Classification of COVID-19 chest X-Ray and CT images using a type of dynamic CNN modification method

Guangyu Jia, Hak-Keung Lam[*], Yujia Xu

Department of Engineering, King's College London, Strand, London, WC2R 2LS, United Kingdom

A B S T R A C T

Understanding and classifying Chest X-Ray (CXR) and computerised tomography (CT) images are of great significance for COVID-19 diagnosis. The existing research on the classification for COVID-19 cases faces the challenges of data imbalance, insufficient generalisability, the lack of comparative study, etc. To address these problems, this paper proposes a type of modified MobileNet to classify COVID-19 CXR images and a modified ResNet architecture for CT image classification. In particular, a modification method of convolutional neural networks (CNN) is designed to solve the gradient vanishing problem and improve the classification performance through dynamically combining features in different layers of a CNN. The modified MobileNet is applied to the classification of COVID-19, Tuberculosis, viral pneumonia (with the exception of COVID-19), bacterial pneumonia and normal controls using CXR images. Also, the proposed modified ResNet is used for the classification of COVID-19, non-COVID-19 infections and normal controls using CT images. The results show that the proposed methods achieve 99.6% test accuracy on the five-category CXR image dataset and 99.3% test accuracy on the CT image dataset. Six advanced CNN architectures and two specific COVID-19 detection models, i.e., COVID-Net and COVIDNet-CT are used in comparative studies. Two benchmark datasets and a CXR image dataset which combines eight different CXR image sources are employed to evaluate the performance of the above models. The results show that the proposed methods outperform the comparative models in classification accuracy, sensitivity, and precision, which demonstrate their potential in computer-aided diagnosis for healthcare applications.

1. Introduction

In the fight against COVID-19, the immediate and accurate screening of infected patients is of great significance. As widely used screening approaches, chest X-Ray (CXR) and computed tomography (CT) play an important role in the diagnosis of COVID-19 cases, especially when viral testing is in short supply.

Studies showed that changes occur in CXR and CT images before the beginning of COVID-19 symptoms for some patients [1,2]. Also, the symptoms of COVID-19 and other lung diseases can be similar in their very early stages [3]. It is crucial to effectively distinguish COVID-19 from other lung diseases during the early stages, otherwise inaccurate diagnoses may expose more people to coronavirus.

The development of deep learning techniques enables end-to-end image classification without manual feature engineering. In the domain of COVID-19 detection, deep learning techniques have been widely adopted for related classification tasks [4–6]. For example, in Ref. [7], Inception net was utilized for COVID-19 outbreak screening

with CXRs [8]. proposed a type of patch-based convolutional neural network (CNN) with a small amount of trainable parameters for COVID-19 diagnosis. The research in Ref. [9] considered flat and hierarchical classification scenarios for COVID-19 identification for more than three classes. In Ref. [10], two algorithms including a deep neural network on the fractal feature of images and a CNN architecture with the direct use of the CXR images were presented. The above research demonstrates that the advantage of deep learning methods for CXR image classification is mainly the capability of capturing the pixel-level information which cannot be obviously noticed by human eyes.

Having reviewed the success of the reported applications, the existing studies on COVID-19 classification also show some limitations and challenges. First, as the number of available training data is limited, data imbalances of different classes are found in plenty of the literature. Deep learning models are unlikely to be well trained by the unbalanced data, and the high accuracy in these circumstances cannot guarantee the effectiveness of COVID-19 detection. Furthermore, after a careful comparison of the images from different data sources, it can be found that
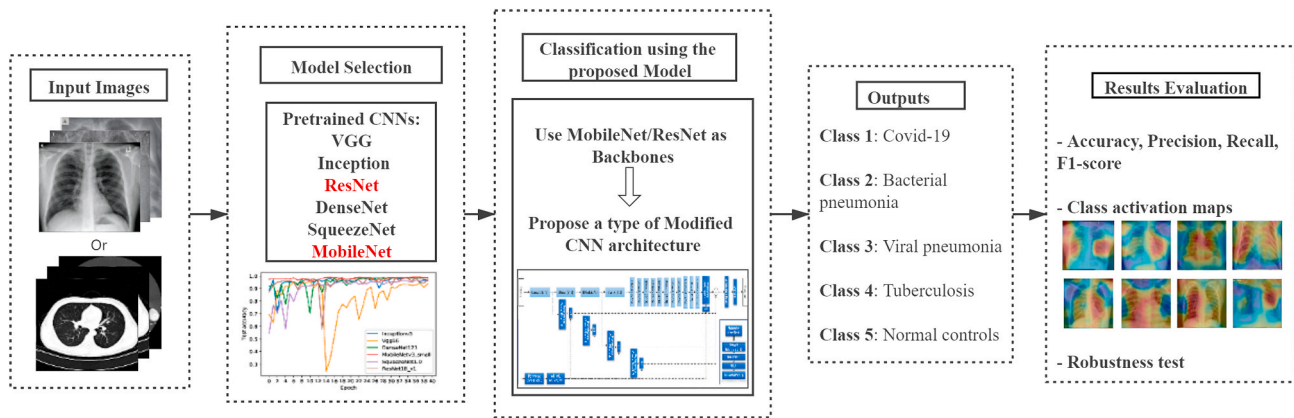
Fig. 1. Diagram of the implementation process.

images in different classes vary in image quality, orientation, brightness, etc. The algorithms might take these into account during classification rather than focusing on the disease related information in images.

In this research, we aim to design a COVID-19 detection system based on deep learning techniques which can be used in the screening process of COVID-19 CXR and CT images. Motivated by the fact that the viral test requires waiting time to obtain the results, the developed end-to-end medical images classification frameworks attempt to accelerate the testing process, alleviate the workload of clinicians in manual images processing, and provide patients with timely results, all of which facilitate effective isolation to further control the spread of coronavirus.

To address the problem of data imbalance, we combined eight different data sources, such that each class in the dataset had a similar number of samples. To establish a more effective and suitable model for COVID-19 classification, we compared the classification results of several widely used CNN architectures which included VGG, ResNet, DenseNet, MobileNet, Inception and SqueezeNet. Among those models, we found that MobileNet achieved the best performance in COVID-19 classification of CXR images, and ResNet achieved high test accuracy with less computational budget in the classification of CT images. Although MobileNet and ResNet possessed satisfactory classification performance, overfitting and gradient vanishing problems still occurred during training.

To overcome these obstacles and further improve the classification performance of deep learning models, we proposed a type of dynamic CNN modification method which combines low-level and high-level features in the original model, such that the model is able to converge faster, be more robust and reach higher classification accuracy. The experimental results demonstrate that the proposed method achieves average 99.7% test accuracy in three-category classification, average 99.9% test accuracy in four-category classification and average 99.6% in five-category classification, which surpass the original MobileNet architecture.

