



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# COVID-19 image classification using deep learning: Advances, challenges and opportunities

Priya Aggarwal<sup>a</sup>, Narendra Kumar Mishra<sup>b</sup>, Binish Fatimah<sup>c</sup>, Pushpendra Singh<sup>d,\*</sup>,  
Anubha Gupta<sup>e</sup>, Shiv Dutt Joshi<sup>b</sup>

<sup>a</sup> The Vehant Technology Pvt. Ltd., Noida, India

<sup>b</sup> The Department of EE, Indian Institute of Technology Delhi, Delhi 110016, India

<sup>c</sup> The Department of ECE, CMR Institute of Technology, Bengaluru, India

<sup>d</sup> The Department of ECE, National Institute of Technology Hamirpur, HP, India

<sup>e</sup> The Department of ECE, IIT-Delhi, Delhi, 110020, India

## ARTICLE INFO

### Index Terms:

COVID-19 detection

Deep learning

Convolutional neural networks

X-ray and CT scan Images

## ABSTRACT

Corona Virus Disease-2019 (COVID-19), caused by Severe Acute Respiratory Syndrome-Corona Virus-2 (SARS-CoV-2), is a highly contagious disease that has affected the lives of millions around the world. Chest X-Ray (CXR) and Computed Tomography (CT) imaging modalities are widely used to obtain a fast and accurate diagnosis of COVID-19. However, manual identification of the infection through radio images is extremely challenging because it is time-consuming and highly prone to human errors. Artificial Intelligence (AI)-techniques have shown potential and are being exploited further in the development of automated and accurate solutions for COVID-19 detection. Among AI methodologies, Deep Learning (DL) algorithms, particularly Convolutional Neural Networks (CNN), have gained significant popularity for the classification of COVID-19. This paper summarizes and reviews a number of significant research publications on the DL-based classification of COVID-19 through CXR and CT images. We also present an outline of the current state-of-the-art advances and a critical discussion of open challenges. We conclude our study by enumerating some future directions of research in COVID-19 imaging classification.

## 1. Introduction

Coronavirus or COVID-19 is a viral disease that was first identified in Wuhan, China, in December 2019 and later spread quickly worldwide [1,2]. It is caused by Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) and has affected millions of people worldwide. COVID-19 infection starts in the throat's mucous membranes and spreads to the lungs through the respiratory tract. COVID-19 is a highly contagious disease; therefore, it is vital to rapidly screen, diagnose and isolate patients to prevent the spread of the disease and accelerate their proper treatment. Diagnosis of COVID-19 infection through medical imaging, such as CXR and CT scans, has been reported to yield accurate results and is being used widely in the screening of the disease [3–5]. However, successful interpretation of results through images faces several challenges due to the very recent development of the disease and similarities

with other pulmonary disorders such as pneumonia [6] (refer to Fig. 1). Due to the complex nature of COVID-19, its accurate diagnosis is a relatively complicated time-taking task that requires the expertise of radiologists to achieve acceptable diagnostic performance.

Control and eradication of COVID-19 depend heavily on isolating the infected and vaccinating the susceptible. At present, the gold standard for COVID-19 detection is the RT-PCR (Reverse Transcription Polymerase Chain Reaction) test; however, it requires more time to process the specimen and generate the result. Also, it has been observed that many patients may test positive for COVID-19 after recovery [7]. Vaccination is known to immunize people against the virus; however, they are still prone to the infection. Developing an effective and safe vaccine with prolonged efficacy is still in progress and will take substantial time. Further, vaccination of the entire global population will also take time due to the constraints on the availability of the vaccine

\* Corresponding author.

E-mail addresses: [priyaaggarwal27@gmail.com](mailto:priyaaggarwal27@gmail.com) (P. Aggarwal), [eez188568@ee.iitd.ac.in](mailto:eez188568@ee.iitd.ac.in) (N.K. Mishra), [binish.fatimah@gmail.com](mailto:binish.fatimah@gmail.com) (B. Fatimah), [spushp@nith.ac.in](mailto:spushp@nith.ac.in), [spushp@gmail.com](mailto:spushp@gmail.com) (P. Singh), [anubha@iitd.ac.in](mailto:anubha@iitd.ac.in) (A. Gupta), [sdjoshi@ee.iitd.ac.in](mailto:sdjoshi@ee.iitd.ac.in) (S.D. Joshi).

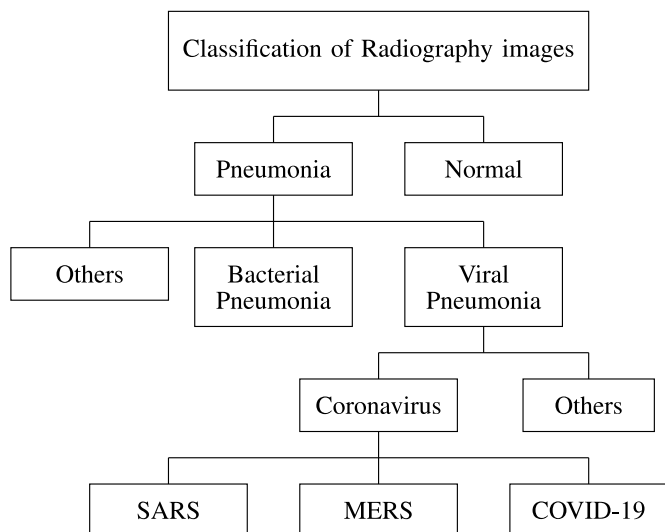
URL: <http://sbilab.iitd.edu.in/> (A. Gupta).

<https://doi.org/10.1016/j.combiomed.2022.105350>

Received 15 November 2021; Received in revised form 10 February 2022; Accepted 22 February 2022

Available online 3 March 2022

0010-4825/© 2022 Elsevier Ltd. All rights reserved.



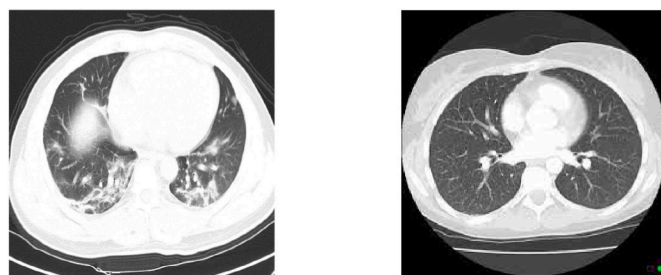
**Fig. 1.** Most common classes considered for labelling of CXR and CT-scan images, where SARS stands for Severe Acute Respiratory Syndrome and MERS stands for Middle East Respiratory Syndrome.

and the geographical spread of the population.

At the same time, efforts are underway for devising quick diagnostic solutions for COVID-19 detection through CXR and CT images that are analyzed routinely by radiologists. The manual diagnosis of COVID-19 is time-consuming, prone to human errors, and needs the assistance of a qualified radiologist. The availability of an expert radiologist is also required because the abnormalities during the early stages of COVID-19

may appear similar to the other pulmonary syndromes of Severe Acute Respiratory Syndrome (SARS) or Viral Pneumonia (VP) that can also pose an impediment to the timely diagnosis and treatment of COVID. As an example, some samples of CXR and CT images of COVID and non-COVID cases are shown in Fig. 2 and Fig. 3. The axial images show bilateral scattered ground-glass opacity with air-space consolidation in the posterior segments of lower lung lobes with the peripheral and subpleural distribution. Since CXR and CT are recommended for various pulmonary abnormalities, any automated solution designed to diagnose COVID-19 should also consider other respiratory disorders to develop a more comprehensive and robust diagnostic system.

The successful application of DL in computer vision and the biomedical domain has encouraged researchers to explore AI-based solutions for COVID-19 detection using CXR and CT-scan images. With the ongoing outbreak of COVID-19, though the research area is nascent but has shown tremendous potential and is progressing fast. Several studies

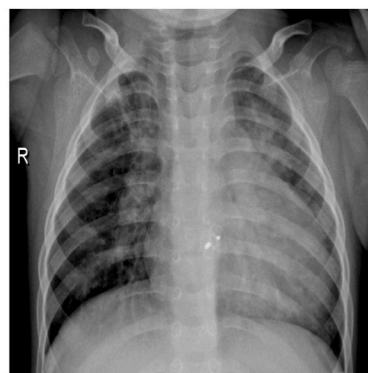


(a) (b)

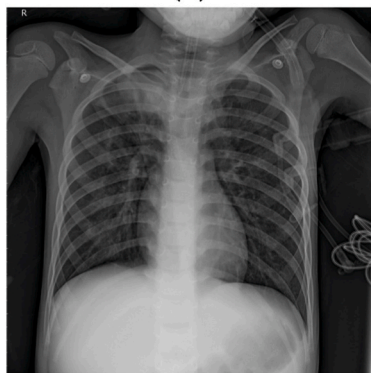
**Fig. 3.** CT-scan images of (3a) a COVID-19 and (3b) a healthy subject.



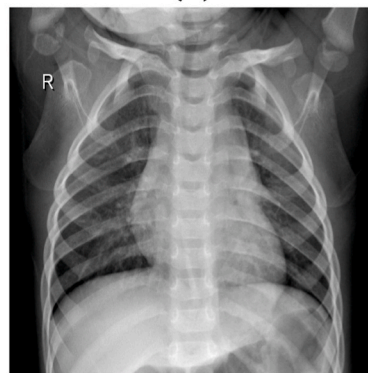
(a)



(b)



(c)



(d)

**Fig. 2.** CXR images of (2a) a COVID-19, (2b) a bacterial pneumonia, (2c) a viral pneumonia, and (2d) a healthy subject.

have been conducted for the automated diagnosis of COVID-19 using DL techniques [5,8]. Typically, the DL-based model consists of a hierarchical structure with a Convolutional Neural Network (CNN) as an important block, where each layer extracts the features pertinent to COVID-19 that can be used to classify COVID-19 images from non-COVID images. Propelled by CNN's automatic feature learning capabilities, deep neural networks-based COVID-19 classification is being widely used.

Of late, detection of COVID-19 using only CXR or CT images has shown potential in developing automated solutions. However, it is important to note that any automated solution for practical application needs a high detection rate and consistent performance over an unseen dataset. Thus, it requires advanced methods that can yield universally acceptable performance. Multimodal data analysis has the potential to yield better performance compared to single-modal data analysis because a DL model can learn robust and accurate features from a large heterogeneous dataset of multiple modalities and hence, can provide better classification performance [9,10]. Multimodal data analysis can be undertaken by considering CXR and CT images, thermal images, cough/speech, and blood samples. Due to the public availability of CXR and CT datasets, several single modal and multimodal data analysis studies have been published recently on COVID-19 detection having advantages, limitations, and challenges. To further increase the pace of research in COVID-19 diagnosis using CXR and CT images, a systematic survey and a comprehensive review of recent literature are required that can assist the researchers in the near future.

Motivated with the above, we present a review of single modal and multimodal DL-based research studies of COVID-19 and introduce an overall pipeline. We also highlight various challenges and limitations in this area and briefly discuss the future scope. Since the development of DL-based methods has been facilitated by the public availability of many CXR and CT datasets, we also present a detailed description of each dataset along with a summary of relevant information in a tabular form to highlight its popularity in the COVID-19 literature and also provide the links for the same.

Since the research in this field has started recently and is progressing fast, it is important to continuously review the developments that can help in catching up with the recent and push towards future developments. In literature, a few survey papers on COVID-19 image classification have been published [11–14] but a majority of these have reviewed a relatively small number of research papers mainly published in 2020. Our review includes a total of 71 research articles. Compared to the other survey papers, we only discuss studies that have used state-of-the-art DL techniques, have reported higher accuracy results, and are mainly published in 2021. In addition, our review is a comprehensive study that includes broad topics, such as DL-based classification pipeline, popular databases for COVID-19 classification, elaborate tables with details on pre-processing and, online data and code availability. Finally, we present a discussion on unique challenges and future directions of DL-based COVID-19 classification. Furthermore, compared to the other review studies that have focused only on either CXR or CT images, we have also covered multimodal works using both CXR and CT images.

A key objective of this review is to summarize notable DL-based studies that can help future researchers in overcoming the challenges to the successful realization of automated solutions for the quick and robust diagnosis of COVID-19. The salient contributions of this study are as follows:

1. It briefs the pipeline of different popular DL-based methods employed in the related studies;
2. It provides details of the widely used COVID-19 datasets available publicly;
3. It presents an overview of data augmentation, pre-processing methodology, and K-Fold cross-validation used in DL approaches along with their code and data availability for reproducing the results. This

information is important because it can help the researchers ascertain the reliability of the studies and can give the required push to further research; and

4. Finally, it suggests possible research directions by discussing unique challenges and future work based on:
  - the contribution percentage of each CNN learning method in the studied papers to find the most popular technique, and
  - the contribution percentage of COVID-19 dataset in the studied papers to target the creation of standard benchmark dataset.

This review paper is organized as follows. Section II presents an overview of a DL pipeline. Section III summarizes the publicly available imaging datasets for COVID-19 diagnosis. A literature review on CXR, CT, and multi-modality-based COVID-19 diagnosis is carried out in Section IV. Challenges with COVID-19 image analysis are presented in section V. Opportunities and future work are discussed in Section VI. Finally, Section VII concludes the study.

## 2. Deep learning-based COVID-19 classification

DL has advanced to a high level of maturity because of three primary factors: 1) availability of a high-end performing Graphics Processing Unit (GPU), 2) advancements in machine learning algorithms, especially CNN, and 3) access to a high volume of structured data. Consequently, DL methods have been very successful in COVID-19 detection using imaging data, whose details are presented next.

### 2.1. Overview of pipeline for COVID-19 image classification

Automated COVID-19 diagnosis with DL algorithms can be performed using data of various imaging modalities. The algorithm may include several steps, including pre-processing, segmentation, feature extraction, classification, performance evaluation, and explainable model prediction. Fig. 4 depicts a generic pipeline of DL-based COVID-19 diagnosis with steps discussed below.

#### 2.1.1. Data pre-processing

Pre-processing involves the conversion of the raw images into an appropriate format for further processing. Medical images collected from different devices vary in size, slice thickness, and the number of scans (e.g., 60 and 70 in CT). Together, these factors generate a heterogeneous collection of imaging data leading to non-uniformity across datasets. Thus, the pre-processing step largely involves resizing, normalization, and sometimes transformation from RGB to grayscale. In CT data, the voxel dimension is also resampled to account for the variation across the datasets that is also known as resampling to an isomorphic resolution [15]. Furthermore, images are improved via smoothing to improve the signal-to-noise ratio so remove the noise.

Another pre-processing step involves the extraction or segmentation of desired regions of interest from an image for the classification task. For example, the lungs are majorly prone to COVID-19 infection. Therefore, for a successful diagnosis, lung regions are segmented from CXR and CT images and fed to the next processing step. It is laborious, tedious, and time-consuming to manually segment the lung area, which also depends heavily on the knowledge and experience of the radiologists. DL-based segmentation techniques such as a few-shot segmentation [16,17] and semantic segmentation [18–20] can automatically identify infected regions providing rapid screening of COVID-19 images. For the segmentation task, there are widely-used segmentation models such as fully convolutional networks (FCN) [21], U-Net [22,23], V-Net [24], and 3D U-Net++ [25]. Sometimes, pixel values are thresholded to obtain a proper range of Hounsfield units (HU) to obtain the lung region. This step is particular to the dataset being used. The lung is an organ filled with air, and the air is the least dense object. Hence, pixel values are thresholded to segment the other non-lung tissue (e.g., skin, bone, or scanner bed) that may negatively impact the analysis.

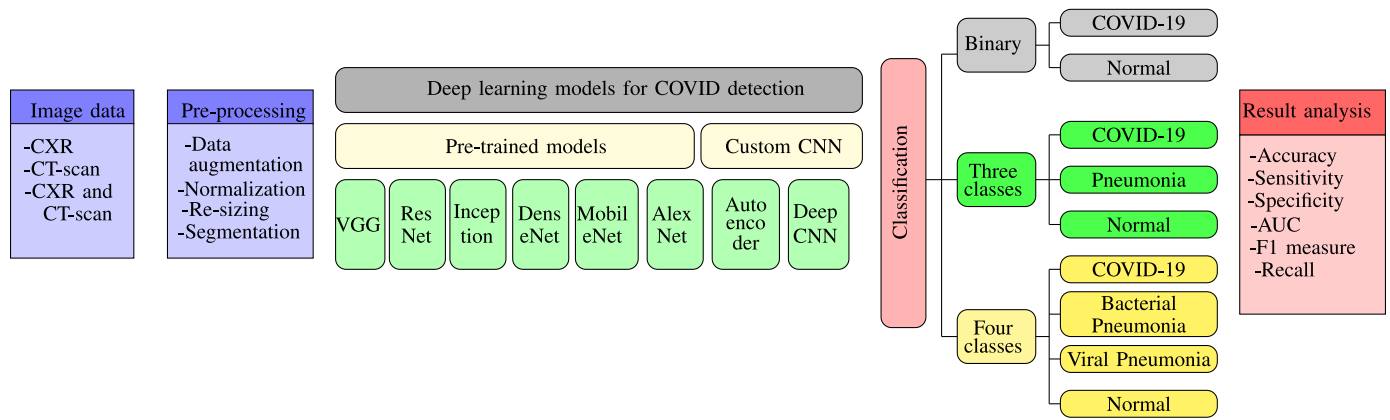


Fig. 4. A work flow of Deep learning based COVID-19 detection pipeline.

