Medical & Biological Engineering & Computing (2023) 61:1257–1297
https://doi.org/10.1007/s11517-022-02758-y

REVIEW ARTICLE

A survey of machine learning-based methods for COVID-19 medical image analysis

Kashfia Sailunaz[1] · Tansel Özyer[2] · Jon Rokne[1] · Reda Alhajj[1,3,4]

Received: 14 March 2022 / Accepted: 22 December 2022 / Published online: 28 January 2023

Abstract
The ongoing COVID-19 pandemic caused by the SARS-CoV-2 virus has already resulted in 6.6 million deaths with more than 637 million people infected after only 30 months since the first occurrences of the disease in December 2019. Hence, rapid and accurate detection and diagnosis of the disease is the first priority all over the world. Researchers have been working on various methods for COVID-19 detection and as the disease infects lungs, lung image analysis has become a popular research area for detecting the presence of the disease. Medical images from chest X-rays (CXR), computed tomography (CT) images, and lung ultrasound images have been used by automated image analysis systems in artificial intelligence (AI)- and machine learning (ML)-based approaches. Various existing and novel ML, deep learning (DL), transfer learning (TL), and hybrid models have been applied for detecting and classifying COVID-19, segmentation of infected regions, assessing the severity, and tracking patient progress from medical images of COVID-19 patients. In this paper, a comprehensive review of some recent approaches on COVID-19-based image analyses is provided surveying the contributions of existing research efforts, the available image datasets, and the performance metrics used in recent works. The challenges and future research scopes to address the progress of the fight against COVID-19 from the AI perspective are also discussed. The main objective of this paper is therefore to provide a summary of the research works done in COVID detection and analysis from medical image datasets using ML, DL, and TL models by analyzing their novelty and efficiency while mentioning other COVID-19-based review/survey researches to deliver a brief overview on the maximum amount of information on COVID-19-based existing researches.

Keywords COVID-19 · Medical image analysis · Machine learning · Deep learning · Transfer learning · Computer tomography

Abbreviations

AI	Artificial intelligence
ANN	Artificial neural network
ARDS	Acute respiratory distress dyndrome
AUC	Area under the curve
AUROC	Area under the receiver operating Characteristic curve
BGWO	Binary grey wolf optimization
BRISK	Binary robust invariant scalable keypoints
CHS	COVID-19 hierarchical segmentation
CLAHE	Contrast limited adaptive histogram equalization
CNN	Convolutional neural network
CT	Computed tomography
CXR	Chest X-ray
DBSCAN	Density-based spatial clustering of applications with noise
DL	Deep learning
DLA	Deep layer aggregation
DT	Decision tree
DTL	Deep transfer learning
EC	Edge computing

✉ Reda Alhajj
alhajj@ucalgary.ca

1 Department of Computer Science, University of Calgary, Calgary, AB, Canada

2 Department of Computer Engineering, Ankara Medipol University, Ankara, Turkey

3 Department of Computer Engineering, Istanbul Medipol University, Istanbul, Turkey

4 Department of Health Informatics, University of Southern Denmark, Odense, Denmark

FCM	Fuzzy C-means
FN	False negative
FOSF	First-order statistical features
FP	False positive
FPR	False positive rate
GAN	Generative adversarial networks
GLCM	Grey level co-occurrence matrix
GRNN	Generalized regression neural network
HOG	Histogram of oriented gradients
HSGO	Hybrid social group optimization
KNN	*k*-nearest neighbor
KPCA	Kernel principal component analysis
LBP	Local binary pattern
LDA	Linear discriminant analysis
LE	Local entropy
LR	Logistic regression
LRP	Layer-wise relevance propagation
L-SVC	Linear support vector classifier
LSTM	Long short-term memory
MADF	Modified anisotropic diffusion filtering
MCC	Matthews correlation coefficient
MERS-CoV	Middle East respiratory syndrome Coronavirus
ML	Machine learning
MSVM	Multi-class support vector machine
NB	Naïve Bayes
PCA	Principal component analysis
PCR	Polymerase chain reaction
PEPX	Projection-expansion-projection-extension
PHEIC	Public Health Emergency of International Concern
PNN	Probabilistic neural network
PPV	Positive predictive value
RF	Random forest
RMSProp	Root mean squared propagation
RNN	Recurrent neural network
ROC	Receiver operating characteristics
ROI	Region of interest
RT-PCR	Reverse transcriptase-polymerase chain reaction
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SEIR	Susceptible-exposed-infected-removed
SGD	Stochastic gradient descent
SGO	Social group optimization
SIR	Susceptible-lovered
SMO	Social mimic optimization
SVC	Support vector classifier
SVM	Support vector machine
SVR	Support vector regression
TL	Transfer learning
TN	True negative
TNR	True negative rate
TP	True positive
TPR	True positive rate
VGG	Visual geometry group
VOC	Variants of concern
WHO	World Health Organization
YOLO	You only look once

1 Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), first reported in December 2019 from Wuhan, China [1, 2], is a novel coronavirus that has caused the ongoing COVID-19 pandemic. It appeared as pneumonia cases with an unidentified cause. The virus spread rapidly causing a large disease outbreak. Due to the severity of the outbreak caused by this virus, the World Health Organization (WHO) declared it a "Public Health Emergency of International Concern (PHEIC)" on January 30, 2020. After the cases outside of China increased almost 13-fold and the number of infected countries increased by 3 times with 118,000 cases in 114 countries, and 4291 deaths, WHO finally declared COVID-19 to be a "Global Pandemic" on March 11, 2020 [3]. On August 02, 2021, the number of cases was 199,558,118 with 4,247,970 deaths and 180,027,978 people recovered [4]. The number of active cases was 15,282,170 where 99.4% were mildly critical and 0.6% were very critical cases. The number of closed cases that had an outcome was 184,275,948 and 4,247,970 (i.e., 2%) died whereas 180,027,978 (i.e., 98%) were discharged after recovery. On January 22, 2020, the number of detected cases was 987 which increased the total to 199,085,178 (almost 200 million) by August 01, 2021. A more recent statistics on November 06, 2022 listed 6,605,691 deaths and 617,384,409 recovered cases among 637,767,096 COVID-infected cases worldwide.

SARS-CoV-2 is a positive-sense single-stranded RNA virus that can cause respiratory, neurological, hepatic, and other diseases [5]. It is the 7th human coronavirus that causes severe pneumonia similar to two other human coronaviruses — SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV) [6]. OC43, NL63, HKU1, and 229E are the four other human coronaviruses that have mild symptoms. Through multiple mutations, new variants of SARS-CoV-2 have evolved that are more transmissible having new symptoms and being more likely to have fatal consequences. The current four variants of concern (VOC) are Alpha, Beta, Gamma, Delta, and Omicron that first appeared in the UK, South Africa, Brazil, India, and South Africa respectively [7]. Alpha and Beta variants were named on December 18, 2020, whereas the Gamma variant was named on January 11, 2021, Delta on May 11, 2021, and Omicron on November 24, 2021. The

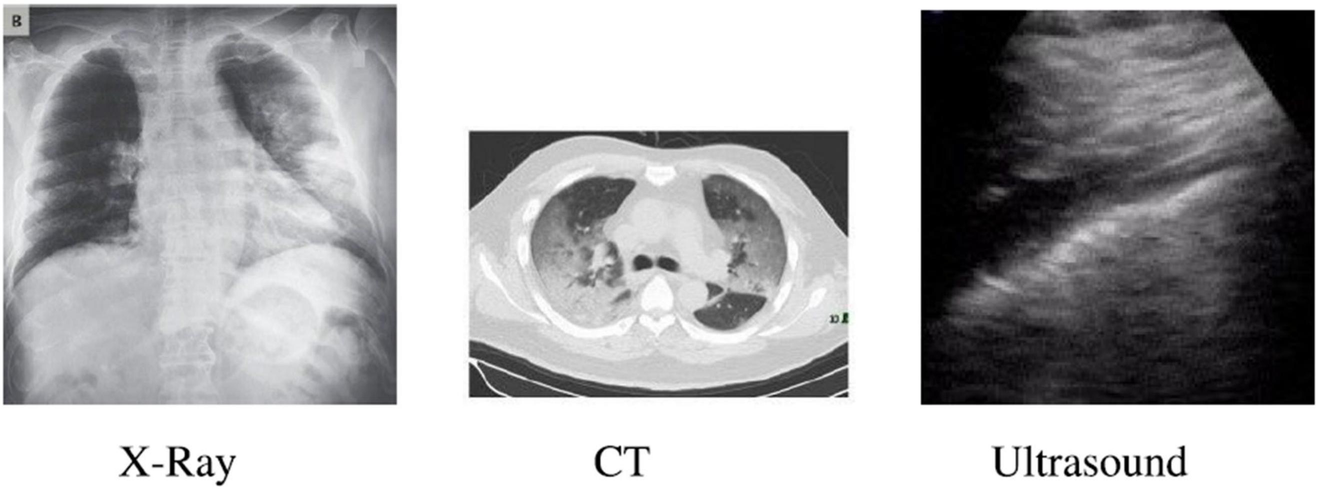

Fig. 1 Sample CXR, CT, and lung ultrasound of COVID-19 patients [8]

subvariants of these variants are due to some descendant lineages of the mutations.

Due to the severity of the disease, early and accurate detection is essential for managing a COVID-19 infection. This requires rapid fast tests for diagnosing the infection. As nucleic acid-based test approaches are among the most reliable for virus detection, the polymerase chain reaction (PCR) method is one of the most popular viral detection methods that achieved higher sensitivity and specificity within minimum time. The reverse transcriptase-PCR (RT-PCR) is the "gold standard" for SARS-CoV-2 detection [9]. The RT-PCR tests the nucleic acids from the upper and lower respiratory specimens like nasopharyngeal or oropharyngeal swabs and sputum. There are not enough RT-PCR testing kits available worldwide and the test itself is time-consuming. Optionally chest imaging techniques can instead be helpful for rapid COVID-19 detection, severity assessment, and patient management since SARS-CoV-2

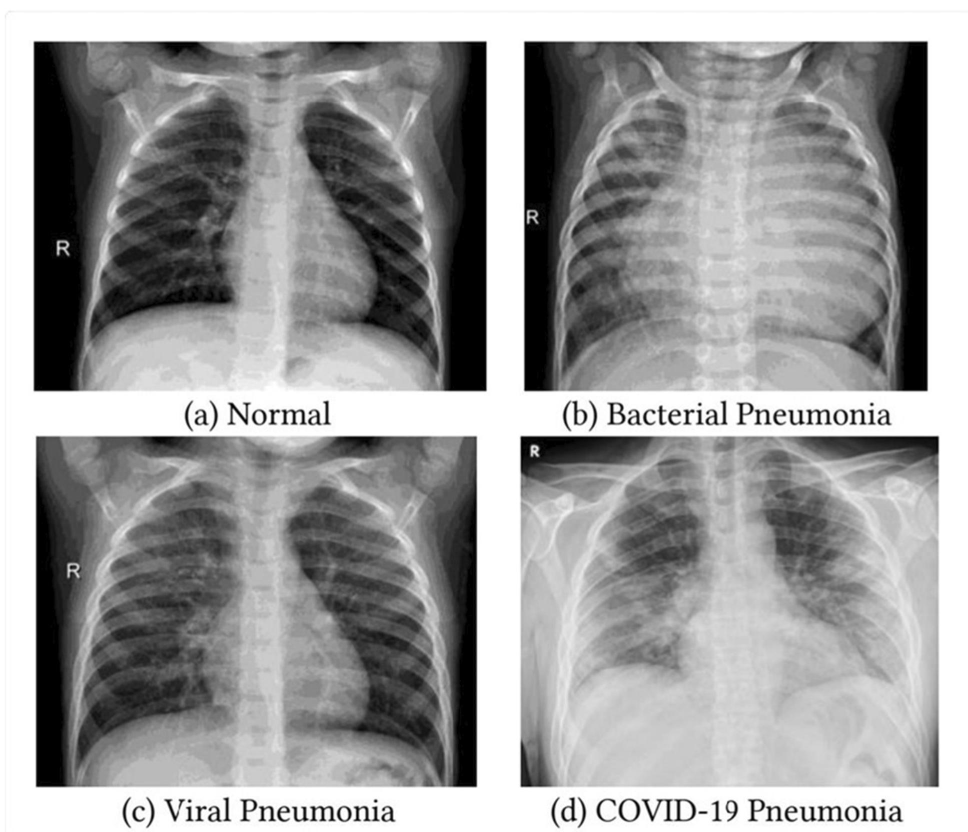

Fig. 2 Sample CXRs of healthy, pneumonia, and COVID-19 patients [11]

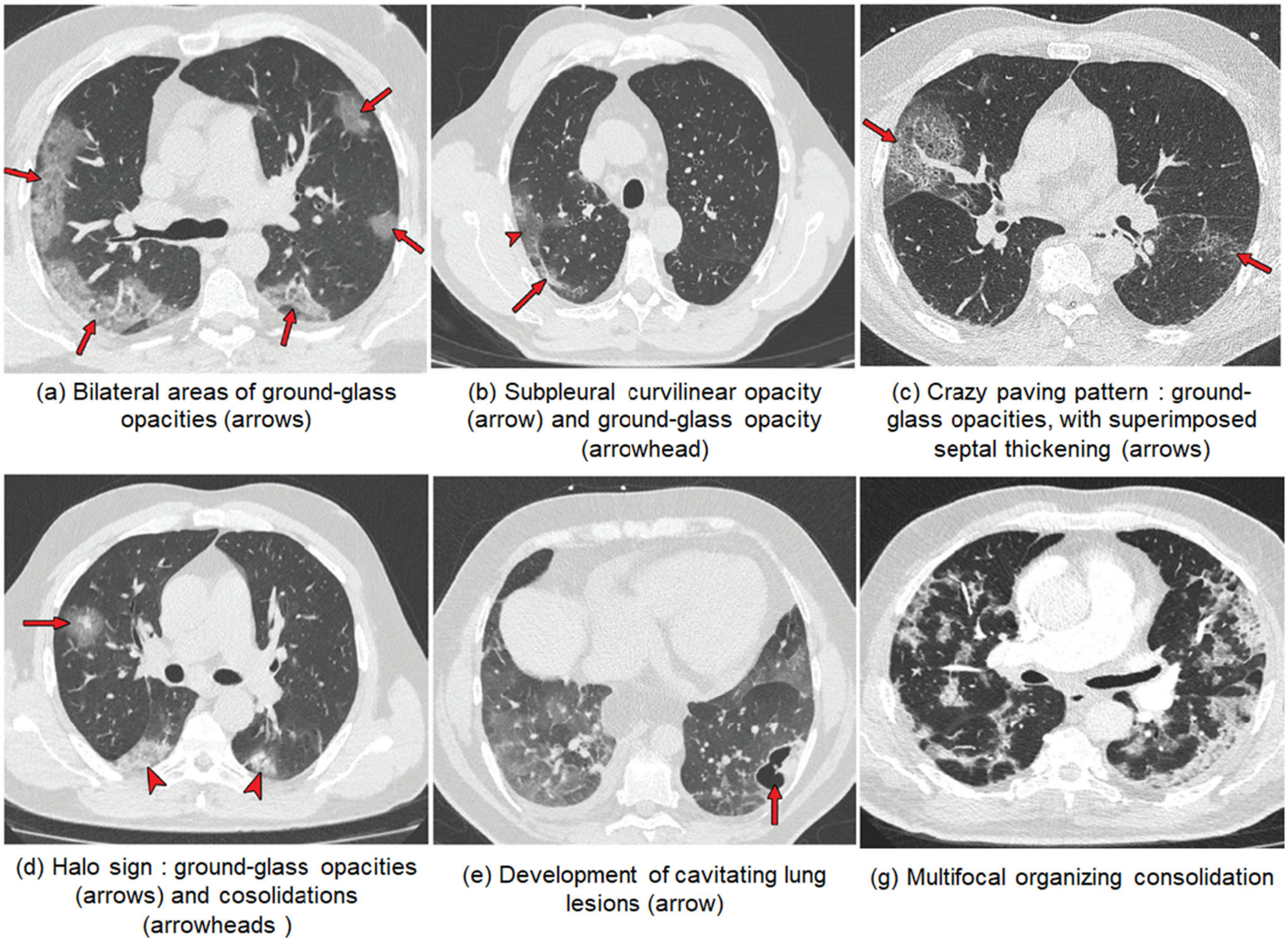

Fig. 3 Sample chest CTs with infected lung regions of COVID-19 patients [16]

normally infects the lung. The imaging techniques of chest radiography (i.e., CXR), chest CT, and lung ultrasounds can therefore be used for chest imaging and diagnosis of COVID-19 patients [10]. Figure 1 shows CXR, CT, and lung ultrasound sample images of COVID-19 patients.

CXRs have been widely used for medical imaging for COVID-19 since X-ray scanners are available in almost every healthcare facility in every country. The cost for CXR is generally lower than for RT-PCR and other testing kits and it requires less processing time [12] compared to the kits. Figure 2 shows sample CXRs of healthy, bacterial pneumonia patient, viral pneumonia patient, and COVID-19 patient. Chest CT produces images with a bit higher sensitivity level than CXR [13]. Most of the experiments using both techniques showed that there are no significant differences between them for COVID-19 diagnosis. CXR and CT also often compliment each other in cases with intermediate findings [14]. Lung ultrasound is another low cost available option for imaging COVID-19 patients. Although it produces low level images, no X-ray radiation is involved [15].