The contributions of this paper are summarized as follows:

- A type of dynamic CNN modification method is proposed in this paper for the detection of COVID-19 cases. The proposed method combines features in different layers of a CNN using weights that are dynamically changed according to the inputs. The modified MobileNet and modified ResNet are respectively applied to the classification of CXR images and CT images for the COVID-19 detection.
- To facilitate effective diagnosis, three classification scenarios are considered in this paper: (a) five-category classification: COVID-19 infection, bacterial pneumonia, viral pneumonia (with the exception of COVID-19), tuberculosis, and normal controls; (b) four-category classification: COVID-19, Non-COVID-19 pneumonia, tuberculosis, and normal controls; (c) three-category classification: COVID-19, Non-COVID-19 infections, and normal controls.

- Comprehensive comparisons are conducted in this research which include the comparisons between the proposed method and two recently published models, i.e., COVID-Net [11] and COVIDNet-CT [12]. Also, six widely used deep learning methods, VGG16 [13], Inceptionv3 [14], ResNet18_v1 [15], DenseNet121 [16], Mobile-Netv3_small [17], and SqueezeNet1.0 [18], are employed for model selection and comparison.
- The proposed models obtain the test accuracy of 99.6% on the 5-class CXR image dataset, 95.0% on the COVIDx dataset [11], and 99.3% on the COVIDx-CT dataset [12]. The dynamic CNN modification method alleviates the gradient vanishing problem, achieves satisfactory sensitivity, precision, and strong robustness in classification.

The implementation process of this paper is shown in Fig. 1, which will be elaborated in the following Sections. The rest of this paper is arranged as follows. Section 2 introduces the background of this research. Data information is introduced in Section 3. Six widely used CNN architectures applied to the COVID-19 classification are illustrated in Section 4. Section 5 presents the proposed dynamic modification method, especially, the modified MobileNet architecture is proposed for the classification of CXR images. Modified ResNet with applications on COVID-19 CT image classification is elaborated in Section 6. The robustness test is then analysed in Section 7. Section 8 makes the conclusion.

2. Background

2.1. COVID-19 related research

Detecting COVID-19 with deep learning techniques is a trending topic and has attracted extensive attention recently. Promising results using advanced CNNs have been published and continue to emerge in this domain. In the detection of COVID-19, CXR images and CT images are the two main types of datasets used for classification. In addition, wearable sensor signals have been used as inputs for COVID-19 detection recently [19].

In the detection of COVID-19, a variety of deep learning models have been designed and achieved success in medical images classification. Researchers in Ref. [1] proposed a type of deep learning method named DarkCovidNet, which achieved 98.08% test accuracy in binary classification and 87.02% test accuracy in three-category classification. In Ref. [11], COVID-Net, a tailored deep learning model derived through generative synthesis [20]) was used for the detection of COVID-19 cases with CXR images. The authors in this research also compiled CXR images from various open sources and made them available to the general public. For the chest CT images, a type of COVIDNet-CT model was proposed in Ref. [12] to identify COVID-19, non-COVID-19 pneumonia and normal cases using a machine-driven design exploration approach
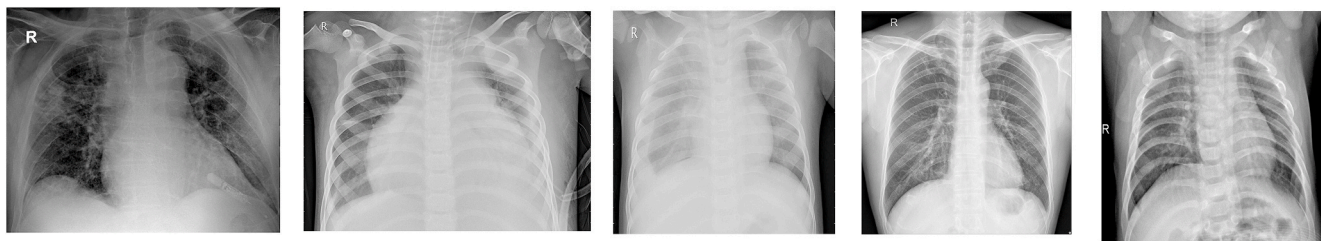
Fig. 2. Example CXR images from data sources [24,36–39].

- COVID-19 CXR images are from the open source GitHub repository ieee8023/covid-chestxray-dataset (DS 1) [36], Actualmed-COVID-chestxray-dataset (DS 2), Fig. 1-COVID-chestx-ray-dataset (DS 3) [37], COVID-19 Radiography Database (DS 4) [24]. Data preparing for COVID-19 images in this paper referred to the code provided in the GitHub repository COVIDx Dataset contributed by Linda Wang et al. [11].
- The images of tuberculosis positive cases are from the dataset [38] which include CXR databases respectively obtained in Shenzhen, China and Montgomery, USA (DS 5), and the data source TB Portals Program, Office of Cyber Infrastructure and Computational Biology (OCICB), National Institute of Allergy and Infectious Diseases (NIAID) (DS 6).
- The bacterial and viral pneumonia CXR images are from Pneumonia Classification Dataset (DS 7) [39].
- The CXR images of normal controls are from COVID-19 Radiography Database (DS 4) [24].

which was similar to the method in Ref. [11]. Another contribution of this research is the introduction of COVIDx-CT, a benchmark CT image dataset derived from CT imaging data which consists of 104,009 images across 1489 patient cases.

However, it must be acknowledged that there are also limitations in COVID-19 detection using deep learning techniques. Firstly, to control the spread of COVID-19 and protect people, many countries have issued a self-isolation policy, and people who begin to have mild symptoms are quarantined without CXR or CT examinations [21]. Thus, in the current CXR and CT image datasets, the number images of patients with severe COVID-19 symptoms is far greater than the number of images of patients suffering from only mild symptoms [21]. Secondly, the related research, especially that was conducted in 2020, suffered from a lack of COVID-19 images and data imbalance. For example, in studies [4,8,21–24], the number of COVID-19 images ranges from 50 to 300, and in some studies the number of data in different categories is much imbalanced. Thirdly, there are no standard criteria in model evaluation, which leads to less effective comparison of the performance among different models. Also, with the increasing number of deep learning methods proposed in this domain, it would be more and more difficult for researchers and health organisations to select the most appropriate classification method for COVID-19 detection [25].

Motivated by these factors, this paper proposes a dynamic CNN modification method which are applied to MobileNetv3 and ResNet18 for the classification of CXR and CT images. The datasets used in this paper are from a combination of different data sources, in which the number of samples is balanced and sufficient for model training. In addition, to provide fair and comprehensive comparison with other methods, we used two benchmark datasets, COVIDx [11] and COVIDx-CT [12], to evaluate the proposed method. In particular, we used the same performance indicators with [11,12] for result comparison, and the same training and testing datasets in comparative studies.

2.2. Modern CNN models

The preliminaries of CNN architectures adopted in this paper are introduced in this section.

As one of the most important and prominent models in deep learning methods, CNNs have shown advantages in many application areas, such as computer vision, speech recognition and medical diagnosis [26–28]. In the ImageNet LSVRC-2012 competition, a revolutionary CNN architecture named AlexNet demonstrated that deep CNNs are able to achieve excellent performance on highly challenging datasets with purely supervised learning [29]. Alex et al. developed a wide range of network settings and training skills, such as ReLU, dropout, pooling and local response normalization, which made it possible to train deep CNNs more

effectively and achieve better performance [29,30].