Of all DL models, U-Net is the most famous architecture for segmentation. It consists of two parts. The first part, considered as encoder, consists of a sequence of two 3x3 convolutional layers followed by a 2x2 Max Pool layer to learn the features at various levels. The second part, considered as a decoder, performs upsampling, concatenation with the correspondingly cropped features from the decoder layer, and two 3x3 convolutional operations. Through decoder operation, it tries to restore the learned feature maps to an image of original input size. U-Net has 23 convolutional layers in total. Karthik et al. (2021) [26] utilized repository-inspired U-Net architecture for segmenting lungs from CXR images.<sup>1</sup> Oh et al. (2020) [27] utilized FC-DenseNet103 architecture for the segmentation of lungs from CXR images and also compared the performance with the U-Net. It was also shown that the segmentation algorithm could be used for small training datasets, and the morphology of the segmentation mask can be used as a discriminatory biomarker. The segmentation scheme was tested on a cross-database to show statistically significant improvement in the segmentation accuracy. Another work by Wang et al. [28] utilized VGG based network for lung segmentation from CXR images. One famous work by Javaheri et al. (2021) [15] utilized BCDU-Net [29] to segment the lung area. This architecture was inspired by U-Net and utilized Bi-directional ConvLSTM along with densely connected convolutions. U-Net has also been used in other studies for lung segmentation [30–34]. Ouyang et al. (2020) [35] considered VB-Net [36], a combination of V-Net [24] and bottle-neck structure for segmentation.

### 2.1.2. Feature extraction and classification

The main step of DL-based COVID-19 diagnosis is feature extraction and classification. DL methods extract features automatically and carry out binary or multiclass classification. Feature extraction can be performed in two ways: using transfer learning with a pre-trained model or a custom CNN model developed from scratch. CNN is the core block of many DL-based neural networks that perform feature extraction from the input images. It consists of several convolutional and pooling layers. Apart from these basic layers, it also consists of several layers of batch normalization and includes Dropout. A schematic representation of a typical CNN is shown in Fig. 5 and is explained below.

1. Convolutional Layer: It consists of learnable filters (or kernels) that are convolved with the input images. It performs an element-wise dot product and sum to provide a number as an element of a matrix, called the feature map. Convolution operation follows two important features: *Local Connectivity* because filter weights are multiplied to a local area of the input image at a time, and *Weight Sharing* because the same filter weights are multiplied to every spatial location of the

input image. Convolutional layers work in a hierarchical manner, where low-level features are extracted in initial layers, and high-level features are extracted in deeper layers. The convolution operation is followed by an activation function (e.g., ReLU) that introduces non-linearity into the network.

2. Pooling Layer: This layer performs dimensionality reduction of the feature maps along the spatial dimension. It reduces the number of learnable parameters and thus, provides a reduction in the computational complexity. Average-Pooling and Max-Pooling are the two dominantly used pooling techniques.
3. Fully-connected Layer: It performs the actual classification task. It consists of several neural network layers. The number of layers and the number of nodes in each layer are called the hyperparameters required to be tuned optimally. It is followed by a softmax layer that provides a class score for every class to an image, similar to the probabilities of belonging to different classes. An input image is classified to the class corresponding to the highest class score.

Pre-trained models are the ones that have already been trained on other datasets by researchers. Generally, these models are trained on large databases such as the ImageNet database of natural images [37]. At first, forcing models to learn general image features is a preventive measure to avoid overfitting and learning domain-specific features. After ImageNet pre-training, the final 1000-node classification layer of the trained ImageNet model is removed and replaced by a  $n$ -node layer, corresponding to the  $n$ -class classification for COVID-19 detection. In transfer learning, learned weights of the pre-trained DL architecture are used as the initial starting point for training the new dataset. A schematic representation of the transfer learning approach is shown in Fig. 6. Transfer learning can be accomplished either by the fine-tuning weights of all the layers or by fine-tuning the weights of a few deeper layers. There are several pre-trained models that are used for COVID-19 diagnosis such as AlexNet [38], different versions of Visual Geometry Group (VGG) [39] and ResNet [40], Inception [41], Xception [42], InceptionResNet [43], DenseNet [44], etc. Apart from these pre-trained models, custom models are also popular for COVID-19 classification training, which implies training a model from scratch without utilizing any pre-trained model.

### 2.1.3. Performance evaluation

The performance of the overall pipeline is assessed by evaluation metrics such as accuracy, sensitivity, specificity, precision, F1-score, Area Under the receiver operating characteristic curve (AUC), and so on. Typically, the data is partitioned into training, validation, and testing sets for the experiment. The training data is used to develop a particular model, while the appropriateness of the training and the model is assessed by monitoring the overfitting or underfitting, respectively, on the validation data at the same time. Finally, the performance

<sup>1</sup> <https://github.com/imlab-uuij/lung-segmentation-2d>.

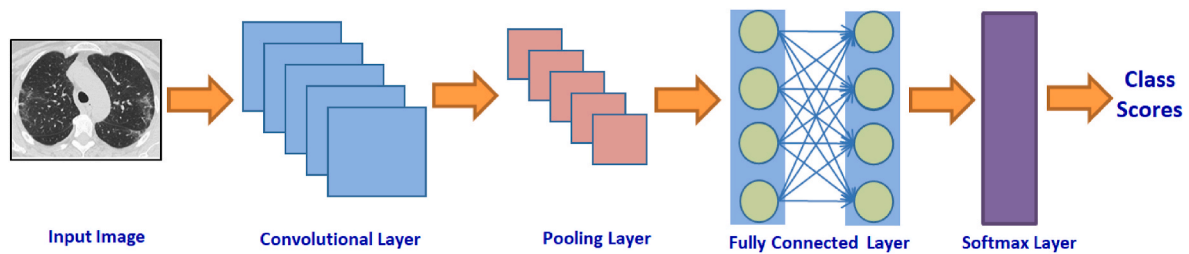


Fig. 5. Schematic representation of a typical Convolutional Neural Network architecture.

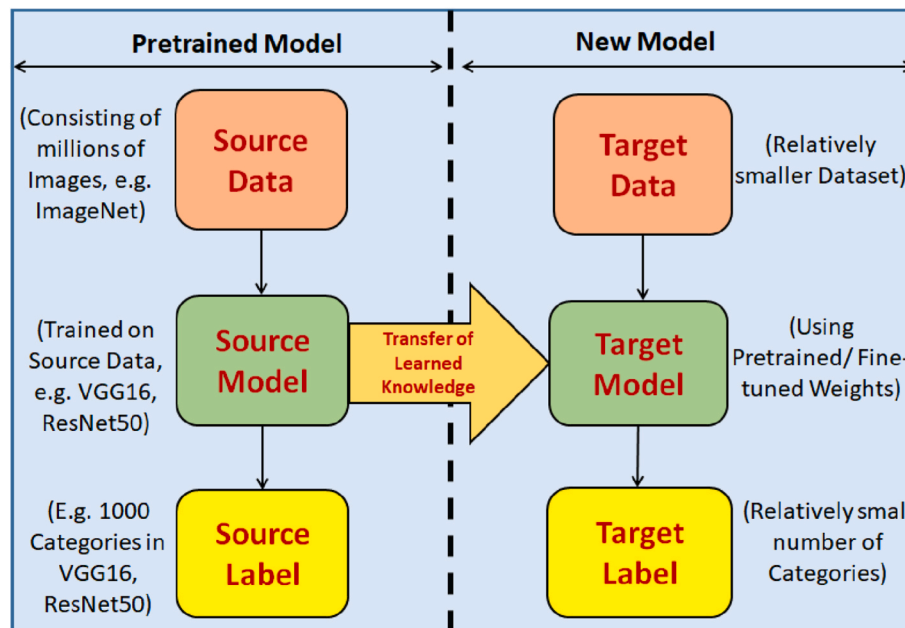


Fig. 6. Schematic representation of Transfer Learning approach.

of the developed model is tested on the unseen test data.

#### 2.1.4. Explanation of model prediction

Deep learning models are trained as black-box classifiers with no evidence of the correctness of the features extracted. Explainable AI is an emerging field that assigns important values to the input image regions leading to the predicted outcome. This assists radiologists in locating abnormalities in the lungs and gives an insight into important spatial areas that are responsible for distinguishing COVID-19 images from others. A few explainable models, including GRAD-CAM and GRAD-CAM++, used for COVID-19 diagnosis, are described in Table 1.

#### 2.2. CT vs CXR

CXR is the most easily accessible and the fastest form of imaging with lesser side effects on the human body. CXR imaging has been traditionally used for the detection of pneumonia and cancer. Although it can detect COVID-19 infection, it fails to provide fine-order details of the infected lungs. CT scan is a more sophisticated technique to evaluate the level of infection in various lobes of the lungs and is used to calculate the CT severity score of the patient. In fact, CXR is a 2D imaging, whereas CT provides 3D scans of organs from various angles. CXR imaging can be used for COVID-19 detection; however, to evaluate the level of severity of the infection, a CT scan is compulsory. This is one of the reasons that multimodal detection of COVID-19 using both CXR and CT scan images can give a better generalization ability to a neural network architecture.

Table 1

Techniques for visual explanation of Deep CNN.

Technique	Details
CAM [45]	Class Activation Mapping is a visual explanation technique for deep convolutional neural networks by providing class-discriminative visualization. The CNN model must be re-trained because it is modified by removing all dense layers and adding a Global Average Pooling layer before the softmax layer.
Grad-CAM [46]	Gradient-CAM is an upgrade of CAM that does not need any architectural change or re-training. It uses the gradient details passing into the last convolutional layer to visualize the significance of each neuron. In an image, if the same class occurs multiple times, it fails to localize objects accurately. Also, it is not able to produce the heat map of the complete object.
Guided Grad-CAM	This technique upsamples the Grad-CAM maps and performs point-wise multiplication with the visualizations from Guided Backpropagation. It provides fine-order and class-discriminative visualization.
Grad-CAM++ [47]	Grad-CAM++ uses more sophisticated backpropagation to overcome issues of CAM and Grad-CAM techniques. It provides better visual explanations of CNN model predictions in terms of better object localization as well as explaining occurrences of multiple object instances in a single image.

### 3. Public imaging datasets for COVID-19 detection

In all, about 35 public datasets (CXR and CT images) have been referred to and used by researchers to validate the algorithms in the articles reviewed in this work. The details are listed in Table 2. Some of

**Table 2**  
Public Imaging Datasets used for COVID-19 Diagnosis.

Reference	Image type	Links	Reference Papers
Ali (2020) [50]	CXR	<a href="https://www.kaggle.com/ahmedali2019/pneumonia-sample-xrays">https://www.kaggle.com/ahmedali2019/pneumonia-sample-xrays</a>	[51]
BIMCV (2020) [52]	CXR	<a href="https://bimcv.cipf.es/bimcv-projects/padchest/">https://bimcv.cipf.es/bimcv-projects/padchest/</a>	[53]
CC-COII database [54]	CT	<a href="http://ncov-ai.big.ac.cn/download?lang=en">http://ncov-ai.big.ac.cn/download?lang=en</a>	[30,55]
Chest Imaging (2020) [56]	CXR	<a href="https://threadreaderapp.com/thread/1243928581983670272.html">https://threadreaderapp.com/thread/1243928581983670272.html</a>	[5,57]
Chung (2020) [58]	CXR	<a href="https://github.com/agchung/Actualmed-COVID-chestxray-dataset">https://github.com/agchung/Actualmed-COVID-chestxray-dataset</a>	[26,57,59–61]
Cohen et al. (2020) [62]	CXR and CT	<a href="https://github.com/ieee8023/covid-chestxray-dataset">https://github.com/ieee8023/covid-chestxray-dataset</a>	[5,9,10,23,26–28,51,53,55,57,59–61,63–90]
COVIDGR [91]	CXR	<a href="https://dasci.es/es/transferencia/open-data/covidgr/">https://dasci.es/es/transferencia/open-data/covidgr/</a>	[91]
Dadario AMV. COVID-19 X-rays	CXR and CT	<a href="http://dx.doi.org/10.34740/KAGGLE/DSV/1019469">http://dx.doi.org/10.34740/KAGGLE/DSV/1019469</a>	[72]
European Society of Radiology [92]	CXR and CT	<a href="https://www.eurorad.org/advanced-search?search=COVID">https://www.eurorad.org/advanced-search?search=COVID</a>	[65]
Gunraj et al. (2020) [93]	CT	<a href="https://www.kaggle.com/hgunraj/covidxct?select=2A_images">https://www.kaggle.com/hgunraj/covidxct?select=2A_images</a>	[94]
Irvin et al. (2019) [95]	CXR	<a href="https://stanfordmlgroup.github.io/competitions/chexpert/">https://stanfordmlgroup.github.io/competitions/chexpert/</a>	[57]
Jaeger et al. [96]	CXR	<a href="https://openi.nlm.nih.gov/faq#faq-tb-coll">https://openi.nlm.nih.gov/faq#faq-tb-coll</a>	[23,27]
JSRT [97]	CXR	<a href="http://db.jsrt.or.jp/eng-01.php">http://db.jsrt.or.jp/eng-01.php</a>	[23,27,70]
Kermany et al. (2018) [48]	CXR	<a href="https://data.mendeley.com/datasets/rscbjbr9sj/2">https://data.mendeley.com/datasets/rscbjbr9sj/2</a>	[23,69,72,75,77,81,89,90,98,99]
Khoong (2020) [100]	CXR	<a href="https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets">https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets</a>	[59]
LIDC-IDRI database [101]	CT	<a href="https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI">https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI</a>	[30]
Montgomery tuberculosis [96]	CXR	<a href="https://www.kaggle.com/raddar/tuberculosis-chest-xrays-montgomery">https://www.kaggle.com/raddar/tuberculosis-chest-xrays-montgomery</a>	[23,27]
Mooney (2017) [49]	CXR	<a href="https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia/version/2">https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia/version/2</a>	[5,10,26,57,60,63–65,67,72,73,76,85,87,88]
MosMedData [102]	CT	<a href="https://mosmed.ai/datasets/covid19_1110/">https://mosmed.ai/datasets/covid19_1110/</a>	[30,103]
Patel et al. (2020) [104]	CXR	<a href="https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia">https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia</a>	[94]
Praveen et al. (2020) [105]	CXR	<a href="https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset">https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset</a>	[27]
Rahman et al. (2020) [106]	CXR	<a href="https://www.kaggle.com/tawsifurrahman/covid19-radiography-database">https://www.kaggle.com/tawsifurrahman/covid19-radiography-database</a>	[5,51,59,61,74,75,107]
Radiology Assistant	CXR and CT	<a href="https://radiologyassistant.nl/chest/covid-19/covid19-imaging-findings">https://radiologyassistant.nl/chest/covid-19/covid19-imaging-findings</a>	[63]
Radiopaedia [108]	CXR and CT	<a href="https://radiopaedia.org/search?lang=us&amp;q=covid&amp;scope=cases">https://radiopaedia.org/search?lang=us&amp;q=covid&amp;scope=cases</a>	[5,9,26,60,79,90,109,110]
RSNA (2020) [111]	CXR	<a href="https://www.kaggle.com/rsna-pneumonia-detection-challenge">https://www.kaggle.com/rsna-pneumonia-detection-challenge</a>	[5,26,28,80,90,109,112]
Sajid [113]	CXR	<a href="https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images">https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images</a>	[59]

**Table 2 (continued)**

Reference	Image type	Links	Reference Papers
Shenzhen [114]	CXR	<a href="https://lhncbc.nlm.nih.gov/LHC-publications/pubs/TuberculosisChestXrayImageDataSets.html">https://lhncbc.nlm.nih.gov/LHC-publications/pubs/TuberculosisChestXrayImageDataSets.html</a>	[23]
SIRM (2020) [115]	CXR and CT	<a href="https://sirm.org/category/senza-categoria/covid-19/">https://sirm.org/category/senza-categoria/covid-19/</a>	[5,26,57,60,65,90,109,110]
SARS-COV-2 CT-Scan (2020) [116]	CT	<a href="https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset">https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset</a>	[9,59,117,118]
Tianchi-Alibaba database [119]	CT	<a href="https://tianchi.aliyun.com/dataset/dataDetail?dataId=90014">https://tianchi.aliyun.com/dataset/dataDetail?dataId=90014</a>	[30]
USCD-AI4H [120]	CT	<a href="https://github.com/UCSD-AI4H/COVID-CT">https://github.com/UCSD-AI4H/COVID-CT</a>	[10,59,117,118,121–123]
Vaya et al. (2020) [124]	CXR and CT	<a href="https://bimcv.cipf.es/bimcv-projects/bimcv-covid19/">https://bimcv.cipf.es/bimcv-projects/bimcv-covid19/</a>	[23,53]
Wang et al. (2017) [125]	CXR	<a href="https://github.com/muhammedtaloo/COVID-19/tree/master/X-Ray">https://github.com/muhammedtaloo/COVID-19/tree/master/X-Ray</a> , <a href="https://www.kaggle.com/nih-chest-xrays/sample">https://www.kaggle.com/nih-chest-xrays/sample</a>	[53,66,68,69,79,83,84,98]
Wang et al. (2020) [126]	CXR	<a href="https://github.com/li-ndawangg/COVID-Net">https://github.com/li-ndawangg/COVID-Net</a>	[112]
Yan et al. (2020) [127]	CT	<a href="https://iee-dataport.org/authors/tao-yan">https://iee-dataport.org/authors/tao-yan</a>	[103]

these datasets contain CXR images and CT-scan images of COVID-19, while others include those of normal subjects and different pulmonary diseases. The reason for using the latter type of datasets is to create more generalizable algorithms that can detect COVID-19 from a pool of more diverse radiography images. We briefly discuss some of these datasets and provide the download links for the dataset in Table 2 for ease of the readers and further research.

### 3.1. CXR dataset

We have included 20 datasets for CXR images, among which Kermany et al. (2018) [48] is the most popular dataset for normal and pneumonia CXR images. This dataset [48] consists of 5856 images, where 2780 images are of bacterial pneumonia, 1493 are of viral pneumonia, and 1583 belong to the normal subjects. The participants were recruited at the Guangzhou Women and Children’s Medical Center. This dataset was used to develop Mooney dataset (2017) [49], which is a CXR dataset available as a Kaggle competition on viral and bacterial pneumonia classification. It consists of 5247 CXR images of normal, viral, and bacterial pneumonia with varying resolution. Out of 5247 CXR images, 3906 images are from different subjects affected by pneumonia (2561 images for bacterial pneumonia and 1345 images for viral pneumonia), and 1341 images are from normal subjects. Wang et al. (2017) dataset [125] has also been used in various studies. This dataset was released by the National Institutes of Health (NIH), having 108,948 CXR images of normal lungs with no lung infection and non-COVID pneumonia cases. A Kaggle dataset with only pneumonia images was developed by Ali [50] which includes viral and bacterial pneumonia CXR images from 53 patients. Radiology Society of North America (RSNA) collaborated with US National Institutes of health, the society of thoracic radiology, and MD.ai to develop a Pneumonia-Detection-Challenge database on Kaggle [111] which includes CXR images of 6002 normal and 20599 pneumonia patients.