Medical images may portray normal (i.e., healthy) representation of any organ or body part, or they may contain some irregularity or deformity (also known as "abnormalities") that indicate some medical condition. As these abnormalities in medical images are not generally present in the images of healthy people, they can be used to differentiate between healthy people and people with medical conditions or diseases. Medical imaging may show various types of abnormalities such as severity and pre-existing lung conditions in COVID-19 patients based on their disease stage. Different types of opacity in lungs are one major indication of the disease. This can include subpleural curvilinear opacity, reticulonodular opacity, and ground-glass opacity (i.e., bilateral, multifocal, peripheral, posterior, medial, and basal). Other common

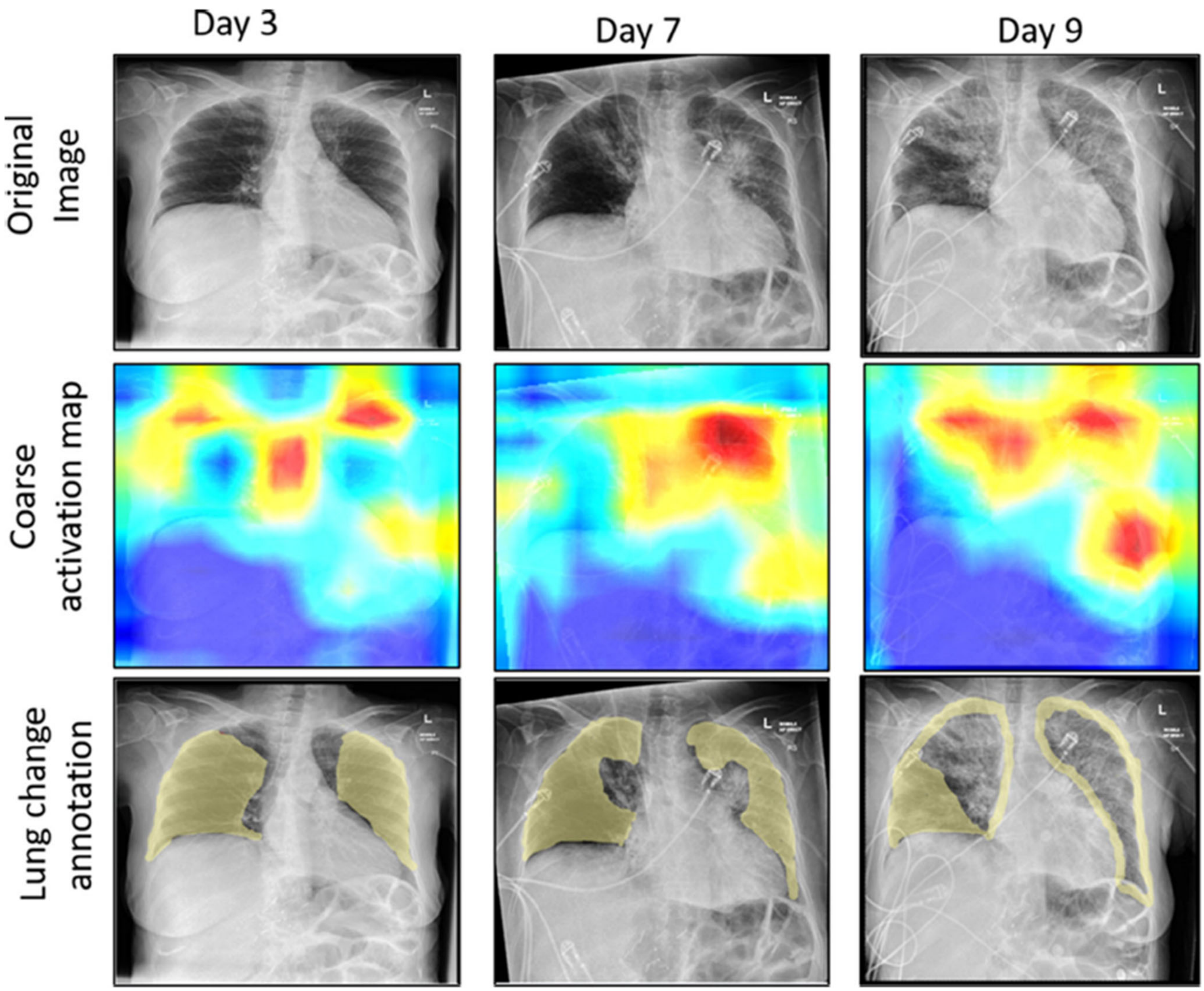

Fig. 4 Sample heatmap visualization of the CXR of a COVID-19 patient at different days using Grad-CAM [17]

abnormalities include septal thickening, consolidation (i.e., pulmonary and air space), cavitating lesions, and crazy-paving appearances with various irregular shapes like tree-in-bud or halo signs [16]. Figure 3 shows some sample chest CT showing infection variations of COVID-19 patients.

The results of medical image analysis (i.e., CXR, CT, and lung ultrasound) for COVID-19 images provide a convenient and non-invasive way of diagnosis and hence, medical professionals have been using them along with the RT-PCR test results (Fig. 4). To further help the medical professionals, analyzing the medical images with AI can be valuable as AI requires significantly less amount of time than the manual process for assessing the results of the analysis. Hence, AI-based tools and techniques have been popular for COVID-19 analysis tasks and achieved high accuracy for (i) detecting if the patient is COVID-19 positive or negative, (ii) classifying COVID-19 cases from healthy, viral pneumonia, and bacterial pneumonia cases, (iii) identifying the severity of the infection, (iv) segmenting the infected regions, (v) tracking the progress of the disease over specific time intervals, and vi) indicating the infected region. Various basic ML, DL, and hybrid models have been implemented and tested on COVID-19 images for these tasks. Some basic COVID analysis applications are shown in Fig. 5. To accelerate the execution of the image processing, some existing efficient DL models trained with various image databases were directly applied through TL. Some researchers also proposed novel ideas and new deep neural network DNN models for the tasks. The experiments done on public or collected COVID-19

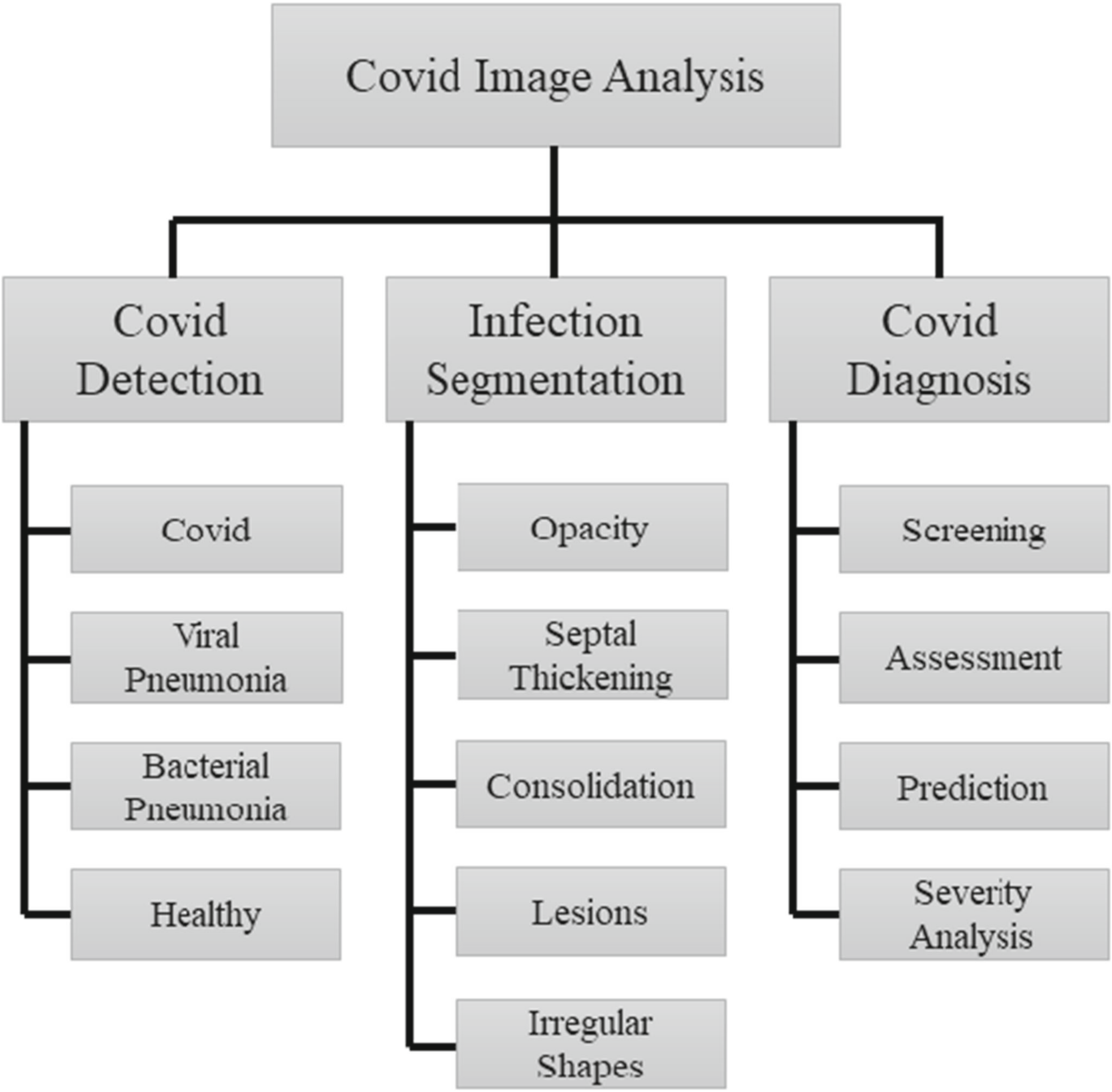

Fig. 5 COVID-19 medical image analysis

image datasets showed impressive performances for most of these approaches and provided some guidelines for new relevant researches.

Within the short period of time since the COVID-19 outbreak, some researchers have developed and published AI-based tools and applications such as LungPrint [18], InferReadTM Solutions [19], Clara COVID-19 [20], CAD4COVID [21], AI4COVID-19 [22], CoViDiag [23], and COVID-19 Assistant Discrimination [24] for COVID-19 data analysis, assessment, and screening. These tools are either works in progress or ready to use. Some visualization-based AI applications have also been used for distinguishing the COVID-19-infected regions of the lungs from the remainder of the lung. Figure 4 shows sample visualizations of heatmaps of infected regions of CXRs of a COVID-19 patient on days 3, 7, and 9 respectively using Grad-CAM [25].

The overall framework of medical image processing for COVID detection and segmentation is shown in Fig. 6. The medical image datasets are used as inputs. CXRs, CTs,

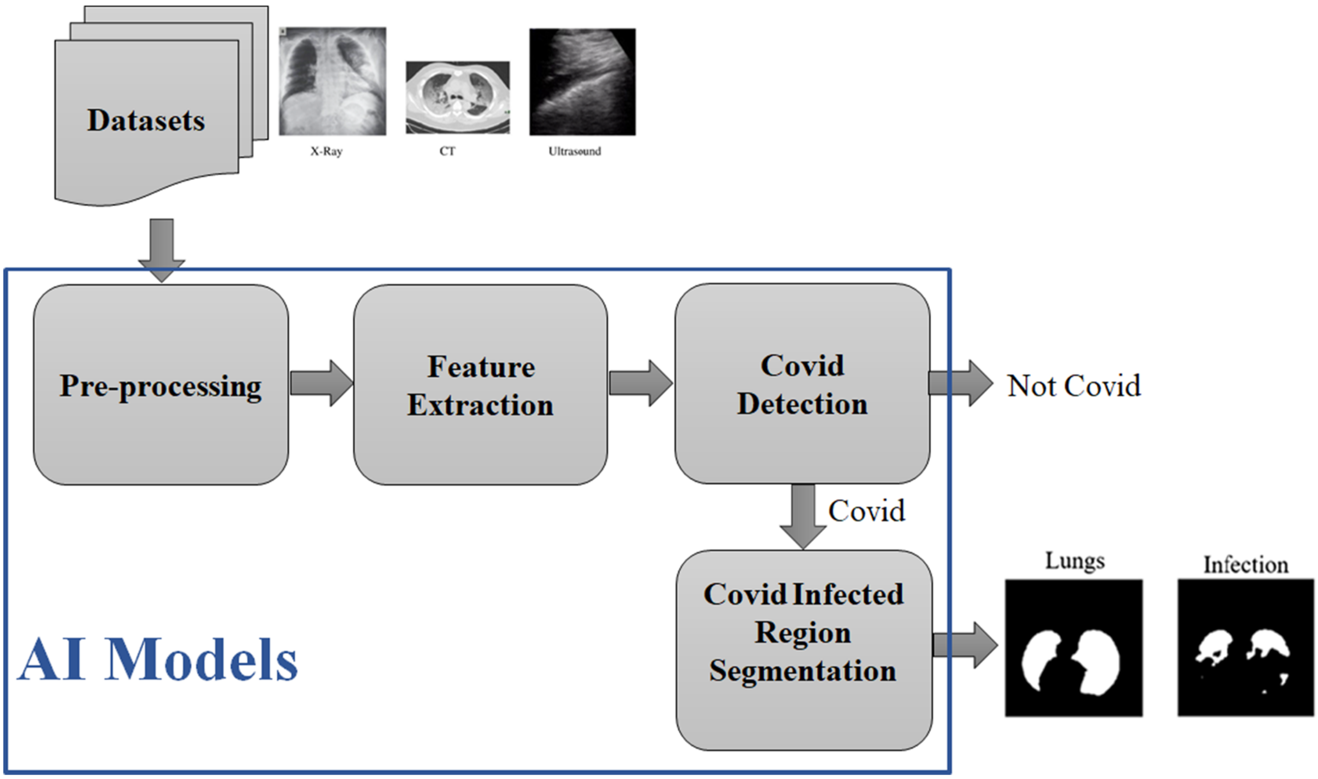

Fig. 6 COVID-19 medical image detection and segmentation workflow

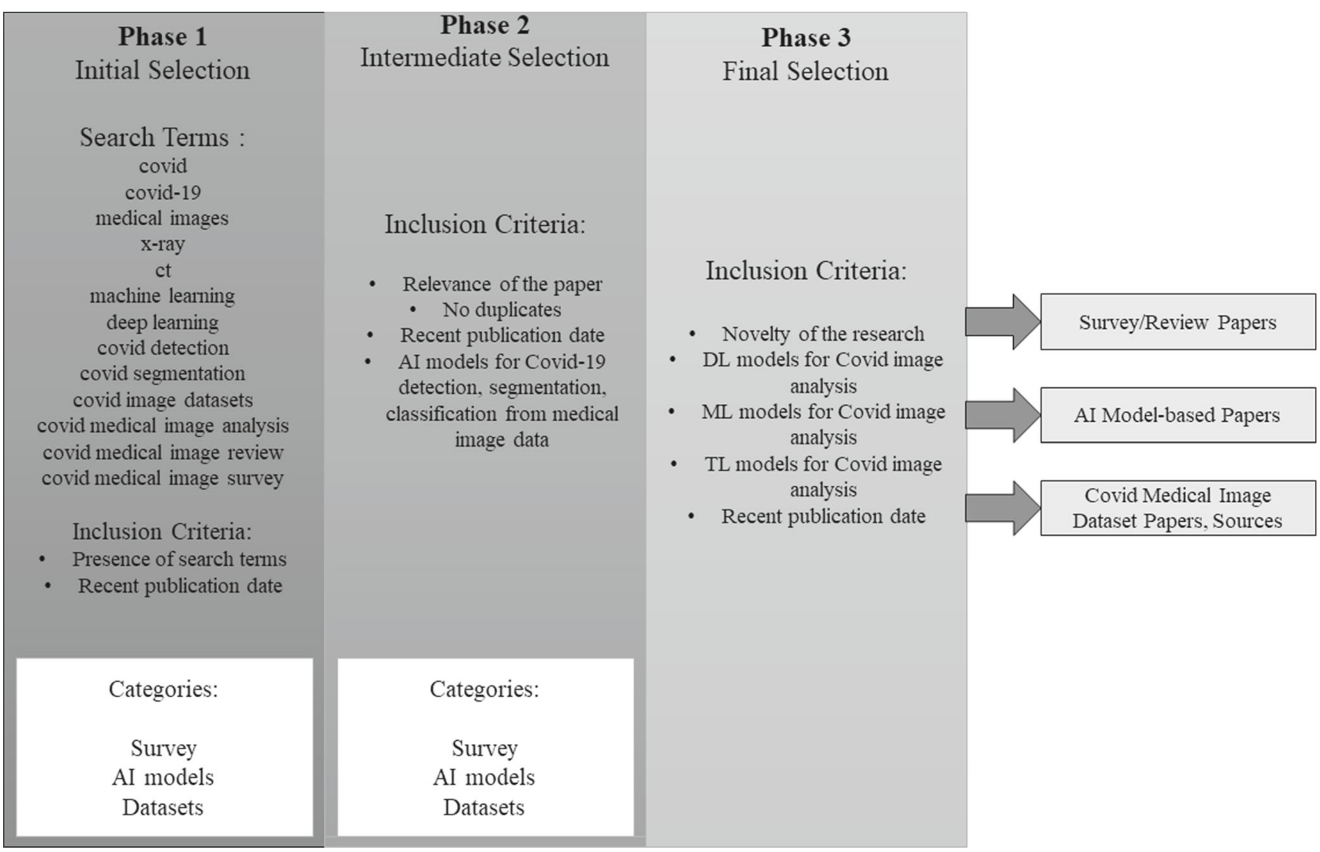

Fig. 7 The paper selection process for this review

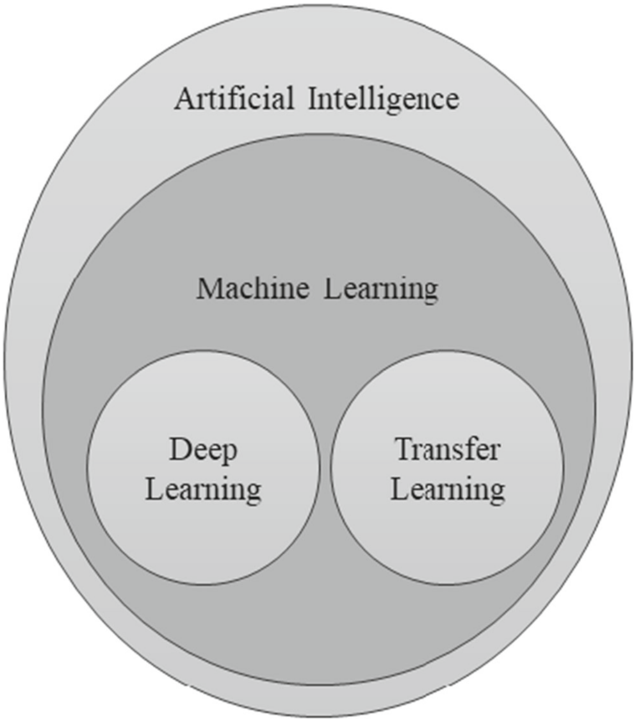

Fig. 8 Relationships between AI, ML, DL, and TL

and ultrasounds are normally used for COVID detection and infected lung region segmentation in combination with various types of AI models/algorithms (i.e., ML, DL, TL, and hybrid). The input images are pre-processed using various algorithms to mostly remove noises and unnecessary parts from the images and enhance the required pixels. Then feature extraction algorithms are used to generate significant feature sets from the images. Based on the image and output requirements, these pre-processing and feature extraction models are chosen from existing image-based algorithms or novel algorithms. The feature sets are then used for classification and segmentation algorithms. The classification process detects if the image is of a COVID-infected lung or not. If it detects COVID, then the next step is segmenting the infected region of the lungs. Segmentation algorithms extract the infected region as outputs. Both detection and segmentation performances are evaluated with standard image-based performance measures. The image detection methods use various classification models to classify the image into healthy, COVID-19, and other disease classes. In most cases, the task is to evaluate if the output class generated by the model matches the class label (i.e., ground truth) of the original data. Accuracy, precision, sensitivity, specificity, and similar scores are computed to evaluate the correctness and quality of the classifier. On the other hand, the image segmentation task generates images with the segmented infection region. So, the segmentation evaluation process checks if the pixels from the ground truth image (i.e., infection region) matches the pixels of the segmented output image. Similarity or dissimilarity metrics like Dice score and Jaccard scores are generally calculated for image segmentation performance evaluation. The definitions and equations of these performance metrics are explained in Section 3.