In recent years, more advanced networks based on AlexNet were created, such as VGG, GoogLeNet, ResNet, DenseNet, MobileNet, SqueezeNet, etc.. In 2013, Network in Network (NiN) structure was proposed in Ref. [31], which introduced 1×1 convolutional layer to act as fully-connected layers on the channels, such that each basic block in the model is like a complete network. VGG is proposed in Ref. [13] which offers a template for using loops and subroutines to design new networks. Taking advantages of repeated convolutional blocks proposed in VGG and the structure of NiN, GoogLeNet combines convolution kernels with different size, uses Inception blocks and employs 1×1 convolutions to reduce channel dimensionality [32]. In 2015, a novel type of CNN architecture named ResNet was proposed in Ref. [33], which profoundly influence network structure design thereafter. ResNet realized identity mapping through inserting shortcut connections between layers, which well-defines the function complexity for adding new residual blocks and obviously improves CNNs' classification performance [33]. DenseNet extended the architecture of ResNet by using concatenation as the cross-layer connections, and make each layer densely connected to the last layer in these connections [16].

The above CNN architectures mainly focus on improving model accuracy. Another stream in this field was developed to improve the training efficiency of CNNs, which focused more on reducing computational budget of CNNs with reasonable compromising of accuracy. MobileNet is a typical model designed in this context. The first version of MobileNet was proposed in 2017, aiming at achieving the resource and accuracy tradeoffs. It uses the depthwise separable convolutions to build light weight deep neural networks [34] and is designed for training on mobile devices. Another commonly used model is SqueezeNet, in which the amount of weights was reduced to 50 times less than AlexNet but the accuracy was kept almost the same as AlexNet [18].

Recently published studies [4,8,21–24,35] have shown that VGG [13], Inception [14], ResNet [15], DenseNet [16], MobileNet [17] and SqueezeNet [18] demonstrated effectiveness in COVID-19 detection. Therefore, these models are employed in this paper for model selection and comparative studies.

3. Data information

Four classes of lung diseases are considered in this paper: COVID-19, bacterial pneumonia, viral pneumonia (except for COVID-19) and tuberculosis. In addition, healthy cases are included as the fifth class. The representative samples of each class are shown in Fig. 2. The dataset used in this paper is the combination across four publicly available data sources (DS):

The distribution of samples across 7 data sources and the split of

Table 1

Sample distributions and the split of training and testing sets (DS denotes data sources).

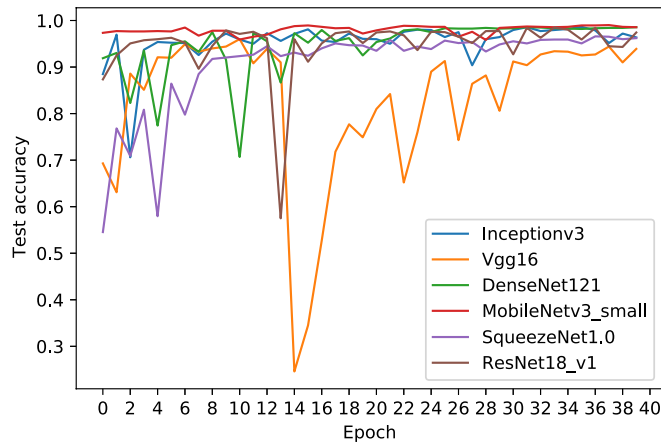| Classes	Total samples	Training	Testing	DS1	DS2	DS3	DS4	DS5	DS6	DS7
COVID-19	1770	1570	200	478	58	35	1199			
Tuberculosis	1436	1236	200					394	1042	
Normal	1341	1141	200				1341			
bacterial pneumonia	1700	1500	200							1700
viral pneumonia	1345	1145	200							1345
Total	7592	6592	1000	478	58	35	2540	394	1042	3045

Fig. 3. Validation accuracy for six CNN models on the Chest X-Ray image dataset.

train, validation and test sets are shown in Table 1.

The classification of COVID-19 cases using CT images is also investigated in this paper. The data information of CT images for COVID-19 detection refers to Section 6.

4. CNNs applied to the COVID-19 detection

In order to obtain the most suitable model for the classification task,

we first employed six widely used CNN architectures which have proven to be successful in computer vision and medical diagnosis.

As mentioned in Section 2, we employed VGG16 [13], Inceptionv3 [14], ResNet18_v1 [15], DenseNet121 [16], MobileNetv3_small [17] and SqueezeNet1.0 [18] for COVID-19 detection in this paper. The model versions are determined by the trade-off of classification performance and computational costs. For example, in our experiments, ResNet18 and ResNet50 reached a similar accuracy in classification. In this case we chose ResNet18 rather than ResNet50 due to its greater parameter efficiency and lower computational cost. Also, the models were selected based on the classification performance reported in the related literature. During the experiments, the above models were set to be the pretrained mode. Therefore, they possess the weights and hyper-parameters fine tuned on ImageNet. As 5 classes are investigated in this paper, the ouput layer in the above models are revised to have 5 nodes.

The classification results are depicted in Fig. 3. As the models are pretrained, almost all models in Fig. 3 achieve above 80% test accuracy after 30 epochs. The test accuracy and the size of parameters of each model are listed in Table 2.

According to the test accuracy shown in Table 2, both DenseNet121 and MobileNetv3_small have a test accuracy of 98.8%. However, from the validation accuracy shown in Fig. 3, it can be viewed that DenseNet121 has a larger oscillation in the training process compared to the MobileNet. In addition, the parameters of DenseNet121 are about four times the size of MobileNet, which indicates that MobileNet has greater parameter efficiency. Therefore, MobileNetv3_small proposed in Ref. [17] is selected as the backbone model for the CXR image

Table 2

Classification performance of six widely used CNN architectures: VGG, Inception, ResNet, DenseNet, SqueezeNet and MobileNet.

