Classification of X-ray images into COVID-19, pneumonia, and TB using cGAN and fine-tuned deep transfer learning models

Tirth Mehta · Ninad Mehendale

Received: 1 December 2020 / Accepted: 28 August 2021 / Published online: 7 October 2021
© Sociedade Brasileira de Engenharia Biomédica 2021

Abstract

Purpose The rapid increase in the spread of the coronavirus disease of 2019 (COVID-19) has led to a need for reliable, effective, and readily available testing on a large scale. While diagnostic testing has been a support to public health, newer technology can be used to provide low-cost and convenient test options for patients. X-ray scanning can be performed to resolve this issue and produce quicker and more precise results. Currently, a radiologist is required to examine these X-ray images. However, deep convolutional neural networks can also be used to perform X-ray examinations and employed for the detection of COVID-19. We propose a Conditional Generative Adversarial Network (cGAN) with a fine-tuned deep transfer learning model to classify chest X-rays into six categories: COVID-Mild, COVID-Medium, COVID-Severe, Normal, Pneumonia, and Tuberculosis.

Methods A total of 1229 images were taken to form a dataset containing six classes corresponding to the six categories. A cGAN was used to increase the number of images. Generative Adversarial Networks (GAN) are used to train a model for generating new images. cGAN is an extension of GAN consisting of a generator and discriminator network that are trained simultaneously to optimize the model. The generated images were then trained using deep transfer learning models such as ResNet50, Xception, and DenseNet-169 to achieve the classification into six classes.

Results The proposed model helped achieve a training and validation accuracy of up to 98.20 % and 94.21 % respectively. The model was able to achieve a test accuracy of 93.67 %. The use of cGAN not only helped to increase the size of the training dataset but it also helped to reduce the problem of over-fitting.

Conclusion The proposed approach will help to diagnose COVID-19 quickly at an early stage.

Keywords COVID-19 · Corona · Deep transfer learning · cGAN

Introduction

The coronavirus ranges from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). It is a wide family of viruses that cause human respiratory infections (Chittora et al. 2021). The infectious disease coronavirus disease of 2019 (COVID-19) is caused by the most recently discovered coronavirus. This novel virus unlike SARS has a lower death rate but faster outspread. About 80 % of the cases are recovered without any medical treatment and only 1 out of 5 COVID-19 patients develop serious illness and difficulty in breathing. The risk of contracting a serious illness is greater for older people and patients with underlying medical conditions. This is the key reason why we need a testing tool that not only determines the existence of COVID-19 but also indicates the seriousness of the disease (Chittora et al. 2021). COVID-19 vaccines have been developed to fight the coronavirus. Although the vaccines are providing very promising results, the issue of manufacturing them effectively for everyone on this planet would require a lot of time. To get into a normal routine and considering the amount of time required before which everyone gets a vaccine, we need an effective, fast, and inexpensive method for early COVID-19 disease detection. We also need a testing process that could detect different conditions like tuberculosis, normal pneumonia, and varying severity of COVID-19.
The most used and gold standard of frontline testing for the COVID-19 is reverse transcription-polymerase chain reaction (RT-PCR) (Xie et al. 2020). Although it is the most used testing procedure, it has limitations. For example, it is very expensive and takes around 4–5 h to get results. Moreover, sampling error and low viral load also affect the results of this test. To overcome these limitations while diagnosing COVID-19, we use radiological image-based techniques like X-rays and CT scans. X-rays are preferred in our proposed study because X-ray machines are extensively available and cost less than CT scans. X-rays also have low ionizing radiation compared to CT scans. Convolutional Neural Networks (CNN)–based deep learning techniques can be used to analyze these radiological images and help simplify the early detection process for the following six categories:

1. Normal
2. Tuberculosis
3. Normal pneumonia
4. COVID-Mild
5. COVID-Medium
6. COVID-Severe

Figure 1 exhibits the concept of the proposed study. It shows how this study can be used to create a very convenient testing process. The steps involved are capturing an X-ray of the patient’s lungs using an X-ray machine, producing an X-ray image, providing the image to the trained model as input, and finally, the trained model will classify the image into one of the six classes.