To verify the performance of the proposed classification algorithm for differentiating COVID-19 CXR images from other pulmonary syndromes, authors have used datasets such as BIMCV (2020) [52], JSRT dataset [97], Irvin et al. (2019) [95] and Jaeger dataset [96,114]. Bustos et al. (2020) [52] introduced a dataset of more than 160,000 CXR images collected from 67,000 subjects at Hospital San Juan (Spain) from 2009 to 2017. This dataset includes CXR for COPD, pneumonia, heart

insufficiency, pulmonary edema, pulmonary fibrosis, emphysema, tuberculosis, and other pulmonary syndromes. 27% of the data was manually labeled by physicians, and the rest was labeled using recurrent neural network. The Japanese Society of Radiological Technology (JSRT) dataset [97] was marked by radiologists for the detection of lung cancer nodules. This dataset contains 247 CXRs from 14 institutions, out of which 154 cases contain nodule markings. In addition, lung masks are also provided that can be used to study the performance of lung segmentation. The dataset of Irvin et al. (2019) [95] includes 224,316 CXRs of 65,240 patients divided into 14 classes, including no findings, enlarged cardiomegaly, lung lesion, lung opacity, and edema; however, no COVID-19 cases were included in this study. The U.S. National Library of Medicine has made two datasets of Postero Anterior (PA) CXR images of various pulmonary diseases with a majority of cases considered for pulmonary tuberculosis (TB) [96]. These two datasets were collected from the Department of Health and Human Services, Montgomery County, Maryland, USA, and Shenzhen No. 3 People's Hospital in China. The former dataset consists of 138 frontal CXR, including 80 normal and 58 TB cases. The Shenzhen dataset consists of 662 CXR, of which 326 are normal and 336 are TB [114].

The above-mentioned CXR datasets do not include COVID-19 infected CXR images and are, thus, insufficient for validating COVID-19 classification algorithms. For this purpose, datasets like Rahman et al. (2020) [106], Chest imaging (2020) dataset [56], Chung dataset [58], COVIDGR dataset [91] have been developed. Rahman et al. (2020) [106] is a COVID-19 CXR database created by a team of researchers from Qatar University (Doha, Qatar) and Dhaka University (Dhaka, Bangladesh) along with the collaborators from Pakistan and Malaysia. The dataset consists of COVID-19 positive, normal, and viral pneumonia CXR images, and it is constantly updated with new CXR images. Chest imaging dataset [56] includes 103 COVID-19 CXR images from a thread reader uploaded by a doctor from Spain. Chung dataset [58] developed in the University of Waterloo Canada, includes 35 COVID-19 and non-COVID-19 CXR images from 31 patients. COVIDGR dataset [91] is a homogeneous and balanced dataset that includes mild, moderate as well as severe cases of COVID-19 and also normal cases. It includes 426 positive and 426 negative PA CXR views. The dataset was developed in collaboration with Hospital Universitario Clínico San Cecilio, Granada, Spain.

There are several publicly available datasets, such as Mooney dataset [49], which do not include new original images but have been developed by collating the data of existing datasets. For example, Wang et al. (2020) [126] developed a dataset, COVIDx, consisting of 13,975 CXR images of 13,870 patients. The dataset was developed using five publicly available datasets, where COVID-19 cases were acquired from Cohen [62], Chung [58], and Rahman [106]. Non-COVID-19 pneumonia cases were acquired from Cohen [62] and RSNA pneumonia detection challenge dataset [111]. Finally, normal cases were collected from RSNA pneumonia detection challenge dataset [111]. The Khoong dataset [100] was constructed using normal and COVID-19 manifested CXR images from Cohen dataset [62], and from [https://github.com/JordanMicahBennett/SMART-CT-SCAN\\_BASED-COVID19\\_VIRUS\\_DETECTOR/](https://github.com/JordanMicahBennett/SMART-CT-SCAN_BASED-COVID19_VIRUS_DETECTOR/). Another such dataset available on Kaggle is Patel [104], which consists of 6432 CXR images from normal, COVID-19, and pneumonia infected subjects acquired from three datasets, namely Cohen dataset [62], Mooney [49], and Chung dataset [58]. To include other pulmonary syndromes, Praveen [105] constructed a dataset consisting of 5800 CXR images from normal, COVID-19 pneumonia, SARS, Streptococcus, and ARDS (acute respiratory distress syndrome). These images have been acquired from Cohen dataset [62]. Sajid [113] consists of 10,000 CXR images created using data augmentation techniques. These images include normal and COVID-19 cases; however, the original source of the dataset has not been mentioned. This data set has been used so far by Ref. [59], where eight different datasets were used to form a COVID-R dataset consisting of 2843 COVID-19 CXR images, 3108 normal, and 1439 viral and bacterial pneumonia manifested CXR

images. These eight datasets include Cohen [62], UCSD-AI4H dataset [120], Chung dataset [58], SARS-COV-2 CT-scan dataset [116], Khoong dataset [100], and Rahman [106].

### 3.2. CT dataset

SARS-COV-2 CT-Scan dataset [116] has 1252 CT scans of 60 patients infected by COVID-19 and 1230 CT scan images of 60 infected patients by pulmonary diseases. CC-CII dataset [54] is a dataset of the CT images collected from cohorts from the China Consortium of Chest CT Image Investigation. Seven hundred fifty CT scans were collected from 150 COVID-19 subjects, and these slices were manually segmented. All CT images are classified into novel coronavirus pneumonia (NCP) due to SARS-CoV-2 virus infection, common pneumonia, and normal controls. LIDC-IDRI dataset [101] includes CT-scan images of 1018 lung cancer with labeled annotated lesions. The dataset was collected in collaboration with seven academic centers and eight medical imaging companies. MosMedData [102] consists of 1110 CT-scans of COVID-19 patients collected between March 1, 2020 and April 25, 2020 in municipal hospital Moscow. Yan et al. [127] published a dataset on IEEE dataport consisting of 416 CT-scan images of 206 COVID-19 patients from two hospitals. The dataset also includes 412 CT-scan images of non-COVID-19 pneumonia patients. UCSD-AI4H dataset [120] consists of 349 CT-scans of 216 COVID-19 patients. The dataset has been confirmed by a radiologist from Tongji hospital. Tianchi-Alibaba database [119] consists of 20 CT scans of COVID-19 patients along with the segmentation of lungs and infections.

The above-mentioned dataset included CT-scan images collected in collaboration with hospitals. There are other publicly available datasets, sometimes available as Kaggle competitions, which have been developed by combining two or more of the original datasets. For example, Gunraj et al. [93], also known as COVIDx CT dataset, is available on kaggle. The first version was released in December 2020, and the second version was released in January 2021. The dataset includes three classes normal, pneumonia, and COVID-19. This dataset is presented in two subsets, "A" and "B", where the former includes cases with confirmed diagnoses and the latter includes all images from "A" and also those which are weakly verified. This dataset was constructed using publicly available datasets like Radiopaedia.org [108], MosMedData [102], CNCB 2019 novel coronavirus resource (2019nCoV) AI diagnosis dataset [128], COVID-19 CT lung and infection segmentation dataset [129], LIDC-IDRI [101], and integrative CT Images and Clinical Features for COVID-19 (iCTCF) [130].

### 3.3. CT and CXR dataset

Cohen et al. [62] is a publicly available dataset consisting of CT-scan and CXR images from 468 COVID-19 patients, 46 bacterial cases of pneumonia, 10 MERS, 5 Varicella, 4 Influenza, 3 Herpes, 16 SARS, 26 Fungal cases, and 59 unknown cases. Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 database [115] consists of COVID-19 positive radiographic images (CXR and CT) with varying resolution. This database is constantly updated with new images. Vaya et al. [124] is a multimodal dataset introduced from the Valencian Region Medical Image Bank (BIMCV) containing chest radiographs of COVID-19 patients, having 2265 chest radiographs belonging to 1311 patients. There are no normal cases in this dataset. Radiopaedia [108] dataset consists of case studies of several diseases, including COVID-19. It provides both CXR and CT images and has been considered as an authentic source of dataset for deep learning-based analysis.

## 4. Recent advances in COVID-19 image analysis

CXR images are generally used as a first-line imaging modality for patients under investigation of COVID-19 and have been analyzed in numerous studies of COVID-19 diagnosis. This imaging is comparatively



inexpensive and is less hazardous to human health owing to being a low radiation modality. Table 3 lists the most relevant state-of-the-art studies in this direction published in recent years.

CT images are processed differently than CXR images. CT data is three-dimensional (3D), consisting of several slices (16, 32, 64, 128, etc.) acquired during the scan. The slice capturing the largest lung region is selected and is often treated as a separate image. Some publicly available datasets consist of only one CT slice per subject. In other cases, all the slices are treated as independent samples for diagnosis that helps in increasing the number of images during training. In the testing phase, majority voting is done to map decisions on multiple slices of a subject to ascertain the class label. In some recent studies, three-dimensional (3D) CT data is utilized with 3D segmentation models and 3D-CNN architectures.

Deep learning is a data-driven approach where classification decisions are made based on the features learned by a model during the training process. During test time, the model assumes that the input has some features similar to the features learned from the training dataset that could be used for decision making. However, if patterns are dissimilar, the model will not be able to classify them accurately, which reduces its generalization ability. Data augmentation is a technique used to overcome this limitation. However, since artificial images generated through data augmentation are from the same training dataset, its scope to improve the diversity or abundance of the features is limited. In such scenarios, a more effective approach towards improving the performance is the augmentation of the actual training dataset through multiple modalities. For detection of COVID-19, a model can achieve superior performance when a multimodal dataset is utilized compared to the single-modal analysis. For example, the performance of the COVID-19 detection model based on only CXR or CT scan images can be further improved by incorporating both kinds of images into the model. Various models are used on CXR imaging, CT imaging, and combining both as described next. These studies can be categorized based on the DL architectures used. Tables 4 and 5 list the most relevant state-of-the-art CT and multimodal based studies respectively.

#### 4.1. Transfer learning work

##### 4.1.1. AlexNet

It is one of the first convolutional networks that performed a large-scale image classification task and revolutionized the application of deep learning. It was the winner of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. It consists of 5 convolutional layers and three dense layers. It is similar to the famous LeNet architecture but incorporates improved techniques of Rectified Linear Unit (ReLU) activation function, dropout, data augmentation, and multiple GPUs. Though several improved neural network architectures have been introduced to date, most are based on or inspired by the AlexNet network.

It has been used in many studies of COVID-19 detection [89,131] which mainly differs in feature selection and training of multiple classifiers in Ref. [131] and image generation using Generative Adversarial Network (GAN) in Ref. [89]. The use of GAN is to tackle the lack of a sufficient dataset for COVID-19 cases, which is one of the most important contributions of this work. Authors in Ref. [89] also compared AlexNet performance with two other deep transfer learning models, GoogleNet and ResNet18. More in-depth details of both these references are presented in Table 3.

##### 4.1.2. VGGNet

VGG stands for Visual Geometry Group of Oxford University. It was the winner of the ILSVRC challenge 2014. This model is simple in architecture but is still very effective in performance. VGG16 and VGG19 architectures consist of 16 and 19 convolutional layers, respectively. VGGNet has a cascade of five convolutional blocks using the fixed kernel sizes of 3X3, where the first two blocks consist of two convolutional

operations in each, and the last three blocks consist of three convolutional operations in each. It is pertinent to mention that a new convolution block along with process improvement techniques (batch normalization and dropout) can be easily added to the standard model that enables the learning of finer features and is more suitable for the newer tasks and improved learning speed/stability.

Few initial studies utilized the VGGNet pre-trained model by adding or fine-tuning few layers for COVID-19 classification using CXR (two-class [64,78] and three-class classification [88,90]), CT [123] and multimodal images [82,109]. Improvements were made by tweaking the pipeline. For example, Abbas et al. (2021) [70] proposed a decompose, transfer and compose method for the classification of CXR images into normal, COVID-19 and SARS classes. First, deep local features are extracted using a pre-trained CNN model, principal component analysis (PCA) is used to reduce the dimensionality of the obtained feature set. The class decomposition layer is then used on the feature matrix to form sub-classes within each class, and each subclass is treated as an independent class. The final classification layer of the pre-trained VGG19 model is adapted to these sub-classes. The parameters of the adopted model are fine-tuned, and finally, the sub-classes are combined to give the predicted label. The final classification is refined using error-correction criteria. This method significantly improved results over the conventional VGG19 pre-trained model. Results were also compared against four different pre-trained models for the three-class classification problem.

Another work by Heidari et al. (2020) [75] combined the original CXR image with two pre-processed images to form a pseudocolor image which is then fed as three input channels for VGG16 pre-training. In Ref. [132], features of the convolutional layers of VGG16 were combined with the attention module (spatial attention module and channel attention module), followed by fully connected layers and softmax layer for COVID-19 classification based on CXR images. Brunese et al. (2020) [84] trained two models using VGG16. The first model discriminates healthy CXR images, and the second model detects COVID-19 from other generic pulmonary diseases.

##### 4.1.3. ResNet

ResNet is the most famous pre-trained model that has been used widely for COVID-19 classification. Generally, it is assumed that the training performance of the model can be increased by adding more convolutional blocks. However, in practice, it has been observed that the performance of the deeper layer models starts decreasing and often returns diminishing results compared to the less deep models. This happens due to the problem of vanishing gradients. ResNet model overcomes this limitation by incorporating skip connections. ResNet consists of a cascade of several residual blocks, wherein the output of each convolutional block is added to the output of the convolution blocks of the deeper stages. ResNet has been used by several authors for the detection of COVID-19 using CXR images [23,27,83,85,91], and CT images [30,32,117,133]. Further, details of these references are given in their respective tables.

Further improvements were made by a few studies, such as the utilization of feature pyramid network along with ResNet for COVID-19 classification using CXR images in Ref. [28], a two-step classification algorithm implementation using CXR images in Ref. [76], wherein first, ResNet50 was used to classify CXR images into healthy and others. Further, ResNet101 was used to separate COVID-19 from the other viral pneumonia class. Authors in Ref. [103] combined multiple image-level ResNet50 predictions to diagnose COVID-19 on a 3D CT volume level. The performance of the proposed method was shown to be better than a single 3D-ResNet model. Authors in Ref. [33] proposed a local attention classification model using ResNet18 as backbone architecture. Ismael and Sengur [63] used SVM with the linear, quadratic, cubic, and Gaussian kernels as a classifier on ResNet50 features for COVID-19 classification using CXR images.

ResNet50 was also used in Ref. [107] along with detecting and

**Table 3**

**Summary of state-of-art DL techniques used for the COVID-19 classification using CXR** Abbreviations: Acc.- Accuracy, BP-Bacterial Pneumonia, C-COVID-19, CAM- Class Activation Maps, CAP- Community Acquired Pneumonia, CN- COVID-19 negative, FPN- Feature Pyramid Network, HU- Hounsfield Units, Infl.- Influenza, LT- Lung Tumor, N-Normal, NF- No Findings, P- Pneumonia, Rad.- Radiologist, SARS- Severe Acute Respiratory Syndrome, Seg.- Segmentation, VP- Viral Pneumonia, Sen.- Sensitivity, Spe.- Specificity.