PerformanceCNNs	VGG16	Inceptionv3	ResNet18_v1	DenseNet121	SqueezeNet1.0	MobileNetv3_small
Test Accuracy (%)	93.9	96.9	96.5	**98.8**	94.9	**98.8**
Params (MB)	537.0	87.4	44.8	28.2	3.0	6.7
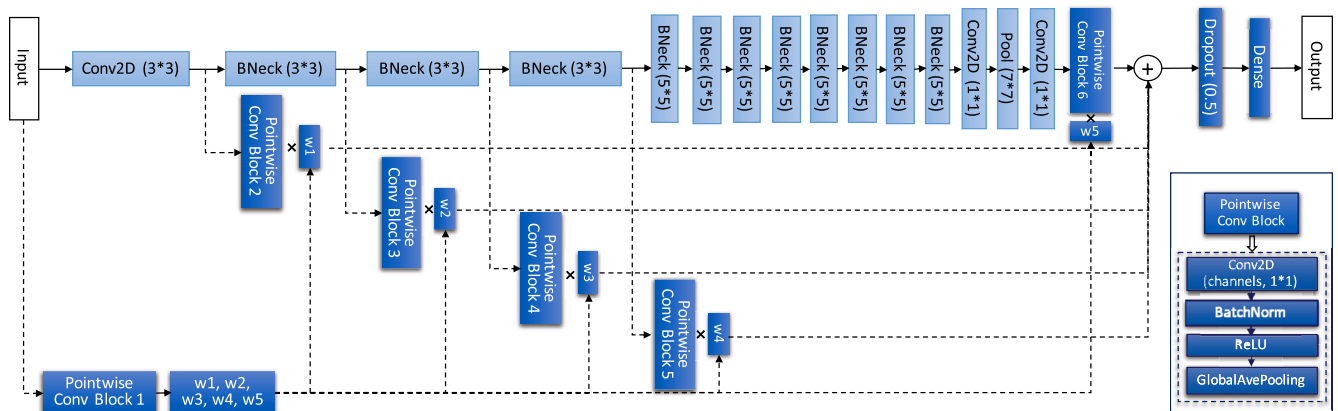
Fig. 4. The structure of the Modified MobileNet. The contents of the original MobileNetv3_small are kept and denoted by light blue, whereas the dark blue blocks represent the modified parts. $w_1, ..., w_5$ are the weights respectively multiplied by the outputs of each pointwise conv block. $\times$ represents element-wise multiplication, $\oplus$ denotes the addition of five outputs of pointwise conv blocks multiplied by the corresponding weights $w_1, ..., w_5$.

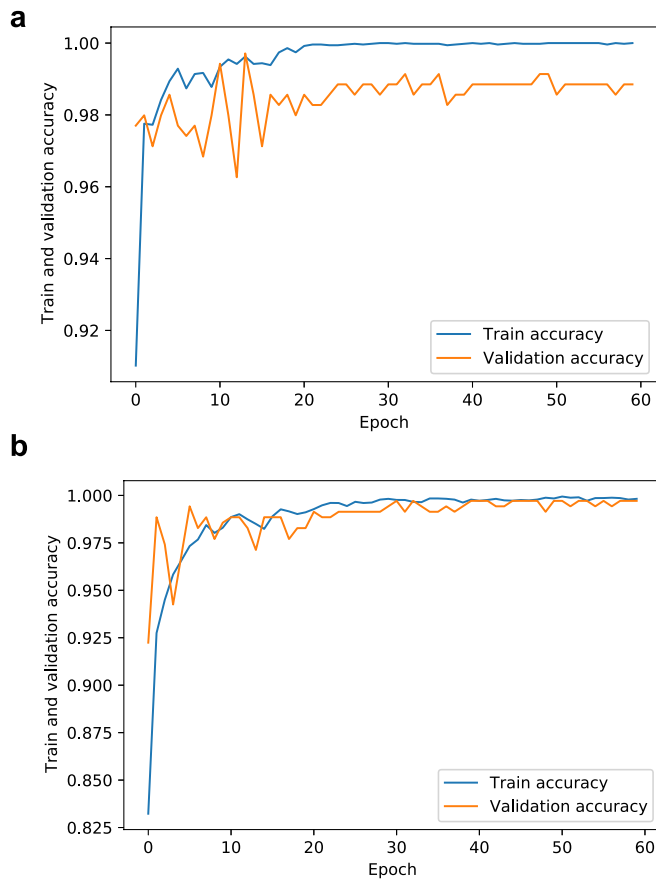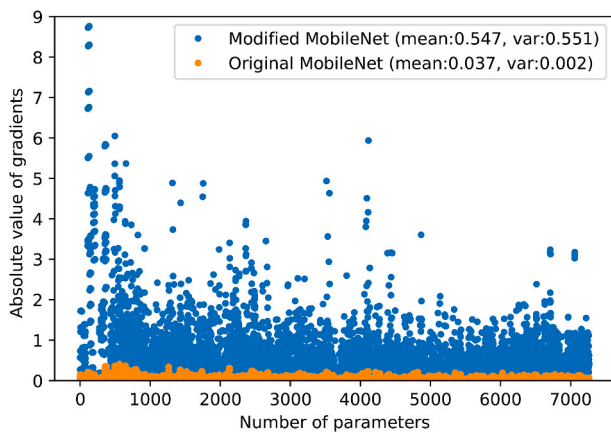**Fig. 5.** Training and validation accuracy of original MobileNetv3_small and modified MobileNetv3_small.

Fig. 6. Gradient values of the first eight convolutional layers in the original and modified MobileNet.

classification in this paper.

However, in the process of training MobileNet, we found that overfitting problem occurred which can be viewed from the divergence of the training and validation accuracy shown in Fig. 5(a). We also found that the gradients in the top four blocks of MobileNet were very small along with the training. To improve the classification performance of MobileNet and address the over-fitting problem, we proposed a type of modified MobileNet for the COVID-19 detection.

5. Modified MobileNet with applications on COVID-19 chest X-Ray image classificaiton

Motivated by the design of ResNet and the channel attention mechanism in SENet [40], we proposed a type of modified MobileNet based on the architecture of MobileNetv3_small in Ref. [17]. In the original MobileNetv3_small, as mentioned above, vanishing gradient and overfitting problems occurred, which are shown in Figs. 5 and 6. Furthermore, there is no residual connection in the top three blocks of MobileNetv3_small as the stride of these three layers is 2. Based on the above observations, we designed a type of modified MobileNet to make the model more adaptive to our dataset.

The structure of the proposed method is depicted in Fig. 4, the corresponding configuration of hyper-parameters is shown in Table 3. In the modified MobileNet, the outputs of the top four blocks of the original model are processed by a pointwise convolution block, which is deigned to reduce the channel dimensionality and keep the ouputs of different layers summable. The outputs of five pointwise conv blocks are multiplied by the corresponding weights (w_1, ..., w_5). As the weights are variable during training, the outputs of pointwise conv blocks are dynamically weighted and added. The weights are in a 1×5 vector which is the output of the first pointwise conv block with five channels fed by the inputs. The configuration of hyper-parameters can be found in Table 3. With this modified CNN architecture, we aim to:

- Combine low-level and high-level information of the original network.
- Address the overfitting problem occurred in the original MobileNet (Fig. 5(a)).
- Solve the gradient vanishing problem during training (Fig. 6).

In essence, the proposed CNN structure is changed dynamically according to the input data. It is because the outputs of pointwise conv blocks (Fig. 4) are multiplied by the corresponding weights w_i where the weights are obtained from a pointwise conv block. As the weights are variable and determined by the inputs, the outputs of pointwise conv blocks are dynamically weighted and combined. Therefore, this CNN modification method provides dynamic combinations of pointwise conv blocks which are adaptive to the inputs.

In the original MobileNet, the output layer is a convolutional layer with 1024 input channels and 1×1 convolution kernels. This design achieved satisfactory classification results in ImageNet with 1000 classes, however, the large amount of parameters leads to extra computational burdens and over-fitting problem for the 5-class COVID-19 classification task. The pointwise conv blocks in the branch circuits are used to first reduce the number of channels through the convolutional layer with 64 channels and 1×1 filters, and secondly reduce the spatial

Table 3
Configuration of hyper-parameters used for training the original and modified MobileNet.