In this study, to classify chest X-rays into the six classes, a Conditional Generative Adversarial Network (cGAN) with a fine-tuned deep transfer learning model for a small dataset is proposed. We use cGAN that involves the conditional generation of images by generative models. These generated images are used to overcome the problem of data imbalance and overfitting. After generating the images, the training data is loaded and various deep convolutional neural network–based pre-trained transfer models are used to achieve the objective of classifying the image. In this study, the transfer learning models used are ResNet-50 (He et al. 2016), ResNet-101 (He et al. 2016), DenseNet-169 (Huang et al. 2017), DenseNet-201 (Huang et al. 2017), and Xception (Chollet 2017). All these models have performed prodigiously in the ImageNet Large Scale Visual Recognition Competition (ILSVRC) annual challenges. These pre-trained models allow us to build a significantly accurate model quickly and with less data.

The advancement in deep learning approaches has enabled researchers to use advanced deep learning techniques to classify or analyze chest X-rays effectively. Three convolutional neural network models, namely ResNet-50, Inception-V3, and Inception-ResNet-v2, were studied by Narin et al. (2020) to detect patients infected with coronavirus pneumonia using X-ray chest radiographs. The study obtained the highest classification accuracy of 98 % using the pre-trained ResNet-50 model. The dataset used was quite small consisting of 100 images in total with 50 COVID-19 patient images and 50 normal images.

In their proposed model, Das et al. (2020) used the deep convolutional neural network along with the Xception model to classify X-ray images into three conditions: COVID-19, COVID-19 pneumonia, and any other infection without COVID-19 or pneumonia. The dataset used consists

![Fig. 1 Concept diagram of the proposed system. X-ray of a patient is taken using an X-ray machine. Furthermore, X-ray images are produced and these images are used to train the model. Finally, the model takes the X-ray image as input and classifies it into one of the 6 categories within seconds](image-url)
of chest X-rays of the above-mentioned three classes with 70% used as training data and the remaining 20% and 10% for testing and validation respectively. The proposed model achieved a training accuracy of 99.5246% and a testing accuracy of 97.4068% which was significantly greater in performance than other competitive models.

Hall et al. (2020) used ResNet-50 and VGG-16 as a pre-trained deep convolutional neural network to diagnose the COVID-19 disease. They fine-tuned the model using 102 chest X-rays of COVID-19 patients and 102 bacterial and viral pneumonia chest X-rays. They obtained an accuracy of 89.2% using the ResNet-50 model. Additionally, they used an ensemble of three classifiers – ResNet-50, VGG-16, and their own small CNN - which increased the overall accuracy to 91.24%.

Apostolopoulos and Mpesiana (2020) analyzed the accuracy of the state-of-the-art convolutional neural network for medical image recognition suggested in recent years. In this study, two datasets were used, one consisting of 1427 chest X-ray images of 224 were of COVID-19, 700 were of normal bacterial pneumonia, and 504 images of normal X-rays. The other dataset consisted of 224 of COVID-19, 714 images of reported bacterial and viral pneumonia, and 504 images of normal X-rays. The models used for transfer learning in this study were VGG-19, MobileNet-V2, Xception, Inception, and Inception-ResNet-v2. The study showed that VGG-19 had better accuracy, but a high false-negative rate compared to MobileNet-V2 whose accuracy was a bit less than VGG-19 but had higher specificity.

Alom et al. (2020) in his study aimed to not only detect the presence of COVID-19 using X-ray and CT scan images but also to segment the regions infected by COVID-19. Inception-Residual Recurrent Convolutional Neural Network (IRRCNN) with Transfer Learning (TL) approach was used to detect COVID-19 and the NABLA-N network model was used to segment infected regions. The model used to detect COVID-19 achieved a testing accuracy of 84.67% for X-ray images and 98.78% for CT scan images. The segmented infected region is then used to measure the probability distribution of the radiological findings of COVID-19 CXR’s localized intensity variations. Their proposed model consists of three parts—lung segmentation, classification model, and probabilistic grad-CAM saliency map for visualization and interpretation. The study compared the results for various conditions and parameters and proved that classification results were the best for the patch-based model with a segmentation mask and patch size of 224 × 224.

