



# Diagnosing and differentiating viral pneumonia and COVID-19 using X-ray images

Hakan Kör<sup>1</sup> · Hasan Erbay<sup>2</sup> · Ahmet Haşim Yurttakal<sup>3</sup>

Received: 8 May 2021 / Revised: 29 January 2022 / Accepted: 4 April 2022 /  
Published online: 27 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Coronavirus-caused diseases are common worldwide and might worsen both human health and the world economy. Most people may instantly encounter coronavirus in their life and may result in pneumonia. Nowadays, the world is fighting against the new coronavirus: COVID-19. The rate of increase is high, and the world got caught the disease unprepared. In most regions of the world, COVID-19 test is not possible due to the absence of the diagnostic kit, even if the kit exists, its false-negative (giving a negative result for a person infected with COVID-19) rate is high. Also, early detection of COVID-19 is crucial to keep its morbidity and mortality rates low. The symptoms of pneumonia are alike, and COVID-19 is no exception. The chest X-ray is the main reference in diagnosing pneumonia. Thus, the need for radiologists has been increased considerably not only to detect COVID-19 but also to identify other abnormalities it caused. Herein, a transfer learning-based multi-class convolutional neural network model was proposed for the automatic detection of pneumonia and also for differentiating non-COVID-19 pneumonia and COVID-19. The model that inputs chest X-ray images is capable of extracting radiographic patterns on chest X-ray images to turn into valuable information and monitor structural differences in the lungs caused by the diseases. The model was developed by two public datasets: Cohen dataset and Kermamy dataset. The model achieves an average training accuracy of 0.9886, an average training recall of 0.9829, and an average training precision of 0.9837. Moreover, the average training false-positive and false-negative rates are 0.0085 and 0.0171, respectively. Conversely, the model's test set metrics such as average accuracy, average recall, and average precision are 97.78%, 96.67%, and 96.67%, respectively. According to the simulation results, the proposed model is promising, can quickly and accurately classify chest images, and helps doctors as the second reader in their final decision.

**Keywords** Convolutional neural networks · Transfer learning · Classification · Chest X-Rays · Pneumonia · COVID-19

---

✉ Hakan Kör  
hakankor@hitit.edu.tr

## 1 Introduction

The lungs are among the primary organs of humans' respiratory systems and are located near the backbone on either side of the heart. Their function in the respiratory system is to extract oxygen from the atmosphere and transfer it into the bloodstream, and to release carbon dioxide from the bloodstream into the atmosphere, in the process of gas exchange. The tissue of the lungs can be attacked by bacteria, viruses, and rarely parasites, as result, a lung infection may occur. The infection might be a sign of various diseases such as pneumonia, and nowadays the world's nightmare coronavirus.

Pneumonia is swelling and possible inflammation of the tissue in one or both lungs. It causes the air sacs, or alveoli, of the lungs to fill up with fluid or pus. Traditionally, pneumonia is diagnosed through reviewing the patient's medical history, physical exam, and other diagnostic tests like a blood test, sputum test, pulse oximetry test. Moreover, the chest X-ray is the main tool in diagnosis. People at high risk of having pneumonia are older than 65 or younger than 2 years of age or already have health problems. The disease can be seen at any age, but children under 2 years of age, people with very low immune systems, and people over 65 years of age are at high risk. According to the data of the World Health Organization (WHO), 450 million pneumonia cases are recorded every year in the world [36].

However, the coronavirus, a commonly used name for the disease caused by a new novel coronavirus caused by SARS-CoV-2, was first seen in Wuhan, China in December 2019. The disease has worsened both human health worldwide and the world's economic health. Also, it has nearly collapsed health systems in many countries, thus it has been declared by WHO as a global pandemic [28, 42]. Indeed, the disease tied to the new coronavirus was originally called novel coronavirus-infected pneumonia (NCIP) and later named COVID-19 (Corona Virus Disease 2019) by WHO [20]. COVID-19 negatively affected many people around the world and caused many deaths at the rate of 6.14% globally [16]. COVID-19 resembles almost 79 percent of SARS-CoV, which caused the SARS epidemic that occurred between 2002–2003, also, COVID-19 is 96 percent similar to coronaviruses seen in bats. The pathologic result of SARS and COVID-19 is diffuse alveolar damage with fibrin-rich hyaline membranes and a few multinucleated giant cells [31], but COVID-19 is more deadly than SARS, the number of deaths in COVID-19 has already overtaken that of in SARS epidemic.

Similar to viral pneumonia (non-COVID-19), the main feature of COVID-19, which is understood to be the result of a new mutation, is easy attachment to lung type 2 alveolar cells in humans [7]. COVID-19 test is not possible due to the absence of the diagnostic kit everywhere, even if the kit exists, its false-negative (giving a negative result for a person infected with COVID-19) rate is high. Also, early detection of COVID-19 is crucial to keep its morbidity and mortality rates low. These reasons were the main motivations to develop other diagnostic methods. It is known that X-rays are used to analyze the health of the lungs of COVID-19 and pneumonia patients. However, a specialist radiologist should analyze chest X-rays or computed tomography (CT) scans to observe structural changes in the chest. With the rise of the pandemic, the need for radiologists has been increased considerably not only to detect COVID-19 but also to identify other abnormalities it caused. Moreover, the diagnosis process is also time-consuming and subject to human error.

On the other hand, advances in computer technology play important roles in medicine as in many areas of life. In medicine, computerized systems assist doctors during interpretation, diagnosis, and treatment as a second reader to reduce medical errors and save time [25]. Convolutional Neural Networks (CNNs) have been a dominant method in medical imaging

used in computer-aided diagnostic systems with successful results in classification and feature extraction problems due to their ability to take advantage of Spatio-temporal correlation structures. There are vast number of successful CNN based studies such as, in medical imaging, COVID-19 image classification [37], detection of breast cancer [29], diagnosis of lung lesions [38], detection of hand osteoarthritis [44] and brain tumor classification [18] are some of these studies. To design a successful CNN model diverse, but labeled data corpus is required. Moreover, the quality of the images and the appropriate network are important components [19, 21, 40]. However, in situations where enough data is not available, data augmentation and transfer learning can be used [23, 27]. Transfer learning allows a pre-trained model to be applied to different problems after some modifications to the model's structure and re-training the modified model for fine-tuning using the problems' dataset. It is exceptionally important where the annotated data is scarce. This is a crucial advantage of transfer learning because it is difficult to obtain adequate annotated medical data [8].