ModelsParameters	Input size	Batch size	Loss function	optimizer[a]	epochs	channels in PCB[b]
Original MobileNet	$299 \times 299 \times 3$	128	softmax cross entropy	adam + sgd	40	–
Modified MobileNet	$299 \times 299 \times 3$	128	softmax cross entropy	adam + sgd	40	5, 64[c]

[a] The first 20 epochs: adam; the second 20 epochs: sgd.
[b] Pointwise Conv Block.
[c] The number of channels in PCB1: 5; PCB2 – PCB6: 64.

Table 4
The validation and test accuracy obtained in the 5-fold cross validation. The test accuracy is obtained using the model achieved the best validation accuracy in the 5-fold cross validation.

Folds	Cross Validation (data partition)					Validation Acc	Test Acc
Fold 1	20%	20%	20%	20%	20%	99.1%	99.6%
Fold 2	20%	20%	20%	20%	20%	98.7%	
Fold 3	20%	20%	20%	20%	20%	98.8%	
Fold 4	20%	20%	20%	20%	20%	99.1%	
Fold 5	20%	20%	20%	20%	20%	99.2%	

Table 5
Test accuracy of the modified MobileNet and original MobileNet in three classification scenarios: 3 categories: COVID-19, Non-COVID-19 infection, healthy; 4 categories: COVID-19, Non-COVID-19 pneumonia, tuberculosis, healthy; 5 categories: COVID-19, bacterial pneumonia, viral pneumonia (with the exception of COVID-19 infection), tuberculosis, healthy.

ClassesModels	Modified MobileNet	Original MobileNet
5 classes	99.6%	98.8%
4 classes	99.9%	99.1%
3 classes	99.7%	99.6%

Table 6
Precision, sensitivity and F1-score obtained for the proposed model on the test set.

Class	Precision	Sensitivity	F1-score	Support
COVID-19	1.000	1.000	1.000	200
Tuberculosis	1.000	1.000	1.000	200
Bacteria Pneumonia	0.995	0.995	0.995	200
Normal Controls	0.985	1.000	0.993	200
Viral Pneumonia	0.995	0.980	0.987	200

modified MobileNet, the mean value and variance of the absolute value of gradients are 0.547 and 0.551. Fig. 6 demonstrates that this type of modification in MobileNet is able to alleviate the gradient vanishing problem which usually happens when networks are deep. Code files of the CNN modification method can be found in the github repository[1].

5.1. Classification results

The training and validation accuracy of the modified MobileNet for the classification of COVID-19, tuberculosis, bacterial pneumonia, viral pneumonia and healthy cases are shown in Fig. 5(b). We adopted 5-fold cross validation for result evaluation. From Table 4, it can be noted that after using the modified MobileNet, the test accuracy reaches 99.6% which is higher than that of the original MobileNet by 0.8%. The overfitting problem is alleviated using the modified MobileNet, which can be seen by comparing Fig. 5(a) and (b). Sensitivity, precision, and F1-score of the original and proposed models are shown in Table 6. It can be noted from Tables 6 and 4 that the proposed model has achieved an accuracy of 99.6% in the five-category classification, and obtained sensitivity, precision and F1-score values of 100% in COVID-19 detection.

To test the effectiveness of the proposed method, we also applied the proposed method to the classification of 3 categories: COVID-19, healthy, and Non-COVID-19 infections; four categories: COVID-19, Non-COVID-19 pneumonia, tuberculosis and normal controls. The results are summarized in Table 5. These two scenarios could be implemented faster and more convenient in practice when rapid diagnosis is required. It can be noted from Table 5 that the proposed method obtains 99.7% test accuracy in three-category classification, 99.9% in four-category classification, and 99.6% in three-category classification, which all outperform the original MobileNet.

To further observe the rationality and effectiveness of the proposed classifier, the class activation mapping (CAM) proposed in Ref. [41] was
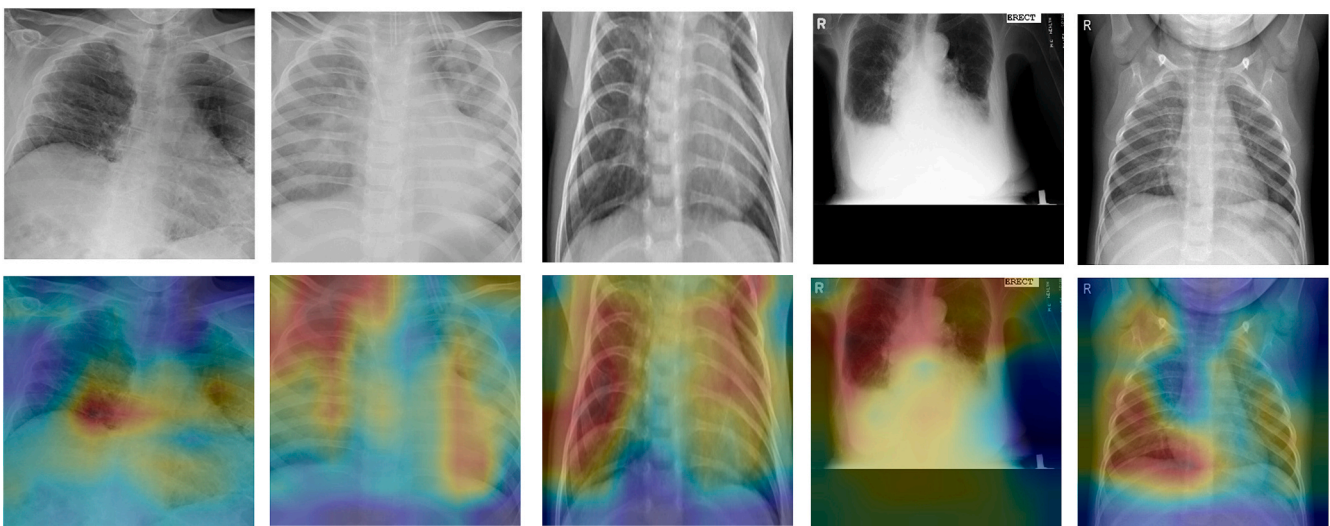

Fig. 7. Class activation maps (CAM) of CXR images in each category. Images in the first row are the original CXR images of each class, the second row lists the corresponding CAM in which the highlighted parts represent the discriminative image regions that the classifier uses to identify that category.

dimension of inputs by global average pooling.

Also, we computed the gradients of weights in convolutional layers which are influenced by the modification structure, and compared them with the gradients in corresponding layers of the original MobileNet. Fig. 6 shows the distribution of absolute values of gradients in the first eight convolutional layers. The number of parameters of corresponding kernels is 7264. The mean value and variance of the absolute value of gradients in the original MobileNet are 0.037 and 0.002. For the

adopted to localize the discriminative image regions. In Fig. 7, the highlighted regions show the areas that the classifier used to identify that category.

---

[1] https://github.com/Guangyu05/COVID-19-CXR-CT-Classification

Table 7
Comparison of the proposed method and COVID-Net [11] on COVIDx: Test accuracy (%) and Parameters (M).