In the studies performed by Maguolo and Nanni (2020), Tartaglione et al. (2020), and Tabik et al. (2020), they have mentioned various issues in recent studies done on COVID-19 detection using deep learning approaches and have tried using different methods to overcome these obstacles. From these studies, the problems in COVID-19 detection using deep learning algorithms can be categorized into the following:

1. There are multiple sources of training and validation data. This may lead to the model being trained to recognize the source of the data rather than the desired class it is meant to classify.
2. There is a bias in the classification result due to different settings and parameters in which the chest X-ray image is captured. The deep model might potentially learn to discriminate based on these biases which would give us results on undesired features.
3. Due to the unavailability of datasets, most papers use very small datasets, and even then, there is data imbalance, with one class represented more than others. This is also one of the major drawbacks of using deep learning for COVID-19 detection.
4. Another factor that has a great impact on the results of these studies is the test set. It has been observed that test dataset of origin different than training and validation dataset gives less accuracy as compared to test dataset of the same origin.

Furthermore, studies mentioned above did not diagnose or detect different degrees/severity of COVID-19 infection. Some studies like Das et al. (2020), Hall et al. (2020), and Apostolopoulos and Mpesiana (2020) proposed a model to classify X-ray images into conditions like pneumonia, COVID-19 and normal. However, none of these methods could detect COVID-19 based on the different severity levels. These systems also failed to detect the possibility of tuberculosis. Though Oh et al. (2020) classified the X-ray images into COVID-19 and tuberculosis as well, it did not make an attempt to find the severity of COVID-19. Similarly, Tabik et al. (2020) proposed a strategy to find the severity of COVID-19 but did not detect other conditions like pneumonia or tuberculosis.

Khalifa et al. (2020) proposed a model that addresses the issue of a limited quantity of data. In this research, to detect chest X-rays as either pneumonia or normal, GAN
and deep transfer learning models are used. The dataset used consists of 5863 X-ray images. Out of these, only 10% are used while training. The remaining 90% of images used for training are the generated images provided by the GAN. The deep transfer learning models used in the study are AlexNet, SqueezeNet, GoogLeNet, and ResNet-18. The model proposed by them achieved an accuracy of 99% using the ResNet-18 model. Though this study is very promising, it has a lot of aspects for improvement. The GAN network used is not conditioned on the label hence, the modes of the generated data are not under control. Also, this study focuses to detect only two conditions, i.e., pneumonia and normal.

In our proposed system, we have aimed to take into account all the factors which cause major hindrances while using the deep learning approach for COVID-19 detection. We have tried to come up with a novel approach that is reliable, inexpensive, widely attainable and can identify a variety of infections as well as their severity. Though the vaccines are being rolled out, we still don’t know to what extent will they be available. Hence till then, we need a streamlined and fast testing process that will eventually help in returning to normalcy.

Methodology

Components and equipment

The entire study is performed using TensorFlow and its libraries such as Keras. Python is used as the fundamental programming language for the entire implementation. The hardware requirements for the experiment were met by using Google Colab. Google Colab provides the following hardware specifications:

- GPU - Nvidia K80 / T4
- GPU Memory - 12GB / 16GB
- GPU Memory Clock - 0.82GHz / 1.59GHz
- Performance - 4.1 TFLOPS / 8.1 TFLOPS
- No. CPU Cores - 2
- Available RAM - 12GB (upgradable to 26.75GB)
- Disk Space - 358GB

Data description

The complete dataset consisted of 1229 chest X-ray images with 42 COVID-Medium images, 40 COVID-Mild images, 36 COVID-Severe images, 348 regular images, 263 tuberculosis images, and 500 pneumonia images. The entire dataset was acquired from a local hospital that contained chest X-ray images of various conditions. Hence, the issue of model learning to discriminate the images based on the source rather than the desired features did not affect this study. The dataset was quite imbalanced and limited. To overcome these hurdles, we used cGAN (Mirza and Osindero 2014) which involves the conditional generation of images by generative models. These generated images were then used to train our model.