The current study offers an automated computerized CNN-based multiclass model to differentiate viral pneumonia, COVID-19, and normal cases using chest radiographs only, without the presence of a specialist. The model takes chest X-ray images and monitors structural differences in the lungs caused by the diseases. The model reaches 98.86% average accuracy, 98.29% average recall, and 98.37% average precision. Moreover, the average false-positive and false-negative rates that significantly impact the pandemic are 0.0085 and 0.0171, respectively. The model was also evaluated with some unused chest images randomly selected from COVID-19 Radiography Dataset [14, 35]. The model accurately predicted 54 of 60 cases, 20 from each case. Thus, the model can be used to assist radiologists as a second reader and is capable of reducing variability in the interpretation of images among the radiologists that may be caused by differences in their experience.

The main advantages and contributions of this research are:

- The proposed model improves the correct detection of pneumonia and differentiating non-COVID-19 pneumonia and COVID-19 and reduces both false-negative and false-positive diagnoses.
- Deep medical imaging is utilized to emphasize the effect of chest X-rays on different cases such as positive COVID-19, non-COVID-19 pneumonia, and normal cases.
- CNNs are capable of extracting radiographic patterns on chest X-ray images to turn into valuable information in medical imaging.
- The advantage of working is the use of pre-trained NasNet-Mobile CNN, an algorithm generated by artificial intelligence. To prevent overfitting, a dropout layer was added to the original network.
- By working on raw data without pre-processing, even images that are difficult to classify were classified accurately and quickly in the shortest time possible. The recommended method is automatic, user-independent.

The rest of the paper is organized as follows. Section 2 overviews the importance of X-rays in monitoring COVID-19 and summarizes studies about COVID-19. Materials and methods utilized in the study are described in Section 3. Section 4 presents simulation results. Comparison to existing studies are given in Section 5.

## 2 Related studies

Recent studies show that coronavirus causes structural effects in the lungs, such as disruption of normal lung structure and transformation of respiratory tissue into fibrotic material

[10], which can be monitored using chest X-rays and CT scans. On the other hand, scientific developments based on deep learning have begun to emerge in medical image analysis, both in extracting features that represent data and in diagnosing diseases. CNNs, which make use of the structural and configurational information available in the images, produce expert-level performances in the fields of medical imaging. For example, Padma and Kumari used convolution 2D techniques applied to open-source data sets of COVID-19. As a result of the study, they proposed a binary classifier that takes chest X-ray images as input to identify COVID-19 with 99% accuracy and 98% verification capability [33]. Moreover, Catak and Sahinbas used 278 positive, 66 negative X-ray images and 39 positive and 30 negative images for the test. In this study, VGG[Visual Geometry Group]16, VGG19, ResNet, DenseNet, and Inception models were used. VGG16 had the highest accuracy rate with 75% success among models [12]. In another study, Shan et al. [39] used VB-Net neural networks to detect COVID-19 infection sites on CT images. The system was trained using 249 COVID-19 patients and validated using 300 new COVID-19 patients. The proposed model gave  $91.6\% \pm 10.0\%$  Dice similarity coefficients between automatic and manual segmentation and an average POI estimation error of 0.3% for the whole lung in the validation dataset [39]. Akram et al. [1] proposed a method that automatically classifies COVID-19 CT images. The method uses features extracted by wavelet transform and fractal texture analysis and applied feature selection with a genetic algorithm. They achieved a 92.6% accuracy rate in the method they used Naive Bayes as a classifier. Khan et al. [26] classified the contrast enhancement preprocessed COVID-19 CT images with the pre-trained DenseNet-201. As a result, they achieved a high accuracy rate of 94.76%. Sahlol et al. [37] proposed a hybrid method using Inception CNN and feature selection algorithm based on marine predators algorithm (MPA) to classify COVID X-ray images. They achieved high accuracy rates of 98.7% and 99.6% in their trials on two different datasets.

Pneumonia has a significant impact on COVID-19 and should be diagnosed urgently with its underlying causes. Mahmud et al. [30], using quite less number of COVID-19 chest X-rays, conducted deep learning-assisted automatic detection of COVID-19 and pneumonia types. They proposed CNN-based architecture, called CovXNet, that uses deep convolution with varying rates of expansion to efficiently extract various features from chest X-rays. At first, chest X-rays of normal and (viral/bacterial) pneumonia patients were used to train CovXNet. Training of this initial phase was carried out by limited chest X-rays, and later with some additional layers, the model trained more with COVID-19, and other pneumonia patients' chest X-ray images. In the proposed method, different forms of CovXNets were designed, trained with X-ray images of various resolutions, and a stacking algorithm was used for further optimization of their estimates. Finally, a gradient-based differential localization was integrated to distinguish the abnormal regions of X-ray images that refer to different types of pneumonia. Extensive experiments using two different datasets, multiple satisfactory detection performance with an accuracy of 97.4% for COVID / Normal, 96.9% for COVID / Viral pneumonia, 94.7% for COVID/Bacterial pneumonia, and 90.2% for multi-class COVID/normal/Viral/Bacterial pneumonias provides [30].

Radiographic patterns in CT chest scans showed higher accuracy compared to RT-PCR detection of COVID-19, which had a relatively low positive detection rate in the early stages, according to WHO. Butt et al. [11] conducted a study comparing multiple convolutional neural network (CNN) models to classify CT samples as COVID-19, Influenza viral pneumonia, or infection-free. In this mentioned study, they developed existing 2D and 3D deep learning models. Combining these with the latest clinical understanding, they compared it with a study that achieved an AUC of 0.996 (95% CI: 0.989-1.00) for non-coronavirus cases, and they calculated 98.2% sensitivity and 92.2% specificity in thoracic CT studies.