In this paper, a review of recent research efforts on AI-based COVID-19 analysis from medical images

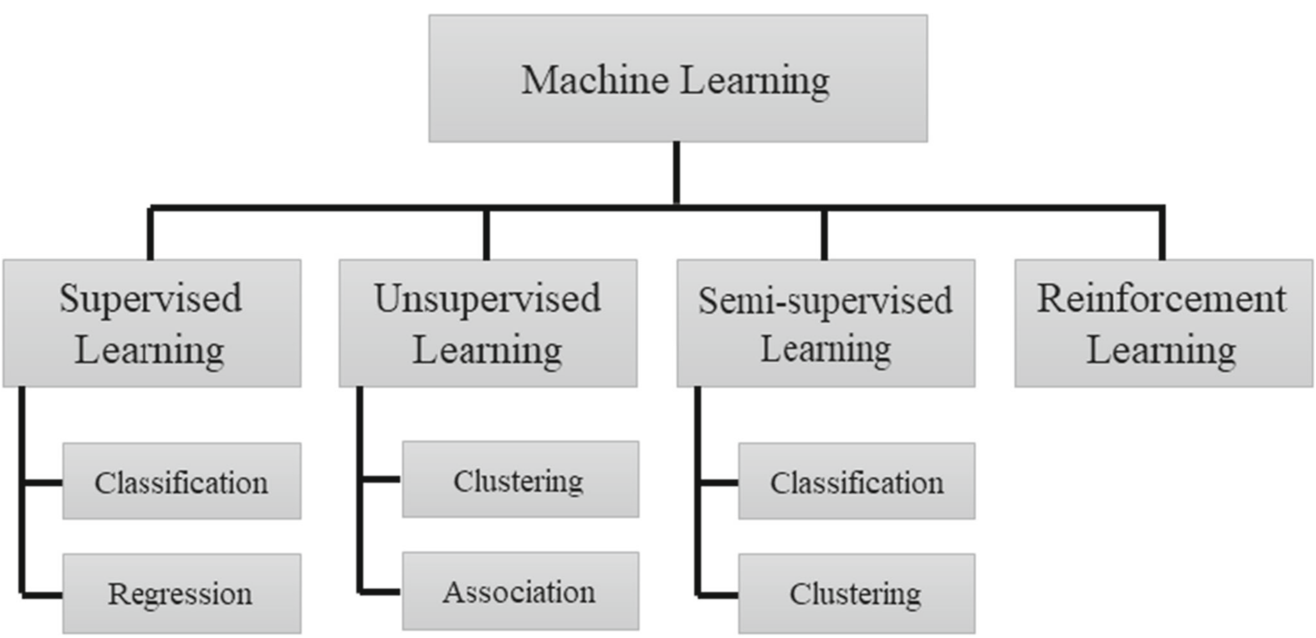

Fig. 9 ML models

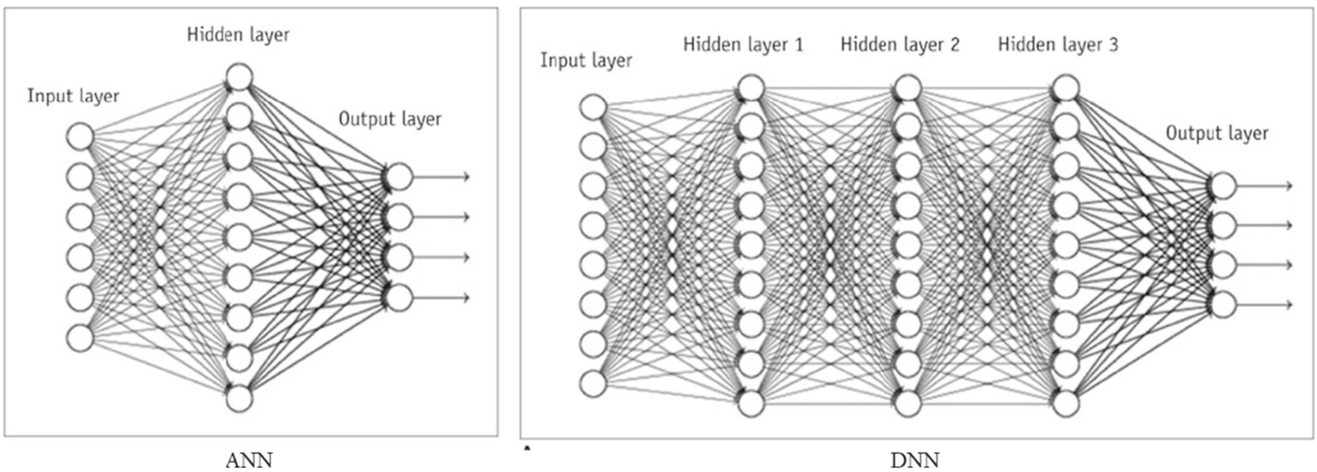

Fig. 10 Artificial neural network (ANN) vs. deep neural network (DNN) [31]

is provided. A few survey papers are also reviewed and summarized along with ML, DL, TL, and hybrid model-based research works on COVID-19 medical image analysis. The mentioned approaches are summarized by including their contributions, novelties, tasks, features, methods, datasets, and performance scores. Some COVID-19 image datasets used in recent relevant approaches are also mentioned with their image characteristics, ground truths, sources, etc. The performance metrics commonly used for COVID-19 image analysis in recent works are defined and explained. Finally, some challenges and future research scopes based on the review of COVID-19 approaches are mentioned to provide indications for research directions.

The publications from January 2020 to August 2022 are considered for this literature review. Publications found in "PubMed" [26] and "Google Scholar" [27] using the following search terms and various combinations of these terms — "covid," "covid-19," "medical images," "x-ray," "ct," "machine learning," "deep learning," "covid detection," "covid segmentation," "covid image datasets," "covid medical image analysis," "covid medical image review," "covid medical image survey." The publications listed in phase 1 were completely based on the inclusion of the search terms and more recent publication date. In phase 2, the listed papers were read and analyzed to select the most relevant ones and to remove duplicates. After removing duplicates and selecting the papers that were based on AI models for COVID-19 detection, segmentation, classification, etc., the final list of papers was created in phase 3. In phase 3, the final papers were selected based on the most recent publications dates and novelty of the research. This process was repeated for the survey papers and the AI model-based papers separately. Among 50 publications collected on the COVID-19 literature review until August 2022, 6 were removed from the list at phase

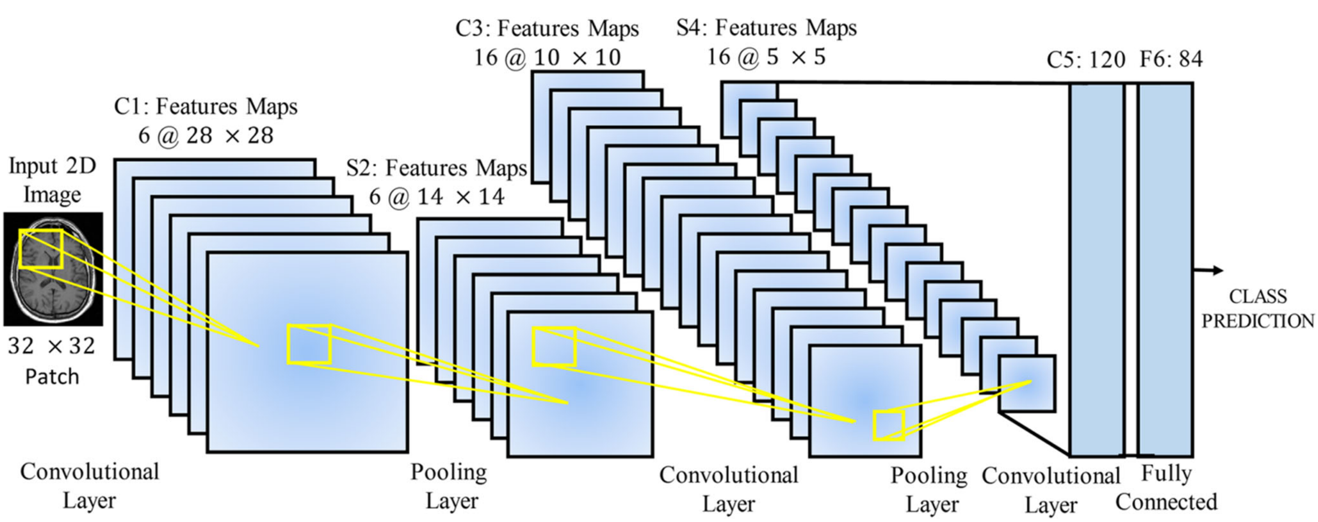

Fig. 11 A sample CNN structure for medical image analysis [32]

1 as they were not completely based on AI-based medical image analysis. A total of 44 survey papers were shortlisted and read in phase 2. Finally, in phase 3, 18 of them were finalized for this paper since they were most relevant to the topic and their publication dates were more recent meaning they included more updated summaries of the research work on COVID-19 medical image analysis. Similarly, for the AI model-based papers, 224 papers were collected in phase 1. After removing the duplicates and irrelevant ones, 50 papers were shortlisted in phase 2 and finally 32 were summarized in phase 3 by checking the novelty of the works. The medical image datasets available for COVID-19 were searched and listed based on the dataset searches and the mentioned datasets in the read papers. Twenty-one datasets that include COVID-19 CXRs, CTs, or ultrasounds (mostly images in different formats) were mentioned with their details in this paper. Figure 7 shows the survey papers, AI model-based papers and dataset selection process.

The major contributions of this paper are as follows:

- providing a summary containing analysis of recent COVID-19 medical image-based researches using AI tools,
- discussing basics of ML, DL, TL models, common evaluation metrics for the topic,
- summarizing existing recent literature reviews on the topic to present the overall research directions on the topic,
- providing a literature review on recent AI model-based researches on COVID-19 medical image analysis,
- listing the methods, datasets, type of data, performances, and major contributions of recent researches on the topic,
- providing necessary resources on available COVID-19 medical image databases,
- listing the challenges related to the research topic to indicate possible future research ideas.

The rest of this paper is outlined as follows. Section 2 shows the basics of AI models and algorithms. Section 3 includes the performance metrics used in COVID-19 detection, classification, and segmentation. Sections 4 and 5 include summaries of some existing works on COVID-19 analysis surveys and various AI models (ML, DL, TL, and hybrid methods) respectively. Section 6 lists some available image datasets for COVID-19 detection methods. Section 7 discusses the scopes and challenges with respect to COVID-19 analysis and Section 8 concludes the paper.

2 AI models

AI is used to minimize human interaction with systems using computers to mimic human intelligence [28]. Medical data analysis is one of the major applications of AI and different ML algorithms have been used for automated systems that use medical data (i.e., image, text, and audio). AI models have been used in disease detection and diagnosis with ML, DL, and TL models. Analyzing large amount of medical data with speed, accuracy, and precision made AI tools and AI-based automated systems popular to healthcare workers and patients. Most healthcare professionals nowadays use some type of AI tool for all phases of patient diagnosis and management. Figure 8 shows the generic structure of AI algorithms.

2.1 Machine learning (ML)

ML is a subset of AI that enables a computer to learn from the given data. ML can be supervised, unsupervised, semi-supervised, and reinforcement learning [29] as shown in Fig. 9. Supervised learning is the learning process where the machine learns to generate outputs from inputs based on training input-output samples. Supervised learning processes use training output labels to learn

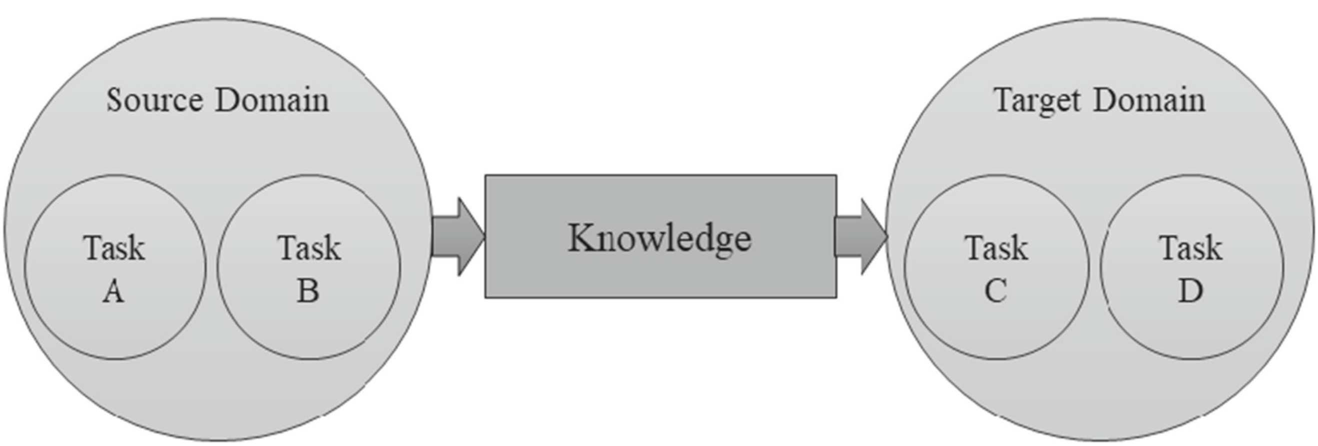

Fig. 12 Transfer learning

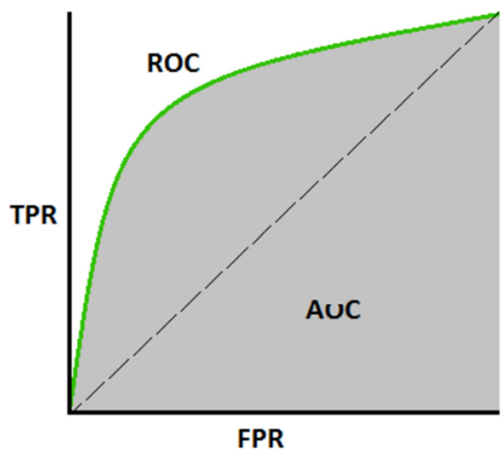

Fig. 13 ROC and AUC [42].

to generate labels for test data for classification and regression tasks. Support vector machine, naïve bayes, linear regression, decision tree, etc. are some efficient classification algorithms [30]. The machine may also learn by extracting features from unlabeled data in what is known as unsupervised ML. Clustering and association algorithms like *k*-means, DBSCAN are data-driven learning algorithms. Semi-supervised learning is a combination of supervised and unsupervised learning process that uses both labeled and unlabeled data to achieve better classification than unsupervised models by utilizing the available data labels in the training process. Reinforcement learning is an environment-driven learning process that learns by rewards and penalties on correct and incorrect predictions respectively.

2.2 Deep learning (DL)

DL is a subset of ML that is implemented with DNNs. The DNN is a special type of ANNs that have multiple hidden layers between the input and output layer as shown in Fig. 10 [31]. ANNs were inspired by the brain of living things and they were created to mimic the learning process of biological brain. As DNN is also an ANN, it also learns complex information by using the layers of neurons connected to each other. Each layer of a DNN transforms the input data or the output of the previous layer into more complex format with different types of feature values. By encoding and decoding the information between layers of the network, the DL models learn and produce outputs. Convolutional neural networks (CNNs), U-Net, GoogleNet, ResNet, DenseNet, and AlexNet are some examples of DNNs.

CNNs are multi-layer perceptron networks that are inspired by biological brain that are able to analyze and detect patterns in image pixels [32]. Due to the efficiency of CNNs in image feature extractions, most of the DL models for image analysis are generated with different combinations of CNNs and other supporting components. As CNN is a DNN, it is constructed with an input layer, an output layer, and multiple hidden layers between them. The layers will include convolution layers that generate feature maps from the input images to pass to the next layer. The convolution layers are normally followed by pooling layers that reduces the dimension of the convolution layers and the fully connected layers connect each neuron of one layer to every neuron of the other layer. The feature maps generated from the input images and passed along the layers while updating the feature maps by collecting more complex information finally produce the classification output for the input image (Fig. 11).

The image analysis DL models have a general format that includes (i) data pre-processing, (ii) data augmentation, (iii) CNN/DNN model, (iv) post-processing, and (v) final prediction/segmentation/classification. Most researches on image analysis for object detection and segmentation from various types of images (i.e., medical images and non-medical images like facial images and tongue images) use similar frameworks with CNN/DNN structures [33, 34]. The pre-processing and post-processing can be as simple as applying some thresholding or as complex as applying some DNNs. They also may include data normalization, model regularization [35], and noise removal. The data augmentation [36] is a popular step in medical image analysis with DNNs as the number of available medical image data is limited and DNNs perform better with large datasets. The data augmentation process uses different methods to modify the existing data to create new data for dataset extension. Various CNN/DNN models are then applied on them to generate outputs.

2.3 Transfer learning (TL)

TL uses the idea of knowledge transfer from one domain to other related domains [37]. Knowledge achieved from a known domain (i.e., source domain) can be utilized to reduce training time, necessity of large amount of labeled data while improving the performance of the model on an unknown but related domain (i.e., target domain). An example is shown in Fig. 12. Tasks A, B, C, and D are different types of tasks from related domains where A and B are source domain tasks and C and D are target domain tasks. The knowledge achieved from tasks A and B is transferred to solve tasks C and D; thus, TL is applied. Deep transfer learning (DTL) provides the scope for disease detection not only from medical images, but also from other non-invasive images like facial images for genetic disorder (i.e., thalassemia and Down syndrome) detection [38]. Although

TL does not always ensure better performance, it definitely generates a comparative alternative approach [39].

Various ML, DL, TL, and hybrid models have their own advantages and limitations. The general workflow of these models is similar. They take medical images as inputs, applies some pre-processing on them to enhance the image features, then train the model with (i.e., supervised) or without (i.e., unsupervised) data labels and ground truth images. The outputs generated by the models detect the class of the image for image detection task, and segment the infected region in segmentation task. The uniqueness of the ML and DL models is at the ANN/DNN structure. By varying the DNN structure, number of levels of neurons, the connections between the neurons, parameters, and weights of parameters, these models can extract different types of implicit and explicit features from the images based on their pixels, textures, intensities, and semantic measurements. Hence, the outputs of the models vary as the outputs depend on the features used to train the system. The TL model outputs may vary based on the similarity and dissimilarity between the source and target domains. Inferring a general decision on the best model is difficult as the performance of the models varies depending on the data, extracted features, model hyperparameters, and the overall structure of the model.

3 Performance metrics

Most approaches for COVID-19 detection from chest image inputs have used some popular classification and segmentation problem-based performance metrics. The evaluations of the proposed or existing systems were mainly analyzed by classification accuracy, loss, true negative rate (TNR) or specificity, positive predictive value (PPV) or precision, true positive rate (TPR) or recall or sensitivity, false positive rate (FPR), F1-score, receiver operating characteristic (ROC), area under the ROC curve (AUC) and Matthews correlation coefficient (MCC), Dice coefficient/score, Jaccard coefficient/index [40, 41].

The values for true positive (TP) and true negative (TN) represent the number of correct predictions of the "Positive" class and "Negative" class respectively. False positive (FP) and false negative (FN) represent the number of incorrect predictions of "Positive" class and "Negative" class respectively. The accuracy value represents the percentage or amount of correctly classified data and loss shows the difference between the prediction and ground truth. The equation for accuracy calculation using the confusion matrix is mentioned in Eq. 1.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

The PPV score shows the ratio of correctly predicted positive values with respect to the total predicted positive values. TPR represents the ratio of correctly predicted positive values to all actual positive values and TNR represents the ratio of correctly predicted negative values to all actual negative values. FPR shows the ratio of predicted values to all actual negative values. F_1-score represents the harmonic mean of PPV and TPR values. The calculations for these parameters are drawn in Eqs. 2, 3, 4, 5, and 6.

$$\text{PPV or Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{TPR or Sensitivity or Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{4}$$

$$\text{TNR or Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{5}$$

$$\text{F}_1 \text{ score} = \frac{2 \,.\, \text{PPV} \,.\, \text{TPR}}{\text{PPV} + \text{TPR}} \tag{6}$$

The ROC curve is a probability curve that plots TPR against false positive rate (FPR) at different decision thresholds. AUC is a summary of the ROC curve and it represents the capability or separability of the classifier to differentiate between classes. The higher the AUC, the better the classifier is at differentiating the elements of different classes. Figure 13 shows a sample of a typical ROC and AUC.

The MCC score represents the quality of classification for a classifier. If the classifier generates 100% correct classifications, then the values of FP and FN are 0s. In that case, MCC score is 1 showing the perfect quality of the classifier. It can be calculated by Eq. 7 as follows.

$$\text{MCC} = \frac{(\text{TP} \,.\, \text{TN}) - (\text{FP} \,.\, \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \tag{7}$$

For Dice score and Jaccard index calculation, let A and B be the output image and the ground truth. Dice score measures the similarity between A and B with Eqs. 8 and 9. Jaccard index calculates the similarity between A and B by incorporating the triangle inequality with Eqs. 10 and 11.

$$\text{Dice score}(A, B) = \frac{2|A \cap B|}{|A| + |B|} \tag{8}$$

$$\text{Dice score}(A, B) = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} = \text{F}_1 \text{ score} \tag{9}$$

$$\text{Jaccard index}(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{10}$$

$$\text{Jaccard index}(A, B) = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} = \text{IoU} \tag{11}$$

4 Survey/review papers

Researchers are trying to find good solutions for detecting, diagnosing, and monitoring the disease as well as developing drugs and vaccines for COVID-19. Although the available knowledge, resources, and systems are very limited with respect to curing the disease completely, researchers from various fields are approaching the problem of finding solutions as soon as possible. The first step for solving any problem is to gather as much information as possible on the problem and information about existing approaches or solutions to the problem. Therefore, it is essential to review existing works in order to understand the problem, provide possible solutions, and review challenges and scopes of research. Some researchers have been working on reviewing or summarizing the problem origin, applied methods, their advantages and limitations, available datasets, available tools, and applications of COVID-19 and presented them in survey/review papers. In this section, a few existing survey papers on COVID-19 approaches using image data were reviewed as follows.