Architectures	Params (M)	Acc (%)
COVID-Net	11.8	93.3
Modified MobileNet	**7.1**	**95.0**

Table 8
Comparison of the proposed method and COVID-Net [11] on COVIDx: **Sensitivity (%)**.

Architectures	COVID-19	Non-COVID19	Normal
COVID-Net	91.0	**94.0**	95.0
Modified MobileNet	**96.0**	92.0	**97.0**

Table 9
Comparison of the proposed method and COVID-Net [11] on COVIDx: **Precision (%)**.

Architectures	COVID-19	Non-COVID19	Normal
COVID-Net	98.9	91.3	90.5
Modified MobileNet	**99.0**	**95.0**	**92.0**

5.2. Result discussion

Comparing Fig. 5(a) and (b), it can be found that the modified structure alleviates the over-fitting problem which occurred in the original MobileNet. The main reasons behind this are analysed as follows: (a) The poinwise conv blocks in the modified MobileNet uses 1×1 convolutional layer with 64 channels as well as global average pooling to reduce the dimensions of the corresponding outputs, which makes the input fed into the last layer to have a size of $64 \times 1 \times 1$ rather than $1024 \times 1 \times 1$ in the original MobileNet; (b) the introducing of a dropout layer after the added outputs from former layers improves the generalisability of the proposed model; (c) this structure is more capable of solving the vanishing gradient problem in back propagation due to the connections with former layers, which is shown in Fig. 6; (d) the proposed structure combines the features obtained from lower layers and higher layers, which provides more comprehensive information for the classification of the final layer.

5.3. Comparison with COVID-Net

To provide comprehensive evaluation of the proposed method, we compared the proposed method with COVID-Net [11], a recently proposed model for the detection of COVID-19 cases using CXR images. To provide fair comparisons, we employed the benchmark dataset COVIDx to evaluate both methods.

COVID-Net [11] is a deep convolutional neural network designed for

the detection of COVID-19 cases on COVIDx, an open access benchmark dataset containing 13,975 CXR images. In Ref. [11], a machine-driven design exploration strategy was used to create COVID-Net. The test accuracy and parameters of the proposed method and COVID-Net are shown in Table 7, and the comparisons of sensitivity and precision are shown in Tables 8 and 9.

From Tables 7–9, it can be concluded that the proposed method outperforms COVID-Net in terms of computational memory, accuracy (1.7% improvement on test accuracy), sensitivity, and precision in COVID-19 detection. It further demonstrates the effectiveness of the proposed method in diagnosing COVID-19 cases with CXR images.

6. Modified ResNet with the application on COVID-19 CT image classification

To further investigate the effectiveness of this type of model modification, we applied the same modification method to ResNet18 [33] and proposed a type of modified ResNet. This version of ResNet is chosen for the reason that it is the residual network with the fewest layers and parameters. The experiment was conducted using a benchmark CT image dataset COVIDx-CT which consists of 104,009 chest CT images across 1489 patients. Example CT images of this dataset are shown in Fig. 8, the information of the dataset can be found in Ref. [12]. Similarly, the comparisons of six CNN architectures are conducted for model selection. The test accuracy, parameters of different models and the training process are shown in Table 10 and Fig. 9.

After comparing the performance of VGG16 [13], Inceptionv3 [14], ResNet18_v1 [15], MobileNetv3_small [17], DenseNet121 [16], and SqueezeNet1.0 [18], we found that both InceptionNetv3 and ResNet18 achieved satisfactory classification performance. Considering the parameter efficiency of ResNet18 and InceptionNetv3, we adopted ResNet18 as the CNN backbone, and modified it with a similar method to that used in the modified MobileNetv3_small (Section 5). The architecture of the modified ResNet18 is depicted in Fig. 10 and the corresponding configuration of hyper-parameters is shown in Table 11.

6.1. Comparison with COVIDNet-CT

In this section, we compared the modified ResNet18 with COVIDNet-CT [12] on the benchmark dataset COVIDx-CT. COVIDNet-CT [12] is tailored for the detection of COVID-19 cases from chest CT images via a deep convolutional neural network.

The results of COVIDNet-CT in the classification of COVID-19, Non-COVID-19 pneumonias, and normal controls are listed in Ref. [42] (COVID-Net CT-2 S (2 A)). To provide fair comparisons, we used the same training and test datasets as that of COVIDNet-CT [12] in our experiment.

The comparative results can be found in Tables 12–14. The proposed method achieves 99.3% test accuracy on the COVIDx-CT dataset which consists of 143,778 training images, 25,486 validation images, and 25,658 test images [12]. The modified ResNet18 possesses higher test
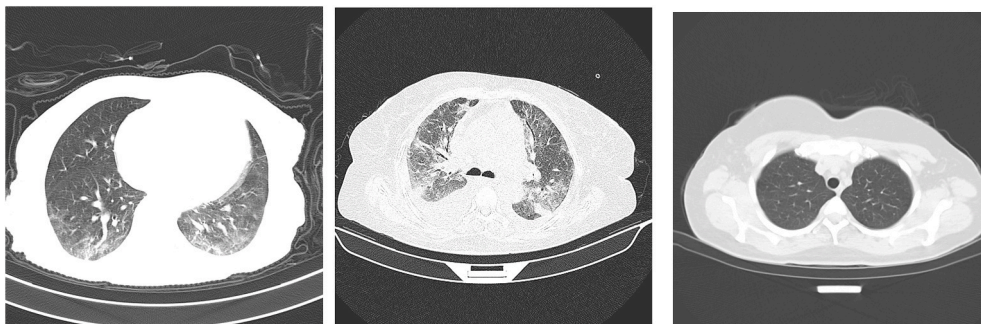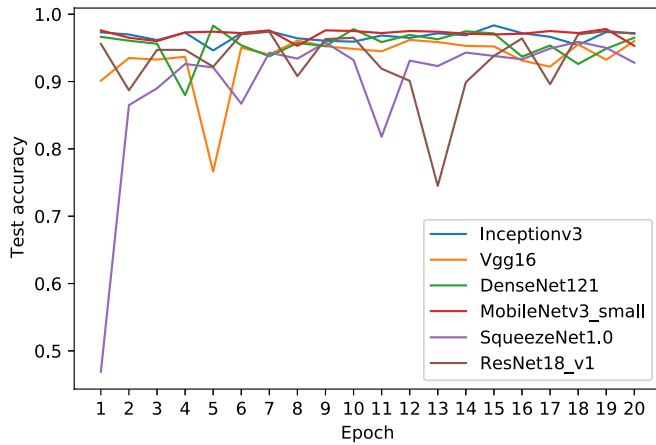

Fig. 8. Example CT Images of three classes [12].

Table 10

Classification performance of six widely used CNN architectures: VGG, Inception, ResNet, DenseNet, SqueezeNet and MobileNet.

