211–252.
- [45] D. Han, Q. Liu, W. Fan, A new image classification method using cnn transfer learning and web data augmentation, *Expert Syst. Appl.* 95 (2018) 43–56.
- [46] N. Aneja, S. Aneja, Transfer Learning Using Cnn for Handwritten Devana- Gari Character Recognition, 2019 arXiv preprint arXiv:1909.08774.
- [47] P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks, *Radiology* 284 (2017) 574–582.
- [48] P. Huang, S. Park, R. Yan, J. Lee, L.C. Chu, C.T. Lin, A. Hussien, J. Rathmell, B. Thomas, C. Chen, et al., Added value of computer-aided ct image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study, *Radiology* 286 (2018) 286–295.
- [49] J.P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong, M. Ghassemi, Covid-19 Image Data Collection: Prospective Predictions Are the Future, 2020 arXiv preprint arXiv:2006.11988.
- [50] O. Gozes, M. Frid-Adar, H. Greenspan, P.D. Browning, H. Zhang, W. Ji, A. Bernheim, E. Siegel, Rapid Ai Development Cycle for the Coro- Navirus (Covid-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring Using Deep Learning Ct Image Analysis, 2020 arXiv preprint arXiv:2003.05037.
- [51] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su, et al., A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia, 2020. Engineering.
- [52] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, et al., A Deep Learning Algorithm Using Ct Images to Screen for Corona Virus Disease (Covid-19), 2020. MedRxiv.
- [53] E. Ayan, H.M. Ünver, Diagnosis of pneumonia from chest x-ray images using deep learning, in: *Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, IEEE, 2019, pp. 1–5.
- [54] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, A. Mittal, Pneumonia detection using cnn based feature extraction, in: *IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, IEEE, 2019, pp. 1–7.
- [55] O. Stephen, M. Sain, U.J. Maduh, D.-U. Jeong, An efficient deep learning approach to pneumonia classification in healthcare, *J. Healthc. Eng.* 2019 (2019).
- [56] J. Melendez, B. van Ginneken, P. Maduskar, R.H. Philipsen, K. Reither, M. Breuninger, I.M. Adetifa, R. Maane, H. Ayles, C.I. Sánchez, A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest x-rays, *IEEE Trans. Med. Imag.* 34 (2014) 179–192.
- [57] S.R. Islam, S.P. Maity, A.K. Ray, M. Mandal, Automatic detection of pneumonia on compressed sensing images using deep learning, in: *IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, IEEE, 2019, pp. 1–4.
- [58] V. Chouhan, S.K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damaševičius, V.H.C. De Albuquerque, A novel transfer learning based approach for pneumonia detection in chest x-ray images, *Appl. Sci.* 10 (2020) 559.
- [59] D. Demner-Fushman, M.D. Kohli, M.B. Rosenman, S.E. Shooshan, L. Rodriguez, S. Antani, G.R. Thoma, C.J. McDonald, Preparing a collection of radiology examinations for distribution and retrieval, *J. Am. Med. Inf. Assoc.* 23 (2016) 304–310.
- [60] M.T. Islam, M.A. Aowal, A.T. Minhaz, K. Ashraf, Abnormality Detection and Localization in Chest X-Rays Using Deep Convolutional Neural Networks, 2017 arXiv preprint arXiv:1705.09850.
- [61] I.D. Apostolopoulos, T.A. Mpesiana, Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks, *Phys. Eng. Sci. Med.* (2020) 1.
- [62] F. Ucar, D. Korkmaz, Covidiagnosis-net: deep bayes-squeezenet based diagnostic of the coronavirus disease 2019 (covid-19) from x-ray images, *Med. Hypotheses* (2020) 109761.
- [63] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of Covid-19 in Chest X-Ray Images Using Detrac Deep Convolutional Neural Network, 2020, p. 13815, arXiv preprint arXiv:2003.
- [64] A.I. Khan, J.L. Shah, M.M. Bhat, Coronet: a deep neural network for detection and diagnosis of covid-19 from chest x-ray images, *Comput. Methods Progr. Biomed.* (2020) 105581.