Alafif et al. [43] recently summarized the existing approaches used for COVID-19 detection, segmentation, diagnosis, challenges, and future research directions in a comprehensive survey paper. They discussed COVID-19 from various perspectives with invasive and non-invasive systems, available tools, drugs and vaccine including a review of the origin of COVID-19, its severity, and current situation. A comparative discussion on manual vs. AI-based COVID analysis and represented the advantages of AI tools due to their accuracy and instant outputs was included. Chest CT and X-rays were discussed with current ML, DL, and hybrid methods using these images for their processing. Some available AI tools and models like CovidAID [44], Clara COVID-19 [45], COVNet [46], and FluSense [47] were also mentioned with their applications for COVID detection, COVID severity classification, COVID ROI segmentation, and patient monitoring. They listed some recent works with their methods, datasets, and performance metrics and most of them used DNNs as Truncated Inception Net, Darknet-19, U-Net, V-Net, VNet-IR-RPN, VB-Net, U-Net++, SqueezeNet, ResNet, DenseNet, etc., TL, and some used popular ML models like SVM and RF. COVID-19 analysis with audio inputs (respiratory sounds like cough, breathing, and voice) was also mentioned by the authors with similar listing. The drug and vaccine development procedure for COVID-19 was explained and AI-based drug development systems that researchers have been working on were included. Finally, the review was concluded with the challenges and scopes of COVID-19 detection, diagnosis, and analysis. The review showed the advantages of using DL models for COVID-19 analysis that were able to produce 90 to 100% accuracy, sensitivity, specificity, and AUC and performed better than radiologists in some cases. It has therefore been shown that AI-based tools and devices can provide rapid outcomes assisting patient diagnosis significantly and thus help healthcare professionals in their fight against COVID-19.

Another recent survey on the application of AI for COVID-19 [48] included the basics of COVID and general overview of AI ML models such as RF, SVM, LR, and XGBoost and DL models such as ANN, CNN, RNN, and LSTM. Ten CNN-based DL models used in recent COVID classification from X-rays and CT scans were summarized and their performances were compared using F-1 score and AUROC with accuracy values of more than 90%. They included the summaries of 7 approaches on COVID-19 severity detection (i.e., moderate, severe, and critical) with both ML and DL methods. The performances were compared with various combinations of CT, blood, and lab test data. Similar information was included on 10 COVID mortality risk assessment methods based on demographic, lab tests, medication, and other data. The limitations of these models included lack of datasets, datasets with limited number of data, missing features in datasets, their applicability to real-time systems and lack of proper performance metrics like FPR/TPR, and lack of discussion. A total of 13 drug repurposing-based methods were also discussed with AI applications, features, and limitation. Finally, the general challenges of AI-based models like generalization, dataset issues, possibilities of variations, and their relevance to the COVID-19 issues were analyzed to provide some guidance for future research efforts.

Nayak et al. [49] also reviewed recent works on ML- and DL-based models for COVID-19 detection and analysis. They collected 795 papers on COVID-19 after applying some relevant keywords and then filtered out 672 deemed not relevant. Finally, 123 papers (36 on ML, 64 on DL, 23 on others) were reviewed for their study. They started with a detail explanation of viruses that had caused severe disease over the last century and then they discussed the coronaviruses and their effects including COVID-19 on human and other animals. They described the statistics on death, recovery, and active cases in different countries; the reason for the spreading of the disease; the advantages of ML models for disease detection, data illustration, analysis, and prediction; and the limitations of available ML methods. Then they discussed the applications of DL models to overcome the challenges of ML models and to provide an easier and more effective way for disease detection, medical image analysis, disease diagnosis, drug and vaccine design, etc. The papers with their datasets, methods, inputs (i.e., text data, image data, and time-series data), and outcomes were summarized and compared. Different supervised and unsupervised ML models (i.e., SVM, RF, KNN, *k*-means, LR, SVR, DT, and ANN) and DL models (i.e., CNN, VGG, ResNet, Inception, LSTM, RNN, and DenseNet) were mentioned with their performances. The data used for COVID-19 research efforts, including clinical data, biomedical data, online data with their contents, applications, advantages, disadvantages, impacts of ML, DL for data and challenges with mention of some popular datasets were discussed. The authors included an extensive critical analysis on existing surveys, popular ML, DL models, publications, data, location, and impacts. Their analysis showed that although 52% of works were done on DL, whereas 29% approaches were based on ML and 19% on other models, only 11% publications of COVID-19 from December 2019 to September 2020 were on ML and DL (7% Dl and 4% ML). LR was the most popular ML model used for COVID-19 analysis and CNN for DL where ResNet, VGG, and Inception were the most popular CNN variations. Among various types of data, 60% COVID-19 research works were on image data (i.e., CT, X-ray, others), 31% were on text data, and the remaining publications were on other data. They also ranked existing works based on different strategies with respect to method, data, and performance. More than 50% of those publications were from Asia, whereas Europe and North America ranked second and third. The authors also included detail analysis on the authors, their countries, and contributions. Five other existing surveys were compared to their review and the impacts of COVID-19 on different sectors and concluded with a summary of their complete analysis.

Lalmuanawma et al. [50] reviewed the ML and AI applications for COVID-19 detection, prediction, contact tracing, and medicine approaches, listed the challenges and future directions for researchers of various fields involved with COVID-19 research in a similar manner. They listed 36 applications for contact tracing developed by different countries and compared 4 papers on screening, and 4 on prediction and forecasting based on methods, datasets, validations, and performance metrics for clinical, mamographic, and demographic datasets.

Mohammad-Rahimi et al. [51] reviewed 105 approaches to COVID-19 from 1827 papers based on some inclusion-exclusion criteria. The papers used various ML and DL models for analyzing CT or X-rays as inputs. ML algorithms like SVM, RF, DT, KNN, and NB and ensemble classifiers were able to achieve more than 90% accuracy in most cases with limited amount of data. ResNet, Inception, NASNetLarge, GoogleNet, CNN, AlexNet, VGG, SqueezeNet, Xception, and MobileNet were applied as DL methods for COVID-19 analysis and the performance metrics scores were close to 100% in many of those methods. Some hybrid models with combinations from ML, DL, and heuristics outperformed individual methods in many cases. They summarized all papers in a tabular comparison containing the data sources, structures, and sizes of the datasets, pre-processing, best model achieved in each research, the accuracy, sensitivity, specificity, and AUC of their models for comparison. Although CT images provide more detail than X-ray images, the review showed that X-rays images were also very efficient for COVID-19 detection and severity classification. One of the major issues of COVID-19 research efforts is limited and inconsistent data with missing annotations. Preparing complete and reliable COVID-19 image datasets was mentioned as necessary future work. Due to the limitation of data, ML models performed very similar as DL models, and also had a higher learning speed that DL methods. Thresholding, morphological operations, histogram equalization, and data augmentation were the most popular pre-processing steps. They also discussed possible future directions with TL as it showed promising performance for COVID-19 approaches.

A recent brief review on AI models for COVID-19 detection, classification, and segmentation from medical images listed 12 infection segmentation researches and 23 AI model researches [52]. Medical imaging and details of CT scan procedure with respiratory disease indications were discussed with respect to COVID patients. The image segmentation framework was explained with 12 recent lung and infection segmentation application that mostly used U-Net, U-Net++, and Inf-Net. The U-Net and Inf-Net architectures were described step by step to explain the segmentation process. Afterwards, 23 papers on COVID classification and segmentation from CXRs and CTs with YOLO, VGG, U-Net, U-Net++, basic DNN, ResNet, GoogleNet, etc. were discussed and their results were compared. Finally, the contributions of image analysis for COVID screening and severity assessment were summarized to conclude the survey.

Another recent survey by Subramanian et al. [53] included a brief summary of recent DL models used in COVID classification and detection from CXR and CT images. They added an elaborated background on CNN and other basic DNN models like VGG, ResNet, MobileNet, Xception, and DenseNet. A brief summary of the existing TL-based researches and novel researches was explained with their architectures and performances. Seven public COVID image datasets used for classification were discussed after a detailed analysis on COVID medical image datasets and their characteristics. The accuracy, precision, sensitivity, specificity, NPV, and F1-scores were defined and the performance of 11 researches with these metrics was compared to discuss the top most models with high efficiencies.

A review on 30 research papers until August 2020 on COVID-19 detection and severity classification with performance analysis was discussed in [54]. The authors discussed the dataset details for each paper, the models used, and their performance metrics scores. ML, DL, and hybrid models were used for feature extraction, classification for COVID/normal/non-COVID cases, classification of different level of severity of COVID, and COVID area segmentation. ML models like LightGBM, SVM, RF, DT, XGBoost, AdaBoost, Bagging, and DL models like Inception, InceptionResNet, IRRCNN, ShuffleNet, ResNet, MobileNet, DenseNet, Xception, attention, U-Net, U-Net++, CNN, and its variations were discussed with image data. These methods showed high performance scores (i.e., higher than 80% in most cases) and up to more than 99% accuracy and AUC for COVID-19 detection and classification tasks. They also showed the comparisons of the datasets, methods, sensitivity, specificity, precision, accuracy, AUC, F1-score of 5 papers on COVID/normal image classification, 13 on COVID/non-COVID classification, 11 on COVID/non-COVID pneumonia classification, and 3 on COVID severity classification. They concluded with the observations that DenseNet, ResNet, and ShuffleNet performed well for classifications, U-Net++ performed better for segmentation tasks, and TL performed well for all tasks.

Alghamdi et al. [55] reviewed 34 DL-based COVID-19 approaches on CXRs from March 2020 to May 2020. As X-rays are more available, less time consuming, and provide precise imaging of organs, using CXRs for COVID-19 diagnosis is very popular. They described the COVID-19 classification task as a multi-class problem that was addressed differently from existing works in that the classes are called COVID-19, healthy, bacterial pneumonia, viral pneumonia, no finding, etc. They explored 13 existing datasets containing X-rays of COVID-19 and other lung diseases with detail descriptions, image specifications, and links. As 71% research papers from the review were on TL of CNN variations, they described TL and its relevance to COVID-19 approaches. Some popular CNN structures used in those approaches like AlexNet, GoogleNet, VGGNet, ResNet, Xception, SENet, DenseNet, MobileNet, ShuffleNet, CapsNet, and autoencoder were discussed with the papers, their contributions, methods, and performances. Twenty-five papers that used these CNN-based DL models and 14 that used novel architectures were summarized and compared according to their datasets, contributions, and performance scores (i.e., accuracy, precision, specificity, sensitivity, AUC, and F1-score). They also discussed the methods used for visualizing the classification decisions in the reviewed papers such as Grad-CAM, CAM, gradients, attribution maps, LRP, GSInquire, and guided backpropagation. They explained challenges such as class imbalance problem, classification uncertainty, COVID severity classification, and dataset quality. Some future directions with proper dataset generation, exhaustive feature extraction, exploring semi-supervised models, GANs and TL, etc. were discussed to guide further methods. Another similar survey on DL-based image analysis for COVID-19 was published recently [56]. Unlike the previous paper, they reviewed works on both chest CT scan and X-rays.

Bhattacharya et al. [57] focused on ML and DL methods for both COVID-19 analysis and other abnormality analysis from medical image data. They discussed COVID-19 in a detailed overview consisting of the origin, current situation, necessity of DL approaches, basics of ML (supervised, unsupervised, semi-supervised) methods, DL models (i.e., RNN, GAN, and CNN), and their applications. They explained medical image analysis with some recent works for abnormality detection, classification, localization, segmentation, and registration on various medical images. Outbreak prediction, tracking spread, diagnosis, treatment, and limitations of COVID-19 with ML and DL models were also discussed. They also listed the datasets, methods, performance metrics, and research challenges of 27 medical image analysis approaches and 16 COVID-based works. They included 3 use cases of DL-based COVID-19 detection and monitoring from China, Canada, and South Korea. Finally, they discussed the challenges and future research directions for the pandemic with AI tools. A similar but more general review was done in [58] for analyzing the applications of AI regarding COVID-19. They provided an elaborate overview of AI applications in COVID-19 research papers regarding tasks like COVID-19 detection, severity classification, prediction, patient data management, patient management, resource management, CT image analysis with ML, DL, TL, and other AI models, analysis of biology-based data, drug discovery, and social behavior monitoring and management. They also discussed and compared the tasks and performances of 7 research works that used AI-based techniques (mainly DL models, some ML, and other models) for COVID-19 analysis from chest CT images and showed that AI was able to achieve 97% accuracy.

A detailed survey on imaging techniques for COVID-19 (i.e., CT and X-ray), segmentation of region of interests (ROIs) from those images, detection of COVID-19, classification, severity detection, diagnosis, monitoring, and COVID-19 image datasets was provided in [59]. The authors reviewed the image-based COVID methods on various ML, DL, and other AI techniques published until 31 March 2020. The workflow for conventional imaging, AI-based imaging, and COVID-19-based-specific AI applications for imaging with proper examples and explanations was included. They discussed basic image segmentation techniques and COVID-19 image segmentation techniques focused on U-Net and U-Net++ models for extracting lung regions and lesions of lungs. Summaries of 11 COVID-19 segmentation approaches with their data modality, subject details, methods, applications, ROI (i.e., lungs or lesions), and highlights of their contributions were provided. They also discussed AI-based techniques for detection, classification, severity measurements, and follow-ups of COVID-19 from images. They summarized 14 approaches on these with their data types, dataset details, tasks, models (i.e., U-Nets, CNNs, and RFs), and performances. Details on 4 COVID-19 image datasets and discussed the challenges and future research scopes for COVID-19 detection, segmentation, and diagnosis were listed.

Another recent survey on COVID-19 approaches that used ML and DL techniques on image and text datasets for diagnosis, prediction, and forecasting and providing a detailed overview of pandemics, COVID-19, existing systems, datasets, challenges, and possible future directions was given in [60]. The authors listed and discussed all pandemics from plague to COVID-19 with timelines and death statistics. Details on COVID-19 origin, transmission, cases, deaths, and their statistics in various countries and timelines were provided. They summarized 5 works on RF, 5 on SVM, and 9 on other ML models like linear regression, *k*-means, XGBoost, and Gabor with problem statements, type of data, used model, publication time, and performance scores. Similar summaries were included for 14 papers on CNN, 8 on LSTM, and 7 on other DL models like GAN and autoencoders. Most of them achieved more than 95% accuracy in their tasks. They also included an overview on the mathematical and statistical models used for the pandemic losses and prediction analysis with 9 papers. A performance analysis using the evaluation scores of 5 ML-based approaches and 13 DL-based approaches was provided. They also included detail statistics on publications on COVID-19, type of publications, countries, journals, type of data, and prediction-classification-forecasting distribution. They concluded with a detail discussion on the challenges of COVID-19 like data limitation, prediction accuracy, lack of advanced tools and applications, and absence of customized systems appropriate for developing countries.

A systematic review of 11 COVID-19 research works from the 36 that were identified as published until the middle of May 2020 was included in [61] after an extensive searching and screening process. The publications were based on various combinations of classification tasks as binary/multi-class/hierarchical classes. The authors also mentioned statistics of the databases and countries of publications for the reviewed papers. Binary, multi-class and hierarchical classes were described in detail with relevance to the existing works and a critical analysis was done on those papers based on dataset types, methods, and case studies of X-rays and CT scans. The challenges of evaluation, trade-offs, and importance of criteria were discussed in detail. An important contribution of this survey was a complete and step by step research proposal using multi-criteria decision analysis based on the reviewed works. A complete methodology starting from identification, data collection, pre-processing, evaluations, methods, and validations was elaborately discussed.

Sufian et al. [62] followed a similar idea in their systematic review on COVID-19 approaches that used DL, DTL, and edge computing (EC) for lung image COVID-19 detection and classification. They discussed each of these methods in detail and with relevance to COVID-19 approaches. Tasks and contributions of 15 papers on DL models, 9 on DTL models, and 7 on EC models for COVID-19 approaches were summarized. They also included the image and textual datasets used in COVID-19 papers and listed 9 datasets and 7 data sources (images and textual) with brief descriptions for each. The current and future challenges that were not only relevant to COVID-19 detection or classification, but also the human resources and other logistic issues were analyzed. They finally proposed a complete framework combining DL, DTL, and EC for COVID-19 detection, diagnosis, hospital management, and social distancing management as a possible future direction for COVID-19 systems.

A survey on COVID-19 based on mathematical models, AI models, and datasets was provided in [63]. The authors reviewed 61 components (19 on mathematical models, 18 on AI, 24 on datasets) published from December 2019 to April 2020 on COVID including journal articles, websites, reports, and fact sheets. They discussed mathematical models used for COVID-19 spread and dynamics analysis with models like susceptible-exposed-infected-removed (SEIR) and susceptible-lovered (SIR) and listed the tasks, models used in the reviewed mathematical model-based approaches. For the AI models mostly used for image data, they provided detailed discussions on detection and classification from CXR and CT data. ML and DL models like CNN, NB, LDA, SVM, RF, DT, and LR providing explanations according to the corresponding research items. Methods, datasets, data details, accuracy, sensitivity, and PPV were compared for the AI-based works and showed more than 99% accuracy. They included descriptions and sources of 24 COVID-19 datasets combining image, text, and value data. Finally, they discussed the advantages and limitations of both mathematical and AI models and proposed some possible future directions.