Panwar et al. [34] studied the visual indicators found in the lungs. They have proposed nCOVnet, a deep learning neural network-based method, which is an alternative rapid screening method that can be used to detect COVID-19 by analyzing patients' x-rays. Dansana et al. [17] VGG-19, Inception-V2, and convolution neural networks method were used for binary classification pneumonia-based transformation of the decision tree model in the X-ray and CT scan images data set containing 360 images. It was concluded that the precision-tuned VGG-19, Inception-V2, and decision tree model performed quite satisfactorily with an increase in training and verification accuracy (91%), except for the Inception-V2 (78%) and decision tree (60%) models [17]. Table 1 summarizes these studies.

**Table 1** Overview of recent studies related to the current study

Model's description	Model's performance	Model's Dataset
VGG16 [12]	Accuracy: 75%	278 positive and 66 negative COVID 19 X-rays and 39 positive and 30 negative images.
ResNet [46]	Accuracy: 86.7%	618 CT samples
Perform segmentation of COVID-19 infection regions from CT scans using deep learning model named as VB-Net [39]	Dice similarity coefficient: 91.6%	249 CT images in training and 300 in validation.
Employ Details Relation Extraction Neural network (DRE-Net) to obtain the image level predictions of corona19 virus form healthy and bacterial pneumonia [43]	Accuracy: 94%	CT scans images (88 patients of COVID-19, 101 patients of bacterial pneumonia, and 86 healthy persons)
2D CNN network [33]	Accuracy: 98%	224 images in two categories covid and normal images from GitHub datasets.
CovXNets [30]	COVID/Normal: Accuracy: 97.4%, COVID/Viral pneumonia: Accuracy: 97.4%, COVID/ Bacterial pneumonia: Accuracy: 94.7% COVID/normal/Viral/Bacterial pneumonia: Accuracy: 90.2%	305 images for each category.
nCOVnet [34]	Accuracy: 97.62%	127 COVID-19 positive patients' X-ray images for training, on contrary, 31 COVID-19 positive patients' X-ray images for testing.
VGG19 [17]	Accuracy: 91%	360 images of COVID-19 patients, 16 images of SARS and 18 images of Streptococcus pneumoniae.
Naive Bayes [1]	Accuracy: 92.6%	3500 COVID-19, 2400 normal chest scans
DenseNet-201 [26]	Accuracy: 94.76%	62 patients
Inception FO-MPA [37]	Inception: Accuracy: 98.7% FO-MPA: Accuracy: 99.6% accuracy	Dataset 1 200 COVID-19 images 1675 negative images Dataset 2 219 COVID-19 images 1341 negative images



**Fig. 1** Original raw images from the dataset left normal, middle viral pneumonia, and right COVID-19

The current study proposes a multiclass computed model to distinguish between viral pneumonia, COVID-19, and normal cases using chest radiographs only. The model takes raw chest X-ray images and tracks structural differences in the lungs caused by the disease. Performance measures such as accuracy, recall, and precision state that the model is promising.

### 3 Materials and methods

**Dataset** The dataset for the study is a collection of Cohen dataset [15] and Kermany dataset [24]. The latter contains normal (i.e. healthy) and viral pneumonia chest radiographs, while the former contains only COVID-19 images. The distribution of images is 131 normal, 130 viral pneumonia, and 129 COVID-19, almost the same number of images in each class. Figure 1 shows sample image from each class.

Later, the dataset was randomly split into two parts using the stratified selection strategy to allocate %70 of the images for training, and %30 for validation. Table 2 shows the training set and validation set numbers.

**Data augmentation** The performance of CNN-based models highly depends on the amount of diverse data available otherwise, the model might not generalize well i.e. the model produces good results on the seen data (training set) but performs poorly on the unseen data (test set). Conversely, Spatio-temporal correlation structures have been proven to improve the effectiveness of models [3–6]. Therefore, the dataset was augmented artificially to increase the model’s effective size, as image data augmentation not only prevents CNN from learning irrelevant patterns, overfitting, and memorizing exact details of training images but also supports temporal and spatial relationships. Herein, a series of online data augmentation such as reflection, translation, and scaling was applied during training.

**Table 2** Train and validation sets statistics

Groups	Normal	Viral Pneumonia	COVID-19	Total
Training	92	91	90	273
Validation	39	39	39	117
Total	131	130	129	390

**Classification model** Classical deep learning models accept inputs in vectorial form; however, especially in medical images, structural information contained in nearby pixels or voxels is important, and vectorization destroys the structural information. CNNs are aiming to take advantage of spatial and configuration information. They have the ability to learn hierarchical features from the data using a backpropagation algorithm. A typical CNN pipeline in image processing contains a pile of convolution, data reduction, and accumulation layers. Convolution layers are used as small-scale detectors to explore the features of an image. Suppose  $I$  represents an input image of size  $m \times n$  in the image domain  $D$ , where  $I(i, j, 3)$  is pixel value at location  $(i, j)$  with  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . Here, the third coordinate represents the number of channels, and for the sake of being more succinct, it is usually omitted, thus,  $I$  is considered as a two-dimensional function on  $D$ . It is worth mentioning that each operation mentioned below is applied to each channel separately. CNNs are composed of a variety of layers such as convolution layers, pooling layers, fully connected layer

Each convolution layer uses filters, say  $F$  of size  $(2p + 1) \times (2q + 1)$ , to output so-called feature maps (aka activation maps) by convolution operation to  $I$ , that is,

$$(I * F)(i, j) = \sum_{r=-p}^p \sum_{s=-q}^q I(i - r, j - s)F(r, s) \tag{1}$$

where  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . Handling the border values is a matter of choice. General approach is padding to obtain an output of size  $m \times n$ . See [22] for details. The convolutional layer is completed by adding bias term  $b(i, j)$   $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$  followed by applying nonlinear activation function  $\sigma(x)$  to each element of the output in (1) to destroy linearity and fix pixel values. Thus, under the assumption that both input and output of (1) having the same size, we obtain

$$S(i, j) = \sigma((I * F)(i, j) + b(i, j)), \quad i = 1, 2, \dots, m \quad \text{and} \quad j = 1, 2, \dots, n. \tag{2}$$

For the choices of nonlinear activation functions, see [2].