Data pre-processing

1. Histogram equalization: The chest X-ray images in the dataset had different dynamic ranges, color contrast, or any other bias present. In order to avoid the model from discriminating the images based on these biases, we applied histogram equalization with an aim to minimize these kinds of biases. Hence, histogram equalization avoided the model from showing unfair results.
2. Lung segmentation: Figure 2 (a) shows one of the images from the dataset of chest X-rays. Segmentation of the lung from the X-ray image is required to remove any text or unwanted markers present in the chest X-rays which might also lead to undesired biases as shown in Fig. 2 (b). U-Net architecture (Ronneberger et al. 2015) was used for segmentation.
2015) was used to segment the lungs and this model was trained using Montgomery County X-ray set and Shenzhen Hospital X-ray set (Jaeger et al. 2014).

3. All the training images were scaled from [0, 255] to [-1, 1] for normalization.

As shown in Fig. 3, a database of chest X-ray images was used to train the cGAN model. The cGAN comprised of a generator that was used to generate images from the original images provided as the input. These generated images along with the real images were fed to the discriminator. The discriminator further compared these images and classified them as real (original images) or fake (generated images). The original, as well as the images generated by the cGAN, were given as an input to the pre-trained CNN. We employed different pre-trained CNN models such as ResNet-50, ResNet-101, DenseNet-169, DenseNet-201, and Xception to perform the classification. The X-ray images were classified as COVID-Medium, COVID-Mild, COVID-Severe, Pneumonia, TB, and Normal.

**Conditional Generative Adversarial Networks (cGAN)**

Generative models generate new samples based on input data fed to the model. GANs generate images of the same distribution as the input images just like the other deep neural networks that are employed to generate images. The generator network in the GAN tricks the discriminator and generates real-looking images. While the discriminator network tries to find the difference between real and fake images. The models are trained in an adversarial manner such that decreased loss in generator comes at the cost of increased loss in discriminator and vice versa.

GANs are very effective in generating new images for a target dataset but it has few drawbacks. It does not allow targeted generation of images and it does not make use of class labels as additional information to improve image synthesis. To overcome these drawbacks, cGAN was introduced by Mirza and Osindero (2014) in which GANs are trained in such a way that both the generator and discriminator models are conditioned on labels.

Overview of cGAN architecture and its various layers is as follows:

1. Generator Network:
   - (a) 2 – Input Layers
   - (b) 2 – Dense Layers
   - (c) 1 – Embedding Layer

![Fig. 3 Process flow diagram of the proposed method. Images are generated with the help of cGAN when it is provided with a database of original X-ray images. The generated images along with the real images are given to the discriminator which classifies the images as real and fake. The original, as well as generated images, are passed through the pre-trained CNN models that classify the X-rays into the following six categories: COVID-Medium, COVID-Mild, COVID-Severe, normal, pneumonia and TB](image-url)
2. Discriminator Network:
(a) 2 – Input Layers
(b) 2 – Dense Layers
(c) 1 – Embedding Layer
(d) 2 – Conv2D Layer (Convolutional Layer with downsampling)
(e) 1 – Dropout Layer
(f) 2 – Leaky ReLu

The detailed architecture of the discriminator network and generator network is shown in Fig. 4 (a) and (b). These figures exhibit the type of layer and how each layer is placed in the network. The input and output feature size of each layer is also shown in Fig. 4 (a) and (b). After we defined the discriminator and generator networks, the model was trained for 200 epochs in which discriminator loss for real images ranged from 0.3 to 0.7 and similarly, discriminator loss for fake images ranged from 0.3 to 0.7, whereas the generator loss was a bit greater than 1 for the most part of the training.