Several researchers have worked on reviews of datasets used in various COVID-19 analyses. A comprehensive study of the publicly available COVID-19 datasets, their applications, and possible future research directions was summarized in [64]. The image datasets included chest CT scans and X-rays, whereas the textual data included case reports on various factors of COVID-19, social network data relevant to COVID-19, and scholarly article information. The authors discussed dataset details like

Table 1 Survey/review papers

Ref.	Topic	Approach	Features	Contributions
[43]	Detection Segmentation Diagnosis AI tool Drugs & Vaccines	ML DL Hybrid	CT X-ray Sounds	Listed AI-tools for COVID-19 detection, segmentation, and patient monitoring. Mentioned DL models like U-Net, U-Net++, V-Net, and VB-Net and ML models like SVM, RF, LR, and MLP used for AI-based COVID detection. Included summaries of existing ML, DL models with datasets, performance metrics methods. Listed the challenges related to data availability, quality, privacy and other similar issues, lack of collaborations between medical and technical professionals, lack of resources, etc. The future directions included various possible approaches from different stages of the government, people, medical, and technology professionals to overcome this pandemic as soon as possible.
[48]	Detection Classification	ML DL Hybrid	CT X-ray Blood Lab tests Demographic Medication Symptoms Comorbidity	Described general overview of various popular ML and DL models. Summarized and compared 10 recent COVID classification models using various CNN-based DL with limitations. Summarized 7 ML and DL models for COVID severity detection with limitations. Summarized 10 ML and DL models for COVID mortality risk assessment with their limitations. Summarized 13 ML and DL models for drug repurposing and epidemic trend detection with limitations. Listed challenges of AI-based models and COVID-19-based approaches with possible future directions.
[49]	Detection Prediction Classification Forecasting Screening & Assessment Diagnosis Others	ML DL Others	CT X-ray Lab tests Demographic Statistics Others	Reviewed 123 papers (36 on ML, 64 on DL, 23 on others) on COVID. Discussed viruses, diseases, corona, COVID-19, COVID-19 statistics, ML and DL in disease detection, prediction, analysis, etc. in detail. Discussed the clinical, biomedical, and online data with their advantages, challenge impacts on COVID-19 research with ML and DL models. Included critical analysis for data, methods, publications, research types, most popular ML and DL models, authors, countries, etc. Compared their ML and DL survey with 5 other ML survey papers. Listed and analyzed the sectors affected by COVID-19.
[50]	Screening Contact tracing Prediction Forecasting Drugs & vaccines	ML DL	Clinical Mamographic Demographic	Reviewed research papers on COVID-19 screening, detection, prediction, contact tracing, drugs and vaccines using basic AI methods containing ML and DL models. Listed 36 applications for contact tracing developed by different countries. Compared 4 papers on screening, and 4 on prediction and forecasting based on methods, datasets, validations, and performance metrics.
[51]	Diagnosis Detection Prediction	ML DL	CT X-ray	Reviewed 105 COVID-19 approaches based on some inclusion and exclusion criteria. Discussed ML, DL, and hybrid models with their data sources, dataset details, models, and performance metrics. Listed limitations and possible future research directions relevant to proper and complete datasets, approaches on X-ray-based models and TL.
[52]	Detection Classification Segmentation Screening Assessment	ML DL	CT X-ray	Reviewed latest applications on COVID detection and segmentation from chest imaging. Discussed CT characteristics of different respiratory issues. Highlighted the details of 12 recent COVID infection segmentation researches on U-Net, U-Net++, and Inf-Net. Discussed details of U-Net and Inf-Net structures foe segmentations. Reviewed 23 AI-based researches for COVID-19 image analysis. Discussed COVID image screening, severity assessment, and public datasets briefly.
[53]	Detection Classification Datasets	DL	CT X-ray	Discussed basic CNN and other DL model architectures in details while mentioning the recent COVID classification researches. Explained TL and novel approaches for COVID classification with brief summaries of existing researches. Compared performances of 11 recent researches that used DL models. Discussed public COVID image datasets and mentioned details of 7 of them.
[54]	Detection Classification Segmentation	ML DL	CT X-ray	Reviewed 30 papers on COVID-19 detection and severity classification using image data. Discussed the papers in detail starting from datasets, subjects of the datasets, pre-processing, feature extraction, classification and segmentation methods, and performances scores. Compared datasets, methods, and performance metrics of 5 papers on COVID/normal image classification, 13 on COVID/non-COVID classification, 11 COVID/non-COVID pneumonia classification, and 3 on COVID severity classification. Observed that DenseNet, ResNet, and ShuffleNet performed well for classification and U-Net++ performed better for segmentation.

Table 2 Survey/review papers

Ref.	Topic	Approach	Features	Contributions
[55]	Detection Classification	DL	X-ray	Reviewed 34 COVID-19 approaches on CXRs between March and May 2020 that used DL models. Discussed the classification task for COVID-19 to define the X-rays as COVID-19, normal, pneumonia, and some other classes. Summarized 13 datasets used in relevant papers with number of images, resolutions, formats, descriptions, COVID existence, and links. Discussed popular CNN-based DL models used in COVID-19 research. Summarized 25 papers that used these CNN-based DL models and 14 that used novel architectures including datasets, contributions, and performance scores like accuracy, precision, specificity, sensitivity, AUC, and F1-score. Explained the methods used for visualization of classification decisions. Included challenges and future directions for COVID-19 approaches.
[57]	Detection Classification Segmentation Localization Registration Contact tracing Drugs & vaccines	ML DL	Medical images COVID images (CT, X-ray)	Provided detail on COVID-19, severity, basics of ML and DL models. Explained ML and DL basics, their applications, medical image analysis fundamentals. Discussed recent works on detection, classification, localization, segmentation, and registration of medical image analysis with ML and DL models. Summarized recent works on DL models for COVID-19 detection, tracking disease, diagnosis, treatment, drug, vaccine, and their limitations. Listed datasets, methods, evaluation metrics, and research challenges of 16 COVID-based papers and 27 medical image analysis papers. Included challenges of existing approaches and datasets. Included 3 use cases of DL based COVID-19 detection and monitoring from China, Canada, and South Korea. Discussed challenges and future research directions for COVID-19 using AI tools.
[58]	Detection Diagnosis Tracking Prediction Drugs Application	ML DL TL Others	CT	Provided detail analysis and overview of AI-based applications for COVID-19 detection, severity classification, prediction, patient management, resource management, various models (ML, DL, TL, hybrids, etc.) for COVID-19 analysis from CT images, performance analysis, biology-based analysis of COVID, drug discovery, social behavior monitoring and management. Compared the methods and performances of 7 AI models that used CT as inputs.
[59]	Imaging Segmentation Detection Classification Diagnosis Monitoring	ML DL Others	CT X-ray	Reviewed image-based COVID-19 approaches until 31 March 2020. Provided detail on the workflows of conventional, AI-based imaging, and AI applications used for scanning patients for COVID-19. Discussed segmentation basics and methods for segmenting lung region and lesions (focused on U-Nets and U-Net++). Summarized 11 approaches on COVID-19 ROI segmentation from lung images with data type, models, applications, target ROIs, and observations. Discussed X-ray and CT-based techniques including classification, severity assessment, follow-ups and summarized 14 approaches with datasets, models, tasks, and results. Included 4 COVID-19 CT and X-ray datasets, research challenges, and future scopes.
[60]	Classification Prediction Forecasting	ML DL	Image Textual	Provided list of pandemics since plague with time period and death statistics. Provided an elaborate analysis on the origin of COVID-19, its transmission, statistics related to case and deaths in different countries, etc. Compared approaches, datasets, and tasks of recent approaches including 19 ML methods (5 on RF, 5 on SVM, and 9 others), 29 DL methods (14 on CNN, 8 on LSTM, 7 others), and 9 statistical and mathematical models. Included performance analysis based on evaluation scores of 5 ML and 13 DL approaches. Provided statistics on publications, type of publications, countries, journals, type of data, and prediction-classification-forecasting distribution. Discussed challenges due to data limitation, prediction accuracy, lack of more advanced tools, and absence of systems appropriate for developing countries.
[61]	Detection Classification Proposed framework	ML DL	CT X-ray	Included a systematic review on 11 COVID-19 approaches that used AI techniques. Discussed COVID-19 classifications as a problem of various combination of classification model — binary, multi-class, hierarchical and explained each in detail. Mentioned statistics of databases and countries of publications. Listed types of datasets, methods, and case studies used in each paper. Included critical analysis and challenges of COVID-19 detection with AI technologies from medical images regarding evaluation, trade-offs, and importance of criteria. Defined all evaluation metrics used in existing approaches. Provided a complete methodology with identification, data collection, pre-processing, evaluations, methods, and validations.

Table 3 Survey/review papers

Ref.	Topic	Approach	Features	Contributions
[62]	Detection Classification Proposed framework	DL DTL Edge Computing	Image Textual	Discussed DL, DTL, and EC models in detail with relevance to COVID-19 approaches. Reviewed existing works on these methods and on COVID-19 related to these methods. Compared the tasks and contributions of 15 papers on DL models, 9 on DTL models, and 7 on EC models for COVID-19. Listed 9 datasets and 7 data sources (images and textual) on COVID-19 and included brief descriptions for each of them. Discussed the challenges of COVID-19-based approaches and proposed a complete framework combining DL, DTL, and EC for COVID-19 detection, diagnosis, hospital and social distancing management.
[63]	Detection Classification Propagation Management Applications Methods Datasets	Mathematical AI (ML,DL, Others)	Image Textual Others	Reviewed 61 approaches including journal articles, reports, websites published from December 2019 to April 2020. Discussed mathematical and AI models for COVID-19 detection, classification, and propagation and compared the methods used in existing works from each category based on tasks, performances, and models.Listed 24 datasets including COVID-19 images (CT, X-ray), text (i.e., demographic and social media), values (i.e., economic and climate), and other types with data type, size, and sources. Discussed advantages, limitations, and possible future research directions for both mathematical and AI models.
[64]	Datasets Applications Methods	ML DL Statistical	Image Textual	Discussed about image (CT, X-ray) and text (case reports, social media posts, scholarly articles) datasets for COVID-19. Described ML, DL, and statistical models that used those datasets and achieved good performance for COVID-19 detection, segmentation, classification, etc. Included comparative lists of 19 image resources and 26 text resources for COVID-19. Discussed challenges and future works for COVID-19 datasets and approaches.
[66]	Datasets Applications Data science Methods Challenges	ML DL Hybrid Evolutionary Others	Image Textual Biomedical	Explored image, textual, biomedical, voice datasets, competition datasets, datasets from developing countries, community-based datasets, other research works, and other statistical datasets on COVID-19. Discussed data science issues relevant to patients and diagnosis. Discussed simulations and models from clinical, social, economical, logistical, and various other perspectives. Reviewed recent works based on datasets, techniques, methodologies for 26 papers on image data, 6 papers on text data, and 7 papers on pharmaceutical data. Included statistics and challenges related to COVID-19 diagnosis.

data collection process, location, number of data, type of data, data annotation, and data availability. They also included descriptions of methods used on those data for COVID-19 detection, segmentation, classification, etc. with their performances. They discussed the textual data containing demographic, economic, transmission, mobility, social media emotion, conversation, and scholarly article analysis. The application, method, data type, and links for 19 resources for images and 26 resources for texts including some Kaggle and Github sources were compared. The challenges regarding medical image and social media data, privacy issues, and authentication with some future research directions were suggested for COVID-19 analysis.

Another similar research on review of clustering algorithms for COVID-19 datasets was presented in [65]. Various textual datasets containing social network data, demographic data, and case report data were discussed that used clustering algorithms as *k*-means, FCM, DBScan, hierarchical clustering, etc. with objectives, datasets, methods, and results of their analysis. Latif et al. [66] prepared a survey paper based on different types of datasets, applications of data science, reviews of current approaches, and challenges of COVID-19. They explored image, textual, biomedical, voice datasets, competition datasets, datasets from developing countries, community-based datasets, other research works, and other statistical datasets on COVID-19. The data science aspects of COVID data such as risk assessments, classifying patients according to different priority criteria, proper screening of each patient, and diagnosis based on all variables were reviewed. They also evaluated various simulations and models used for COVID-19 diagnosis, contact tracing, imposing social interventions by applying social distancing, managing rumor propagation, planning resources and logistics accordingly, improving patient care, managing innovations of vaccines and other new medicines and treatments, controlling economic interventions, and applying all of those in different ways for developed and developing countries. The methods, datasets, and techniques (i.e., ML, DL, and hybrid) of 26 recent approaches on image data, 6 on text data, and 7 on pharmaceutical data were summarized. They included some statistics based on COVID-based publications, research topics, and other pandemics and listed the challenges of COVID-19 approaches like data limitation, time constraint, authentication, and security, necessity for multidimensional approaches and data and more concrete approaches for developing countries.

Various researchers reviewed COVID image analysis from different perspectives. Some survey papers focused on the general framework of the image analysis tasks and mentioned existing works for each step of those tasks. Some reviews were focused on the papers on COVID image analysis and tried to provide brief summaries of recent works in the field. Some researchers organized their literature review to combine all known information on AI-based COVID analysis. In this section, a brief idea of the existing literature is provided with the major tasks they performed, the general approaches they used, and their summarized contributions. The goal of this section is to provide an overall idea of each discussed review paper to relevant researchers so that they have some indications of combined information resources on these tasks available currently. Tables 1, 2, and 3 show lists of summaries of the contributions mentioned in the review/survey papers on COVID-19 analysis.

5 AI model-based papers

COVID-19 approaches on medical image datasets mostly used DL and TL models. Some methods included popular ML techniques or hybrid models. As CNN models perform better in medical image analysis than other techniques, various popular CNN models were used in most of the recent works. Although the necessity and urgency of a complete system for COVID-19 screening encouraged the researchers to use pre-trained CNNs on COVID-19 datasets for rapid output generation, some researchers focused on proposing novel frameworks with ML, DL algorithms and achieved promising outputs. Most recent works chose the detection and classification task, but few papers also provided methods for infected region segmentation and visualization. The detection and classification tasks were used on medical images like CXR, CT, and ultrasounds as binary (i.e., COVID-19 or normal), 3 classes (i.e., COVID-19, normal, pneumonia), and 4 classes (i.e., COVID-19, normal, viral pneumonia, bacterial pneumonia) tasks. The visualization task-based works mostly highlighted infected lung regions. The segmentation tasks segmented the lung regions and the infected parts of the lungs. Some recent AI-based approaches for COVID-19 detection, classification, segmentation, and visualization are summarized below.

Chandra et al. [67] recently proposed a two phase COVID-19 screening system for classifying X-ray images into COVID-19-infected, COVID-19-suspected, and normal images. After pre-processing (i.e., resizing, denoising, and normalizing) the images, they used 4 data transformation techniques (i.e., sharpening, Gaussian blur, brightness modification, and contrast modification) for data augmentation. Then a binary grey wolf optimization (BGWO) algorithm was used to select the best features from 8196 features (i.e., 8 first-order statistical features (FOSF), 88 grey level co-occurrence matrix (GLCM), and 8100 histogram of oriented gradients (HOG) features). Decision tree (DT), support vector machine (SVM) with 3 different kernel functions in 3 separate models, *k*-nearest neighbor (KNN), naïve bayes (NB), and ANN were trained in two phases (phase I — normal/abnormal, phase II — normal/abnormal COVID-19/abnormal pneumonia). Finally, majority voting was used to generate the final output. The proposed method achieved 98.06% accuracy in phase I and 93.41% in phase II and the results were comparable to 5 and 6 similar COVID-19 approaches for 2 and 3 class problems respectively.

Another similar ML-based COVID-19 detection and classification research was proposed recently in [68]. They explained the origin, transmission, statistics, and other basics of COVID in detail. Then they proposed a Ml approach for COVID-19 X-rays to classify them into COVID-19 and non-COVID-19 classes that outperformed the accuracy of DL models. After some basic pre-processing were done on the input image, the luminance value for each pixel was determined by the luma transform to extract 10,000 features from each image. Then the hybrid social group optimization (HSGO) algorithm was used to select the best set of features (i.e., 116 features) for each image as it achieved the highest accuracy with lowest number of features compared to SGO, principal component analysis (PCA), and kernel PCA (KPCA). The feature sets were applied as inputs to 5 ML classifiers — KNN, DT, random forest (RF), support vector classifier (SVC), and linear SVC (L-SVC). The SVC classifier achieved the highest accuracy of 99.65% outperforming 12 other DL and bio-inspired algorithms for COVID-19 classification.

A novel idea called one-shot learning was proposed and implemented in [69]. To use all necessary information from a limited amount of data of a multi-class problem, the model extracted the best samples from each class by ranking them based on their discrimination ability among classes. Then samples were chosen randomly from the best sample sets to create clusters for the one-shot learning of ensemble classifier. They proposed an ensemble model combining generalized regression neural network (GRNN) and probabilistic neural network (PNN) classifiers. After extracting raw features by the one-shot learning model, GRNN and PNN were individually applied for best performing samples that were used for clustering. The enhanced features were then used for the ensemble model of GRNN and PNN to produce final classification outputs. They used 2, 3, and 4 classes for COVID, normal, viral pneumonia, and bacterial pneumonia and achieved 100%, 80%, and 65% accuracies on average respectively. Their experiments on COVID-19 X-ray dataset showed performances comparable to the popular DNNs AlexNet, GoogleNet, and ResNet18.

A 23-layer CNN was proposed in [70] for COVID-19 CT image classification with a comparative analysis of ML, DL, and texture analysis methods for the task. They included summaries of 7 studies from 2020 on COVID-19/non-COVID-19 classification from CT images with their datasets, methods, test models, and results. At the pre-processing stage of the proposed approach, they applied GLCM, local binary pattern (LBP), and local entropy (LE) on each original image and then used basic data augmentation on these and the original images to increase the dataset 5, 10, and 20 times. The augmented datasets for their proposed 23-layer CNN containing 1 input layer, 5 blocks of convolution, batch normalization, Relu and max-pooling, 1 fully connected layer, 1 softmax layer, and 1 output layer was then applied. They experimented on their proposed model with various combinations of data augmentations (i.e., 5, 10 , 20 times) and data combinations (i.e., original, GLCM, LBP, LE) with 2- and 10-fold cross-validation. They applied 2 DL models (AlexNet and MobileNetV2) and 2 ML models (SVM and KNN) with similar data augmentations and combinations to compare their model. They also compared the model performance with 7 existing COVID-19 CT image classification and achieved comparable performances.

Ismael et al. [71] proposed a novel COVID-19 detection model for CXR images using a hybrid model containing ML, DL, and TL methods. After resizing the input images, pre-trained ResNet18, ResNet50, ResNet101, VGG16, and VGG19 models were fine tuned and used for deep feature extractions from the images. A SVM model with linear, quadratic, cubic, and gaussian kernel functions was used for training on the extracted deep features. They also experimented with eight local texture descriptor features to compare their affects on the classification. ResNet50 features with linear kernel SVM achieved the highest accuracy with 94.7% for classifying the images into COVID-19 and normal categories in minimum time (i.e., 48.9 s).

Another similar approach combining ML, DL, and Sobel filtering for COVID-19 detection from X-ray images was proposed in [72]. They reviewed 25 existing approaches on COVID-19 detection from X-ray and CT images with DL models and proposed the scopes of their work. They then generated their own dataset containing verified annotations of 333 X-rays collected at Omid Hospital in Tehran from February 2020 to April 2020. A total of 256 of those images were normal and 77 were COVID positive. The Sobel filtering for detecting the edges from the images was then applied and then they were pre-processed for two CNN models — one with sigmoid output layer and another with SVM output layer that used 10-fold cross-validation. They showed that using Sobel filtering improved the performance of the network and achieved 99.02% accuracy. Their model was also tested with 6 public datasets and generated high accuracy. The proposed method outperformed most of the results of 11 recent DL-based COVID-19 detection approaches.

A novel pre-processing method with fuzzy and stacking techniques was used in a hybrid model in [73] combining ML and DL methods for classifying images into COVID-19, normal, or pneumonia classes. Each input image was reconstructed with fuzzy color algorithm for RGB values and the reconstructed images were added to the original data using stacking algorithm to reduce noise and improve image quality. The complete image set was then fed separately into 2 pre-trained DL models named SqueezeNet and MobileNetV2 for feature extraction. Then a social mimic optimization (SMO) algorithm was used on each feature set to select the best set of features that were then combined to create the feature set of the image. Finally, a multi-class SVM used these features to classify the images into COVID-19, normal, or pneumonia classes. Various combinations of the images and networks showed that the proposed model with fuzzy color and stacking algorithms achieved the highest accuracy of 99.27% compared to the model with features from the individual networks. The system was also computationally cost effective as it used fewer parameters and features.

Elkorany et al. [74] proposed a very similar hybrid of ML, DL, and TL models named COVIDetection-Net combining two popular pre-trained CNN models for deep feature extraction and a multi-class SVM (MSVM) for classification of images into COVID, normal, viral pneumonia, and bacterial pneumonia classes (2, 3, and 4 classes respectively) with those features. Pre-trained ShuffleNet and SqueezeNet were applied on the datasets separately to generate 544 and 1000 features respectively. All 1544 deep features were then fed to a MSVM for 2, 3, and 4 class classification. The comparisons between the results of COVIDetection-Net and individual Shufflenet and SqueezeNet showed that combining features from both improved the performances for all classes. A comparison between the proposed model and two popular COVID-19 detection models Coronet[75] and CovXNet[76] showed 2 to 7% improvements in COVIDetection-Net. They also compared their model with 18 other approaches on COVID-19 detection including 2, 3, or 4 class classifications proving COVIDetection-Net achieved the highest accuracy 100%, 99.72%, and 94.44% for 2, 3, and 4 classes respectively.