- [65] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K.N. Platani- otis, A. Mohammadi, Covid-caps: A Capsule Network-Based Framework for Identification of Covid-19 Cases from X-Ray Images, 2020 arXiv preprint arXiv: 2004.02696.
- [66] N. Gianchandani, A. Jaiswal, D. Singh, V. Kumar, M. Kaur, Rapid covid- 19 diagnosis using ensemble deep transfer learning models from chest radiographic images, *J. Ambient Intell. Hum. Comput.* (2020) 1–13.
- [67] D. Singh, V. Kumar, V. Yadav, M. Kaur, Deep neural network-based screening model for covid-19-infected patients using chest x-ray images, *Int. J. Pattern Recogn. Artif. Intell.* (2020) 2151004.
- [68] D. Singh, V. Kumar, M. Kaur, Densely Connected Convolutional Networks- Based Covid-19 Screening Model, *Applied Intelligence*, 2021, pp. 1–8.
- [69] J. Flusser, T. Suk, Character recognition by affine moment invariants, in: *International Conference on Computer Analysis of Images and Patterns*, Springer, pp. 572–577.
- [70] E.J.d.S. Luz, P.L. Silva, R. Silva, L. Silva, G. Moreira, D. Menotti, To- Wards an Effective and Efficient Deep Learning Model for Covid-19 Patterns Detection in X-Ray Images, *CoRR*, 2020.
- [71] S. Khobahi, C. Agarwal, M. Soltanalian, Coronet: A Deep Network Architecture for Semi-supervised Task-Based Identification of Covid-19 from Chest X-Ray Images, *medRxiv*, 2020.
- [72] S. Rajaraman, J. Siegelman, P.O. Alderson, L.S. Folio, L.R. Folio, S.K. Antani, Iteratively Pruned Deep Learning Ensembles for Covid-19 Detection in Chest X-Rays, 2020 arXiv preprint arXiv:2004.08379.
- [73] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G.J. Soufi, Deep-covid: Predicting Covid-19 from Chest X-Ray Images Using Deep Transfer Learning, 2020 arXiv preprint arXiv:2004.09363.
- [74] M. Farooq, A. Hafeez, Covid-resnet: A Deep Learning Framework for Screening of Covid 19 from Radiographs, 2020 arXiv preprint arXiv:2003.14395.
- [75] M.E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z. B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, et al., Can Ai Help in Screening Viral and Covid-19 Pneumonia?, 2020 arXiv preprint arXiv: 2003.13145.
- [76] T. Rahman, M.E. Chowdhury, A. Khandakar, K.R. Islam, K.F. Islam, Z.B. Mahbub, M.A. Kadir, S. Kashem, Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray, *Appl. Sci.* 10 (2020) 3233.
- [77] A. Brulet, G. Llorca, L. Letrilliart, Medical wikis dedicated to clinical practice: a systematic review, *J. Med. Internet Res.* 17 (2015) e48.
- [78] S. Rajaraman, S. Candemir, I. Kim, G. Thoma, S. Antani, Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs, *Appl. Sci.* 8 (2018) 1715.
- [79] E. Neri, F. Coppola, A. Larici, N. Sverzellati, M. Mazzei, P. Sacco, G. Dalpiaz, B. Feragalli, V. Miele, R. Grassi, Structured reporting of chest ct in covid-19 pneumonia: a consensus proposal, *Insights Imag.* 11 (2020) 1–9.
- [80] T.M. Mitchell, *Mach. Learn.* 1 of 1 (1997).
- [81] Z.N.K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J. Lu, Brain tumor classification for mr images using transfer learning and fine-tuning, *Comput. Med. Imag. Graph.* 75 (2019) 34–46.
- [82] Z.N.K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J. Lu, Content-based brain tumor retrieval for mr images using transfer learning, *IEEE Access* 7 (2019) 17809–17822.