Transfer learning applied to deep convolutional neural networks

The advancement of deep convolutional neural networks has significantly decreased the error rates during image classification. Models such as AlexNet, ResNet, and DenseNet were built to classify the ImageNet dataset and showed outstanding performance in the ImageNet Large Scale Visual Recognition Competition (ILSVRC). These pre-trained models can be used to classify other datasets as well. These models are also widely used for the purpose of medical image classification. One of the major drawbacks of medical image classification is the limited amount of...
Table 1 Training, validation and testing accuracy for models trained without the generated images

| Architecture | Training accuracy (%) | Validation accuracy (%) | Testing accuracy (%) |
|--------------|-----------------------|-------------------------|---------------------|
| ResNet-50    | 95.11                 | 92.56                   | 92.33               |
| ResNet-101   | 91.69                 | 90.06                   | 89.5                |
| Xception     | 89.25                 | 90.08                   | 88.26               |
| DenseNet-201 | 90.93                 | 91.74                   | 91.33               |
| DenseNet-169 | 92.27                 | 91.74                   | 92.13               |

The highest training accuracy of 95.11% was achieved using ResNet-50 while the lowest training accuracy of 89.25% was achieved using the Xception model. Similarly, ResNet-50 showed the highest validation and testing accuracies. The lowest validation accuracy was achieved using ResNet-101 while the lowest testing accuracy was achieved using Xception.

Table 2 Training, validation and testing accuracy for models trained with the help of generated images

| Architecture | Training accuracy (%) | Validation accuracy (%) | Testing accuracy (%) |
|--------------|-----------------------|-------------------------|---------------------|
| ResNet-50    | 98.2                  | 94.21                   | 93.67               |
| ResNet-101   | 95.78                 | 93.39                   | 91.27               |
| Xception     | 96.02                 | 95.87                   | 90.91               |
| DenseNet-201 | 97.27                 | 95.87                   | 92.91               |
| DenseNet-169 | 97.42                 | 95.87                   | 90.91               |

ResNet-50 showed highest training and testing accuracy while Xception, DenseNet-201, DenseNet-169 showed highest validation accuracy.

Results

A comparison of the performance of the network without and with cGAN is shown in the Tables 1 and 2 respectively. Figure 5 (a) shows the validation accuracy obtained by the different deep learning models used. Each model was trained thrice, once without generated images, once with generated images, and once with lung segmentation along with generated images. When the models were trained without generated images, the ResNet-50 model showed the highest validation accuracy of 92.56%. When the generated images were used, the Xception, DenseNet-201, and DenseNet-169 showed the highest validation accuracy of 95.87%, whereas when lung segmentation and generated images were both used together the DenseNet-201 showed the highest validation accuracy of 90.91%. As shown in Fig. 5 (b), ResNet-50 achieved the highest accuracy of 95.11% when the model was trained without the help of generated images while Xception showed the lowest accuracy of 89.25%. ResNet-50 also showed the highest training accuracy of 98.2% when the model was trained using generated images while ResNet-101 showed the lowest training accuracy of 95.78%. The training accuracy of 92.81% was achieved when the DenseNet-201 was trained with segmented masks and generated images. DenseNet-201 showed the highest accuracy while ResNet-101 achieved a training accuracy of 84% when it was trained using segmented masks and...
generated images. Figure 5 (c) shows the testing accuracy when the models were trained without the generated images was highest for the ResNet-50 model which showed a 92.33 % accuracy. The ResNet-50 achieved the highest testing accuracy of 93.67 % even when the models were trained with the generated images, whereas the DenseNet-201 showed the highest testing accuracy of 89.76 % when the models were trained using both the generated images and lung segmentation. It is clearly seen from the graphs in Fig. 5 (a), (b), and (c) that using the generated images for training has not only helped increase training accuracy but also validation and testing accuracy. Training accuracy increased by almost 1–7 %. Similarly, validation accuracy increased by 1–6 % and testing accuracy by 1–4 %. The testing accuracy decreased for DenseNet-169 but the loss is not significant. The greatest increase in accuracy is seen in the Xception model. As seen in Table 3, using segmented masks for training the model has led to a decrease in the accuracy of the model significantly. There is a 3–12 % drop in accuracy and ResNet-101 has shown the worst drop among all the models.