Bhattacharyya et al. [77] proposed a segmentation and classification system for COVID-19 X-ray images to classify them into normal, COVID, and pneumonia classes. A conditional GAN model was used for lung segmentation to minimize the ROI for classification. After segmenting the lungs from the X-rays, different keypoint extraction methods like SIFT, BRISK, and *k*-means were used for keypoint extract and DL models were used for deep feature extractions. The feature sets were then applied for the final classification. The proposed model was compared to similar 5 recent DL-based classification and outperformed all of them with 96.6% accuracy. The A Bayesian DL model was used in [78] for classifying CXRs into the same three classes — normal, COVID, pneumonia with 96% classification accuracy on similar datasets. They applied the proposed model on a combination of a COVID X-ray dataset and a pneumonia X-ray dataset. The combined dataset was divided into training, testing, and validation sets. A basic 5 layer CNN model was used to extract the feature set from the X-rays and then the Bayesian optimizer was used to tune the hyperparameters to achieve the highest performance for the CNN for the 3 class classifier.

A novel feature fusion model for COVID-19 detection, classification, and segmentation was proposed in [79] that combined 3780 histogram-oriented gradient (HOG) features and 4096 features extracted by a pre-trained CNN (VGG19) automatically. The best 1186 features selected by maximum entropy method were used for COVID-19 classification with another pre-trained VGG19 model. A modified anisotropic diffusion filtering (MADF) algorithm was used for noise removal from COVID-19 positive images. Then the effected lung regions were segmented with the watershed algorithm. The classification achieved 99.49% accuracy that outperformed 19 other DL-based COVID-19 detection models.

A TL-based model was proposed in [80] to overcome the limitations of existing COVID-19 detection models with the application of Haralick features for texture analysis. The input images were pre-processed for histogram equalization, Weiner filter, and ROI cropping. Then the Haralick texture features were extracted from the pre-processed images for classification. Pre-trained DL models Resnet50, VGG16, and InceptionV3 were applied on a combination of X-ray and CT images collected from various sources to classify them into 4 classes: normal, COVID-19, viral pneumonia, and bacterial pneumonia. They also provided visualization for the infected lung region using Grad-CAM. The proposed model was tested with individual and combined Haralick features to analyze their effects on the classification and was discussed in detail. The model outperformed 9 existing COVID-19 classification approaches using ML, DL, and TL models with 93% accuracy.

Wang et al. [81] proposed one of the pioneering state-of-the-art DL models for COVID-19 detection called COVID-Net based on a projection-expansion-projection-extension (PEPX) design with human-machine collaboration strategy. They also generated one of the largest COVID-19 X-ray benchmark datasets called COVIDx with 13,975 CXRs from 13,870 patients. They designed a two-step network that included one stage with 1 convolution layer, 16 PEPX blocks, 1 flatten layer, 1 fully connected layer, and a softmax output layer, and second stage with 4 convolution layers for long range connectivity. Each PEPX block contained 5 layers for projection, expansion, depth-wise convolution, second stage projection, and extension for incorporating all features accurately. The proposed model was compare to VGG-19 and ResNet-50 and outperformed them both by 3 to 10% while achieving 93.3% accuracy.

Another novel 22-layer CNN-based COVID detection model named CoroDet was proposed in [82]. CoroDet contained 9 2D convolutional layers, 9 max pooling layers, 1 flatten layer, 2 dense layers, and 1 layer with a leaky Relu activation function. CoroDet was able to classify X-ray and CT images into 2, 3, and 4 classes as COVID-10, normal, and non-COVID pneumonia (non-COVID viral pneumonia and non-COVID bacterial pneumonia). The model was tested on of the largest COVID image dataset COVID-R and achieved 99.1%, 94.2%, and 91.2% accuracy for 2, 3, and 4 class problems respectively. COVID-R was generated by collecting, combining, and modifying X-ray and CT images from eight public datasets and was one of the major contributions of their research. The proposed method was compared to 10 similar COVID-19 classification methods and outperformed all by at least 1% to at most 17%.

Narin et al. [83] proposed a TL CNN-based COVID-19 detection model for classifying CXRs into 4 classes as COVID-10, normal, non-COVID viral pneumonia, and non-COVID bacterial pneumonia with 3 binary classification models. They used 5 pre-trained CNNs: ResNet50, ResNet101, ResNet152, InceptionV3, and Inception-ResNetV2 using 5-fold cross-validations on 3 datasets. The highest accuracy of 96.1 to 99.7% was achieved with ResNet50. The comparison between their proposed method and 10 similar approaches on 2, 3, and 4 class COVID-19 detection showed that their model outperformed most similar ML and DL approaches for COVID-19 detection from X-ray images.

Another TL-based model was proposed in [84] using an EfficientNet pre-trained on ImageNet dataset to optimize the computational cost (i.e., time and memory) of the COVID-19 screening system. They used the EfficientNet with two structures: flat (classification without considering the relationship between classes) and hierarchical (considering the taxonomy of classes). Two different dataset divisions were used that combined 3 COVID-19 X-ray datasets to train and test on normal and cross-dataset environments. Both flat and hierarchical models achieved 93% accuracy with more than 30% higher computational efficiency and 100% COVID-19 positive prediction rate. These characteristics made the approach efficient for multiple platforms, devices, and applicable for developing countries.

Das et al. [85] proposed an ensemble DL-based approach for COVID-19 X-ray classification with DenseNet201, Resnet50V2, and Inceptionv3. They used 7 COVID X-ray datasets from different sources to train and test their model. A pre-processed dataset was used to train and test the models separately and the test images were used for loss minimization. The weights for each model were calculated by using 5-fold cross-validation on the test data. Then the ensembler used the average of the weights for 3 models to generate the final output label (i.e., COVID positive or COVID negative). They compared the outputs of each individual networks, their concatenation, and the proposed approach to show that the proposed model achieved the highest accuracy of 91.62% and outperformed the others by at least 1%. They also included a detail comparative analysis between their method and 8 other recent works according to dataset, evaluation method, accuracy, sensitivity, specificity, classifier, and summary of approaches.

Another DL-based ensemble model named EDL-COVID was proposed in [86]. They designed the model by combining snapshots of a pioneer open source deep CNN for COVID-19 called COVID-Net [81]. The authors used multiple snapshots of the same COVID-Net at the same training execution to reduce the computation cost of training and testing multiple different networks. To overcome the lack of diversity problem of snapshots from the same network, they applied cosine annealing learning rate schedule for aggressive learning rates. After generating 6 models from the COVID-Net, the ensembler applied calculations on the class probabilities to provide an average and generated the final classification results. They tested all individual models and the ensemble one on COVID-19 X-ray images and compared the results to show that the ensemble output performance was higher than individual models in most cases.

A similar DL ensemble model of 3 popular CNNs (AlexNet, GoogleNet, ResNet18) with the same name (i.e., EDL-COVID) was proposed in [87] combining CNNs, TL, and majority voting. They discussed the clinical values of lung CT images for COVID-19 detection and listed the summaries of 6 relevant approaches. Detailed definitions and frameworks of TL, ensemble models, AlexNet, GoogleNet, and ResNet were provided. A dataset containing normal, lung tumor, and COVID-19 CT images was pre-processed and fed into pre-trained AlexNet, GoogleNet, and ResNet18 separately with 5-fold cross-validation. An ensembler used the outputs from 3 classifiers and applied a relative majority voting to produce the final classification output. The experimental results showed more than 99% classification accuracy with an optimized detection time proved that the ensemble model performed better than individual DL models in terms of accuracy and detection time.

A comparative analysis between 3 CNN models for classifying X-rays into COVID-19, normal, or pneumonia classes was presented in [88]. They mentioned that the novelty of the work was proposing a Leaky Relu activation function instead of Relu for all 3 models. They implemented Inception Net V3, XCeption net, and ResNeXt models with Leaky Relu and tested them with Kaggle COVID-19 X-ray repository data. A detailed analysis on the comparisons of the models was included and they concluded that XCeption net performed best for COVID-19 X-ray classification. Nayak et al. [12] worked on a similar research by comparing 8 popular CNN models on X-rays for COVID-19 detection. They applied pre-trained AlexNet, VGG-16, GoogleNet, MobileNet-V2, SqueezeNet, ResNet-34, ResNet-50, and Inception-V3 with various batch sizes (i.e., 8, 16, 32), optimizers (i.e., Adam, SGD, RMSProp, Adadelta), learning rates, and epochs to compare and choose the best network with the best set of parameters. They concluded with the decision that ResNet-34 achieved the highest accuracy of 98.33 with a sensitivity score of 100 that outperformed 8 recent similar research works.

Degerli et al. [89] not only proposed a DL-based COVID-19 detection and segmentation model, but also provided the largest COVID-19 X-ray dataset so far named "Qata-COV19" that was also the first dataset with COVID-19 segmentation ground truth mask images generated by a human-machine collaborative method. They used a 2-stage architecture that applied manual segmentation on a randomly chosen subset and trained DL-models on this to produce segmentations. Then MDs chose best segmentations and they were used at stage 2 for training and cross-validation. Three different DNNs named U-Net, U-Net++, and Deep layer aggregation (DLA) were used with four pre-trained encoders (i.e., DenseNet-121, CheXNet, Inception-v3, and ResNet-50) with two settings as frozen or non-frozen encoder. All 24 combinations of these networks were then used with two different dataset divisions and the results were compared to find the best model. U-Net, U-Net++ with DenseNet-121 achieved the best performance with more than 99% accuracy for both detection and segmentation.

A novel 14 layer CNN models was proposed in [90] for classifying chest CT and X-rays into normal, pneumonia, or COVID-19. They built the CNN model with 9 convolution layers, 1 max pooling layer, and 4 fully connected layers with proper hyper-parameters. Their model was tested on 3 public datasets (i.e., Cohen CXR, RSNA, Radiopaedia) and 1 local dataset from the radiology department of BVHB, Pakistan, to compare the similarities and differences of COVID-19 traits between local and public datasets. The proposed model achieved 96.68% that either outperformed or had similar scores as 7 existing COVID-19 classification approaches on ML, DL, and hybrid models. The model was also trained on a large dataset and is currently being used at the radiology department of BVHB, Pakistan.

Ozturk et al. [91] recently proposed a novel DL model called DarkCovidNet for COVID-19 X-ray detection and classification into 2 and 3 classes (i.e., COVID-19, no findings, pneumonia). The proposed model is based on the DarkNet-19 model using the YOLO (you only look once) [92] system for object detection. They generated a 19-layer model with DarkNet and convolution blocks including batch normalization, max pooling, and LeakyRelu activation with gradually increasing filters as 8, 16, 32, etc. They applied a 5-fold cross-validation and provided a visualization of the heat map with Grad-CAM. An expert radiologist observed the performance and results of the proposed network and commented on DarkCovidNet. They mentioned the model as an outstanding model for COVID-19 detection, a sensitive model for pneumonia detection, a helpful model with heatmap visualization with the drawback of incorrect detection for unclear lung images or X-rays of acute respiratory distress syndrome (ARDS) patients that produced diffused images. The network was compared to 9 other COVID-19 detection models and outperformed all for binary classification with 98.08% accuracy and achieved comparable performance for 3 class problems with 87.02% accuracy.

TL-based COVID-19 detection using multimodal images was discussed in [8] to identify pneumonia in the lungs and detect the pneumonia type (i.e., COVID-19 or non-COVID). Their research was also one of the very few approaches to COVID-19 detection that used lung ultrasound images with X-rays and CT scans. A detail comparative analysis on popular CNN models for COVID-19 classification was provided and 8 CNN models were evaluated. They evaluated pre-trained VGG16, VGG19, Resnet50, Inception V3, Xception, InceptionResNet, DenseNet, and NASNetLarge and compared their performances to choose the best model for COVID classification with minimum tuning requirement to apply TL task. The models showed F1-scores varied from 0.63 to 0.99 for test data and VGG19 performed better compared to others considering all types of data and ultrasound images and provided more accurate results than the other images. They also proposed a pre-processing framework to minimize the noises and quality imbalance and maximize the lung area visibility. After reading the color images, they were converted to grey scale and normalized with N-CLAHE were then converted back to color images and resized for augmentation and training. They discussed the different experiments done on the models with varying parameters and their effects on classification results. A comparative analysis provided guidelines for future COVID-19 detection approaches on the performances of popular CNNs on various image types.

Zebin et al. [17] proposed a framework for COVID-19 detection, synthetic data generation, and progress monitoring using DL-based models, generative adversarial networks (GANs), and activation mappings. They discussed the COVID-19 datasets, sources, and their distributions according to percentages of labels. They applied a cycle GAN algorithm for 5000 iterations on the normal images from the datasets to generate realistic synthetic COVID-19 images to solve data limitation issue and added 100 synthetic data to the training set. Then pre-trained VGG16, ResNet50, and EfficientNetB0 were used for feature extraction and classification of the images into normal, COVID-19, and non-COVID pneumonia classes. They also used Grad-CAM to generate heatmap visualizations for the COVID-infected lung region of the same patient after intervals to monitor the progress of the disease. The proposed framework was compared to 7 COVID-19 approaches and achieved comparable performance with accuracy varying between 0.88 and 0.96.

Xu et al. [93] proposed a mask-attention base DNN called MANet for COVID X-ray classification. The input dataset was created by combining 3 different datasets to achieve 5 class labels — normal, COVID, TB, viral pneumonia, bacterial pneumonia. A basic ResUNet was used to segment the lungs from the images and then four different DNN (i.e., ResNet34, ResNet50, VGG16, InceptionV3) were used with and without the mask attention layers. The performance evaluation clearly showed that the mask-attention models achieved around 2% better performance and ResNet50 with MA achieved the highest classification accuracy of 97%. An ensemble model with a very similar idea was proposed in [94] for similar COVID-19 X-ray segmentation and classification. A ResUNet model was used to segment the lungs from the X-rays to optimize the ROI for the classifier. The segmented lungs were then applied as inputs to an ensemble model containing ResNet, VGG, and DensNet. A majority voting method was used on the outputs of them to finalize the classifier output to detect COVID-19-infected lungs from the X-rays. The Grad-CAM tool was used for visualization of the lung infections. The proposed model achieved around 77% classification accuracy with 95% dice score.

A soft attention-based U-Net model was used for COVID infection lesion segmentation from chest CT images in [95]. As U-Net [96] is one of the most popular DNNs for medical image processing, basic U-net and various

Table 4 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[67]	Detection Classification	BGWO SVM DT KNN NB ANN MajorityVoting	FOSF GLCM HOG	Phase I - Accuracy: 98.062AUC: 0.956 Phase II - Accuracy: 91.329AUC: 0.831	COVID-Chestxray Montgomery NIH ChestX-ray14	X-ray	Provided and evaluated 3 hypothesis on data augmentation, SVM, and ensemble model. Summarized DL methods, classification type, dataset size, type, and performance metrics efficiently for 16 publication on COVID-19 detection and classification from 2020. Applied a binary grey wolf optimization (BGWO) algorithm to select best features among 8196 features. Augmented dataset using sharpening, contrast management, brightness modification, and Gaussian blur technique. Applied majority voting on the results of 7 ML models.
[68]	Detection Classification	HSGO SGO KNN DT SVC L-SVC PCA KPCA	Luminosity	Accuracy - 99.65 Precision - 99.66 Sensitivity - 99.65 F1-score - 99.65	Kaggle COVID-19 radiography DB	X-ray	Extracted 10,000 luminosity level features from each pre-processed image. Used HSGO algorithm to select best 116 features from each image after comparing the performances of HSGO, HGO, PCA, and KPCA for feature selection. Applied 5 ML models (KNN, DT, RF, SVC, L-SVC) on the selected feature sets with SVC having highest scores for COVID/non-COVID classification. Outperformed 12 DL and bio-inspired COVID-19 classification approaches.
[69]	Detection Classification	Clustering GRNN PNN Ensemble	ANN	Accuracy - 2 class: 100 3 class: 80 4 class: 65	Cohen chest X-ray	X-ray	Introduced the idea of one-shot learning model. Proposed a one-shot cluster-based learning model for CXR classification into 2, 3, or 4 classes (normal, COVID, bacterial pneumonia, viral pneumonia). Ranked samples from each class according to discrimination characteristics and created clusters from randomly chosen best samples for one-shot learning. Used GRNN and PNN on raw features individually and on enhanced features using an ensemble model later. Achieved comparable results as AlexNet, GoogleNet, and ResNet-18.
[70]	Detection Classification	CNN KNN SVM	LBP LE GLCM DNN	2-fold, 10-fold: Accuracy - 94.73, 95.99 Sensitivity - 91.97, 94.04 Specificity - 98.91, 99.01 F1-score - 90.58, 92.84 AUC - 98.88, 99.03	Cohen UCSD-AI4H LIDC-IDRI	CT	Proposed a novel 23-layer CNN for COVID-19/non-COVID-19 classification task on chest CT images. Applied GLCM, LBP, and LE on original images for texture analysis and data increment. Included a comparative analysis for ML (SVM, KNN), DL (AlexNet, MobileNetV2), and texture-based models for COVID image classification with various data augmentation type, data augmentation amount, and cross-validations. Achieved comparable performances as 7 existing CT image-based COVID-19 classification approaches.

Table 5 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[71]	Detection Classification	CNN SVM TL	DNN	Accuracy - 94.7	Kaggle chest X-ray	X-ray	Used pre-trained ResNet18, ResNet50, ResNet101, VGG16, and VGG19 for feature extraction and used those features as inputs to a SVM for classification (i.e., COVID or normal). Applied linear, quadratic, cubic, and gaussian kernel functions to the SVM to compare performances. Achieved highest accuracy in minimum time with ResNet50 and SVM after extracting and comparing effects of eight local texture descriptors.
[72]	Detection Classification	CNN SVM Sobel	DNN	Accuracy - 99.02 Sensitivity - 100 Specificity - 95.23	Generated (private)	X-ray	Reviewed 25 DL-based COVID-19 approaches. Generated a new COVID-19 X-ray dataset with verified annotation labels for 333 X-rays containing 77 COVID positive and 256 normal images collected from February 2020 to April 2020 at Omid Hospital, Tehran. Applied Sobel filtering for edge detection from X-rays. Implemented a CNN with sigmoid output layer and another CNN with SVM output layer with 10-fold cross-validation. Tested proposed system with 6 other public datasets and achieved high accuracy. Compared proposed method with 11 approaches on COVID-19 detection using DL methods and outperformed most of them.
[73]	Detection Classification	CNN Fuzzy Color Stacking SMO SVM TL	DNN	Accuracy - 98.25 Sensitivity - 97.04 Specificity - 97.17 Precision - 97.18 F1-score - 97.09	Cohen chest X-ray Kaggle radiography	X-ray	Proposed a novel hybrid model using a new pre-processing method for COVID-19 classification. Pre-processed images by applying fuzzy color algorithm based on RGB values of the pixels and then added the with original data using stacking algorithm. Used pre-trained SqueezeNet and MobileNetV2 individually on the pre-processed data to extract 1000 features per image. Applied SMO on each feature set to select best features. Combined best features from each DL as inputs for a SVM for final classification. Achieved high performance with mentioned pre-processing and feature selection methods.
[74]	Detection Classification	CNN (ShuffleNet, SqueezeNet) MSVM TL	DNN	Accuracy - 2 class: 100 3 class: 99.72 4 class: 94.44 Recall - 94.45 Specificity - 98.15 Precision - 94.42 F1-score - 94.4 AUC - 0.98	Cohen chest X-ray Kaggle Chest X-ray	X-ray	Proposed a novel method for COVID-19 detection called COVIDetectionNet for 2, 3, and 4 class classification (COVID, normal, viral pneumonia, bacterial pneumonia). Extracted 1544 deep features using pre-trained ShuffleNet (544 features) and SqueezeNet (1000 features) separately on the pre-processed X-ray images. Applied multi-class SVM on those features for classification. Outperformed individual ShuffleNet and SqueezeNet. Outperformed Coronet and CovXNet by 2 to 7%. Outperformed 18 recent approaches on 2, 3, or 4 class COVID-19 detection and classification problems.