PerformanceCNNs	VGG16	Inceptionv3	ResNet18_v1	DenseNet121	SqueezeNet1.0	MobileNetv3_small
Test Accuracy (%)	97.80	**98.59**	**98.47**	96.30	94.38	96.79
Param (MB)	537	87.3	44.8	28.2	2.95	6.74

Fig. 9. Validation accuracy for commonly used CNN models on the CT image dataset.

accuracy (1.4% higher than COVIDNet-CT) and outperforms COVIDNet-CT in sensitivity and precision in identifying three categories. However, the advantage of COVIDNet-CT is its higher parameter efficiency in terms of much fewer parameters in the network.

7. Robustness test

Robustness analysis aims to evaluate the model's capability to resist input perturbations. To test the robustness of the model, we add different levels of Gaussian noise to the test data and feed the contaminated data into the classifiers. In this section, we adopt modified MobileNet in Section 5 to analyse the robustness of the proposed method. Assume that the inputs are denoted by a which has the size of $s_1 \times s_2 \times 3$. To test the robustness of the model, we utilized two forms of

Table 12

Comparison between the proposed method and COVIDNet-CT [12] on COVIDx-CT: Test accuracy and Parameters.

Architectures	Params	Acc (%)
COVIDNet-CT	447.6K	97.9
Modified ResNet	44.9 M	**99.3**

Table 13

Comparison between the proposed method and COVIDNet-CT [12] on COVIDx-CT: **Sensitivity (%)**.

Architectures	COVID-19	Non-COVID19	Normal
COVIDNet-CT	95.7	98.1	98.9
Modified ResNet	**99.2**	**99.8**	**99.1**
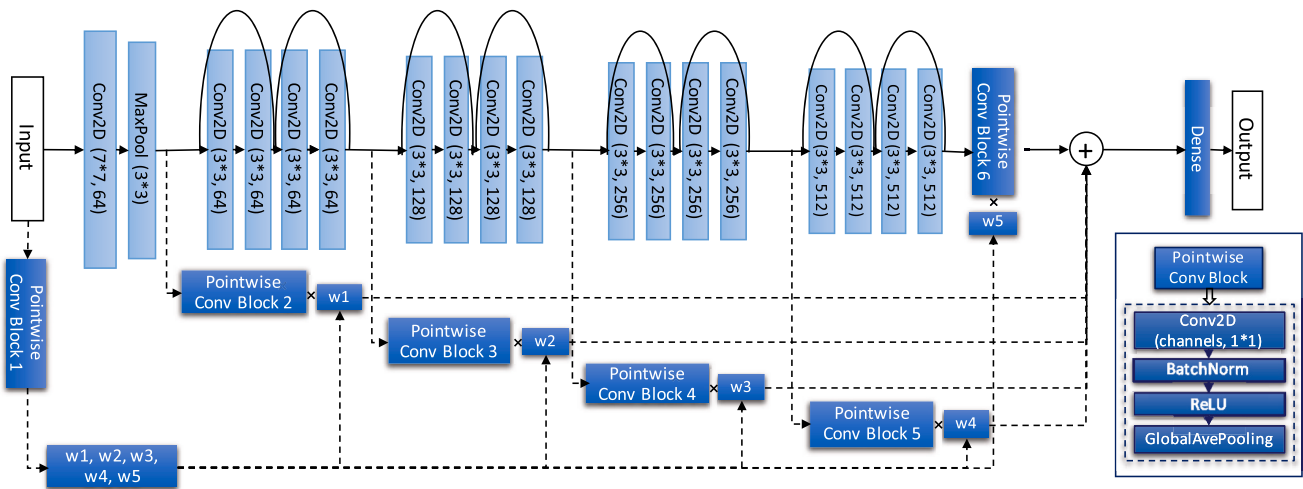
Fig. 10. The structure of the Modified ResNet18. The contents of the original ResNet18 are kept and denoted by light blue, whereas the dark blue blocks represent the modified parts. $w_1, ..., w_5$ are the weights respectively multiplied by the outputs of each pointwise conv block. $\times$ represents element-wise multiplication, $\oplus$ denotes the addition of five outputs of pointwise conv blocks multiplied by the corresponding weights $w_1, ..., w_5$.

Table 11

Configuration of hyper-parameters used for training the original and the modified ResNet.

ModelsParameters	Input size	Batch size	Loss function	optimizer	epochs	channels in PCB[a]
Original ResNet	$256 \times 256 \times 3$	128	softmax cross entropy	adam	20	–
Modified ResNet	$256 \times 256 \times 3$	128	softmax cross entropy	adam	20	5, 32[b]

[a] Pointwise Conv Block.
[b] The number of channels in PCB1: 5; PCB2 – PCB6: 32.

Table 14

Comparison between the proposed method and COVIDNet-CT [12] on COVIDx-CT: **Precision (%)**.

Architectures	COVID-19	Non-COVID19	Normal
COVIDNet-CT	96.4	97.0	99.3
Modified ResNet	**98.6**	**99.4**	**99.7**

Table 15

Robustness analysis of noise form (1) on test dataset.

Noise LevelModel	Original MobilNet	Modified MobileNet
$m = 1$	97.72%	98.58%
$m = 3$	95.45%	97.22%
$m = 5$	73.92%	87.95%
$m = 7$	32.62%	69.37%
$m = 10$	19.60%	46.53%

Table 16

Robustness analysis of noise form (2) on test dataset.

Noise LevelModel	Original MobilNet	Modified MobileNet
$n = 0.5,\ \ m = 2$	83.0%	87.8%
$n = 0.7,\ \ m = 2$	92.4%	95.7%
$n = 0.9,\ \ m = 2$	97.7%	98.5%
$n = 1.1,\ \ m = 2$	98.2%	99.1%
$n = 1.3,\ \ m = 2$	97.3%	97.6%

perturbations. The first one is the additive form:

$$\bar{a} = a + m \cdot \mathrm{randn}(s_1, s_2, 3), \tag{1}$$

where $\bar{a}$ denotes the contaminated data, m denotes the noise levels which are positive values, randn() represents the function of Gaussian noise and '·' denotes the element-wise multiplication. The noise level is respectively chosen as $m = 1, 3, 5, 7, 10$.

The second form is the combination of multiplicative and additive noise:

$$\bar{a} = n \cdot a + m \cdot \mathrm{randn}(s_1, s_2, 3), \tag{2}$$

where n denotes the multiplicative noise level, other denotations are the same as equation (1). The following levels of noise are considered: $n = 0.5, 0.7, 0.9, 1.1, 1.3; m = 2$.

Note that through data perturbations, the pixel values of input images may exceed [0, 255]. We limited the pixel values in [0, 255] by clipping the exceeded values. The robustness results of two forms of noise on test dataset are shown in Tables 15 and 16.

From Tables 15 and 16, it can be concluded that the proposed model possesses stronger robustness compared to the original model, which demonstrates that the proposed method has stronger generalisation ability to accept wider range of data and its suitability to clinical applications when having inevitable data perturbations.

8. Conclusion

In this paper, a type of dynamic CNN modification method is proposed for the classification of two COVID-19 CXR image datasets and a CT image dataset. The proposed method establishes connections between different layers of the original CNN architecture through pointwise convolution blocks, which achieve dynamic combinations of different layers. Six widely used deep learning algorithms, as well as two recently published models specifically designed for COVID-19 detection, are employed and compared with the proposed method. Three scenarios of the classification problem are investigated using the proposed

method. The results are analysed through test accuracy, sensitivity, precision, robustness test, and class activation maps. The modified CNN architecture demonstrates satisfactory classification performance in our comparative study, which shows its potential to be applied in clinical settings for computer-aided diagnosis of COVID-19 positive cases. Future work may investigate the impacts of image qualities on COVID-19 detection due to the difference of image sources.