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
Abbas et al. [70]	Classes:3C/N/SARS 105/88/11	Augmentation, contrast enhancement	VGG19 with class decomposition and composition	✓	✓	×	97.4	98.2	96.3	handled the class-imbalance problem using the proposed architecture
Abraham and Nair [72]	Classes:2C/CN 453/497	Resized to different dimensions	Features extracted from multi-CNNs (Squeezenet, Darknet-53, MobilenetV2, Xception, Shufflenet); feature	×	✓	×	91.2	98.5	–	Correlation-based feature selection; bilinear interpolation for resizing; three RGB channels processing with single grayscale image being replicated to all the three channels
Afshar et al. [67]	Classes2: C/CN 71/7		selection and Bayesnet classifier				97.4	98.6	–	
	Classes:2C/CN (The number of images are not mentioned)	Resized to 224 × 224	Custom CNN	✓	✓	×	98.3	80.0	98.6	4 convolutional layers and 3 Capsule layers; modified the loss function to handle the class-imbalance problem
Agrawal and Choudhary [65]	Classes:2C/N 1143/1 345	Augmentation; resized to 224 × 224; normalization	Custom CNN	×	✓	✓	99.2	99.2	99.2	FocusNet [144] inspired CNN architecture having combination of multiple convolutional, 3 residual, and 2 squeeze-excitation blocks in between; evaluation by weighted
	Classes:3C/N/P 1143/1 345/1345						95.2	95.2	95.6	F1-score; handled the class-imbalance problem by oversampling technique such as SMOTE; validation done on two separate datasets
Al-Bawi et al. [88]	Classes:3C/N/VP 310/654/864	None	VGG16	✓	✓	×	95.3	98.5	98.9	Replaced last fully connected layer with 3 new convolutional layers
Apostol et al. [90]	Classes:3C/N/BP 224/504/700	Resized to 200 × 266, black background of	VGG19	×	✓	×	93.5	92.8	98.7	Fixed feature extractor with modification only in the last layer
	Classes:3C/N/P 224/504/714	1:1.5 ratio was added to avoid distortion					96.8	98.7	96.5	
Brunese et al. [84]	Classes:3C/ Pulmonary disease/ N 250/2 753/3 520	Resized to 224 × 224	VGG16, Grad CAM	×	✓	×	97.0	91.0	96.0	Fixed feature extractor with fine tuning of only last layers; added few layers like average pooling, flatten, dense, and dropout layers; two binary classifiers- training one for healthy and pulmonary, and the other for COVID and rest
Chowdhury et al. [5]	Classes:3C/N/VP 423/423/423	Augmentation; resized to 224 × 224; normalization	DenseNet201, activation mapping	×	✓	×	97.9	97.9	98.8	Investigation of features of deep layers
Das et al. [57]	Classes:2C/CN 538/468	Resized to 224 × 224, Normalization	Weighted averaging: DenseNet201 Resnet50V2 Inceptionv3	✓	✓	✓	91.6	95.1	91.7	Development of a Graphical User Interface (GUI)-based application for public use
DeGrave et al. [53]	Classes:2C/CN 408/30 805	Augmentation; resized to 224 × 224	DenseNet121, interpretation by expected gradient & CycleGAN	✓	✓	×	–	–	–	Classifier training on 15 classes; comparison of results using AUC
Dhiman et al. [85]	Classes:2C/N 50/50	Resized to 280 × 280	ResNet101	×	✓	✓	100	100	98.9	Analysis of segmented chest area; computational time analysis of multiple architectures; use of J48 decision tree classifier; fine-tuning using a multi-objective spotted hyena optimizer
Ezzat et al. [73]	Classes:2C/N 99/207	Augmentation; resized to 180 × 180; normalization	DenseNet121; Grad-CAM	×	✓	×	98.38	98.5	98.5	Hyper-parameters optimization using gravitational search algorithm
Gupta et al. [74]	Classes:3C/N/P 361/365/362	Augmentation, fuzzy color image enhancement and	Integrated stacked multiple CNNs (ResNet101, Xception,	×	✓	×	99.1	–	–	Both image enhancement and denoising

(continued on next page)

Table 3 (continued)

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
Hammoudi et al. [145]	Classes:2C/NC 361/727	stacking it with original Resized 224 × 224 × 3	InceptionV3, MobileNet, and NASNet), Grad-CAM				99.5	–	–	Measures were presented to associate survival chance with COVID-19 using risk factors like comorbidity, age, and infection rate indicator; Predicted patients' health status.
	Classes:4C/N/VP/BP 1493/1 493/1493/1 493	Resized to 310 × 310	DenseNet169	×	✓	×	99.1	–	–	
Heidari et al. [75]	Classes:3C/N/P 415/2 880/5 179	Augmentation, histogram equalization, bilateral low-pass filtering, pseudo-color image generation	VGG16	×	✓	×	94.5	98.4	98.0	handled class-imbalance problem by class weighting; removal of diaphragm regions; three channel processing; addition of 3 fully connected layers in the end
Hemdan et al. [78]	Classes:2C/N 25/25	Resized to 224 × 224	VGG19	×	✓	×	90.0	–	–	One hot encoding on the labels of the dataset i.e. '1' for COVID-19 and '0' for all other images in the dataset
Ismael and Sengur [63]	Classes:2C/N 180/200	Augmentation; resized to 224 × 224, grayscale image copied three times to form RGB image	ResNet50 with SVM	×	✓	×	94.7	91.0	98.9	No fine-tuning of ResNet50; analysis of eight well-known local texture descriptors of images
Islam et al. [60]	Classes:3C/N/P 1525/1 525/1 525	Augmentation; resized to 224 × 224	Custom CNN with LSTM, heatmaps	×	✓	✓	99.4	99.1	98.9	12 convolutional layers with 1 fully connected layer and 1 LSTM layer
Jain et al. [76]	Classes:2C/CN 440/1 392	Augmentation, resized to 640 × 640, normalization	ResNet50, ResNet101, Grad-CAM	×	✓	✓	97.2	–	–	Training of 2 two-class classification networks
Karthik et al. [26]	Classes:4C/N/BP/VP 558/10 434/2780/1 493	Augmentation; resized to 256 × 256	U-Net; custom CNN; interpretation analysis by class saliency maps, guided backpropagation, & Grad-CAM	×	✓	✓	97.9	99.8	–	Channel-shuffled dual-branched CNN comprising of three types of convolutions: (1) depth-wise separable convolution, (2) grouped convolution and (3) shuffled grouped convolution; augmentation done with distinctive filters learning paradigm
Keles et al. [98]	Classes:3C/N/VP 210/350/350	Augmentation; resized to 224 × 224	Custom CNN	×	✓	×	97.6	98.7	98.7	One input convolutional layer followed by 2 residual type blocks and 3 fully connected layers
Khan et al. [87]	Classes:4C/N/BP/VP 284/310/330/327	Resized to 224 × 224, resolution of 72 dpi	XceptionNet	✓	✓	✓	89.6	90.0	96.4	handled the class-imbalance problem by undersampling
	Classes:3C/N/P 284/310/657						95.0	95.0	97.5	
	Classes:2C/N 284/310						99.0	98.3	98.6	
	Classes:3C/N/P 157/500/500						90.2	–	–	
Loey et al. [89]	Classes:4C/N/BP/VP 69/79/79/79	Augmentation; resized to 512 × 512; normalization	GoogleNet	×	✓	×	80.6	80.6	–	Image generation using Generative Adversarial Network (GAN)
	Classes:3C/N/BP 69/79/79		AlexNet				85.2	85.2	–	
	Classes: 2C/N 69/79		AlexNet				100	100	–	
Luz et al. [112]	Classes:3C/N/P 189/8 066/5 521	Augmentation; normalization	EfficientNet; activation mapping	✓	✓	×	93.9	96.8	–	Hierarchical classification; use of swish activation; computational cost analysis by multiply-accumulate (MAC) operations
Mahmud et al. [99]	Classes:2C/N 305/305	Resized to 256 × 256, 128 × 128, 64 × 64, and 32 × 32; normalization	Stacked Custom CNN, Grad-CAM	✓	✓	✓	97.4	96.3	94.7	Multiple residual and shifter units comprising of both depthwise dilated convolutions along with pointwise convolutions; training on multiple resized

(continued on next page)

Table 3 (continued)

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
	Classes:2C/VP 305/305						87.3	88.1	85.5	input images followed by predictions combining using meta learner
	Classes:2C/BP 305/305						94.7	93.5	93.3	
	Classes:3C/VP/BP 305/305/305						89.6	88.5	87.6	
	Classes:4C/N/VP/BP 305/305/305/305						90.2	90.8	89.1	
Madaan et al. [77]	Classes:2C/N 196/196	Augmentation; resized to 224 × 224	Custom CNN	×	✓	×	98.4	98.5	–	5 convolutional layers along with a rectified linear unit as an activation function
Narayanan et al. [23]	Classes:2C/CN 2504/6 807	Thresholding; grayscale, resized to 256 × 256; local contrast enhancement	U-Net; ResNet50; CAM	×	✓	✓	99.3	91.0	99.0	handled the class-imbalance problem by novel transfer-to-transfer learning; replaced last FC layer with two more fully connected layers
Nayak et al. [83]	Classes:2C/N 203/203	Augmentation, normalization	ResNet34	×	✓	×	98.3	–	–	Fine tuning of all the layers
Oh et al. [27]	Classes:4 N/BP/TB/VP and C were considered as one class)	Data type casting to float 32; histogram equalization; gamma correction; resized to 256 × 256	FC-DenseNet103 for segmentation; patch-based CNN based on ResNet18; use of Grad-CAM	×	✓	×	88.9	83.4	96.4	Morphological analysis of lung area; evaluation of segmentation performance; peculiar pre-processing steps to remove heterogeneity across then dataset
Ozturk et al. [66]	Classes:2C/N 127/500	Resized to 256 × 256	Modified Darknet-19	✓	✓	✓	98.1	95.1	95.3	Multiple Darknet layers having one convolutional layer followed
Panwar et al. [64]	Classes:3C/N/P 127/500/500						87.0	85.4	92.2	by batch normalization and leaky ReLU operations
	Classes:2C/N 142/142	Augmentation; resized to 224 × 224	VGG16	×	✓	×	88.1	97.6	78.6	Utilized first 18 Imagenet pre-trained VGG16 layers and added 5 new different layers (average pooling, flatten, dense, dropout and dense) on the top
Pereira et al. [79]	Classes:7 N/C/SARS/MERS/Pneumocystic/Streptococcus/Varicella 1000/90/11/10/11/12/10	None	Fusion of texture-based features and InceptionV3 features; classification using late fusion of multiple standard classifiers	✓	✓	×	95.3	–	–	handled the class-imbalance problem by re-sampling; multiclass and hierarchical classification
Pham et al. [61]	Classes:2C/N 403/721	Resized to 227 × 227	SqueezeNet	✓	✓	×	99.8	100	99.8	Features visualization of different layers
	Classes:2C/N 438/438						99.7	99.5	99.8	
Rahimzadeh and Attar [80]	Classes:3C/N/P 180/8 851/6 054	Resized to 300 × 300, augmentation	XceptionNet concatenated with ResNet50V2	✓	✓	✓	91.4	87.3	93.9	handled the class-imbalance problem by training multiple times on resampled data
Sakib et al. [71]	Classes:3C/N/P 209/27 228/5794	Augmentation using GANs	Custom CNN	×	✓	×	93.9	–	–	Analysis of different optimization algorithms; 5 convolutional layers along with exponential linear unit as an activation function
Sitaula et al. [132]	Classes:5C/N/BP/VP/NF (exact segregation is not given)	Resized to 150 × 150	VGG16	✓	✓	×	79.6	89.0	92.0	Leveraged both attention and convolution modules in the 4th pooling layer of VGG-16 for identifying deteriorating lung regions in both local and global levels of CXR images
Tabik et al. [91]	Classes:2 N/C 426/426	Class-inherent transformation method using GANs	U-Net, ResNet50, Grad-CAM	×	✓	✓	76.2	72.6	79.8	Quantified COVID-19 in terms of severity levels so to build triage systems; Replaced last layer; fine-tuned all the layers; use of class-inherent transformation network to increase discrimination capacity; fusion of twin CNNs
Togacar et al. [51]	Classes:3C/N/P 295/65/98	Resized to 224 × 224; Data	Feature extraction using MobileNetV2 and	✓	✓	✓	98.2	97.0	99.2	Image quality improvement using fuzzy technique

(continued on next page)

Table 3 (continued)

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
Toraman et al. [68]	Classes:3C/N/P 1050/1050/1050 Classes:2C/N 1050/1 050	restructured and stacked with the Fuzzy Color technique Augmentation; resized to 128 × 128	SqueezeNet; processed using the social mimic optimization method; classified using SVM Custom CNN	×	✓	✓	84.2	84.2	91.8	4 convolutional layers and 1 primary capsule layer
Ucar et al. [69]	Classes:3C/N/P 66/1 349/3 895	Augmentation; normalization; resized to 227 × 227	Bayes-SqueezeNet; activation mapping	×	✓	×	98.3	–	99.1	Handled the class-imbalance problem by multi scale offline augmentation; evaluation of proposed method using multiple metrics such as correctness, completeness and Matthew correlation coefficient; computational time analysis
Wang et al. [126]	Classes:3C/N/P	Augmentation; image cropping; resized to 480 × 480	Custom CNN; interpretation by GSInquire [146]	✓	✓	×	93.3	91.0	98.9	Multiple projection-expansion-projection-extension blocks; different filter kernel sizes ranging from 7 × 7 to 1 × 1
Wang et al. [28]	Classes:3C/N/CAP 225/1 334/2 024	Augmentation; resized to 224 × 224	VGG based Segmentation; ResNet with feature pyramid network	×	✓	×	93.7	90.9	92.6	Handled the class-imbalance with multi-focal loss function; residual attention network for localizing infected pulmonary region

removing noise from images using the top-2 smooth loss function. In Ref. [35], authors considered an ensemble classifier using two 3D ResNet-34 architectures for CT-scan images. The prediction scores obtained from the two ResNets were linearly combined, where the weights were decided according to the ratios of the pneumonia infection regions and the lung. CT-scan images of CAP and COVID-19 patients were collected from 8 hospitals, and the images were segmented to obtain the lung regions using the VB-Net [134] with a refined attention module that provided interpretability and explainability to the model. VB-Net was designed by adding bottleneck layers to a V-Net to integrate feature map channels. The role of the attention module was twofold. First, it learned all important features for classification, and second, it gave the 3D class activation mapping. The images were normalized voxel-wise, and the window/level scaling was performed to enhance the contrast of images. The ResNet architecture was trained using dual sampling to compensate for the unbalanced dataset.

Li et al. [31] utilized a 3D ResNet50 model to differentiate COVID-19 from CAP. Before fine-tuning this model, the lung was segmented from 3D CT images using a U-Net-based segmentation method. Also, the framework could extract both two-dimensional local and 3D global representative features. Wu et al. [135] used a joint classification and segmentation approach termed JCS using 1,44,167 chest CT scans, which is one of the largest CT-scan datasets used in the literature. The dataset includes scans from 400 COVID-19 patients and 350 non-COVID subjects. Of these, 3855 chest CT images of 200 patients have been annotated with fine-grained pixel-level labels of opacifications, lesion counts, opacification areas, and locations, thus benefiting various diagnosis aspects. A Res2Net was used in this work for classification, and image mixing was used to avoid over-fitting. Segmentation was performed using an encoder-decoder module, and an Enhanced Feature Module (EFM) was used with VGG-16 in the encoder. Feature maps acquired from different stages were fused to predict the side-output of each stage. An attention mechanism was used to filter relevant features. The output from the last stage, which gave the final prediction value, had the same resolution as the input CT image.

#### 4.1.4. Inception or GoogleNet

The idea of inception network [41] is to use several filter sizes

instead of choosing a particular filter size. The feature maps are concatenated at the output so that the network learns about the combination of required filter sizes. It cascades several inception modules, where each module consists of a concatenation of outputs from 1x1 convolution, 3x3 convolution, 5x5 convolution, and pooling operation. It has additional side branches also. Inception network has three more versions with improved performance. InceptionV2 and InceptionV3 have been proposed in the same paper [136] and InceptionV4 is explained in Ref. [43]. InceptionV2 replaces the 5x5 convolution operation with two 3x3 convolution operations to avoid information loss and uses factorization methods to achieve performance improvement. InceptionV3 contains all the features of InceptionV2 in addition to RMSprop optimizer, batch normalization, regularization, and 7x7 factorized convolution. In Ref. [79], multiple texture-based features were extracted from the CXR images, such as local binary pattern, elongated quinary pattern, local directional number, locally encoded transform feature histogram, binarized statistical image features, local phase quantization, and oriented basic image features. These features were combined with the features learned by the InceptionV3 network. These features were resampled to handle the problem of the unbalanced dataset. Five popular machine learning classifiers, including K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Multilayer Perceptrons (MLP), and Decision Trees (DT), were used to predict class labels. The predicted values were combined by considering the sum of the prediction probabilities obtained for each label by each learner, the product of the prediction probabilities obtained for each label by each learner, and also as the majority vote. Wang et al. (2021) [137] modified the Inception-V3 pre-trained model in the end and used it for the classification of CT images.

#### 4.1.5. DenseNet

DenseNet can be understood as the extension of ResNet50 architecture, where each layer receives additional input from all the preceding layers rather than a skip connection from a single previous layer. It transfers its output to all the subsequent convolutional layers for concatenation. Thus, each layer is said to obtain “collective knowledge” from all the preceding convolutional layers. DenseNet is being utilized in a few studies in the literature and has shown good performance

**Table 4**

**Summary of state-of-art DL techniques used for the COVID-19 classification using CT** Abbreviations: Acc.- Accuracy, BP-Bacterial Pneumonia, C-COVID-19, CAM-Class Activation Maps, CAP- Community Acquired Pneumonia, CN- COVID-19 negative, FP- Fungal Pneumonia, FPN- Feature Pyramid Network, HU- Hounsfield Units, Inlu.- Influenza, LT- Lung Tumor, MP- Mycoplasma Pneumonia, N-Normal, NF- No Findings, P- Pneumonia, Rad.- Radiologist, SARS- Severe Acute Respiratory Syndrome, Seg.- Segmentation, VP- Viral Pneumonia, Sen.- Sensitivity, Spe.- Specificity.