After that, during the pooling layer, a downsampling operation is applied to each feature map. With an equal dimensional pooling of size  $k$ , the windowed area of size  $k \times k$  over the feature map  $S$  is summarized. The most commonly used pooling function  $\rho$  is max pooling since it preserves detected features. The output of the pooling layer, say  $W$ , also depends on the stride value  $s$ , the amount of window moved after each pooling operation, and the padding  $t$ . Thus,

$$W(i, j) = \rho(S, k, s) \tag{3}$$

where  $i = 1, 2, \dots, \lfloor \frac{m+2t-k}{s} \rfloor + 1$  and  $j = 1, 2, \dots, \lfloor \frac{n+2t-k}{s+1} \rfloor + 1$ . We define  $u = \lfloor \frac{m+2t-k}{s} \rfloor + 1$  and  $v = \lfloor \frac{n+2t-k}{s} \rfloor + 1$ . We also define  $\bar{p} = 2p + 1$  and  $\bar{q} = 2q + 1$ . Then, one can represent the result of (1), (2) and (3) in implicit form as

$$W = \Gamma(I_{m \times n}, F_{\bar{p} \times \bar{q}}, b_{m \times n}, S, k, s; *, \sigma, \rho) \tag{4}$$

which is downsampled feature map for the filter  $F$  and is an input to the next layer. We call this sort of combination of layers as a CNN block. If we ask a CNN block to produce  $r$  such outputs we have to provide  $r$  filters, which can be formulated by

$$W^r = \Gamma(I_{m \times n}, F_{\bar{p} \times \bar{q}}^r, b_{m \times n}^r, S^r, k, s; *, \sigma, \rho). \tag{5}$$

Note that (5) produces total of  $r$  downsampled feature maps, one map corresponding to each filter. Moreover, if the initial input image set has  $w$  images then (5) produces  $r \cdot w$  downsampled feature maps.

An efficient CNN model is composed of many of these kinds of CNN blocks, say  $l$ , where the output of one is the input to the next one. We can formulate the following recurrence relation

$$W_i^{r_i} = \Gamma(W_{i-1}^{r_{i-1}}, F_{\tilde{p}_i \times \tilde{q}_i}^{r_{i-1}}, b_{m_i}^{i-1}, S_{i-1}, k_{i-1}, s_{i-1}; *, \sigma, \rho) \tag{6}$$

where  $W_0^{r_0} = I_{m \times n}^w$  and  $i = 1, 2, \dots, l$ .

Later, all of the downsampled feature maps  $W_l^i$ ,  $i = 1, 2, \dots, r_l$  are stretched and concatenated to a single column vector, say  $x$ , and fed into one or more fully connected layers. We denote these layers by  $f$ . The output size of the last fully connected layer is equal to the number of classes, say  $c$ , in the target data. The softmax activation function  $\theta$  normalizes the output so that all output values are positive and sum to one. Then, the predicted value for the input image  $I$  with label  $y$  is given by

$$\hat{y} = \max_{1 \leq i \leq c} \theta(f(x)). \tag{7}$$

During the training process two variables lost function  $\ell$  of  $y$  and  $\hat{y}$  is computed at each data in validation dataset  $\Omega$  of size  $v$  and the following minimization problem

$$\min_{F_{\tilde{p}_i \times \tilde{q}_i}^{r_i}, b_{m_i}^i, 0 \leq i \leq l} \sum_{j=1}^v \ell(y_j, \hat{y}_j). \tag{8}$$

is solved.

This is a good place to point out that when the model is trained from scratch, the recurrence relation (6) is initiated by assigning random values to the contents of the filters  $F_{\tilde{p}_0 \times \tilde{q}_0}^{r_0}$ . However, with the idea of using other’s experiences, so-called transfer learning, we prefer to transfer filters from pre-exists models such as Xception [13], VGG-19 [41], AlexNet [27], and NASNet-Mobile [48]. We constructed models on each of these, but NASNet-Mobile received the highest accuracy. Thus, the proposed classification model was built on NASNet-Mobile. There are 5,326,716 learnable parameters in NASNet-Mobile.

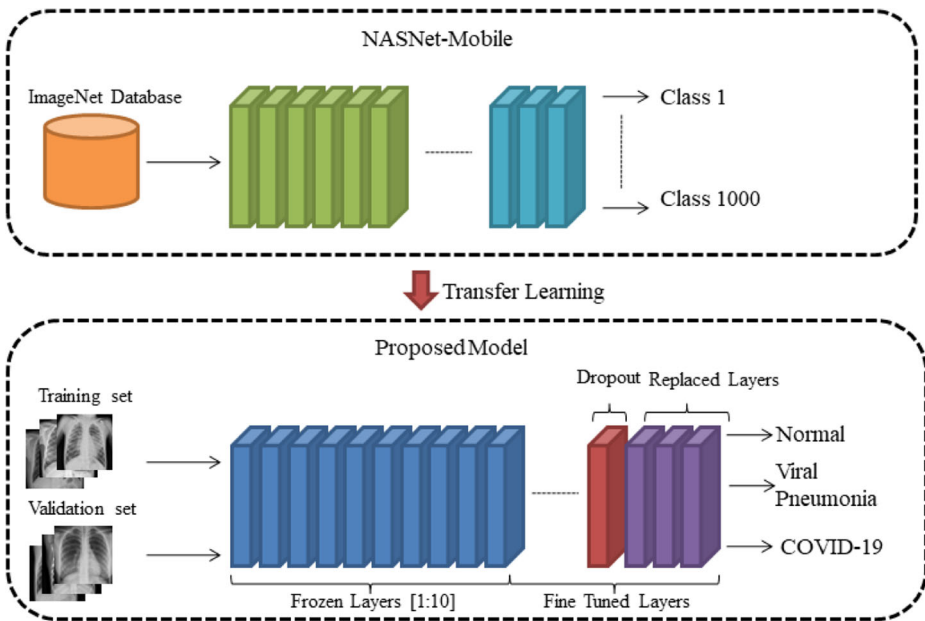
NASNet-Mobile, a state-of-the-art pre-trained network, is a convolutional neural network that had trained on 1.2 million training images from the ImageNet database. The model has a total of 913 layers. The image input size for the network is  $224 \times 224 \times 3$  [48].