Declaration of interests

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] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, Automated detection of covid-19 cases using deep neural networks with x-ray images, Comput. Biol. Med. (2020) 103792.

[2] 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.

[3] D. Verma, C. Bose, N. Tufchi, K. Pant, V. Tripathi, A. Thapliyal, An efficient framework for identification of tuberculosis and pneumonia in chest x-ray images using neural network, Procedia Computer Science 171 (2020) 217–224.

[4] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of Covid-19 in Chest X-Ray Images Using Detrac Deep Convolutional Neural Network, Applied Intelligence, 2020, pp. 1–11.

[5] J. Zhang, Y. Xie, Y. Li, C. Shen, Y. Xia, Covid-19 Screening on Chest X-Ray Images Using Deep Learning Based Anomaly Detection, arXiv preprint arXiv:2003.12338.

[6] H. Burdick, C. Lam, S. Mataraso, A. Siefkas, G. Braden, R.P. Dellinger, A. McCoy, J.-L. Vincent, A. Green-Saxena, G. Barnes, et al., Prediction of respiratory decompensation in covid-19 patients using machine learning: the ready trial, Comput. Biol. Med. 124 (2020) 103949.

[7] D. Das, K. Santosh, U. Pal, Truncated inception net: covid-19 outbreak screening using chest x-rays, Physical and engineering sciences in medicine (2020) 1–11.

[8] 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.

[9] R.M. Pereira, D. Bertolini, L.O. Teixeira, C.N. Silla Jr., Y.M. Costa, Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios, Comput. Methods Progr. Biomed. (2020) 105532.

[10] S. Hassantabar, M. Ahmadi, A. Sharifi, Diagnosis and detection of infected tissue of covid-19 patients based on lung x-ray image using convolutional neural network approaches, Chaos, Solitons & Fractals 140 (2020) 110170.

[11] L. Wang, Z.Q. Lin, A. Wong, Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, Sci. Rep. 10 (1) (2020) 1–12.

[12] H. Gunraj, L. Wang, A. Wong, Covidnet-ct: a tailored deep convolutional neural network design for detection of covid-19 cases from chest ct images, Front. Med. 7.

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

[14] 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, 2016, pp. 2818–2826.

[15] K. He, X. Zhang, S. Ren, J. Sun, Identity mappings in deep residual networks, in: European Conference on Computer Vision, Springer, 2016, pp. 630–645.

[16] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

[17] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, et al., Searching for mobilenetv3, in: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 1314–1324.

[18] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, Squeezenet: Alexnet-Level Accuracy with 50x Fewer Parameters and¡ 0.5 Mb Model Size, arXiv preprint arXiv:1602.07360.

[19] S. Hassantabar, N. Stefano, V. Ghanakota, A. Ferrari, G. N. Nicola, R. Bruno, I. R. Marino, N. K. Jha, Coviddeep: Sars-Cov-2/covid-19 Test Based on Wearable Medical Sensors and Efficient Neural Networks, arXiv preprint arXiv:2007.10497.

[20] A. Wong, M. J. Shafiee, B. Chwyl, F. Li, Ferminets: Learning Generative Machines to Generate Efficient Neural Networks via Generative Synthesis, arXiv preprint arXiv:1809.05989.

[21] I.D. Apostolopoulos, S.I. Aznaouridis, M.A. Tzani, Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases, J. Med. Biol. Eng. 40 (2020) 462–469.

[22] E. E.-D. Hemdan, M. A. Shouman, M. E. Karar, Covidx-net: A Framework of Deep Learning Classifiers to Diagnose Covid-19 in X-Ray Images, arXiv preprint arXiv: 2003.11055.

[23] A. Narin, C. Kaya, Z. Pamuk, Automatic Detection of Coronavirus Disease (Covid-19) Using X-Ray Images and Deep Convolutional Neural Networks, arXiv preprint arXiv:2003.10849.

[24] 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? IEEE Access 8 (2020) 132665–132676.

[25] O. Albahri, A. Zaidan, A. Albahri, B. Zaidan, K. H. Abdulkareem, Z. Al-Qaysi, A. Alamoodi, A. Aleesa, M. Chyad, R. Alesa, et al., Systematic review of artificial intelligence techniques in the detection and classification of covid-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects, Journal of infection and public health.

[26] M. Ahmadi, S. Jafarzadeh-Ghoushchi, R. Taghizadeh, A. Sharifi, Presentation of a new hybrid approach for forecasting economic growth using artificial intelligence approaches, Neural Comput. Appl. 31 (12) (2019) 8661–8680.

[27] S. Dorosti, S.J. Ghoushchi, E. Sobrakhshankhah, M. Ahmadi, A. Sharifi, Application of gene expression programming and sensitivity analyses in analyzing effective parameters in gastric cancer tumor size and location, Soft Computing 24 (13) (2020) 9943–9964.

[28] M. Ahmadi, A. Sharifi, S. Khalili, Presentation of a developed sub-epidemic model for estimation of the covid-19 pandemic and assessment of travel-related risks in Iran, Environ. Sci. Pollut. Control Ser. (2020) 1–9.

[29] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[30] S. Igarashi, Y. Sasaki, T. Mikami, H. Sakuraba, S. Fukuda, Anatomical classification of upper gastrointestinal organs under various image capture conditions using alexnet, Comput. Biol. Med. 124 (2020) 103950.

[31] M. Lin, Q. Chen, S. Yan, Network in Network, arXiv preprint arXiv:1312.4400.

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

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

[34] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications, arXiv preprint arXiv:1704.04861.

[35] M. Ilyas, H. Rehman, A. Naït-Ali, Detection of Covid-19 from Chest X-Ray Images Using Artificial Intelligence: an Early Review, arXiv preprint arXiv:2004.05436.

[36] J.P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong, M. Ghassemi, Covid-19 image data collection: prospective predictions are the future, arXiv 2006.11988. URL, https://github.com/ieee8023/covid-chestxray-dataset.

[37] Figure 1 covid-19 chest x-ray dataset initiative. URL https://github.com/agchung/Figure1-COVID-chestxray-dataset.

[38] S. Jaeger, S. Candemir, S. Antani, Y.-X.J. Wáng, P.-X. Lu, G. Thoma, Two public chest x-ray datasets for computer-aided screening of pulmonary diseases, Quant. Imag. Med. Surg. 4 (6) (2014) 475.

[39] D. Kermany, K. Zhang, M. Goldbaum, et al., Labeled optical coherence tomography (OCT) and chest x-ray images for classification, Mendeley data 2 (2).

[40] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.

[41] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning deep features for discriminative localization, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2921–2929.

[42] COVIDx-CT dataset–results. https://github.com/haydengunraj/COVIDNet-CT#results.