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
Ardakani et al. [133]	Classes:2C/CN 510/510	Gray-scale conversion; affected region resized to 60 × 60,	ResNet101	×	✓	×	99.6	100	99.3	Kolmogorov-Smirnov test to check the normality of all quantitative data; evaluation of age and gender distributions among COVID-19 and non-COVID-19 groups by two-tailed independent sample t-test and chi-square test, respectively. Copying of same image to the three RGB channels; padding to alter size without resizing; t-distributed stochastic neighbor embedding visualization of feature vectors
Alshazly et al. [117]	Classes:2C/CN 1252/1 230	Augmentation	ResNet101; Grad-CAM	×	✓	✓	99.4	99.1	99.6	
Arora et al. [118]	Classes:2C/CN 349/463	Augmentation; Residual dense network to improve the resolution	DenseNet201	×	✓	✓	92.9	93.7	92.2	Image resolution improvement by residual dense block
	Classes:2C/CN 349/463		MobileNet	×	✓	×	94.1	96.1	–	
El-Kenawy et al. [131]	Classes:2C/CN 1252/1 230	None	AlexNet	×	✓	×	100	100	–	Feature selection using Guided Whale Optimization Algorithm (WOA) and voting on multiple classifiers output resampled along three axes (z, y, x) to account for the variety of voxel dimensions; 10 convolution layers with five max-pooling layers along with two fully connected layers; filtering of CT slices (2D) to remove non-lung tissue (e.g., skin, bone, or scanner bed) and denoising of CT slices
	Classes:2C/CN 334/794			×	×	×	79.0	81.0	77.3	
Javaheri et al. [15]	Classes:3C/N/CAP 111/109/115	HU-based filtering; min-max normalization; resized to 128 × 128	BCDU-Net & 3D-CNN	✓	×	×	86.6	90.9	100	Task-specific fusion block to get a 3D CT prediction from slice-level prediction; t-SNE visualization; Attention region identification by binarizing the output of Guided Grad-CAM and their phenotype feature analysis among different classes
Jin et al. [30]	Classes:4C/N/CAP/I 3084/3 562/2296/83	Resampling to 1 × 1 × 1 mm voxel; HU-based filtering; normalization; resized to 224 × 224	U-Net; ResNet152; & Guided Grad-CAM	✓	✓	✓	–	94.1	95.5	Statistical analysis of training and test cohorts
Li et al. [31]	Classes:3C/N/CAP 1292/1 325/1735	None	U-Net, 3D-ResNet50 & heatmaps	✓	×	×	–	90.3	94.7	Modifications in the last layer by addition of fully connected; batch normalization and dropout layers
Mishra et al. [110]	Classes:2C/N 400/400	Resized to 224x224	ResNet50	×	✓	✓	99.6	99.6	99.6	Handled the class-imbalance problem by dual sampling training; refined the attention of training network using attention module
	Classes:3C/N/P 400/400/250	Resized to 138 × 256 × 256; dual sampling; contrast enhancement; normalization	VB-Net for segmentation, two 3D ResNet34 with online attention module	×	×	×	87.5	86.9	90.1	
Ouyang et al. [35]	Classes:2C/CAP 3389/1 593	Resized to 138 × 256 × 256; dual sampling; contrast enhancement; normalization	VB-Net for segmentation, two 3D ResNet34 with online attention module	×	×	×	87.5	86.9	90.1	Handled the class-imbalance and noisy data problem using top-2 smooth loss
Pathak et al. [107]	Classes:2C/N 413/439	None	ResNet50	×	✓	✓	93.1	–	–	SqueezeNet inspired architecture
Polsinelli et al. [121]	Classes:2C/CN 449/386	Augmentation	Custom CNN	×	✓	✓	85.1	87.6	81.9	Majority voting of multiple parallel trained CNNs
Serte et al. [103]	Classes:2C/N 90/49	Resized to 256 × 256	majority voting on multiple ResNet50 trained on single slice	×	✓	×	96.0	100	96.0	Fine tuning of all the layers; changed dimensions of last 2 fully connected layers
Shah et al. [123]	Classes:2C/CN 349/463	Resized to 128 × 128	VGG19	✓	✓	×	94.5	–	–	
				✓	×	×	93.0	–	93.0	

(continued on next page)

Table 4 (continued)

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
Song et al. [143]	Classes:3C/N/BP 88/86/100	Lung region segmentation through OpenCV	ResNet50 & feature pyramid network							Use of attention module to learn the importance part of an image
Turkoglu [122]	Classes:2C/N 349/397	Augmentation; image scaling; resized to 224 × 224	DenseNet201 with multiple kernels extreme learning machine classifiers	×	✓	✓	98.4	98.2	98.4	Analysis on multiple different activation functions
Wang et al. [32]	Classes:2C/N 723/413	Augmentation; thresholding; normalization; resized to 256 × 256	3D U-Net++ & ResNet50	✓	×	×	–	97.4	92.2	Both slice and intensity level normalization; lung and lesion segmentation; evaluation of segmentation using dice coefficient
Wang et al. [139]	Classes:5C/BP/VP/MP/FP 924/271/29/31/11	Normalization; resized to 48 × 240 × 360	DenseNet121-FPN for segmentation; DenseNet based architecture for classification	✓	×	×	81.2	78.9	89.9	Trained on CT-EGFR dataset to predict EGFR mutation status using the lung-ROI; Built multivariate Cox proportional hazard (CPH) model to predict the hazard of patient needing a long hospital-stay time to recover; Visualized suspicious lung area and feature patterns; Evaluated by calibration curves and Hosmer-Lemeshow test; Prognostic analysis using Meier analysis and log-rank test
Wang et al. [137]	Classes:2C/CN 325/740	Resized to 229 × 229	Lung area segmentation; InceptionV3	×	×	×	89.5	87.0	88.0	Copied of gray scale image three times to form RGB image; Fixed feature extractor with modification in only last FC layer
Wang et al. [142]	Classes:2C/N 320/320	Augmentation; gray scale conversion; histogram stretching; image cropping; resized to 256 × 256	Custom CNN; Grad-CAM	×	×	×	97.1	97.7	96.5	Feature fusion of CNN (with 7 convolutional layers and 2 fully connected layers) and graph convolution network. CNN is used to extract image-level features and graph convolutional network (GCN) to extract relation-aware features among images; Used rank-based average pooling
Wu et al. [135]	Classes:2C/N 67 505/75 541	None	Explainable joint classification and segmentation network (Res2Net for classification; VGG16 for segmentation); activation mapping	✓	✓	×	–	95.0	93.0	Released large scale COVID-CS dataset with both patient and pixel-level annotations (helped to focus more on the decisive lesion areas of COVID-19 cases); computational time analysis; evaluation of segmentation using dice coefficient; alleviated overfitting by image mixing; detailed ablation analysis
Xu et al. [33]	Classes:3C/N/I 189/145/194	Augmentation, HU-based filtering	3D-Segmentation, Attention ResNet18 & Noisy-OR Bayesian function based voting	×	×	×	71.8	76.5	68.9	Proposed a local attention classification model using ResNet18 as backbone architecture; used image patch vote and noisy-OR Bayesian function based vote for voting a region and enhancement
Zheng et al. [34]	Classes:2C/N 313/229	Augmentation; HU-based filtering; resized to 224 × 336	U-Net & 3D-CNN	✓	×	×	90.1	84.0	98.2	Residual blocks and 3D CNN layers; CT volume and its 3D lung mask as an input
Zhou et al. [141]	Classes:3C/N/LT 2500/2 500/2500	Normalization, resized to 64 × 64	Ensemble modelling (majority voting) with AlexNet, GoogleNet, ResNet18	×	×	✓	99.1	99.1	99.6	Training time analysis and evaluation by Matthews correlation coefficient

compared to the other pre-trained networks. For example, Alshazly et al. (2021) [117] tested the efficacy on two datasets [116,120], and obtained good results with the DenseNet201 pre-trained model on one of the datasets compared to the other pre-trained models.

Chowdhury et al. [5] showed the superiority of the DenseNet201 pre-trained model over the other six pre-trained models for a three-class classification problems using CXR images. Comparison of the activation maps of different classes obtained from the convolutional layers

provided insight into the image regions contributing to classification. In Ref. [53], DenseNet-121 architecture is used to classify COVID-19 patients from control. First, a set of 1024 higher-level features are extracted using the DenseNet-121 trained on ImageNet, and then a logistic regression is fitted to these features. Finally, interpretation is done using the *Expected Gradients* [138]. To further identify the features used by the ML model to differentiate between COVID-19-positive and COVID-19-negative datasets, a GAN is trained to transform

**Table 5**

**Summary of state-of-art DL techniques used for the COVID-19 classification using Multimodality** Abbreviations: Acc.- Accuracy, BP-Bacterial Pneumonia, C-COVID-19, CAM- Class Activation Maps, CAP- Community Acquired Pneumonia, CN- COVID-19 negative, FP- Fungal Pneumonia, FPN- Feature Pyramid Network, HU- Hounsfield Units, Infl.- Influenza, LC- Lung Cancer, LT- Lung Tumor, MP- Mycoplasma Pneumonia, N-Normal, NF- No Findings, P- Pneumonia, Rad.- Radiologist, SARS- Severe Acute Respiratory Syndrome, Seg.- Segmentation, VP- Viral Pneumonia, Sen.- Sensitivity, Spe.- Specificity.

Ref.	Dataset	Pre-processing	Architecture	Code	Data	K-Fold	Performance reported			Critical Observations
							Acc.	Sen.	Spe.	
Hilmizen et al. [9]	Classes:2 CT: C/N 1257/1 243 CXR: C/N 1257/1 243	Resized to 150 × 150	Ensembling of ResNet50 and VGG16	×	✓	×	99.8	99.7	100	Concatenation of CT and X-ray features extracted using two separate models; only binary classification; good reference for multimodal
Ibrahim et al. [109]	Classes:4 N/C/P/LC 3500/4 320/5856/20 000	Augmentation; resized to 224 × 224; normalization	VGG19	×	✓	×	98.1	98.4	99.5	Used mixed dataset of CT and X-ray images; implemented four different architectures; randomness is observed in learning curves
Irfan et al. [55]	Classes:3 CT: C/N/P 1000/600/700 CXR: C/N/P 1200/500/1 000	Noise removal	Custom CNN + LSTM	×	✓	✓	-	95.5	-	Used a mixed dataset of CT and CXR; performance learning curves are not shown
Kamil MY [82]	Classes:2 CT: C 23 CXR: C/N 172/805	None	VGG19	×	✓	×	99.0	97.4	99.4	Unbalanced dataset in terms of CT vs CXR; Combined training of both type of images; randomness in performance learning curves
Mukherjee et al. [10]	Classes:2 CT: C/N 168/168 CXR: C/N 168/168	Resized to 100 × 100	Custom CNN	×	✓	✓	96.3	97.9	94.6	Balanced dataset; three convolutional layers followed by three fully connected layers; validation loss curve is not shown
Thakur et al. [94]	Classes:3 CT: C/N/P 2035/2 119/2 200 CXR: C/N/P 1200/1 341/2 200 Classes:2 CT: C/N 2035/2 119 CXR: C/N 1200/1 341	None	Custom CNN	×	✓	×	98.3	98.2	-	Used a mixed dataset; proposed deep learning architecture is missing; performance learning curves are missing
		None	Custom CNN	×	✓	×	99.6	95.6	-	

COVID-19-negative radiographs to resemble COVID-19-positive radiographs and vice versa. This technique should capture a broader range of features than saliency maps because the GANs are optimized to identify all possible features that differentiate the datasets. In Ref. [73], authors used DenseNet121 to classify CXR images. A gravitational search algorithm (GSA) was used to find the optimum hyperparameters for the network. This was compared with the performance of DenseNet121 and Inception-v3 with manual hyperparameter tuning, and the GSA performance was shown to be superior. The authors used gradient-weighted class activation mapping (Grad-CAM) for explainability.

Wang et al. (2021) [139] consists of a sequence of three separate parts for automatic lung segmentation, non-lung area suppression, and COVID-19 diagnostic and prognostic analysis. It proposes DenseNet121-FPN for lung segmentation in chest CT image, and DenseNet-like structure for COVID-19 diagnostic and prognostic analysis Turkoglu in Ref. [122] used multiple kernels extreme learning machine (ELM) based DenseNet201 to detect COVID-19 cases from CT scans. The transfer learning approach was used because the available COVID-19 datasets were insufficient to train the CNN models effectively. ELM based on majority voting was used to generate the final prediction of the CT scan, and results of ReLU-ELM, PReLU-ELM, and Tanh ReLU-ELM were compared.

**4.1.6. XceptionNet**

XceptionNet, developed by Google, stands for extreme version of Inception. It achieves better performance than InceptionV3 by introducing depthwise convolution and pointwise convolution. Khan et al. [87] used XceptionNet to propose a COVID-19 detection algorithm. The performance learning curve for four classes is shown only for fold-4 that gives the best accuracy among all folds. However, the curves show randomness in the learning and overfitting. The optimization of the model towards good fit may reduce the achieved accuracy and other metrics levels.

**4.1.7. MobileNet**

It uses depthwise separable convolution from the Xception network with the aim to reduce model complexity and parameters that would compress the network and improve the speed. It has been developed considering mobile and embedded-based DL applications. Arora et al. (2020) [118] used MobileNet architecture along with residual dense neural network to detect COVID-19 from the CT-scan images. Results were compared with multiple other pre-trained architectures and MobileNet has shown better performance.

**4.1.8. SqueezeNet**

It has been developed with the aim of smaller networks having fewer parameters that can easily fit into the applications with low memory and bandwidth requirements. It achieves the goal by decreasing the number of input channels and replacing 3 × 3 filters with 1 × 1 filters. SqueezeNet is one of the light networks that has been used for COVID-19 classification in CXR-based studies [69]. In this study, the author proposes a SqueezeNet-based architecture using Bayesian optimization for embedded and mobile systems. It is shown that the model size of the proposed network is 77.31 times smaller than the AlexNet. Pham [61] used AlexNet, GoogleNet, and SqueezeNet to classify CXR images into 2-classes and 3-classes to detect COVID-19 cases. Six datasets were constructed using publicly available images to consider balanced and unbalanced binary and multiclass scenarios with normal, COVID-19, and pneumonia cases. The algorithm’s efficacy in distinguishing between COVID and non-COVID cases, COVID, and normal cases are illustrated. Also, different cases of train and test data split are considered. Polsinelli et al. (2020) [121] proposed a light CNN architecture based on SqueezeNet to classify CT images into COVID and non-COVID. The authors used a publicly available dataset to train the network and utilized a separate dataset for testing. The proposed CNN architecture outperformed the conventional SqueezeNet.



#### 4.1.9. EfficientNet

It is a neural network architecture that uniformly scales the three dimensions, viz., depth (number of hidden layers), width (number of channels), and resolution (input image resolution) using compound coefficient for improved efficiency and accuracy. In one recent work, Luz et al. (2021) [112] utilized EfficientNet for three-class classification problem and also evaluated results on cross-site datasets. Though the original dataset is highly imbalanced, consisting of 7966 Normal images and 152 COVID-19 images, it has adopted data augmentation to undertake the analysis on a balanced dataset. It provides a valuable analysis by proposing a model with relatively 5 to 30 times lesser parameters and hence, reduced memory requirements.

#### 4.1.10. Hybrid model

The hybrid model consists of an ensemble of the aforementioned models. Several works, such as [57,72,74] can be found in this direction. [72] has used several pre-trained as feature extractors and correlation-based feature selection using CXR images. The proposed model has been validated on two separate public dataset and has shown promising results. [57] has separately trained three transfer learning architectures (DenseNet201, Resnet50V2, and Inceptionv3) and combined them using weighted average ensembling to predict COVID-19. [74] has stacked many pre-trained models (ResNet101, Xception, InceptionV3, MobileNet, and NASNet), and extracted features are concatenated together before feeding them to the dense layer for the classification task. However, these methods add complexity experienced in pre-training multiple models. To classify CXR images into three classes; normal, COVID-19 and pneumonia, authors in Ref. [80] concatenated features obtained using XceptionNet and ResNet50V2. The concatenation was done to obtain features with both inception-based layers and residual-based layers. After concatenation, the feature set was passed through a convolutional layer and, further, given to a classifier. The network was trained with eight different subsets of a balanced dataset. The performance of the concatenated network was compared with XceptionNet and ResNet50V2 individually, and only marginal improvement was observed with the proposed method. Togacar et al. (2020) [51] stacked original images with images pre-processed using the fuzzy color technique. Features were extracted from these stacked images using MobileNetV2 and SqueezeNet. Social Mimic optimization method [140] was used to process the features, and a support vector machine classifier was used for classifying the images into three classes.

In [141], an ensemble DL model is proposed using three pre-trained models: AlexNet, GoogleNet, and ResNet using CT-scan images. Ensembling is performed using majority voting. The performance of the proposed method was observed to be superior compared to the three individual pre-trained networks. Hilmizen et al. [9] fed the CXR images to VGG16 and CT scan images to ResNet50 for feature extraction, which were concatenated before providing to the dense layers for classification.

**4.1.10.1. Custom CNN work.** ImageNet pre-trained models are not sufficient for classifying medical images because they are trained using natural images. Since medical and natural images are different in many aspects, some studies trained new and special deep CNNs from scratch. These studies majorly either adopted simpler few stacked convolutional layers in Refs. [10,71,77,94] or adopted advanced layers such as residual blocks in Refs. [98,99]. Furthermore improvements are made by utilizing advanced architectures such as residual blocks with squeeze-excitation blocks in Ref. [65], channel-shuffled dual-branched CNN in Ref. [26], newly lightweight residual projection-expansion-projection-extension blocks in Ref. [126], Long Short Term Memory (LSTM) with CNN in Refs. [55,60], 3D CNN in Ref. [15], 3D CNN with residual blocks in Ref. [34], capsule network inspired architecture in Refs. [67,68], and Graphs Convolution Network (GCN) with CNN in Ref. [142]. More in-depth details of all these

references are presented in Tables below. *Summary:* In this paper, we have reviewed a total of 71 COVID-19 detection studies, based on the imaging modality used, i.e., 23 CT image studies, 42 CXR image studies, and six studies using both CT and CXR images. We observed that transfer learning had been efficiently used to detect COVID-19 from chest CT and CXR images. Of all studies, 57 (80% of the reviewed systems) used transfer learning with pre-trained weights, and only 14 used custom CNN. Fig. 7b shows the number of published papers using various DL architectures. ResNet is the most popular architecture used by 28% of the reviewed articles, followed by custom CNN and VGGNet.

Since the transfer learning approach offers several advantages, it is a preferred choice in many studies. In general, training a model from scratch requires high computational power and larger datasets. The primary issue in training deeper models from scratch is learning a large number of parameters using a limited number of (available) training samples that lead to overfitting. Also, it is quite time-consuming to decide parameters and architectures from scratch. A pre-trained model with transfer learning facilitates faster convergence with network generalization. Thus, we observe that many studies on DL-based COVID-19 detection models using CXR, CT, and multimodal have used the transfer learning approach.

## 5. Unique challenges in COVID-19 image analysis

In the last section, we discussed several works on COVID-19 image analysis. Although the performance of the proposed algorithms seems promising, there are certain shortcomings that must be addressed. We now present a discussion on some of the challenges and gaps in this area.

### 5.1. Reproducibility and code availability

Reproducibility of DL-based models has emerged as one of the major challenges in the literature. Results can be ascertained if only the dataset and the details of the model architecture and training hyperparameters are made available. Also, the open-source availability of code helps in reproducing the results and in devising further improvements. Some of the works based on CXR and CT-scan image classification have provided their codes [15,30–32,34,51,57,66,67,70,80,87,99,112,132,135,139,143]. However, none of the papers using multimodal architecture have provided codes in the open-source domain. Almost all the papers that derived the dataset from multiple sources have provided details of individual sources. However, most of them have not provided the link of their consolidated dataset except a few studies such as [30,32,67,79,91]. It is important to note that many authors who have provided their codes have also not provided their dataset in the public domain. A study by the authors in Ref. [94] have not provided details of their architecture, although it is based on a custom CNN.