Table 6 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[77]	Detection Classification Segmentation	GAN *k*-means CNN VGG DenseNet BRISK TLRF SVM XG Boost	DNN Keypoints	Accuracy - 96.60Sensitivity - 95Specificity - 97.4	Cohen chest X-rayPneumonia X-ray	X-ray	Proposed a novel model combining multiple DL and TL models for COVID X-ray classification into 3 classes — COVID, normal, pneumonia. Used conditional GAN for lung segmentation and applied keypoint extraction algorithms and DNN for feature extraction. Outperformed similar recent researches using InceptionV3, ResNet, AlexNet, SVM, and MobileNet.
[78]	Detection Classification	Bayesian CNN	CNN	Accuracy - 96	COVID-19 Radiography Pneumonia Database	X-ray	Proposed a novel Bayesian optimization-based CNN for COVID detection. Used Bayesian optimizer to refine the hyperparameters of CNN to classify the images. Used a combination of COVID X-ray dataset and a pneumonia X-ray dataset.Achieved comparable performance compared to 10 similar CNN-based DL models.
[79]	Detection Classification Segmentation	CNN MADF Watershed TL	HOG CNN	Accuracy - 99.49 Specificity - 95.7 Sensitivity - 93.65	Kaggle Cohen chest X-ray Chest X-rays	X-ray	Proposed a feature fusion DNN model combining HOG and CNN features. Extracted 3780 HOG features and 4096 CNN features for each pre-processed image. Selected best features by maximum entropy-based feature selection and finalized 1186 features for each image. Applied a pre-trained VGG19 for CNN feature extraction and another for classification (i.e., COVID or normal). Segmented effected lung regions by watershed segmentation. Outperformed 19 existing DL-based COVID-19 detection.
[80]	Detection Classification	CNN TL	Haralick	Accuracy - 93	Combined multiple Github RSNA Google NIH Mendeley	X-ray CT	Proposed a TL model with Haralick features for a 4 class COVID classification. Pre-processed images with histogram equalization, Weiner filter, and ROI cropping. Extracted Haralick features from pre-processed images for 3 pre-trained DL models — Resnet50, VGG16, InceptionV3. Outperformed 9 existing ML-, DL-, and TL-based COVID-19 classification approaches. Provided visualization of infected region using Grad-CAM.

Table 7 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[81]	Detection Classification	CNN (COVID-Net)	DNN	Accuracy - 93.3 Sensitivity - 91 PPV - 98.9	COVIDx	X-ray	Proposed one of the first novel deep CNN-based open source COVID-19 detection model for X-ray data using human-machine collaboration based on a PEPX design. Generated one of the largest COVID-19 X-ray datasets named COVIDx containing 13,975 CXRs from 13,870 patient cases. Designed COVID-Net with two stages of CNN layers that included first stage with 1 convolution layer, 16 PEPX blocks (each block with 5 layers for projection, expansion, depth-wise convolution, second stage projection and extension), 1 flatten layer, 1 fully connected layer and softmax output layer, and second stage with 4 convolution layers for long range connectivity. Outperformed VGG-19 and ResNet-50.
[82]	Detection Classification	CNN (CoroDet)	DNN	Accuracy - 2 class: 99.1 3 class: 94.2 4 class: 91.2	COVID-R	X-ray CT	Proposed a novel 22-layer CNN called CoroDet. Detected and classified images into 2, 3, and 4 classes (i.e., COVID-19, normal, non-COVID viral pneumonia, non-COVID bacterial pneumonia). Prepared the biggest X-ray and CT images dataset of COVID by combining images from eight public COVID datasets. Outperformed 10 existing COVID-19 approaches for COVID detection and classification.
[83]	Detection Classification	CNN TL	DNN	Accuracy - 96.1 - 99.7	Cohen chest X-ray ChestX-ray8 Chest X-Ray Images(Pneumonia)	X-ray	Proposed 3 binary classifiers to detect 4 classes as normal, COVID, viral pneumonia, and bacterial pneumonia. Applied pre-trained ResNet50, ResNet101, ResNet152, InceptionV3, and Inception-ResNetV2 with 5-fold cross-validations separately on the data. Compared the performance scores to conclude that ResNet50 was the best classifier for COVID-19. Outperformed few similar COVID-19 detection approaches that used ML and DL models.
[84]	Detection Classification	EfficientNet TL	DNN	Flat: Accuracy - 93.9 Sensitivity - 96.8 PPV - 100 Hierarchical: Accuracy - 93.5 Sensitivity - 80.6 PPV - 100	RSNA Cohen chest X-ray COVIDx HCV-UFPR	X-ray	Proposed a novel model for COVID, normal, or pneumonia classification based on EfficientNet family. Used EfficientNet pre-trained on ImageNet dataset and fine tuned for COVID screening. Assessed the classification problem with both flat subclass and hierarchical classification. Applied 2 different data setups for training and testing. Included the first cross-dataset analysis on COVID-19. Developed a low computational cost (time and memory) network with high performance scores that achieved more than 30% computational efficiency compared to similar approaches.

variations of U-Net have been used in medical image segmentation researches for last last few years with high performance and accuracy. Recent COVID image segmentations have been using U-Net variants for different infection region segmentation. This research applied soft attention at every layer of U-net to extract the implicit features so as to enhance the segmentation performance. They used 3 COVID segmentation datasets and extended them by applying spatial, color, and noise augmentation. The proposed model achieved 98.2% accuracy and 76.3% dice score while outperforming basic U-Net, U-Net++, SD-UNet, and basic attention U-net.

Punn et al. [97] proposed CHS-Net, a hierarchical segmentation model for COVID CT datasets to extract the infection region. The used two cascaded residual attention inception U-Net (RAIU-Net) to incorporate semantic information. The U-Net used a residual inception model with hybrid pooling function combining max pooling and spectral pooling, and spectral spatial and depth attention based skip connections. The proposed model segmented the lung mask and then the infection region with 95% accuracy. A few-shot U-Net model was proposed in [98] for similar infection segmentation from CT images. The model was trained with few data and then a medical professional provided feedback on the segmented outputs. The feedbacks were then included in the training process to refine the model. The proposed model was compared to CNN, FCN, and U-Nets and achieved higher accuracy and performance. Another infection segmentation model was proposed in [99] as BS Net (i.e., boundary guided semantic learning network. A dual branch multi level feature extraction and aggregation framework was defined and implemented to incorporate high level features for enhancing the segmentation performance. They were able to achieve 85% dice scores with 84.9% sensitivity and 86.7% precision.

There are other existing researches on COVID image analysis for COVID detection, classification, and segmentation. Although almost all types of ML, DL algorithms have been already used for these tasks, the DL models with complex CNN structures, U-Nets, ResNets, InceptionV3, DenseNets, attention-based networks, and their variations were able to achieve better performances for COVID X-ray/CT classification and segmentation. The hybrid, ensemble, and TL models also generated comparative performance scores. The performance metrics used in different researches were different and the datasets of these researches had various resources. Hence, comparing the performance of these models containing different datasets and different metrics would not represent a fair comparison. Tables 4, 5, 6, 7, 8, 9, 10, and 11 tabulate summaries of these approaches with their research task, methods, features, datasets, data type, performance, and contributions.

6 Datasets

As COVID-19 was discovered at the end of 2019 and it is still an ongoing pandemic, datasets containing COVID-19 patient information (i.e., images, texts, and gene) are very limited. The available datasets also have limitations such as unavailable labels or ground truth values, not proper annotations, not valid evaluation, etc. Among the few available COVID-19 image datasets, the majority include CXRs, some include CT scans and very few contain lung ultrasounds. Tables 12 and 13 show list of few available COVID-19 image datasets with their image types, tasks, ground truth availability and labels, source URLs, etc. Most of these datasets are still growing and few of them have overlapping images. As most of these datasets are still being updated, their specifications may vary on a future date. The X-rays, CTs, MRIs, and ultrasounds are mostly 2D or 3D images of various formats like .png, .jpg, and .dcm and only few of them have recently started including dynamic files. The text data is mostly the metadata, case descriptions, and sometimes annotations of the data.

7 Challenges and future research scopes

Researchers have been trying to find effective solutions for COVID-19 detection, diagnosis, and analysis using various AI-based tools and applications. They have encountered a few issues while working on automated systems for COVID-19 and have mentioned these limitations and possible solutions. Other few additional challenges were identified while reviewing the existing approaches for this survey. Some major challenges and suggestions for future research are mentioned below.

7.1 Challenges

7.1.1 Authenticity of the systems

Due to the emergency pandemic situation, many researchers have been proposing various AI-based systems for COVID-19 detection and analysis. As the medical professionals and patients need accurate automated systems as soon as possible, proper screening for the proposed models is needed to ensure the authenticity of the approaches. Researchers should provide clear justifications and information that can be used to reproduce the results and check the accuracy of the claims of the researchers.

7.1.2 Time constraint

As the COVID-19 pandemic is ongoing, the necessity of accurate automated systems that are able to detect, classify,

Table 8 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[85]	Detection Classification	CNN Ensemble	DNN	Accuracy - 91.62 Sensitivity - 95.09 Specificity - 88.33 AUC - 91.71	Twitter ChestImaging SIRM Kaggle COVID-chestxray Cohen chest X-ray Case studies CheXpert	X-ray	Proposed a novel ensemble model with DenseNet201, Resnet50V2, and Inceptionv3 to achieve better performance. Trained each model individually and used test data for loss minimization. Used test data to calculate weights of each model with 5-fold cross-validation and applied the average weight of all 3 in the ensembler block. Achieved higher performance than each individual model and their concatenation. Provided a detail comparative analysis of proposed approach and 8 recent approaches with their datasets, evaluation method, accuracy, sensitivity, specificity, classifier, and summary of approaches.
[86]	Detection Classification	CNN Ensemble (EDL-COVID)	DNN	Accuracy - 95 Sensitivity - 96 PPV - 94.1	COVIDx	X-ray	Proposed an ensemble method by combining multiple snapshots models of COVID-Net named EDL-COVID. Reduced computational cost of training multiple models by using snapshots of the same model. Applied cosine annealing learning rate schedule to add aggressive learning rates to introduce diversity between the 6 snapshots of the same training model. Generated average class probabilities at the ensembler to provide final classification output.
[87]	Detection Classification	CNN Ensemble TL Majority Voting (EDL-COVID)	DNN	Accuracy - 99.054 Sensitivity - 99.05 Specificity - 99.6 F_score - 98.59 MCC - 97.89	Collected	CT	Proposed a TL-based ensemble model for chest CT classification of 3 classes — normal, lung tumor, and COVID. Discussed clinical values of CT, TL, ensemble model, AlexNet, GoogleNet, and ResNet in detail. Applied pre-trained AlexNet, GoogleNet, and ResNet18 separately with 5-fold cross-validation on the dataset. Used relative majority voting at ensembler to assess classification outputs from 3 individual CNNs and generate final classification output. Achieved high performance with low detection time for ensemble model compared to individual ones.
[88]	Detection Classification	CNN	DNN	Accuracy - 97.97	Kaggle repository	X-ray	Classified X-rays as COVID-19, pneumonia, or normal using 3 popular CNN models for a comparative analysis of their performances on COVID-19 detection. Implemented Inception Net V3, XCeption net, and ResNeXt separately and used the pre-processed augmented data to train and test for comparing the DNN models. Proposed novelty by applying Leaky Relu instead of Relu. Showed XCeption net had the highest performance scores.

Table 9 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[12]	Detection Classification	CNN TL	DNN	Accuracy - 98.33 Sensitivity - 100	Cohen chest X-ray COVID-19	X-ray	Classified X-rays as normal or COVID with 8 popular CNN models for providing a comparative analysis on COVID-19 detection with DL models. Applied pre-trained AlexNet, VGG-16, GoogleNet, MobileNet-V2, SqueezeNet, ResNet-34, ResNet-50, and Inception-V3 on pre-processed and augmented datasets separately with batch sizes 8, 16, and 32. Experimented with different batch sizes, epochs, learning rates, and optimizers. Declared ResNet-34 as the best method as it outperformed 8 recent COVID-19 approaches.
[89]	Detection Segmentation	CNN TL	DNN	Segmentation, detection: Sensitivity - 81.72, 94.96 Specificity - 99.93, 99.88 Precision - 84.74, 96.40 F1-score - 83.20, 95.24 F2-score - 82.31, 95.24 Accuracy - 99.85, 99.73	QaTa-COV19	X-ray	Provided the largest COVID-19 X-ray benchmark dataset and the first dataset with COVID-19 segmentation ground truth. Used two-stage human-machine collaboration by random manual segmentation, training DL-models, selecting correct masks by MDs, training, cross-validation. Implemented U-Net, U-Net++, and DLA with pre-trained encoders DenseNet-121, CheXNet, Inception-v3, and ResNet-50 and tested all combinations of 3 DNNs and 4 encoders with 2 options (frozen or non-frozen encoder). Achieved highest performances with U-Net and UNet++ models using DenseNet-121 encoder.
[90]	Detection Classification	CNN	DNN	Accuracy - 96.68 Sensitivity - 96.24 Specificity - 95.65	Cohen chest X-ray RSNA Radiopaedia BVHB	X-ray CT	Proposed a novel 14-layer CNN model (9 convolution, 1 max pooling, and 4 fully connected layers) for classifying images into normal, COVID-19, and pneumonia classes. Provided higher or comparable performance scores as 7 recent ML, DL, and hybrid methods for COVID-19 classification on CT and X-rays. Used a local dataset from BVHB to observe the differences of local traits of COVID-19. Proposed model is being used at the radiology department of BVHB, Pakistan, after training on a larger dataset.
[91]	Detection Classification	CNN (DarkCovidNet)	DNN	Accuracy(3class) - 87.02 Accuracy(2class) - 98.08 Sensitivity - 95.13 Specificity - 95.3 Precision - 98.03 F1-score - 96.51	Cohen chest X-ray ChestX-ray8	X-ray	Proposed a COVID-19 classification model named DarkCovidNet based on DarkNet using YOLO. Designed a 19-layer CNN containing DarkNet layers, convolution layers, max-pooling, LeakyRelu, batch normalization with increasing filters 8, 16, and 32. Generated heatmap using GradCAM. Evaluated by an expert radiologist who listed the contributions and limitations of the system. Performed well for COVID-19, but made incorrect classifications for some pneumonia cases or unclear images. Outperformed 9 COVID-19 detection models for binary classification.

Table 10 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[8]	Detection Classification	CNN TL	DNN	Precision: 84–100	Cohen chest X-ray UCSD-AI4H ChestX-ray8 POCOVID-Net	X-ray CT Ultrasound	Provided a comparative analysis on CNN and TL models for COVID-19, pneumonia, and normal class classification. Proposed a novel pre-processing method with color to greyscale to color conversion with N-CLAHE normalization. Compared popular CNNs as VGG16, VGG19, Resnet50, Inception V3, Xception, InceptionResNet, DenseNet, and NASNetLarge by applying pre-trained models on COVID-19 datasets and comparing them. Decided that ultrasound images achieved better accuracy than X-ray and CT and VGG19 was the best model for COVID-19 detection with that multimodal image data.
[17]	Detection Progression Visualization	CNN TL CycleGAN	DNN	Accuracy: 88–97	Cohen chest X-ray Kaggle ChestXray Synthetic Data	X-ray	Proposed a COVID-19 detection, disease progression and visualization model with TL. Applied cycle GAN to generate realistic synthetic COVID-19 data from healthy images after 5000 iterations and added them for model training. Used pre-trained VGG16, ResNet50, and EfficientNetB0 on the pre-processed data for feature extraction and classification. Used Grad-CAM for highlighting infected regions.
[93]	Segmentation Detection Classification	ResUNet VGG ResNet InceptionV3	DNN	Accuracy: 96.03–97.06 IOU: 89.07–93.49	Customized	X-ray	Proposed a mask-attention (MA) approach for segmenting and classifying COVID X-rays into 5 classes — normal, COVID, TB, bacterial pneumonia, viral pneumonia. Used ResUNet to segment the lungs regions only for classification. Applied a mask based soft attention on the classifiers to improve performance of the basic DL models. Used Grad-CAM for infection visualization.
[94]	Segmentation Detection Classification	ResUNet VGG ResNet DenseNet Ensemble	DNN	Accuracy: 77–91 Dice - 95 Sensitivity: 78–92	COVID-19 Chest X-Ray Dataset COVIDX Customized	X-ray	Proposed a multi-class ensemble COVID classification and segmentation model for COVID X-ray images. Applied a ResUNet model to segment the lungs as ROI from X-rays. Used majority voting on an ensemble model with ResNet, VGG, and DenseNet for classifying the segmented images into COVID, normal, and pneumonia classes. Used Grad-CAM for visualizing the lung infections. Outperformed one multi-class and three 2-class COVID classifiers from recent researches.

Table 11 Recent approaches on COVID-19

Ref.	Type	Approach	Features	Performance	Dataset	Data type	Contributions
[95]	Segmentation Severity classification	U-Net Attention	DNN	Accuracy - 98.2 Dice - 76.3	SIRM Radiopaedia COVID-19 CT segmentation	CT	Proposed an attention base U-Net for COVID-infected region segmentation. Applied soft attention to extract implicit features from the images at each layer of U-Net to enhance the performance. Used boundary loss function to manage small lesion segmentation. Used spatial, color, and noise augmentation to extend dataset. Outperformed basic U-Net, U-Net++, SD-UNet, and attention U-Net.
[97]	Segmentation	RAIU-Net	DNN	Accuracy - 95 Precision - 71.2 Specificity - 96.1 Recall - 80.4 Dice - 75.6 Jaccard - 74.2	COVID-19 CT segmentation COVID-19 CT Lung and infection Segmentation	CT	Proposed a hierarchical segmentation with 2 cascaded residual attention inception U-Net (RAIU-Net) for COVID infection segmentation from CT images. Applied residual inception, hybrid of max pooling and spectral pooling, attention-based skip connection for better performance. Used Grad-CAM++ for segmentation visualization. Outperformed 6 recent researches using U-Net, VGG, and ResNet variations for segmenting the infections.
[98]	Segmentation	U-Net	DNN	Accuracy - 99 Precision - 99 AUC - 96.8	Radiopaedia Customized	CT	Proposed a few-shot U-Net with dynamic learning process. Trained the model with few shots and refined the parameters. Applied the feedback of domain experts on the current segmented output and used that as the input for the next phases. Mentioned several ML, DL models for COVID image analysis. Discussed detail steps and showed segmentation outputs with basic CNN, FCN, and U-Net. Compared proposed model with CNN, FCN, and U-Nets.
[99]	Segmentation	BS Net Res2net	DNN	Dice - 85.1 Sensitivity - 84.9 Precision - 86.7	COVID-19 CT segmentation COVID-19 CT Lung and infection Segmentation	CT	Proposed a boundary guided semantic learning network (BS net) for Covid infection segmentation. Used a dual branch approach to include high level features in the segmentation by incorporating semantic connections. Implemented 3 phase feature processing — five multilevel features in phase 1, semantic attention mask to refine the features in phase 2, and dual branch to aggreagte the high level features. Explained all level feature extraction, aggregation in details. Outperformed UNet UNet++, Attention-UNet, and ResNet-UNet in COVID CT infection segmentation.