**Discussions**

In our work, we have addressed all the major issues faced by deep learning approaches that were previously proposed for the detection of various conditions using chest X-ray images. We put forth a solution to tackle each hurdle. This was necessary to get unbiased and fair results. First of all,
Table 3  Training, validation and testing accuracy for models trained using segmentation mask and generated images

| Architecture     | Training accuracy (%) | Validation accuracy (%) | Testing accuracy (%) |
|------------------|-----------------------|-------------------------|----------------------|
| ResNet-50        | 91.57                 | 87.26                   | 88.27                |
| ResNet-101       | 84                    | 81.07                   | 80.61                |
| Xception         | 92.46                 | 88.43                   | 88.02                |
| DenseNet-201     | 92.81                 | 90.91                   | 89.76                |
| DenseNet-169     | 92.24                 | 88.43                   | 88.47                |

DenseNet-201 achieved the highest training, validation, and testing accuracy. ResNet-101 achieved the lowest training, validation, and testing accuracy.

having a dataset of various classes from the same source of origin negated the problem of the model discriminating the images based on the source rather than the desired features. Secondly, bias in the classification result due to different settings and parameters in which the chest X-ray image is captured is minimized by using histogram equalization and lung segmentation. Finally, the major issue of small and imbalance datasets was taken care of by using cGAN.

The results show poor performance for models trained on the segmented mask because of the error induced during the segmentation of the images. Certain images are either over-segmented or under-segmented. These under and over segmented images are classified incorrectly by the model which leads to a decrease in the accuracy. As we have used a single source to collect our dataset, we can rely on the result obtained by the model without using any segmented mask.

As it is seen in Table 4 that none of the studies in literature classified the images in as many conditions as we have proposed. Our proposed model provides the best accuracy for classifying the X-ray images into 6 conditions. We not only propose to detect COVID but also aim to detect the severity of COVID.

Table 4 Comparison between existing methodologies and the proposed system

| Method                         | Accuracy (%) | Number of images | Number of classes |
|--------------------------------|--------------|------------------|-------------------|
| Narin et al. (2020)            | 98           | 100              | 2                 |
| Das et al. (2020)              | 97.4         | -                | 3                 |
| Hall et al. (2020)             | 91.24        | 455              | 2                 |
| Apostolopoulos and Mpesiana (2020) | 96.78      | Dataset 1: 1427  | 3                 |
| Khalifa et al. (2020)          | 99           | 5863             | 2                 |
| Alom et al. (2020)             | 84.67        | 5216             | 2                 |
| Tabik et al. (2020)            | 61.80        | 852              | 4                 |
| Oh et al. (2020)               | 88.9         | 502              | 5                 |
| Proposed method                | 93.67        | 1229             | 6                 |

The existing systems are used to detect different conditions from the X-ray images. The existing methods cannot categorize the COVID-19 based on its severity.
Conclusion