### 5.2. Unbalanced dataset

It is noted that most of the dataset used for binary class or multiclass classification for COVID-19 diagnosis is highly unbalanced. The skewness in the dataset can introduce bias in the performance of a trained model. These unbalanced datasets pose a major challenge to the AI researchers because the collection of a sufficient number of quality images, especially at the initial stage of the pandemic, was indeed difficult.

For example, as listed in Table. 5, the author in Ref. [82] has used only 23 CT and 172 CXR images as compared to 805 normal images. The author in Ref. [109] has used 20,000 lung cancer images as compared to 3500 normal images. In order to handle class imbalance with small COVID-19 data, larger penalties were associated with the mis-classification error in COVID-19 cases in Ref. [67]. In Refs. [74,87], random sampling was used to select a balanced multiclass dataset from an unbalanced larger dataset. However, this method reduced the size of the dataset. Pereira et al. [79] investigated the effect of different re-sampling methods such as ADASYN, SMOTE, SMOTE-B1, SMOTE-B2,

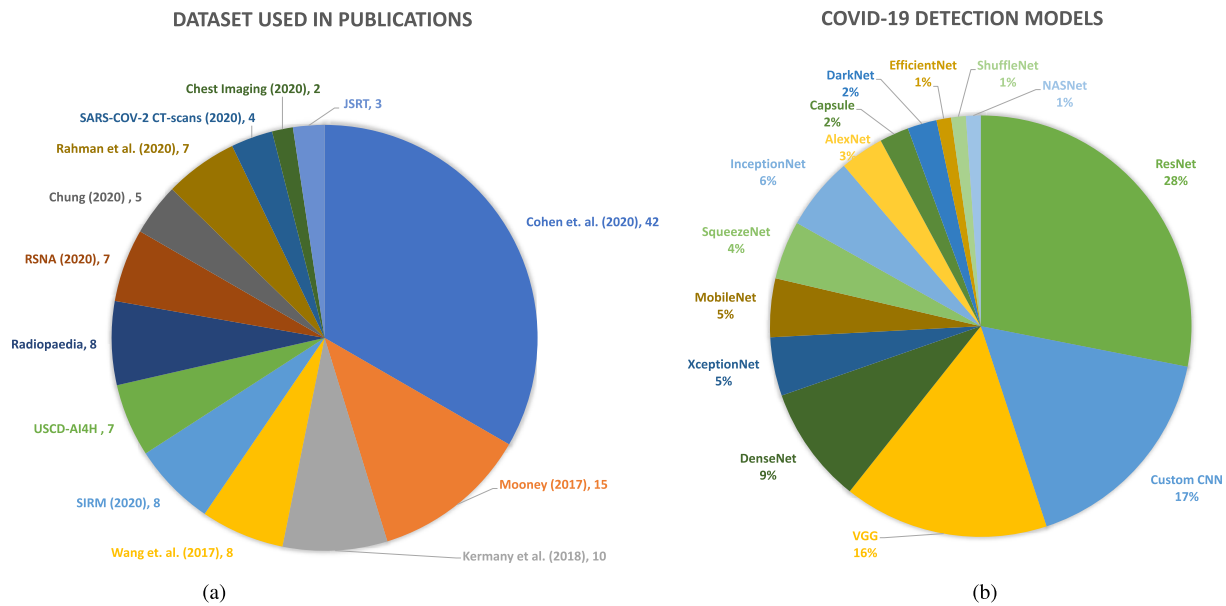


Fig. 7. (7a) shows the number of publications using most popular datasets for validating COVID-19 detection models, and (7b) shows the number of published papers using various deep learning architectures.

AllKNN, ENN, RENN, Tomek Links (TL), and SMOTE + TL on the performance of the proposed classification algorithm. In Ref. [26], data augmentation, weighted-class batch sampling, as well as stratified data sampling were used to obtain an equal number of samples for each class in every training batch. In Ref. [131], dataset balancing was carried out using the SMOTE algorithm. Authors in Ref. [23] addressed the class-imbalance problem owing to the limited set of COVID-19 CXR images by proposing a novel transfer-to-transfer learning approach, where a highly imbalanced training dataset was broken into a group of balanced minisets, followed by transferring the learned (ResNet50) weights from one set to another for fine-tuning.

### 5.3. Data augmentation

Data Augmentation is employed to increase the size of the training dataset by transforming the images of the original dataset in multiple ways. This enables the learning of diverse features during the training process and reduces the overfitting of the model. Two important techniques of data augmentation have been observed in the reviewed literature. First, several variations such as flip, rotate, skew, translation, random clipping, tilting, and scaling have been introduced to the original dataset, increasing the number of training samples. Second, inbuilt software libraries (e.g., Keras ImageDataGenerator function) have been utilized that introduce random variations in the training dataset during each iteration without increasing the number of samples. The range of random variations is a hyperparameter that needs to be fine-tuned for a given problem. Authors in Refs. [28,32,63–65,68,70,75–77,83,98,109,112,118,122,142] have used the first method, while the authors in Refs. [9,26,26,33,34,74,82,110,121] have used the second technique of data augmentation.

There is a third method of data augmentation by generating synthetic images using a Generative adversarial network (GAN). For example, authors in Ref. [71] have used GAN and generic data augmentation techniques to increase the dataset size. Authors in Refs. [69,117] have added Gaussian noise and used brightness variations to augment the images. One study also used Gamma correction to augment images [142]. Authors in Refs. [30,103,107,141,143] have not incorporated data augmentation that could have definitely improved performance of the proposed model.

### 5.4. Quality of images

Medical images are generally low contrast images. Hence, efforts are made to increase the contrast of these images so that the images are better transformed to the feature space while they traverse through a DL model. Moreover, broad heterogeneity in the quality of images captured at different sites using different imaging devices causes potential bias in image analysis. This challenge emphasizes the need for improving image quality as a pre-processing step. Contrast enhancement techniques are generally used in the literature for enhancing the quality of images and making them visually more appealing. A few studies carried out histogram modification of images for contrast enhancement [27,70,75,142]. Authors in Ref. [23] utilized local contrast enhancement on thresholded grayscale CXR images for enhancement and also for removing any text from the image. Authors in Ref. [75] removed the diaphragm region from the CXR images and applied bilateral lowpass filtering to the original images. Others normalized their images before feeding to a neural network [15, 32, 57, 65, 83, 112]. Authors in [15, 33] applied HU-based filtering on raw CT images for removing the redundant parts. Some used gamma correction to control the brightness of images used [27]. Authors in Ref. [118] used residual dense network (RDNN) to enhance the quality of CT-Scan images through super-resolution. The performance of the model deteriorated for low-quality images in Ref. [66]. A large number of non-infected regions or background have been separated in [31, 33–35] using 3D CNN segmentation model based on U-Net [22].

### 5.5. Transfer learning architecture

Transfer learning has been used either as a fixed feature extractor (where the weights of the convolutional layers of the pre-trained architectures are used without alteration) or weights of the few or all convolutional layers of the model are fine-tuned or retrained. The choice of an approach depends upon the size and similarity of the training dataset of the given problem to the dataset used in the training of the original transfer learning model. Since weights of most of the standard DL models (used for transfer learning) are learned over 1000 classes of the ImageNet dataset consisting of natural images, these DL models may not be completely relevant for the classification of CT or CXR images. Hence, it is recommended to employ transfer learning by retraining the

weights of a few convolutional layers. Several studies in papers [23, 30–32, 35, 63, 69, 72, 75, 79, 83–85, 87, 103, 107, 112, 117, 122, 123, 131, 137, 139, 141, 143] have used transfer learning models as fixed feature extractor only. Also, it is important to note that very few studies such as class decomposition with VGGNet [70], attention with VGGNet [132], feature pyramid network with ResNet [28] have proposed architectural changes in the proposed model that is very much required not only to achieve better classification capability but also to have faster and stable learning.

In [67], authors used a publicly available CXR dataset for common thorax diseases to pre-train their Capsule network, unlike other works where natural images from the ImageNet dataset have been used. They also demonstrated the superiority of their method as compared to the latter. In Ref. [87], a pre-trained XceptionNet was retrained end-to-end, while authors in Ref. [99] fine-tuned a custom CNN stacked architecture pre-trained on CXR images of normal and pneumonia cases. In Ref. [66], a 19 layer CNN has been developed using DarkNet as the classifier for YOLO real-time objective detection system.

### 5.6. Performance learning curves

Training and validation curves of accuracy and loss function provide a visual assessment of the three aspects of the training/trained model. First, it indicates how rapidly the model learns the objective function in terms of the number of iterations. Second, it informs how well the problem has been learned in terms of underfit/overfit/good fit of the model. Underfitting is shown by the low training accuracy, while overfitting of the model is indicated by a substantial gap between the training and validation curves. The good fit of the model is represented by higher training accuracy and convergence between training and validation curves. Third, there could be random fluctuations or noisy behavior in the training/validation loss curves. This could be due to a number of reasons, including the small size of the dataset compared to the model capacity, need of regularization, feature normalization, etc. Hence, depiction of learning curves is important in research studies as has been done in Refs. [10, 15, 26, 27, 51, 55, 60, 63–66, 68, 69, 75, 76, 78, 82–84, 87, 98, 110, 112, 118, 132, 137].

### 5.7. Stratified K-fold cross-validation

When the dataset is small, as is the case in the medical imaging domain, cross-validation is an important technique to assess the robustness of a model. Here, every sample of the dataset is used once as the test sample. The complete dataset is divided into  $k$  number of folds. In the literature study, very few studies have been undertaken to incorporate  $K$ -fold cross-validation. Authors in Refs. [35, 55, 57, 60, 65, 66, 117, 141] have used 5-fold and authors in Refs. [10, 23, 121, 122] have used 10-fold cross-validation. It is important to note that although the authors in Refs. [10, 35, 122] implemented  $K$ -fold cross-validation, details about the outcome of each fold has not been discussed. For a small dataset, this is a highly recommended training strategy.

### 5.8. Distinction from other viral diseases

During the COVID-19 pandemic, it has been observed that people were being infected symptomatically as well as asymptotically, where the latter is less contagious. A CXR or CT scan is taken at a later stage to determine the degree of infection so that proper medication can be advised to a patient. In such scenarios, it becomes imperative to differentiate not only between COVID-19 versus healthy but also between COVID-19 and the other viral diseases such as pneumonia that affect human organs in a similar manner. The development of an efficient and optimal AI-based solution to specifically and exclusively detect COVID-19 is still a prime challenge.

In one study [110], detection of COVID-19 from CT scans achieved accuracy of more than 99% while classifying from normal images.

However, the performance of the same model degrades considerably when the multiclass classification was undertaken, including pneumonia images. The same was observed in another study [66] with an accuracy drop of 11% after adding pneumonia samples. In Ref. [76], an accuracy of 97.2% was obtained for detecting COVID-19 cases from non-COVID cases, including healthy, bacterial pneumonia, and non-COVID viral pneumonia using a two-step detection algorithm using ResNet pre-trained architectures. The authors also investigated the performance of ResNet101 in detecting COVID-19 in the presence of other pulmonary diseases such as edema, cardiomegaly, atelectasis, consolidation, and effusion. In this case, the performance of ResNet101 was found to be inferior to the rest of the networks for the COVID-19 class. SARS patient samples along with healthy ones are also considered in one study for three-class classification using CXR images [70]. Authors in Ref. [27] developed an algorithm to detect healthy cases, bacterial pneumonia, viral pneumonia, and tuberculosis. Here, the COVID-19 cases were included in the viral pneumonia class that could be further detected using RT-PCR or CT-scan. Pediatric cases were excluded from this study to prevent the network from learning age-related data. Authors in Ref. [79] considered seven classes, including COVID-19, SARS, pneumocystis, streptococcus, varicella, MERS, and normal cases. In another work by Wang et al. [139], COVID cases are classified against several versions of pneumonia such as bacterial pneumonia, viral pneumonia, mycoplasma pneumonia, and viral pneumonia. Authors in Refs. [141, 143] have undertaken three-class classification by adding lung tumor and bacterial pneumonia.

In a few studies such as [23, 57, 67, 117], CXR images are classified into COVID and non-COVID, where the latter included normal, bacterial pneumonia cases, non-COVID viral pneumonia cases, and other pulmonary syndromes.

### 5.9. Generalization

Generalization is the ability of a DL model to perform well on an unseen dataset. A model to classify dog and cat trained using black cats only may not perform well when tested on white-colored cats. This requires the training of the model on a diverse dataset. Apart from the dataset, generalization ability can also be ascertained through the choice of hyperparameters of a network that cater to the high variance (overfitting) and high bias (underfitting) issues. Regularization, dropout, batch normalization, early stopping are some techniques that can be incorporated to achieve better generalization abilities.

To demonstrate the generalization ability of the proposed network, a few works like [57, 72, 75, 117] have demonstrated the performance of their proposed model on more than one dataset. Wang et al. (2021) [137] utilized Inception-V3 pre-trained model for transfer learning. Performance was evaluated on CT datasets from two different sites. Results on the same site test dataset achieved a total accuracy of 89.5% with a specificity of 0.88 and sensitivity of 0.87. Results on the test dataset from different sites (trained and tested on different site data) showed an accuracy of 79.3% with a specificity of 0.83 and sensitivity of 0.67.

### 5.10. Use of explainable AI

Convolutional neural network-based architecture possesses automatic feature extraction capability leading to the representation of DL models as a black box. To achieve wider acceptability of the automated solutions, it becomes imperative to have an interpretability and behavioral understanding of the model. The transparency and explainability of AI solutions are very critical in the medical domain, especially when used for life-threatening COVID-19 diseases. In the reviewed literature, some studies have utilized various interpretation methods with the most used one being Grad-CAM [26–28, 30, 31, 74, 76, 84, 91, 99, 117, 142] and CAM [23, 69, 121]. As illustrated, GRAD-CAM and CAM methods work in a similar manner, using heat maps as being used in a

few studies [35,60,66], while the other methods work differently and highlight the affected area in a different manner. For example, Karthik et al. (2021) [26] visualized the infected areas detected by the proposed work using saliency maps, guided backpropagation, and Grad-CAM. Authors in Ref. [117] also used *t*-distributed Stochastic Neighbor Embedding (t-SNE) plot to visualize the clustering of COVID and non-COVID cases. In Ref. [139], visualization of the suspicious lung area along with the visualization of feature patterns extracted by convolutional layers of the proposed architecture is done to understand the inference of a DL system.

### 5.11. Lack of comparison

Literature lacks the comparison among methods on the same data [98]. Instead of considering different datasets while evaluating and training any new model, methods should be trained on the same data for comparison. Again, this poses the need to create larger and more heterogeneous datasets that can be used to train both large and small neural networks.

It is pertinent to mention that a few authors, such as in Refs. [15,26,28,57,60,65–69,72,74,75,77,83,84,98,99,112] have compared the performance of various state-of-the-art algorithms using different datasets, which is not very informative as the performance metrics obtained in each case may be data-dependent. Some works such as [27,63,87,117,132,142] used the codes available for the existing publications or the same dataset to present a comparison. However, the benchmarking is still very limited. For example, in Refs. [27,63,87], authors compared the results obtained using their proposed algorithm with only one existing methodology on the same data.

### 5.12. Multimodal architecture

Multimodal studies undertaken using both CXR and CT have shown great potential in learning various features and improved performance. Further, it has been observed that most of the studies used a single sequential architecture that is trained on a mix of CXR and CT datasets. It is expected that the model would perform better by employing two parallel feature extractors, one each for CT and CXR, respectively. These separately extracted features can be combined before feeding to the classification (Dense) layer. In this regard, [9] uses two separate transfer learning models to extract features from CT and CXR images and achieves improved performance than any individual model alone.

## 6. Opportunities and scope for future work

Based on the literature review presented above, we provide some suggestions for future researchers. Some of these suggestions are apparent from the above discussion, while some entail the existing scenarios in the COVID era.

### 6.1. Availability of structured and authentic dataset

During the study of literature, it is observed that a one-to-one performance comparison between two reference papers cannot be undertaken due to lack of uniformity in the datasets and the performance metrics used. It is worth noting that the current public datasets have a limited number of images for the training and testing of AI algorithms. This necessitates the creation of one or more big, authentic, publicly available quality datasets that can be used, compared, or evaluated against by future researchers. For ease of research, we are presenting the most popularly used datasets in Fig. 7a. Table 2 includes details of 35 COVID-19 datasets, and we have selected the most-cited datasets for making this figure.

### 6.2. Generalization of a detection model

From the literature, it has been learned that the datasets used by researchers are highly unbalanced. It raises concerns about the generalizability of the trained models on the prospective patients' data. Some studies utilized a few methods for combating unbalancing problems, such as dual sampling in Ref. [35] and SMOTE in Ref. [131]. However, a vast majority of work has suffered from the challenge of unbalanced data. Secondly, any model developed for detecting COVID-19 should perform the same with the claimed accuracy on the unseen/prospective subjects' data or data of a different hospital. Thus, we believe that a cross-dataset study is of paramount importance to ascertain the generalizability of the model with respect to variation in images across sites. To the best of our knowledge, cross-data evaluation is conducted in only a few studies [53,112]. For the successful classification of a new test image, it is assumed that this new test image will consist of features similar to those learned during the training of the classification model. It necessitates the creation of a global training dataset that includes/captures major features. Furthermore, a proper benchmarking of different methods (or cross-method analysis) on the same dataset should be carried out to ascertain the efficacy of the proposed methods.

### 6.3. Multimodality scope

A viral infection affects different parts of a body with different severity that leads to multiple symptoms. The accuracy of detection or diagnosis of a disease depends on the effectiveness of identifying and measuring the symptoms or patterns. Different diagnostic tools are used to identify these symptoms, measured at varying degrees and levels. Accumulation of patterns from various modalities can provide diverse features compared to the individual variables that can be utilized to learn a DL model better. For the detection of COVID-19, besides CXR and CT scan images, cough and thermal images can be used to augment the detection capabilities of the model. Any model can have practical application if it has a high degree of generalization ability, and multimodal data analysis provides a better approach towards its achievement.