The initial construction of the model was NASNet-Mobile. Last three layers named ‘predictions’, ‘predictions\_softmax’, ‘ClassificationLayer\_predictions’ replaced with the appropriate layers, (i.e. fullyConnectedLayer, softmaxLayer, classificationLayer). Local hyperparameters were defined to the final fullyConnectedLayer such that ‘WeightLearnRateFactor’ was assigned 10, and ‘BiasLearnRateFactor’ assigned 10. A dropout layer with dropout probability 0.5 was added just before the final fullyConnectedLayer so only randomly selected neurons enter the following layer. The first 10 layers was kept frozen during training which corresponds to 2018 learnable parameters, however, higher layers were fine-tuned. Figure 2 shows the graphical abstraction of the model.

Before fine-tuning the model, both the training and validation sets went through a sequence of online data augmentation. Each image was reflected horizontally with a 50% probability. Also, all images were translated both horizontally and vertically with translation distance picked randomly from a continuous uniform distribution within the interval  $[-30, 30]$ . Moreover, horizontal and vertical scaling was applied using a scale factor selected randomly from a continuous uniform distribution within the interval  $[0.9, 1.1]$ .

Table 3 shows additional hyperparameters and their assigned values along with frozen weights.





**Fig. 2** Graphical abstraction of the classification model. The model was built on NASNET-Mobile. The classification layer was updated, the dropout layer was added, and the initial layers were kept unchanged

### 4 Simulation results

The implementation, as well as training of the proposed transfer model, was carried out using the programming environment MATLAB®. The model was trained using 273 chest radiography, among them 92 were normal, 90 were COVID-19, and 91 were viral pneumonia. The training accuracy is checked at the beginning and during training at every five epochs via the validation set. The validation set contained 117 chest images distributed evenly among classes normal, viral pneumonia, and COVID-19. Furthermore, the losses for

**Table 3** Model’s training hyperparameters, frozen weights, training elapsed time

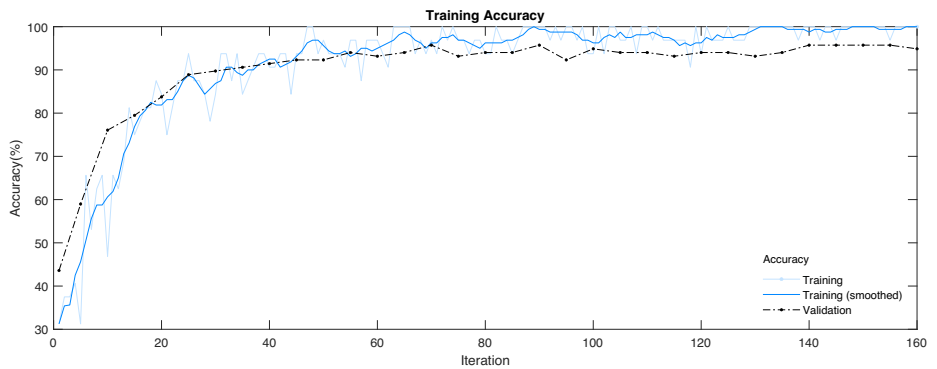
Optimizer	Stochastic Gradient Descent
Momentum	0.900
Initial Learn Rate	1.0000e – 03
L2 Regularization	1.0000e – 04
Gradient Threshold Method	L2 Norm
Gradient Threshold	Inf
Validation Frequency	5
Validation Patience	Inf
Maximum Epochs	20
Mini Batch Size	32
Shuffle	Every Epoch
Frozen Weights (layers)	1:10
Training Time (elapsed)	57 min 11 sec

training and validation were measured using the training and the validation sets, respectively. Figure 3 outlines the training progress, but Fig. 3a shows plots of accuracy for both training and validation, and Fig. 3b shows the same for the loss.

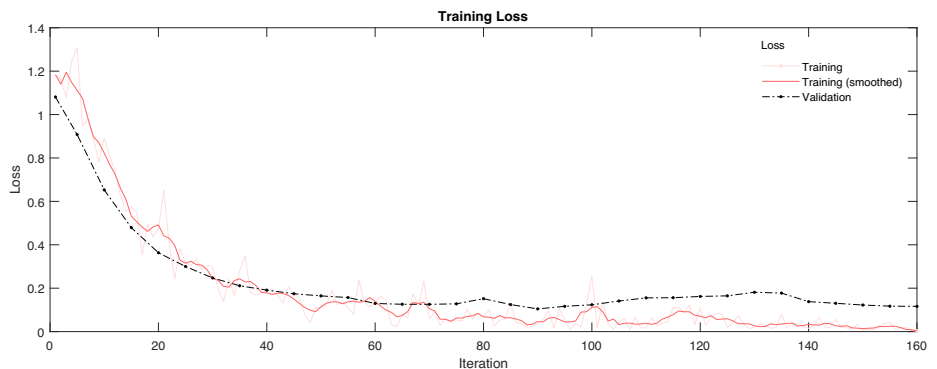
Figure 4 shows the confusion matrix for the validation test. The model correctly predicted 115 of 117 cases with only two mispredictions with two viral pneumonia cases predicted as normal.

Moreover, the efficiency of the transfer model was evaluated using performance metrics such as accuracy, recall, precision, F1 score, false-positive rate, false-negative rate, false discovery rate, false omission rate, critical success index, and Matthews correlation coefficient. These performance metrics are evaluated using the confusion matrices given in Fig. 4, and Table 4 presents the evaluation results along with their macro average values. Note that the model uses only pixel information. The macro average accuracy of the model was 0.9886, while the macro average recall and precision were 0.9829 and 0.9837. However, both false-positive and false-negative rates for COVID-19 were 0.0000.

As mentioned above, throughout our experiments, we constructed models by modifying Xception, VGG-19, and AlexNet, but they received the highest accuracy of 94.87, 93.16, and 94.94, respectively.