- [83] S. Vesal, N. Ravikumar, A. Davari, S. Ellmann, A. Maier, Classification of breast cancer histology images using transfer learning, in: *International Conference Image Analysis and Recognition*, Springer, pp. 812–819.
- [84] S. Rajaraman, S.K. Antani, Modality-specific deep learning model ensembles toward improving tb detection in chest radiographs, *IEEE Access* 8 (2020) 27318–27326.
- [85] C. Garbin, X. Zhu, O. Marques, Dropout vs. Batch Normalization: an Empirical Study of Their Impact to Deep Learning, *Multimedia Tools and Applications*, 2020, pp. 1–39.
- [86] L. Hussain, T. Nguyen, H. Li, A.A. Abbasi, K.J. Lone, Z. Zhao, M. Zaib, A. Chen, T. Q. Duong, Machine-learning classification of texture features of portable chest x-ray accurately classifies covid-19 lung infection, *Biomed. Eng. Online* 19 (2020) 1–18.
- [87] S. Basu, S. Mitra, N. Saha, Deep learning for screening covid-19 using chest x-ray images, in: *IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, 2020, pp. 2521–2527.
- [88] F. Cavallo, A. Moschetti, D. Esposito, C. Maremmani, E. Rovini, Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning, *Park. Relat. Disord.* 63 (2019) 111–116.
- [89] L. Abenavoli, P. Cinaglia, F. Luzzza, I. Gentile, L. Boccuto, Epidemiology of coronavirus disease outbreak: the Italian trends, *Rev. Recent Clin. Trials* 15 (2020).
- [90] M.A. Johansson, N.G. Reich, L.A. Meyers, M. Lipsitch, Preprints: an underutilized mechanism to accelerate outbreak science, *PLoS Med.* 15 (2018), e1002549.
- [91] R.S.T. de Menezes, R.M. Magalhaes, H. Maia, Object recognition using convolutional neural networks, in: *Artificial Neural Networks, International Open*, 2019.
- [92] A.S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on mri, *Z. Med. Phys.* 29 (2019) 102–127.
- [93] A. Helwan, R. Abiyev, Shape and texture features for the identification of breast cancer, in: *Proceedings of the World Congress on Engineering and Computer Science*, volume vol. 2, pp. 19–21.

- [94] A. Howard, A. Zhmoginov, L.-C. Chen, M. Sandler, M. Zhu, Inverted Residuals and Linear Bottlenecks: Mobile Networks for Classification, Detection and Segmentation, 2018.
- [95] 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, pp. 770–778.
- [96] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks?, in: Advances in Neural Information Processing Systems, pp. 3320–3328.
- [97] H.-C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mol- lura, R. M. Summers, Deep convolutional neural networks for computer- aided detection: cnn architectures, dataset characteristics and transfer learning, *IEEE Trans. Med. Imag.* 35 (2016) 1285–1298.
- [98] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient Convolutional Neural Net- Works for Mobile Vision Applications, 2017 arXiv preprint arXiv:1704.04861.
- [99] C. Wang, D. Chen, L. Hao, X. Liu, Y. Zeng, J. Chen, G. Zhang, Pulmonary image classification based on inception-v3 transfer learning model, *IEEE Access* 7 (2019) 146533–146541.
- [100] Q. Yang, Y. Liu, T. Chen, Y. Tong, Federated machine learning: concept and applications, *ACM Trans. Intell. Syst. Technol. (TIST)* 10 (2019) 1–19.
- [101] N.-N. Liu, J.-C. Tan, J. Li, S. Li, Y. Cai, H. Wang, Covid-19 pandemic: experiences in China and implications for its prevention and treatment worldwide, *Curr. Cancer Drug Targets* 20 (6) (2020) 410–416.
- [102] S. Mittal, Retracted article: ensemble of transfer learnt classifiers for recognition of cardiovascular tissues from histological images, *Phys. Eng. Sci. Med.* (2019) 1, 1.
- [103] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, *Cell* 172 (2018) 1122–1131.