### 6.4. Explainable AI

An explanation of how a DL model has reached a certain conclusion is crucial for ensuring trust and transparency, especially when one deals with identifying life-threatening COVID-19 disease. In order to be sure of the decision, doctors would like to know how AI decides whether someone is suffering from a disease by analyzing CXR and CT scan images. In this paper, we survey some existing techniques used for explaining the interpretability of DL models trained for COVID-19 classification. There is a need to explore more methods of Explainable AI for COVID-19 diagnosis as used in other applications [147,148].

### 6.5. Semi-supervised and reinforcement learning

Annotation of medical images is one of the laborious works due to the shortage of radiologists and technicians who can label the images. Deep learning has a great power to extract features from images, but its performance depends heavily on the size of labeled training data. However, one can still train deep networks by utilizing semi-supervised and reinforcement learning methods that consider a mixture of unlabelled and limited labeled data for training deep models. This can address the problem of highly imbalanced data, one of the major challenges in COVID-19 image analysis, if arising of the difficulties in labeling/annotations.

### 6.6. Severity of disease

It is important to not only predict COVID-19 but also the degree of its severity in a patient for deciding appropriate treatment. Tabik et al. [91]

classified COVID-19 positive CXR images into severe, moderate, and mild classes based on the severity level of the disease. A more comprehensive understanding of the severity of the disease can aid doctors in curing this disease carefully. In all, future improvements would require the collection of hundreds or thousands of labeled image data of severe COVID-19 and other pneumonia data. The dataset should be collected considering geographic diversity in mind, which will help to increase its applicability worldwide. In addition, future work should also be considered in the direction of identifying the infected pulmonary region as belonging to the left lung, right lung, or bi-pulmonary region. One study has already been done in this direction by employing a residual attention network as a basic block [139].

### 6.7. Generic suggestions on COVID research

Besides the above suggestions based on the AI/ML work in COVID, a few more suggestions are in order, as discussed below.

#### 6.7.1. Study on regional variation

It has been noted that the COVID-19 virus is highly mutant, and several variants have evolved over the course of time. Hence, a scaling-up of diagnostic capabilities of AI-based automated solutions quickly and widely will be critical for diagnosing new variants of COVID-19, in decision making, and in choosing the way ahead. Regional variations in the impact of the virus on human organs may be studied. This can assist in a better understanding of the identification of a robust global/local solution.

#### 6.7.2. Regulation and transparency

Global solution needs global participation. As this pandemic has affected every corner of humanity, any strategy or measure to handle the crisis relies on a wider public acceptance. In order to have a better public trust, it is required that the decisions and information be transparent and available openly, especially when things are related to people's health. Any vaccine or medicine development program needs a high degree of acceptance of public confidence and should be in the common interest. At the same time, international legislation and regulatory bodies will play a crucial role in ensuring the needs of individuals, preserving intellectual rights, and resolving disputes. It is also required to ensure accessibility, availability, and affordability to everyone.

## 7. Conclusion

This study presents a comprehensive review of the work regarding COVID-19 diagnosis based on CXR and CT image analysis using deep learning. The state-of-the-art DL techniques for the CXR, CT, and multimodal data diagnosis are presented in Tables 3–5, respectively. Publicly available datasets used in these reviewed studies are summarized in Table 2. We discussed challenges associated with the current DL approaches for the COVID-19 diagnosis. It is important to note that each study in the literature has shown potential in automated detection of COVID-19 and at the same time faced challenges or lacked in analysis and evaluation of the proposed solutions from several points of view. We are of a considered opinion that consolidation of the important observations can act as a benchmark and assistance to the researchers while developing DL-based automated and efficient COVID-19 detection solutions. Some of the important findings from this study are as follows. This review indicates significant utilization of DL methods for the automatic identification of COVID-19 cases from other pulmonary diseases and normal groups. Despite so many studies being undertaken, the majority of the research has been carried out on either CXR or CT image analysis. Further, most studies utilized smaller datasets and also lacked comparative analysis with the existing research. It is further noted that codes and data are not available for many studies, posing challenges in ascertaining the utility of the methods in clinical settings. Although efforts are now being made to show the interpretability of the DL model

via visual saliency on the CXR or CT images, these methods are still in the early stages.

In order to assist the clinicians in hospitals with respect to COVID-19 diagnosis and cure, the upcoming trends in this area require cross-data evaluation (i.e., testing on the unseen dataset of a different hospital) and comparison of cross-methods or benchmarking of the most recent methods. Availability of codes and data in the public space should be required with any research paper so that future researchers/clinicians can deploy and test the methods in actual hospital settings. Efforts should be made to consolidate some bigger public, comprehensive, and diverse datasets having multi-modality data of COVID-19 collected from multiple sites. This would allow the researchers to develop more reliable methods and also enable benchmarking of methods. Interpretability of AI methods should be demonstrated and validated with the help of expert radiologists. It is highly recommended that clinicians, radiologists, and AI engineers work together to evolve interpretable and reliable DL solutions that can also be deployed with ease in the hospitals. Otherwise, despite umpteen number of global efforts, it will take time to utilize these technologies in hospitals to assist mankind.

## Declaration of competing interest

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

## Acknowledgments

The authors would like to express their sincere appreciation to the editor and anonymous reviewers for their comments and valuable suggestions which have helped in improving the manuscript significantly.

## References

- [1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, A novel coronavirus from patients with pneumonia in China, 2019, *N. Engl. J. Med.* 382 (8) (2020) 727–733.
- [2] J.F.-W. Chan, S. Yuan, K.-H. Kok, K.K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R.W.-S. Poon, H.-W. Tsoi, S.K.-F. Lo, K.-H. Chan, V.K.-M. Poon, W.-M. Chan, J.D. Ip, J.-P. Cai, V.C.-C. Cheng, H. Chen, C.K.-M. Hui, K.-Y. Yuen, A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster, *Lancet* 395 (10223) (Feb 2020) 514–523.
- [3] G. Rubin, C. Ryerson, L. Haramati, N. Sverzellati, J. Kanne, S. Raouf, N. Schluger, A. Volpi, J. Yim, I. Martin, D. Anderson, C. Kong, T. Altes, A. Bush, S. Desai, O. Goldin, J. Mo Goo, M. Humbert, Y. Inoue, H. Kauczor, F. Luo, P. Mazzone, M. Prokop, M. Remy-Jardin, L. Richeldi, C. Schaefer-Prokop, N. Tomiyama, A. Wells, A. Leung, The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner society, *Radiology* 296 (1) (Jul. 2020) 172–180, publisher Copyright: © RSNA, 2020; American College of Chest Physicians, published by Elsevier Inc.
- [4] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, W. Ji, Sensitivity of chest CT for COVID-19: comparison to RT-PCR, *Radiology* 296 (2) (2020) E115–E117.
- [5] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z. B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N.A. Emadi, M.B.I. Reaz, M.T. Islam, Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access* 8 (2020) 132 665–132 676.
- [6] L. Orioli, M.P. Hermans, J.-P. Thissen, D. Maiter, B. Vandeleene, J.-C. Yombi, COVID-19 in diabetic patients: related risks and specifics of management, *Ann. Endocrinol.* 81 (2) (2020) 101–109.
- [7] E. Frederick, New research reveals why some patients may test positive for COVID-19 long after recovery. <https://wi.mit.edu/news/new-research-reveals-why-some-patients-may-test-positive-covid-19-long-after-recovery>, last Accessed: 2021-11-02.
- [8] X. Mei, H.-C. Lee, K.-y. Diao, M. Huang, B. Lin, C. Liu, Z. Xie, Y. Ma, P.M. Robson, M. Chung, A. Bernheim, V. Mani, C. Calcagno, K. Li, S. Li, H. Shan, J. Lv, T. Zhao, J. Xia, Q. Long, S. Steinberger, A. Jacobi, T. Deyer, M. Luksza, F. Liu, B.P. Little, Z.A. Fayad, Y. Yang, Artificial intelligence-enabled rapid diagnosis of patients with COVID-19, *Nat. Med.* 26 (8) (Aug 2020) 1224–1228.
- [9] N. Hilmizen, A. Bustamam, D. Sarwinda, The multimodal deep learning for diagnosing COVID-19 pneumonia from chest CT-scan and X-ray images, in: 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, (ISRITI), 2020, pp. 26–31.

- [10] H. Mukherjee, S. Ghosh, A. Dhar, O. Sk, K. Santosh, K. Roy, Deep neural network to detect COVID-19: one architecture for both CT scans and chest X-rays, *Appl. Intell.* 51 (05 2021) 1–13.
- [11] S. Bhattacharya, P.K. Reddy Maddikunta, Q.-V. Pham, T.R. Gadekallu, S. R. Krishnan, S. C.L. Chowdhary, M. Alazab, M. Jalil Piran, Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey, *Sustain. Cities Soc.* 65 (2021) 102589.
- [12] M. Alzubaidi, H.D. Zubaydi, A.A. Bin-Saleem, A.A. Abd-Alrazaq, A. Ahmed, M. Househ, Role of deep learning in early detection of COVID-19: scoping review, *Comput. Methods Progr. Biomed. Upd.* 1 (2021) 100025.
- [13] W. Hariri, A. Narin, Deep neural networks for COVID-19 detection and diagnosis using images and acoustic-based techniques: a recent review, *Soft Comput.* 25 (Aug 2021) 15345–15362. <https://doi.org/10.1007/s00500-021-06137-x>.
- [14] F.M. Shah, S.K.S. Joy, F. Ahmed, T. Hossain, M. Humaira, A.S. Ami, S. Paul, M.A. R.K. Jim, S. Ahmed, A comprehensive survey of COVID-19 detection using medical images, *SN Comput. Sci.* 2 (6) (Aug 2021) 434.
- [15] T. Javaheri, M. Homayounfar, Z. Amoozgar, R. Reiazi, F. Homayounieh, E. Abbas, A. Laali, A.R. Radmard, M.H. Gharib, S.A.J. Mousavi, O. Ghaemi, R. Babaei, H. K. Mobin, M. Hosseinzadeh, R. Jahanban-Esfahlan, K. Seidi, M.K. Kalra, G. Zhang, L.T. Chitkushev, B. Haibe-Kains, R. Malekzadeh, R. Rawassizadeh, CovidCTNet: an open-source deep learning approach to diagnose covid-19 using small cohort of CT images, *npj Digit. Med.* 4 (1) (Feb 2021) 29.
- [16] M. Abdel-Basset, V. Chang, H. Hawash, R.K. Chakraborty, M. Ryan, FSS-2019-nCov: a deep learning architecture for semi-supervised few-shot segmentation of COVID-19 infection, *Knowl. Base Syst.* 212 (2021) 106647.
- [17] A. Vouloudimos, E. Protopapadakis, I. Katsamenis, A. Doulamis, N. Doulamis, Deep learning models for COVID-19 infected area segmentation in CT images, in: *The 14th Pervasive Technologies Related to Assistive Environments Conference. Plus 0.5em Minus 0*, Association for Computing Machinery, 4emNew York, NY, USA, 2021, pp. 404–411.
- [18] D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, L. Shao, Inf-Net: automatic COVID-19 lung infection segmentation from CT images, *IEEE Trans. Med. Imag.* 39 (8) (2020) 2626–2637.
- [19] D. Selvaraj, V. Arunachalam, V. Mahesh, A.N. Joseph Raj, An integrated feature frame work for automated segmentation of COVID-19 infection from lung CT images, *Int. J. Imag. Syst. Technol.* 31 (2020) 11.
- [20] D. Müller, I. Soto-Rey, F. Kramer, Robust chest CT image segmentation of COVID-19 lung infection based on limited data, *Inform. Med. Unlock.* 25 (2021) 100681.
- [21] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
- [22] O. Ronneberger, P. Fischer, T. Brox, U-Net, Convolutional networks for biomedical image segmentation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention. Plus 0.5em Minus 0*, 4emSpringer, 2015, pp. 234–241.
- [23] B.N. Narayanan, R.C. Hardie, V. Krishnaraja, C. Karam, V.S.P. Davuluru, Transfer-to-transfer learning approach for computer aided detection of COVID-19 in Chest Radiographs, *A&I* 1 (4) (2020) 539–557.
- [24] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net, Fully convolutional neural networks for volumetric medical image segmentation, in: *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.
- [25] Z. Zhou, M.M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, UNet++: a nested U-Net architecture for medical image segmentation, in: D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J.M.R. Tavares, A. Bradley, J.P. Papa, V. Belagianni, J.C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, A. Madabhushi (Eds.), *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Plus 0.5em Minus 0.4emCham*, Springer International Publishing, 2018, pp. 3–11.
- [26] R. Karthik, R. Menaka, H. M, Learning distinctive filters for COVID-19 detection from chest X-ray using shuffled residual CNN, *Appl. Soft Comput.* 99 (2021) 106744.
- [27] 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.
- [28] Z. Wang, Y. Xiao, Y. Li, J. Zhang, F. Lu, M. Hou, X. Liu, Automatically discriminating and localizing COVID-19 from community-acquired pneumonia on chest X-rays, *Pattern Recogn.* 110 (2021) 107613.
- [29] R. Azad, M. Asadi-Aghbolaghi, M. Fathy, S. Escalera, Bi-Directional ConvLSTM U-Net with densely connected convolutions, in: *2019 IEEE/CVF International Conference on Computer Vision Workshop, ICCVW*, 2019, pp. 406–415.
- [30] C. Jin, W. Chen, Y. Cao, Z. Xu, Z. Tan, X. Zhang, L. Deng, C. Zheng, J. Zhou, H. Shi, J. Feng, Development and evaluation of an artificial intelligence system for COVID-19 diagnosis, *Nat. Commun.* 11 (1) (Oct 2020) 5088.
- [31] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, J. Xia, Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy, *Radiology* 296 (2) (Aug 2020) E65–E71.
- [32] B. Wang, S. Jin, Q. Yan, H. Xu, C. Luo, L. Wei, W. Zhao, X. Hou, W. Ma, Z. Xu, Z. Zheng, W. Sun, L. Lan, W. Zhang, X. Mu, C. Shi, Z. Wang, J. Lee, Z. Jin, M. Lin, H. Jin, L. Zhang, J. Guo, B. Zhao, Z. Ren, S. Wang, W. Xu, X. Wang, J. Wang, Z. You, J. Dong, AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical AI system, *Appl. Soft Comput.* 98 (2021) 106897.
- [33] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su, et al., A deep learning system to screen novel coronavirus disease 2019 pneumonia, *Engineering* 6 (10) (2020) 1122–1129.
- [34] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, X. Wang, Deep Learning-Based Detection for COVID-19 from Chest CT Using Weak Label, *medRxiv*, 2020.
- [35] X. Ouyang, J. Huo, L. Xia, F. Shan, J. Liu, Z. Mo, F. Yan, Z. Ding, Q. Yang, B. Song, F. Shi, H. Yuan, Y. Wei, X. Cao, Y. Gao, D. Wu, Q. Wang, D. Shen, Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia, *IEEE Trans. Med. Imag.* 39 (8) (2020) 2595–2605.
- [36] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, Y. Shi, Abnormal lung quantification in chest CT images of COVID-19 patients with deep learning and its application to severity prediction, *Med. Phys.* 48 (4) (2021) 1633–1645.
- [37] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, ImageNet: a large-scale hierarchical image database, in: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.
- [38] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, *Adv. Neural Inf. Process. Syst.* 25 (2012) 1097–1105.
- [39] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 09 2014 *arXiv* 1409.1556.
- [40] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2016, pp. 770–778.
- [41] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, 2015, pp. 1–9.
- [42] F. Chollet, Xception: deep learning with depthwise separable convolutions, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [43] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, *AAAI*, 2017.
- [44] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2017, pp. 2261–2269.
- [45] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning deep features for discriminative localization, in: *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2016, pp. 2921–2929.
- [46] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: Visual explanations from deep networks via gradient-based localization, in: *2017 IEEE International Conference on Computer Vision, ICCV*, 2017, pp. 618–626.
- [47] A. Chattopadhyay, A. Sarkar, P. Howlader, V.N. Balasubramanian, Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks, in: *2018 IEEE Winter Conference on Applications of Computer Vision, WACV*, 2018, pp. 839–847.
- [48] Labeled optical coherence tomography (OCT) and chest X-Ray images for classification. <https://data.mendeley.com/datasets/rsclbr9sj/2>. (Accessed 15 August 2021).
- [49] Chest X-ray images (pneumonia). <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia/version/2>. (Accessed 15 August 2021).
- [50] Pneumonia sample xrays. <https://www.kaggle.com/ahmedali2019/pneumonia-sample-xrays>. (Accessed 15 August 2021).
- [51] M. Toğaçar, B. Ergen, Z. Cömert, COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches, *Comput. Biol. Med.* 121 (2020) 103805.
- [52] A. Bustos, A. Pertusa, J.-M. Salinas, M. de la Iglesia-Vaya, Padchest: a large chest x-ray image dataset with multi-label annotated reports, *Med. Image Anal.* 66 (2020) 101797.
- [53] A.J. DeGrave, J.D. Janizek, S.-I. Lee, AI for Radiographic Covid-19 Detection Selects Shortcuts over Signal, *Nature Machine Intelligence*, May 2021.
- [54] AI diagnosis,” <http://ncov-ai.big.ac.cn/download?lang=en>, last Accessed: 2021-08-15.
- [55] M. Irfan, M.A. Iftikhar, S. Yasin, U. Draz, T. Ali, S. Hussain, S. Bukhari, A. S. Alwadie, S. Rahman, A. Glowacz, F. Althobiani, Role of hybrid deep neural networks (HDNNs), computed tomography, and chest X-rays for the detection of COVID-19, *Int. J. Environ. Res. Publ. Health* 18 (6) (2021).
- [56] COVID-19 CXR (all SARS-CoV-2 PCR+). <https://threadreaderapp.com/thread/1243928581983670272.html>. (Accessed 15 August 2021).
- [57] A.K. Das, S. Ghosh, S. Thunder, R. Dutta, S. Agarwal, A. Chakrabarti, Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network, *Pattern Anal. Appl.* (Mar 2021) 1111–1124.
- [58] Actualmed-covid-chestxray-dataset. <https://github.com/agchung/Actualmed-COVID-chestxray-dataset>. (Accessed 15 August 2021).
- [59] E. Hussain, M. Hasan, M.A. Rahman, I. Lee, T. Tamanna, M.Z. Parvez, CoroDet: A Deep Learning Based Classification for COVID-19 Detection Using Chest X-Ray Images,” *Chaos*, vol. 142, Solitons & Fractals, 2021, p. 110495.
- [60] M.Z. Islam, M.M. Islam, A. Asraf, A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images, *Inform. Med. Unlock.* 20 (2020) 100412.
- [61] T. Pham, Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning? *Health Inf. Sci. Syst.* 9 (2020) 11.
- [62] J.P. Cohen, P. Morrison, L. Dao, COVID-19 Image Data Collection, 2020.
- [63] A.M. Ismael, A. Şengür, Deep learning approaches for COVID-19 detection based on chest X-ray images, *Expert Syst. Appl.* 164 (2021) 114054.
- [64] H. Panwar, P. Gupta, M.K. Siddiqui, R. Morales-Menendez, V. Singh, Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet, *Chaos, Solit. Fractals* 138 (2020) 109944.
- [65] T. Agrawal, P. Choudhary, FocusCovid: automated COVID-19 detection using deep learning with chest X-ray images, *Evolv. Syst.* (2021) 1–15.