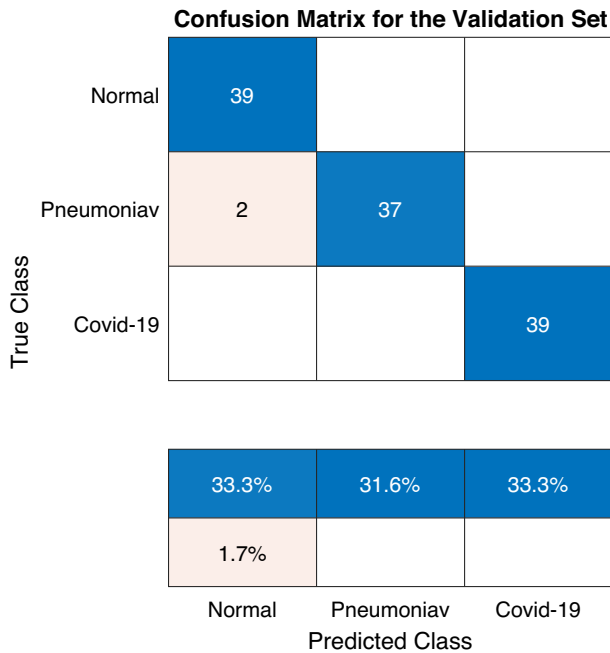


(a) Training curves for accuracy. Both Training and validation curves state that the model's accuracy improves throughout the training process.



(b) Training curves for the loss. Both Training and validation curves state that the model's loss degrades throughout the training process.

**Fig. 3** Transfer learning using pretrained NASNet-Mobile with fine-tuning



**Fig. 4** NASNet-Mobile confusion matrix for the validation set. The rows correspond to the true class and the columns correspond to the predicted class. The bottom table shows the model’s true and false predictions ratio to the validation set size in terms of percentages

The model was further tested with some unused chest images, namely, the images that had used neither used in training nor validating. The test set contained a total of 60 images, with 20 from each class. The images were randomly selected from COVID-19 Radiography Dataset [14, 35]. Figure 5 presents the confusion matrix for the test set. The model accurately predicted 58 of 60 cases in which one COVIT-19 case predicted as normal, and one normal case predicted as viral pneumonia.

**Table 4** NASNet-Mobile performance statistics for the validation set

Measure	Normal	Pneumoniav	COVID-19	Macro-Average
Accuracy	0.9829	0.9829	1.0000	0.9886
Recall	1.0000	0.9487	1.0000	0.9829
Precision	0.9512	1.0000	1.0000	0.9837
F1 Score	0.9750	0.9737	1.0000	0.9829
Negative Predictive Value	1.0000	0.9750	1.0000	0.9916
False Positive Rate	0.0256	0.0000	0.0000	0.0085
False Negative Rate	0.0000	0.0513	0.0000	0.0171
False Discovery Rate	0.0488	0.0000	0.0000	0.0162
False Omission Rate	0.0000	0.0250	0.0000	0.0083
Critical Success Index	0.9512	0.9487	1.0000	0.9666
Matthews Correlation Coefficient	0.9627	0.9617	1.0000	0.9748

**Confusion Matrix for the Test Set**

True Class	Normal	19	1	
	Pneumoniav		20	
	Covid-19	1		19

31.7%	33.3%	31.7%
1.7%	1.7%	
Normal	Pneumoniav	Covid-19

Predicted Class

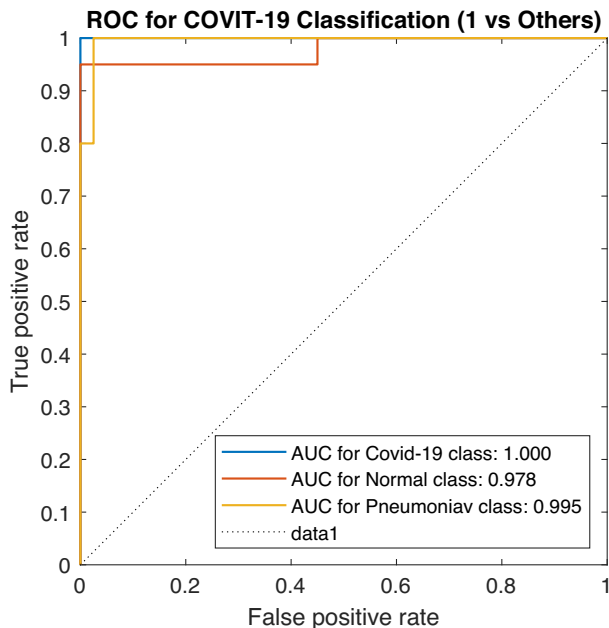
**Fig. 5** NASNet-Mobile confusion matrix for the test set. The rows correspond to the true class and the columns correspond to the predicted class. The bottom table shows the model's true and false predictions ratio to the test set size in terms of percentages

Moreover, Fig. 6 shows the Receiver operative characteristics (ROC) curve for the model. It is the ROC of one versus others. Observe that the class prediction probability value is high, that is the model is robust.

## 5 Discussion

According to the World Health Organization data for December 2020, there have been more than 79 million cases and more than 1.7 million deaths worldwide since the beginning of the pandemic due to the epidemic that has spread all over the world (WHO, 2020). The restrictions applied in the manufacturing industry had serious negative effects on the global supply chain. Currently, with the high rate of continuing spread of COVID-19 worldwide, the state of emergency is declared in many countries: schools are being closed, online work for both the government and private agencies are promoted. As economists warn, the epidemic could cost \$1.1 trillion globally.

It is known that RT-PCR testing of 2019-nCoV RNA is capable of making the definitive diagnosis of COVID-19 from Influenza-A viral pneumonia patients. However, in some cases, nucleic acid testing has some disadvantages, including delays in results, relatively low detection rate, and shortage of supplies. Moreover, in the early stages of COVID-19, some patients may already have positive pulmonary imaging findings but lack sputum and therefore may receive negative RT-PCR test results from nasopharyngeal swabs. These patients are not diagnosed as suspected or confirmed cases, isolated, treated, and become potential sources of infection.



**Fig. 6** NASNet-Mobile ROC curve for the test set

Mahmud et al. [30] identified the distinctive localization of abnormal areas on X-rays that might help to diagnose variations of the clinical features of pneumonia. It is stated that this result obtained with a small number of images can be improved by increasing the number of images. Panwar et al. [34] pointed out that economic differences may cause limitations in testing and that rapid determinations made through such images can help hospital administration and medical professionals. Dansana et al. [17] drew attention to the importance of early detection of covid19 and analyzed the data with 3 different models. As a result, they achieved an accurate prediction success between 60 percent and 91 percent.