X-ray machines are widely present in each country and are quite accessible and inexpensive. Hence, building a testing process based on the X-ray images would solve the problems faced during COVID tests like RT-PCR. The proposed study shows that deep learning techniques can be effectively utilized to provide a reliable, accessible, and inexpensive testing process to diagnose not only COVID-19 infection from X-ray images. It can also be used to determine the severity of the COVID-19 infection and other types of diseases like pneumonia and tuberculosis. The proposed cGAN model consists of a generator network that is trained along with the discriminator network to generate images conditioned on class labels. These generated images along with the training images are used to fine-tune transfer learning models (ResNet, Xception, DenseNet). These fine-tuned models are then used to diagnose the specific disease from the X-ray images. Using cGAN, generated images along with the original dataset are used for training. Our model achieved maximum training, validation, and testing accuracy of 98.2%, 94.21%, and 93.67%, respectively. The ResNet-50 model showed the best overall accuracy. The Xception model had shown the greatest increase in the accuracy with the generated images. Considering the limited data and computational resources, our study has shown remarkable results compared to other related works in terms of the number of diseases or infections our model can predict. To start a normal life, we definitely require a testing process that could determine the severity, as well as the type of infections in a cheap, reliable, and accessible way, and our proposed model is a big step in that direction.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42600-021-00174-z.

Acknowledgements We would like to thank Ms. Hetal from Research Innovation Incubation Design Labs (RiiDL) and the Somaiya Initiatives For Research and Consultancy (SIRAC).

Author contribution Conceptualization was done by Ninad Mehendale (NM). All the experiments/code executions were performed by Tirth Mehta (TM) and NM. The formal analysis was performed by TM and NM. Manuscript writing—original draft preparation was done by TM. Review and editing was done by NM. Visualization work was carried out by TM and NM.

Code availability All the codes used in this study are provided in the supplementary material.

Declarations

Ethics approval All authors consciously assure that the manuscript fulfills the following statements: (1) This material is the authors’ own original work, which has not been previously published elsewhere. (2) The paper is not currently being considered for publication elsewhere. (3) The paper reflects the authors’ own research and analysis in a truthful and complete manner. (4) The paper properly credits the meaningful contributions of co-authors and co-researchers. (5) The results are appropriately placed in the context of prior and existing research.

Consent to participate This article does not contain any studies with animals or humans performed by any of the authors. The informed consent was taken from all the participants whose X-ray images were used in the manuscript. All the necessary permissions were obtained from Institute Ethical committee and concerned authorities.

Consent for publication Authors have taken all the necessary consents for publication from participants wherever required.

Conflict of interest The authors declare no competing interests.

References

Alom MZ, Rahman M, Nasrin MS, Taha TM, Asari VK. Covid_mntnet: Covid-19 detection with multi-task deep learning approaches. arXiv:2004.03747. 2020.

Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med: 1. 2020.

Chittora D, Meena BR, Sharma K. Updates of COVID-19. Rese Biomed Eng. 2021.

Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 1251–1258.

Das NN, Kumar N, Kaur M, Kumar V, Singh D. Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays. Irbm. 2020.

Hall LO, Paul R, Goldgof DB. Finding covid-19 from chest x-rays using deep learning on a small dataset. arXiv:2004.02060. 2020.

He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–778.

Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–4708.

Jaeger S, Candemir S, Antani S, Wang YXJ, Lu PX, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quantitative Imaging Med Surgery. 2014;4(6):475.

Khalifa NEM, Taha MHN, Hassaniene AE. Detection of coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset. arXiv:2004.01184. 2020.

Maguolo G, Nanni L. A critic evaluation of methods for covid-19 automatic detection from x-ray images. arXiv:2004.12823. 2020.

Mirza M, Osindero S. Conditional generative adversarial nets. arXiv:1411.1784. 2014.

Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv:2003.10849. 2020.

Oh Y, Park S, Ye JC. Deep learning covid-19 features on cxr using limited training data sets. IEEE Trans Med Imaging. 2020;39(8):2688–2700.

Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference
on medical image computing and computer-assisted intervention; 2015. p. 234–241.

Tabik S, Gómez-Ríos A, Martín-Rodríguez JI, Sevillano-García I, Rey-Area M, Charte D. COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images. IEEE J Biomed Health Informatics. 2020;24(12):3595–3605.

Tartaglione E, Barbano CA, Berzovini C, Calandri M, Grangetto M. Unveiling covid-19 from chest x-ray with deep learning: a hurdles race with small data. Int J Environ Res Public Health. 2020;17(18):6933.

Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. Radiology: 200343. 2020.

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