- [66] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U. Rajendra Acharya, Automated detection of COVID-19 cases using deep neural networks with X-ray images, *Comput. Biol. Med.* 121 (2020) 103792.
- [67] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K.N. Plataniotis, A. Mohammadi, COVID-CAPS: a capsule network-based framework for identification of COVID-19 cases from X-ray images, *Pattern Recogn. Lett.* 138 (2020) 638–643.
- [68] S. Toraman, T.B. Alakus, I. Turkoglu, Convolutional capsnet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks, *Chaos, Solit. Fractals* 140 (2020) 110122.
- [69] F. Ucar, D. Korkmaz, COVIDDiagnosis-net: deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images, *Med. Hypotheses* 140 (2020) 109761.
- [70] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, *Appl. Intell.* 51 (2) (Feb 2021) 854–864.
- [71] S. Sakib, T. Tazrin, M.M. Fouda, Z.M. Fadlullah, M. Guizani, DL-CRC: deep learning-based chest radiograph classification for COVID-19 detection: a novel approach, *IEEE Access* 8 (2020) 171 575–171 589.
- [72] B. Abraham, M.S. Nair, Computer-aided detection of COVID-19 from X-ray images using multi-CNN and Bayesnet classifier, *Biocybern. Biomed. Eng.* 40 (4) (2020) 1436–1445.
- [73] D. Ezzat, A.E. Hassanien, H.A. Ella, An optimized deep learning architecture for the diagnosis of COVID-19 disease based on gravitational search optimization, *Appl. Soft Comput.* 98 (2021) 106742.
- [74] A. Gupta, Anjum, S. Gupta, R. Katarya, InstaCovNet-19: a deep learning classification model for the detection of COVID-19 patients using Chest X-ray, *Appl. Soft Comput.* 99 (2021) 106859.
- [75] M. Heidari, S. Mirniaharikandehi, A.Z. Khuzani, G. Danala, Y. Qiu, B. Zheng, Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms, *Int. J. Med. Inf.* 144 (2020) 104284.
- [76] G. Jain, D. Mittal, D. Thakur, M.K. Mittal, A deep learning approach to detect COVID-19 coronavirus with X-ray images, *Biocybern. Biomed. Eng.* 40 (4) (2020) 1391–1405.
- [77] V. Madaan, A. Roy, C. Gupta, P. Agrawal, A. Sharma, C. Bologna, R. Prodan, XCOVNet: Chest X-Ray Image Classification for COVID-19 Early Detection Using Convolutional Neural Networks, *New Generation Computing*, 02 2021.
- [78] 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 (2020) preprint arXiv:2003.11055*.
- [79] R.M. Pereira, D. Bertolini, L.O. Teixeira, C.N. Silla, Y.M. Costa, COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios, *Comput. Methods Progr. Biomed.* 194 (2020) 105532.
- [80] M. Rahimzadeh, A. Attar, A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2, *Inform. Med. Unlock.* 19 (2020) 100360.
- [81] S. Varela-Santos, P. Melin, A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays using texture features and neural networks, *Inf. Sci.* 545 (2021) 403–414.
- [82] M. y kamil, A deep learning framework to detect COVID-19 disease via chest X-ray and CT scan images, *Int. J. Electr. Comput. Eng.* 11 (02 2021) 844–850.
- [83] S.R. Nayak, D.R. Nayak, U. Sinha, V. Arora, R.B. Pachori, Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study, *Biomed. Signal Process Control* 64 (2021) 102365.
- [84] L. Brunese, F. Mercaldo, A. Reginelli, A. Santone, Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays, *Comput. Methods Progr. Biomed.* 196 (2020) 105608.
- [85] G. Dhiman, V. Chang, K.K. Singh, A. Shankar, ADOPT: automatic deep learning and optimization-based approach for detection of novel coronavirus COVID-19 disease using X-ray images, 0, no. 0, *J. Biomol. Struct. Dyn.* (2021) 1–13. PMID: 33475019.
- [86] A. Khadidos, A.O. Khadidos, S. Kannan, Y. Natarajan, S.N. Mohanty, G. Tsaramirsis, Analysis of COVID-19 infections on a CT image using deepsense model, *Front. Public Health* 8 (2020) 751.
- [87] A.I. Khan, J.L. Shah, M.M. Bhat, CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest X-ray images, *Comput. Methods Progr. Biomed.* 196 (2020) 105581.
- [88] A. Al-Bawi, K. Al-Kaabi, M. Jeryo, A. Al-Fatlawi, CCBlock: an effective use of deep learning for automatic diagnosis of COVID-19 using X-ray images, *Res. Biomed. Eng.* (Nov 2020) 1–10.
- [89] M. Loey, F. Smarandache, N.E.M. Khalifa, Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning, *Symmetry* 12 (4) (2020).
- [90] I.D. Apostolopoulos, T.A. Mpesiana, COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks, *Phys. Eng. Sci. Med.* 43 (2) (Jun 2020) 635–640.
- [91] S. Tabik, A. Gómez-Ríos, J.L. Martín-Rodríguez, I. Sevillano-García, M. Rey-Area, D. Chartre, E. Guirado, J.L. Suárez, J. Luengo, M.A. Valero-González, P. García-Villanova, E. Olmedo-Sánchez, F. Herrera, COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images, *IEEE J. Biomed. Health Inform.* 24 (12) (2020) 3595–3605.
- [92] Covid cases. <https://www.eurorad.org/advanced-search?search=COVID>. (Accessed 15 August 2021).
- [93] C.T. COVIDx. [https://www.kaggle.com/hgunraj/covidxct?select=2A\\_images](https://www.kaggle.com/hgunraj/covidxct?select=2A_images). (Accessed 15 August 2021).
- [94] S. Thakur, A. Kumar, X-ray and CT-scan-based Automated Detection and Classification of COVID-19 Using Convolutional Neural Networks (CNN), *Biomedical Signal Processing and Control*, 2021, p. 102920.
- [95] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Illcus, C. Chute, H. Marklund, B. Haghighi, R. Ball, K. Shpanskaya, et al., Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison, in: *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 2019, pp. 590–597, 1.
- [96] S. Jaeger, S. Candemir, S. Antani, Y.-X.J. Wang, P. Lu, G. Thoma, Two public chest X-ray datasets for computer-aided screening of pulmonary diseases, *Quant. Imag. Med. Surg.* 4 6 (2014) 475–477.
- [97] Digital image database. <http://db.jsrt.or.jp/eng-01.php>. (Accessed 15 August 2021).
- [98] A. Keles, M.B. Keles, A. Keles, COV19-CNNNet and COV19-ResNet: Diagnostic Inference Engines for Early Detection of COVID-19, *Cognitive Computation*, Jan 2021.
- [99] T. Mahmud, M.A. Rahman, S.A. Fattah, 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 (2020) 103869.
- [100] COVID-19 xray dataset (train & test sets), last Accessed: 2021-08-15, <https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets>.
- [101] LIDC-IDRI. <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>. (Accessed 15 August 2021).
- [102] MosMedData: COVID19\_1000 dataset. <https://mosmed.ai/en/>. (Accessed 15 August 2021).
- [103] S. Serte, H. Demirel, Deep learning for diagnosis of COVID-19 using 3D CT scans, *Comput. Biol. Med.* 132 (2021) 104306.
- [104] Chest X-ray (Covid-19 & Pneumonia), Dataset contains chest x-ray images of Covid-19, Pneumonia and normal patients. <https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia>. (Accessed 15 August 2021).
- [105] CoronaHack -chest X-Ray-dataset, classify the X Ray image which is having Corona. <https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset>. (Accessed 15 August 2021).
- [106] COVID-19 radiography database. <https://www.kaggle.com/tawsfurrahman/covid19-radiography-database>. (Accessed 15 August 2021).
- [107] Y. Pathak, P. Shukla, A. Tiwari, S. Stalin, S. Singh, P. Shukla, Deep Transfer Learning Based Classification Model for COVID-19 Disease, *IRBM*, 2020.
- [108] Covid cases. <https://radiopaedia.org/search?lang=us&q=covid&scope=cases>. (Accessed 15 August 2021).
- [109] D.M. Ibrahim, N.M. Elshennawy, A.M. Sarhan, Deep-chest: multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases, *Comput. Biol. Med.* 132 (2021) 104348.
- [110] N.K. Mishra, P. Singh, S.D. Joshi, Automated detection of COVID-19 from CT scan using convolutional neural network, *Biocybern. Biomed. Eng.* 41 (2) (2021) 572–588.
- [111] RSNA pneumonia detection challenge. <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge>. (Accessed 15 August 2021).
- [112] E. Luz, P. Silva, R. Silva, L. Silva, J. Guimarães, G. Miozzo, G. Moreira, D. Menotti, Towards an Effective and Efficient Deep Learning Model for COVID-19 Patterns Detection in X-Ray Images, *Research on Biomedical Engineering*, Apr 2021.
- [113] COVID-19 patients lungs X ray images 10000. <https://www.kaggle.com/nabuelsajid917/covid-19-x-ray-10000-images>. (Accessed 15 August 2021).
- [114] Tuberculosis chest X-ray image data sets. <https://lhncbc.nlm.nih.gov/LHC-publications/pubs/TuberculosisChestXrayImageDataSets.html>. (Accessed 15 August 2021) last.
- [115] COVID cases. <https://sirm.org/category/senza-categoria/covid-19/>. (Accessed 15 August 2021).
- [116] E. Soares, P. Angelov, S. Biaso, M.H. Froes, D.K. Abe, SARS-CoV-2 CT-scan Dataset: A Large Dataset of Real Patients CT Scans for SARS-CoV-2 Identification, *medRxiv* (2020).
- [117] H. Alshazly, C. Linse, E. Barth, T. Martinetz, Explainable COVID-19 detection using chest CT scans and deep learning, *Sensors* 21 (2) (2021).
- [118] V. Arora, E.Y.-K. Ng, R.S. Leekha, M. Darshan, A. Singh, Transfer learning-based approach for detecting COVID-19 ailment in lung CT scan, *Comput. Biol. Med.* 135 (2021) 104575.
- [119] Tianchi competition. <https://tianchi.aliyun.com/competition/entrance/231601/information?lang=en-us>. (Accessed 15 August 2021).
- [120] J. Zhao, Y. Zhang, X. He, P. Xie, Covid-CT-dataset: a CT Scan Dataset about COVID-19, 2020 *arXiv preprint arXiv:2003.13865*.
- [121] M. Polsinelli, L. Cinque, G. Placidi, A light CNN for detecting COVID-19 from CT scans of the chest, *Pattern Recogn. Lett.* 140 (2020) 95–100.
- [122] M. Turkoglu, COVID-19 Detection System Using Chest CT Images and Multiple Kernels-Extreme Learning Machine Based on Deep Neural Network, *IRBM*, 2021.
- [123] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, N. Mehendale, Diagnosis of COVID-19 using CT scan images and deep learning techniques, *Emerg. Radiol.* 28 (3) (Jun 2021) 497–505.
- [124] M. de la Iglesia Vaya, J.M. Saborit, J.A. Montell, A. Pertusa, A. Bustos, M. Cazorla, J. Galant, X. Barber, D. Orozco-Beltrán, F. García-García, M. Caparrós, G. González, J.M. Salinas, BIMCV COVID-19+: a Large Annotated Dataset of RX and CT Images from COVID-19 Patients, 2020.
- [125] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2017, pp. 3462–3471.

- [126] 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) (Nov 2020), 19549.
- [127] Datasets & analysis," <https://iee-dataport.org/authors/tao-yan>, last Accessed: 2021-08-15.
- [128] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Cha, W. Liang, C. Wang, K. Wang, L. Ye, M. Gao, Z. Zhou, L. Li, J. Wang, Z. Yang, H. Cai, J. Xu, L. Yang, G. Wang, Clinically applicable AI system for accurate diagnosis, quantitative measurements and prognosis of COVID-19 pneumonia using Computed Tomography, *Cell* 181 (05 2020).
- [129] M. Jun, G. Cheng, W. Yixin, A. Xingle, G. Jiantao, Y. Ziqi, Z. Minqing, L. Xin, D. Xueyuan, C. Shucheng, W. Hao, M. Sen, Y. Xiaoyu, N. Ziwei, L. Chen, T. Lu, Z. Yuntao, Z. Qiongjie, D. Guoqiang, H. Jian, COVID-19 CT Lung and Infection Segmentation Dataset, Apr. 2020, <https://doi.org/10.5281/zenodo.3757476> [Online]. Available:.
- [130] W. Ning, S. Lei, J. Yang, Y. Cao, P. Jiang, Q. Yang, J. Zhang, X. Wang, F. Chen, Z. Geng, L. Xiong, H. Zhou, Y. Guo, Y. Zeng, H. Shi, L. Wang, Y. Xue, Z. Wang, Open resource of clinical data from patients with pneumonia for the prediction of COVID-19 outcomes via deep learning, *Nat. Biomed. Eng.* 4 (2020) 1–11, 12.
- [131] E.-S.M. El-Kenawy, A. Ibrahim, S. Mirjalili, M.M. Eid, S.E. Hussein, Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images, *IEEE Access* 8 (2020) 179 317–179 335.
- [132] C. Sitaula, M.B. Hossain, Attention-based VGG-16 model for COVID-19 chest X-ray image classification, *Appl. Intell.* 51 (5) (May 2021) 2850–2863.
- [133] A.A. Ardakani, A.R. Kanafi, U.R. Acharya, N. Khadem, A. Mohammadi, Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks, *Comput. Biol. Med.* 121 (2020) 103795.
- [134] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, Y. Shi, Abnormal lung quantification in chest CT images of COVID-19 patients with deep learning and its application to severity prediction, *Med. Phys.* 48 (4) (Mar 2021) 1633–1645.
- [135] Y.-H. Wu, S.-H. Gao, J. Mei, J. Xu, D.-P. Fan, R.-G. Zhang, M.-M. Cheng, JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation, *IEEE Trans. Image Process.* 30 (2021) 3113–3126.
- [136] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR), 2016, pp. 2818–2826.
- [137] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, et al., A Deep Learning Algorithm Using CT Images to Screen for Corona Virus Disease (COVID-19), *European radiology*, 2021, pp. 1–9.
- [138] G. Erion, J.D. Janizek, P. Sturmfels, S.M. Lundberg, S.-I. Lee, Improving Performance of Deep Learning Models with Axiomatic Attribution Priors and Expected Gradients, *Nature Machine Intelligence*, May 2021.
- [139] S. Wang, Y. Zha, W. Li, Q. Wu, X. Li, M. Niu, M. Wang, X. Qiu, H. Li, H. Yu, et al., A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis, *Eur. Respir. J.* 56 (2) (2020).
- [140] S. Balochian, H. Balochian, Social mimic optimization algorithm and engineering applications, *Expert Syst. Appl.* 134 (2019) 178–191.
- [141] T. Zhou, H. Lu, Z. Yang, S. Qiu, B. Huo, Y. Dong, The ensemble deep learning model for novel COVID-19 on CT images, *Appl. Soft Comput.* 98 (2021) 106885.
- [142] S.-H. Wang, V.V. Govindaraj, J.M. Górriz, X. Zhang, Y.-D. Zhang, COVID-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network, *Inf. Fusion* 67 (2021) 208–229.
- [143] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Zha, J. Shen, Y. Chong, Y. Yang, Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, *IEEE ACM Trans. Comput. Biol. Bioinf* (1–1) (2021).
- [144] C. Kaul, S. Manandhar, N.E. Pears, FocusNet: an attention-based fully convolutional network for medical image segmentation, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 455–458.
- [145] K. Hammoudi, H. Benhabiles, M. Melkemi, F. Dornaika, I. Arganda-Carreras, D. Collard, A. Scherpereel, Deep Learning on Chest X-Ray Images to Detect and Evaluate Pneumonia Cases at the Era of Covid-19, 04 2020 *arXiv preprint arXiv: 2004.03399*.
- [146] Z.Q. Lin, M.J. Shafiee, S. Bochkarev, M.S. Jules, X.Y. Wang, A. Wong, Do Explanations Reflect Decisions? a Machine-Centric Strategy to Quantify the Performance of Explainability Algorithms, 2019 *arXiv preprint arXiv:1910.07387*.
- [147] P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, A.I. Explainable, A review of machine learning interpretability methods, *Entropy* 23 (1) (2021).
- [148] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bénéto, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI, *Inf. Fusion* 58 (2020) 82–115.