Although WHO strongly recommends testing as many people as possible, it has not been fulfilled by many countries due to the apparent lack of resources/staff and the lack of RT-PCR testing. However, the use of CT imaging with artificial intelligence methods might help detect COVID-19 by offering a rapid alternative and thus helping to limit the disease spread.

Developments in computer technology in recent years have been facilitating the acquisition of high-resolution images and the processing of these images. These developments have revolutionized the field of medical imaging and artificial intelligence, thus, so many successful studies have been conducted in the field of medical imaging. Especially, transfer learning-based CNN models achieved expert-level performances even on the nowadays world's nightmare COVID-19 there exists successful studies. Below we name some of them.

Apostolopoulos and Mpesiana [9] designed transfer learning-based CNN model for automatic detection of COVID-19 disease. Their dataset contained X-ray images of normal, common bacterial pneumonia, and COVID-19 diseases. They modeled VGG-19 based CNN, and achieved the highest accuracy of 93.48%, while the sensitivity of 92.85%, and specificity of 98.75%.

**Table 5** NASNet-Mobile performance statistics for the test set

Measure	Normal	Pneumonia	COVID-19	Macro-Average
Accuracy	0.9667	1.0000	0.9667	0.9778
Recall	0.9500	1.0000	0.9500	0.9667
Precision	0.9500	1.0000	0.9500	0.9667
F1 Score	0.9500	1.0000	0.9500	0.9667
Negative Predictive Value	0.9750	1.0000	0.9750	0.9833
False Positive Rate	0.0250	0.0000	0.0250	0.0167
False Negative Rate	0.0500	0.0000	0.0500	0.0333
False Discovery Rate	0.0500	0.0000	0.0500	0.0333
False Omission Rate	0.0250	0.0000	0.0250	0.0167
Critical Success Index	0.9048	1.0000	0.9048	0.9365
Matthews Correlation Coefficient	0.9250	1.0000	0.9250	0.9500

Narin et al. [32] proposed three transfer learning-based deep learning models, based on ResNet50, InceptionV3 and Inception-ResNetV2, to diagnose COVID-19. In the study, 50 COVID-19 and 50 normal cases chest X-ray images were included. ResNet50 based model reached the highest classification performance accuracy of 98.0%. The accuracies of InceptionV3 and Inception-ResNetV2 were 97.0% accuracy and 87%, respectively.

Zhang et al. [47] proposed ResNet-based transfer model for detecting COVID-19 using X-rays. The dataset contained 70 COVID-19 cases and 1008 non-COVID-19 pneumonia cases. Sensitivity and specificity were 96.0% and 70.7%, respectively.

In another study by Wang et al. [45], a deep convolution neural network-based model (COVID-Net) was designed to detect COVID-19 cases. The dataset was composed of 5941 chest X-ray images in total, among them 1203 were normal, 931 bacterial pneumonia, 660 viral pneumonia, and 45 COVID-19. COVID-Net achieved 83.5% test accuracy.

Just by looking at performance metrics, when compared with the above studies, the model proposed in this study receives the highest validation accuracy and the highest validation recall and validation precision. However, even though the performance of the model was slightly lower in the test images, it is still competitive.

## 6 Conclusion

Herein, a fine-tuning CNN model was developed to characterize chest X-ray images as viral pneumonia, COVID-19, and normal cases. Performance metrics for both validation and test sets such as accuracy, recall, and precision along with the others presented in Tables 4 and 5 state that the model is encouraging. For example, the model's macro-average accuracy rate on the training set is 0.9886 and 0.9778. Moreover, macro-average recall is 98.29% on the training set and 96.67% on the test set. Also, the model's precision on the training set is 98.37%, on the test set 96.67%. We can conclude that the model is suitable for use in a laboratory environment. As an outcome of this study, the use of CNN technology can assist physicians in their final decisions and increase accuracy in the decision-making process.

**Funding** No funds, grants, or other support was received.

## Declarations

**Ethics approval** Informed consent is not required for the study.

**Conflict of Interests** The authors declare that they have no conflict of interest.

## References

- Akram T, Attique M, Gul S, Shahzad A, Altaf M, Naqvi SSR, Damaševičius R, Maskeliū nas R (2021) A novel framework for rapid diagnosis of covid-19 on computed tomography scans. *Pattern Analysis and Applications*, 1–14
- Albelwi S, Mahmood A (2017) A framework for designing the architectures of deep convolutional neural networks. *Entropy* 19(6):242
- Ali A, Zhu Y, Chen Q, Yu J, Cai H (2019) Leveraging spatio-temporal patterns for predicting citywide traffic crowd flows using deep hybrid neural networks. In: 2019 IEEE 25Th international conference on parallel and distributed systems (ICPADS). IEEE, pp 125–132
- Ali A, Zhu Y, Zakarya M (2021) A data aggregation based approach to exploit dynamic spatio-temporal correlations for citywide crowd flows prediction in fog computing. *Multimed Tools Appl*, 1–33
- Ali A, Zhu Y, Zakarya M (2021) Exploiting dynamic spatio-temporal correlations for citywide traffic flow prediction using attention based neural networks. *Inf Sci* 577:852–870
- Ali A, Zhu Y, Zakarya M (2022) Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction. *Neural Netw* 145:233–247
- Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF (2020) The proximal origin of sars-cov-2. *Nat Med* 26(4):450–452
- Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK (2018) Medical image analysis using convolutional neural networks: a review. *J Med Syst* 42(11):1–13
- Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med* 43(2):635–640
- Bussani R, Schneider E, Zentilin L, Collesi C, Ali H, Braga L, Volpe MC, Colliva A, Zanonati F, Berlot G et al (2020) Persistence of viral rna, pneumocyte syncytia and thrombosis are hallmarks of advanced covid-19 pathology. *EBioMedicine* 61(103):104
- Butt C, Gil J, C D, Babu B (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Appl Intell* April(2020):1–7
- Catak FO, Şahinbaş K (2021) Human-in-the-loop enhanced COVID-19 detection in Transfer Learning based CNN Models
- Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258
- Chowdhury ME, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, Islam KR, Khan MS, Iqbal A, Al Emadi N et al (2020) Can ai help in screening viral and covid-19 pneumonia? *IEEE Access* 8:132,665–132,676
- Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M (2020) Covid-19 image data collection: Prospective predictions are the future. [arXiv:2006.11988](https://arxiv.org/abs/2006.11988)
- Coronavirus 2020-nCoV C. (2020) The center for systems science and engineering (csse) at Johns Hopkins university (jhu)
- Dansana D, Kumar R, Bhattacharjee A, Hemanth DJ, Gupta D, Khanna A, Castillo O (2020) Early diagnosis of covid-19-affected patients based on x-ray and computed tomography images using deep learning algorithm. *Soft Comput*, 1–9
- Deepak S, Ameer P (2019) Brain tumor classification using deep cnn features via transfer learning. *Comput Biol Med* 111(103):345
- Dreyer KJ, Geis JR (2017) When machines think: radiology’s next frontier. *Radiology* 285(3):713–718
- Gorbalenya AE, Baker SC, Baric R, Groot RJD, Drosten C, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM, Neuman BW et al (2020) Severe acute respiratory syndrome-related coronavirus: The species and its viruses—a statement of the coronavirus study group
- Greenspan H, Van Ginneken B, Summers RM (2016) Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans Med Imaging* 35(5):1153–1159

22. Hamey LG (2015) A functional approach to border handling in image processing. In: 2015 International conference on digital image computing: Techniques and applications (DICTA). IEEE, pp 1–8
23. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size. arXiv:1602.07360
24. Kermany D, Zhang K, Goldbaum M et al (2018) Labeled optical coherence tomography (oct) and chest x-ray images for classification Mendeley data 2(2)
25. Khairat S, Marc D, Crosby W, Al Sanousi A (2018) Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 6(2):e24
26. Khan MA, Hussain N, Majid A, Alhaisoni M, Bukhari SAC, Kadry S, Nam Y, Zhang YD (2021) Classification of positive covid-19 ct scans using deep learning. Computers Materials and Continua 66(3)
27. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inform Process Syst 25:1097–1105
28. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR (2020) Severe acute respiratory syndrome coronavirus 2 (sars-cov-2) and coronavirus disease-2019 (covid-19): The epidemic and the challenges. Int J Antimicrob Agents 55(3):105,924
29. Li H, Giger ML, Huynh BQ, Antropova NO (2017) Deep learning in breast cancer risk assessment: evaluation of convolutional neural networks on a clinical dataset of full-field digital mammograms. J Med Imaging 4(4):041,304
30. Mahmud T, Rahman MA, Fattah SA (2020) Covxnet: A multi-dilation convolutional neural network for automatic covid-19 and other pneumonia detection from chest x-ray images with transferable multi-receptive feature optimization. Comput Biol Med 122(103):869
31. Mason RJ (2020) Pathogenesis of covid-19 from a cell biology perspective
32. Narin A, Kaya C, Pamuk Z (2021) Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Anal Applic, 1–14
33. Padma T, Kumari CU (2020) Deep learning based chest x-ray image as a diagnostic tool for covid-19. In: 2020 International conference on smart electronics and communication (ICSECC). IEEE, pp 589–592
34. Panwar H, Gupta P, Siddiqui MK, Morales-Menendez R, Singh V (2020) Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. Chaos, Solitons Fractals 138(109):944
35. Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Kashem SBA, Islam MT, Al Maadeed S, Zughair SM, Khan MS et al (2021) Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. Comput Biol Med 132(104):319
36. Ruuskanen O, Lahti E, Jennings LC, Murdoch DR (2011) Viral pneumonia. The Lancet 377(9773):1264–1275
37. Sahlol AT, Yousri D, Ewees AA, Al-Qaness MA, Damasevicius R, Abd Elaziz M (2020) Covid-19 image classification using deep features and fractional-order marine predators algorithm. Scientific Reports 10(1):1–15
38. Setio AAA, Ciompi F, Litjens G, Gerke P, Jacobs C, Van Riel SJ, Wille MMW, Naqibullah M, Sánchez CI, Van Ginneken B (2016) Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 35(5):1160–1169
39. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D, Shi Y (2020) Lung infection quantification of covid-19 in ct images with deep learning. arXiv:2003.04655
40. Shen D, Wu G, Suk HI (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19:221–248
41. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
42. Sohrabi C, Alsafi Z, O’neill N, Khan M, Kerwan A, Al-Jabir A, Iosifidis C, Agha R (2020) World health organization declares global emergency: a review of the 2019 novel coronavirus (covid-19). Int J Surg 76:71–76
43. Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang Z, Chen J, Wang R, Zhao H, Zha Y et al (2021) Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images. IEEE/ACM Transactions on Computational Biology and Bioinformatics
44. Üreten K., Erbay H, Maraş HH (2020) Detection of hand osteoarthritis from hand radiographs using convolutional neural networks with transfer learning. Turkish Journal of Electrical Engineering & Computer Sciences 28(5)
45. Wang L, Lin ZQ, Wong A (2020) Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports 10(1):1–12
46. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Ni Q, Chen Y, Su J et al (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 6(10):1122–1129
47. Zhang J, Xie Y, Li Y, Shen C, Xia Y (2020) Covid-19 screening on chest x-ray images using deep learning based anomaly detection. arXiv:2003.12338 27



48. Zoph B, Vasudevan V, Shlens J, Le QV (2018) Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 8697–8710

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Hakan Kör<sup>1</sup> · Hasan Erbay<sup>2</sup> · Ahmet Haşim Yurttakal<sup>3</sup>

Hasan Erbay  
herbay@thk.edu.tr

Ahmet Haşim Yurttakal  
ahyurttakal@aku.edu.tr

<sup>1</sup> Department of Computer Engineering, Engineering Faculty, Hitit University, Çorum, Turkey

<sup>2</sup> Computer Engineering Department, Engineering Faculty, University of Turkish Aeronautical Association, 06790 Etimesgut Ankara, Turkey

<sup>3</sup> Computer Engineering Department, Engineering Faculty, Afyon Kocatepe University, 03204 Erenler Afyon, Turkey