Table 12 Image datasets used in COVID-19 approaches

Dataset	Type	Task	Abnormality	GT	URL
IEEE8023 [100, 101]	X-ray Text	Classification	COVID-19 Few other diseases	Viral Bacterial Fungal Lipoid Aspiration Unknown	www.github.com/ieee8023/covid-chestxray-dataset
Kaggle COVIDX rays [102]	CT X-ray	Classification	COVID-19 Non-COVID	COVID-19 Non-COVID	www.kaggle.com/andrewmvd/convid19-x-rays
COVID-19 Radiography Database [103–105]	X-ray	Classification	COVID-19 Few other diseases	Normal Lung opacity Viral pneumonia COVID-19	www.kaggle.com/tawsifurrahman/covid19-radiography-database
COVIDx Dataset [106]	X-ray	Classification	COVID-19 Pneumonia	Normal COVID-19 Pneumonia	github.com/lindawangg/COVID-Net/blob/master/docs/COVIDx.md
COVIDx CT Dataset [107, 108]	CT	Classification	COVID-19 Pneumonia	Normal COVID-19 Pneumonia	www.kaggle.com/dataset/c395fb339f210700ba392d81bf200f766418238c2734e5237b5dd0b6fc724fcb/version/1
COVIDx CT-2 Dataset [107, 109]	CT	Classification	COVID-19 Pneumonia	Normal COVID-19 Pneumonia	www.kaggle.com/hgunraj/covidxct
UCSD-AI4H [110, 111]	CT	Classification	COVID-19	Covid Non-COVID	www.github.com/UCSD-AI4H/COVID-CT/tree/master/Images-processed
NCCID [112]	CT X-ray MRI Text	Classification Diagnosis Assessments	COVID-19	COVID-19 Non-COVID	nhsx.github.io/covid-chest-imaging-database/
COVID-19 image data [113]	CT X-ray Text	Classification	COVID-19 Few other diseases	COVID-19 MERS SARS ARDS	www.kaggle.com/bachrr/covid-chest-xray
CNCB [114, 115]	CT Text	Classification	COVID-19 Pneumonia	Normal COVID-19 Pneumonia	ncov-ai.big.ac.cn/download?lang=en

Table 13 Image datasets used in COVID-19 approaches

Dataset	Type	Task	Abnormality	GT	URL
COVID-19 [116]	X-ray	Classification	COVID-19 Pneumonia	Normal COVID-19 Pneumonia	github.com/muhammedtalo/COVID-19/tree/master/X-Ray%20Image%20DataSet
COVID-19 Lung ultrasound dataset [117]	Ultrasound	Classification	COVID-19 Pneumonia	Normal COVID-19 Bacterial pneumonia Viral pneumonia	github.com/jannisborn/covid19_ultrasound/tree/master/data
COVID-19 Chest X-Ray Dataset [94, 118]	X-ray	Classification Segmentation	COVID-19 Pneumonia	Normal COVID-19 Bacterial pneumonia Viral pneumonia	darwin.v7labs.com/v7-labs/covid-19-chest-x-ray-dataset
QaTa-COV19 [89, 119]	X-ray	Segmentation	COVID-19	Infection mask	www.kaggle.com/aysendegerli/qatacov19-dataset
COVID-19 CT segmentation dataset [120]	CT	Segmentation	COVID-19	Ground-glass Consolidation Pleural effusion	medicalsegmentation.com/covid19/
COVID-19 CT Lung and Infection Segmentation Dataset [121, 122]	CT	Segmentation	COVID-19	Left lung Right lung Infection	zenodo.org/record/3757476#.YPELROhKhPZ
COVID-19-CT-Seg-Benchmark [121, 123, 124]	CT	Segmentation	COVID-19	Lung Infection Few other organs	gitee.com/junma11/COVID-19-CT-Seg-Benchmark/tree/master#datasets
Google Drive [125]	CT	Segmentation	COVID-19	Infection mask Multiclass infection mask Edges	drive.google.com/file/d/1bbKAqUuk7Y1q3xsDSwP07oOXN_GL3SQM/view
BSTI [126]	CT	Visualization	COVID-19	Various views	www.bsti.org.uk/training-and-education/covid-19-bsti-imaging-database/
CoronaCases [127]	CT	Visualization	COVID-19	Various views	coronacases.org/
Twitter Chest Imaging [128]	CT	Visualization	COVID-19	COVID-19	twitter.com/chestimaging
SIRM [129]	CT Text	Diagnosis	COVID-19	COVID-19	sirm.org/covid-19/

diagnose, and monitor COVID-19 patients is acute. The urgency of working COVID-19 analysis systems that can be immediately implemented for COVID-19 patients is a very real issue. Another time constraint issue of the automated systems is the execution or analysis time. To stop the rapid progress of the pandemic, the automated systems are required to process patient data as quickly as possible but with correct results.

7.1.3 Lack of a complete automated system

Most approaches have been focusing on detection of the disease or segmentation of the ROIs. To properly manage the severity of the pandemic, complete automated systems are needed that will be able to separate COVID-19 patients from healthy or any other respiratory disease patients, then asses the severity of the disease, detect the exact infected regions, provide diagnosis, monitor patient progress, and provide a complete and comprehensive report to the medical professionals.

7.1.4 Amount of FPs and FNs

Detecting COVID-19 accurately is necessary to treat the patients; however, the accuracy scores of proposed or existing systems are not good enough to address the severity of the disease. Each automated system needs to check and minimize the number of false positives and false negatives detected by their systems. COVID-19 death rates and contamination rates are very clear indications that the errors in detection of COVID-19 can have serious impact on the ongoing pandemic. So, AI systems with zero false negative and false positive classifications should be one of the main objectives of the automated systems.

7.1.5 Applicability to real-time data

Researchers from different countries have been working on COVID-19-based systems using ML, DL, TL, and other methods. But most approaches were tested on a very limited amount of data. These AI-based systems need to be implemented and tested with real-time patient data in hospitals and other medical care facilities treating COVID-19 patients so that they can be fine tuned according to real-time data and can be used in large-scale practical scenarios.

7.1.6 Lack of customized systems

The statistics and news of COVID-19 from all over the world have already shown that the same systems, resources, and measures are not applicable for every country. Developed and developing countries need different types of setups and patients of different variants of COVID-19 require different care. These variations regarding countries, their resources, available healthcare systems, accessibility to treatments, variants of COVID-19, and their severity should be addressed while designing automated systems for various countries.

7.1.7 Lack of available datasets

One of the major limitations of COVID-19 analysis is the limited amount of available datasets. Some of the datasets mentioned in Section 6 are combinations of some other datasets. Lack of sufficient unique datasets of CXRs, CTs, and lung ultrasounds of COVID-19 patients limits the complete and accurate applications of DL models.

7.1.8 Limitations of available datasets

The available COVID-19 datasets have some limitations. Most datasets have very few images from a limited amount of patients. The image qualities, dimensions, modalities, and consistencies are not similar among most of the images of the same dataset. The lack of variations in the dataset images including all possible abnormal scenarios of COVID-19-infected lungs is absent. Another severe issue with the existing datasets is the lack of ground truths or labels of images. The datasets that included proper labels also need profound verification about the authentication and accuracy of their labeling.

Most of the limitations of the current researches are related to the time constraint. The urgency of the pandemic led to automated systems using TL models for faster disease detection and segmentation with a limited amount of verified data. Researchers did not have sufficient time to experiment with unique ideas for providing more concrete novel approaches. To provide a non-invasive automated COVID-19 detection and segmentation system, most researchers applied few popular ML, DL, and TL models that generated outputs with high performance scores previously on other medical or non-medical images. This process neglected the analysis of features unique to COVID-19 medical images and their affect on the classification and segmentation tasks. The time constraint also affected the dataset verification. The limitation of correctly annotated and verified datasets from patients of different age, gender, ethnicity, location, etc. made it difficult to provide general and customized automated systems for COVID detection. The time constraint raised some concerns about the validity of the researches. Currently, there are hundreds of proposed automated systems tested with small amount of various verified and unverified datasets that are not tested on real-time data to establish their authenticity.

7.2 Future research scopes

The challenges mentioned in the previous subsection can also provide some future research scopes on COVID-19 medical image analysis. As the efficiency of the detection, segmentation, and classification of COVID-19 infections is directly related to the availability, quality, completeness, and amount of medical image data, one of the major future research goal can be generating new benchmark datasets by combining, refining, and labeling the available datasets and adding more recent and precise data to them. Medical professionals or radiologists can provide correct data annotation label (i.e., healthy, COVID, and other diseases) for classification datasets and ground truth images with segmented infection regions for segmentation datasets. More advanced datasets may contain the specific types of infections, their severity, etc. Feedbacks from medical professionals to verify these datasets are also needed for future researches. Another possible research idea can be to develop more novel algorithms or systems for COVID-19 detection and segmentation. Due to the time constraint, most of the current researches are based on known ensemble methods, hybrid methods, TL models, or variations of popular DL/ML models. Although they work well for the current limited datasets, more original researches are needed in this field to provide some unique perspectives on the solutions to extract and apply the unique features of COVID images. More original researches incorporating geographical, environmental, and other specific constraints for different countries can also lead to more practical outcomes or automated systems for non-invasive COVID-19 detection.

8 Conclusion

Researchers from all fields have been trying to contribute in different ways to fight and win the war against the SARS-CoV-2 virus. Researchers have been trying to develop tools and applications to aid the medical professionals. In this paper, a review on the recent researches on COVID-19 medical image analysis is presented. Basic tasks like COVID detection, COVID-infected lung region segmentation, and severity assessment are solved using CXRs, CTs, etc. and the AI-based systems applied for these tasks recently have been analyzed and summarized here to provide a complete idea about the current researches. Although some researchers worked with basic ML models like SVM, RF, DT, and KNN, most recent methods applied DL-based methods like ResNet, Inception, Xception, MobileNet, U-Net, and U-Net++ for COVID-19 image analysis achieving high accuracy. TL models are also very popular in recent approaches and have been applied to incorporate the advantages of popular DL networks for COVID-19 analysis without spending time to implement and train existing DL models from scratch. Some researchers have provided some novel ideas to analyze the tasks from different perspectives and showed comparable performances. A large number of papers showed more than 99% accuracy for COVID-19 detection and classification on the available datasets and most of the performance scores were higher than 80% or 90%. Despite having hundreds of publications on COVID-19 medical image analysis with almost 100% accuracy, the requirements of more extensive, customized, and novel experiments on this field is still there due to the challenges mentioned Section 7. This is an interesting research area having societal and scientific impact. It is important to work on approaches that will reduce the false negatives to the most possible minimum, though zero is the preferred target to avoid further spread of the virus.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. World Health Organization (2022) Origins of the SARS-CoV-2 virus, Available at https://www.who.int/health-topics/coronavirus/origins-of-the-virus (2022/09/09)
2. World Health Organization (2021) Who-convened global study of origins of SARS-CoV-2: China part
3. World Health Organization (2022) Who director-general's opening remarks at the media briefing on COVID-19 - 11 March 2020, Available at https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing on-covid-19—11-march-2020 . Accessed 09 Sep 2022
4. Worldometer (2022) COVID-19 coronavirus pandemic, Available at https://www.worldometers.info/coronavirus/ (2022/11/06)
5. Jin Y, Yang H, Ji W, Wu W, Chen S, Zhang W, Duan G (2020) Virology, epidemiology, pathogenesis, and control of COVID-19. Viruses 12(4):372
6. Ciotti M, Ciccozzi M, Terrinoni A, Jiang WC, Wang CB, Bernardini S (2020) The COVID-19 pandemic. Crit Rev Clin Lab Sci 57(6):365–388
7. World Health Organization (2022) Tracking sars-cov-2 variants, Available at https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/. Accessed 09 Sep 2022
8. Horry MJ, Chakraborty S, Paul M, Ulhaq A, Pradhan B, Saha M, Shukla N (2020) COVID-19 detection through transfer learning

using multimodal imaging data. IEEE Access 8:149 808–149 824
9. Tahamtan A, Ardebili A (2020) Real-time RT-PCR in COVID-19 detection: issues affecting the results. Expert Rev Mol Diagn 20(5):453–454
10. Akl EA, Blažić I, Yaacoub S, Frija G, Chou R, Appiah JA, Fatehi M, Flor N, Hitti E, Jafri H, Jin ZY (2021) Use of chest imaging in the diagnosis and management of COVID-19: a who rapid advice guide. Radiology 298(2):E63–E69
11. News-Medical.Net (2022) Transfer learning exploits chest-xray to diagnose COVID-19 pneumonia, Available at https://www.news-medical.net/news/20201218/Transfer-learning-exploits-chest-Xray-to-diagnose-COVID-19-pneumonia.aspx. Accessed 09 Sep 2022
12. Nayak SR, Nayak DR, Sinha U, Arora V, Pachori RB (2021) Application of deep learning techniques for detection of COVID-19 cases using chest x-ray images: a comprehensive study. Biomed Signal Process Control 64:102365
13. Borakati A, Perera A, Johnson J, Sood T (2020) Diagnostic accuracy of x-ray versus CT in COVID-19: a propensity-matched database study. BMJ Open 10(11):e042946
14. Sverzellati N, Ryerson CJ, Milanese G, Renzoni EA, Volpi A, Spagnolo P, Bonella F, Comelli I, Affanni P, Veronesi L, Manna C (2021) Chest x-ray or CT for COVID-19 pneumonia? Comparative study in a simulated triage setting. European Respiratory Journal
15. Zhang Y, Xue H, Wang M, He N, Lv Z, Cui L (2021) Lung ultrasound findings in patients with coronavirus disease (COVID-19). Am J Roentgenol 216(1):80–84
16. Kwee TC, Kwee RM (2020) Chest CT in COVID-19: what the radiologist needs to know. RadioGraphics 40(7):1848–1865
17. Zebin T, Rezvy S (2021) COVID-19 detection and disease progression visualization: deep learning on chest x-rays for classification and coarse localization. Appl Intell 51(2):1010–1021
18. VIDA (2022) Lung intelligence for COVID-19, Available at https://hub.vidalung.ai/covid-19. Accessed 09 Sep 2022
19. InferVision (2022) A.I. Solutions, Available at https://global.infervision.com/product/5/. Accessed 09 Sep 2022
20. NVIDIA (2022) Clara COVID-19, Available at https://ngc.nvidia.com/catalog/collections/nvidia:claracovid19. Accessed 09 Sep 2022
21. Thirona (2022) Artificial intelligence to screen for COVID-19 on CT- and x-ray images, Available at https://thirona.eu/cad4covid/. Accessed 09 Sep 2022
22. Ai4Networks (2022) An artificial intelligence powered app for detecting COVID-19 from cough sound, Available at https://www.ai4networks.com/covid19.php. Accessed 09 Sep 2022
23. Innobiochips (2022) Covidiag, Available at http://www.innobiochips.fr/applications/covidiag. Accessed 09 Sep 2022
24. Li S (2022) COVID-19 assistant discrimination, Available at http://lishuyan.lzu.edu.cn/COVID2019_2/. Accessed 09 Sep 2022
25. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision, pp 618–626
26. NCBI (2022) Pubmed, Available at https://pubmed.ncbi.nlm.nih.gov/. Accessed 09 Sep 2022
27. Google (2022) Google Scholar, Available at https://scholar.google.com/. Accessed 09 Sep 2022
28. Hamet P, Tremblay J (2017) Artificial intelligence in medicine. Metabolism 69:S36–S40
29. Sarker IH (2021) Machine learning: algorithms, real-world applications and research directions. SN Computer Science 2(3):1–21
30. Abirami RS, Kumar GS (2022) Comparative study based on analysis of coronavirus disease (COVID-19) detection and prediction using machine learning models. SN Computer Science 3(1):1–8
31. Lee J, Jun S, Cho Y, Lee H, Kim GB, Seo JB, Kim N (2017) Deep learning in medical imaging: general overview. Korean J Radiol 18(4):570–584
32. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK (2018) Medical image analysis using convolutional neural networks: a review. J Med Syst 42(11):1–13
33. Zhao M, Chang CH, Xie W, Xie Z, Hu J (2020) Cloud shape classification system based on multi-channel CNN and improved FDM. IEEE Access 8:44 111–44 124
34. Xu Q, Zeng Y, Tang W, Peng W, Xia T, Li Z, Teng F, Li W, Guo J (2020) Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network. IEEE J Biomed Health Inform 24(9):2481–2489
35. Zheng Q, Yang M, Yang J, Zhang Q, Zhang X (2018) Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process. IEEE Access 6:15 844–15 869
36. Zheng Q, Yang M, Tian X, Jiang N, Wang D (2020) A full stage data augmentation method in deep convolutional neural network for natural image classification. Discrete Dyn Nat So 2020:4706576
37. Pan SJ, Yang Q (2009) A survey on transfer learning. IEEE Trans Knowl Data Eng 22(10):1345–1359
38. Jin B, Cruz L, Gonçalves N (2020) Deep facial diagnosis: deep transfer learning from face recognition to facial diagnosis. IEEE Access 8:123 649–123 661
39. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H, He Q (2020) A comprehensive survey on transfer learning. Proc IEEE 109(1):43–76
40. Seliya N, Khoshgoftaar TM, Hulse JV (2009) A study on the relationships of classifier performance metrics. In: 2009 21st IEEE International Conference on Tools with Artificial Intelligence, IEEE, pp 59–66
41. Taha AA, Hanbury A (2015) Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging 15(1):1–28
42. Narkhede S (2022) Understanding AUC - ROC curve, Available at https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5. Accessed 09 Sep 2022
43. Alafif T, Tehame AM, Bajaba S, Barnawi A, Zia S (2021) Machine and deep learning towards COVID-19 diagnosis and treatment: survey, challenges, and future directions. Int J Environ Res Public Health 18(3):1117
44. Mangal A, Kalia S, Rajgopal H, Rangarajan K, Namboodiri V, Banerjee S, Arora C (2020) CovidAID: COVID-19 detection using chest x-ray, arXiv preprint arXiv: 2004.09803
45. NVIDIA (2022) Clara COVID-19, Available at https://ngc.nvidia.com/catalog/collections/nvidia:claracovid19. Accessed 09 Sep 2022
46. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K (2020) Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology 296(2):E65–E71
47. Hossain FA, Lover AA, Corey GA, Reich NG, Rahman T (2020) Flusense: a contactless syndromic surveillance platform for influenza-like illness in hospital waiting areas. Proc ACM Interact Mob Wearable Ubiquitous Technol 4(1):1–28

48. Islam M, Poly TN, Alsinglawi B, Lin MC, Hsu MH, Li YCJ (2021) A state-of-the-art survey on artificial intelligence to fight COVID-19. J Clin Med 10(9):1961
49. Nayak J, Naik B, Dinesh P, Vakula K, Rao BK, Ding W, Pelusi D (2021) Intelligent system for COVID-19 prognosis: a state-of-the-art survey. Appl Intell 51(5):2908–2938
50. Lalmuanawma S, Hussain J, Chhakchhuak L (2020) Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: a review, Chaos, Solitons & Fractals, p 110059
51. Mohammad-Rahimi H, Nadimi M, Ghalyanchi-Langeroudi A, Taheri M, Ghafouri-Fard S (2021) Application of machine learning in diagnosis of COVID-19 through x-ray and CT images: a scoping review. Front Cardiovasc Med 8:185
52. Zhang F (2021) Application of machine learning in CT images and x-rays of COVID-19 pneumonia. Medicine 100(36):e26855
53. Subramanian N, Elharrouss O, Al-Maadeed S, Chowdhury M (2022) A review of deep learning-based detection methods for COVID-19. Comput Biol Med 143:105233
54. Ozsahin I, Sekeroglu B, Musa MS, Mustapha MT, Ozsahin DU (2020) Review on diagnosis of COVID-19 from chest CT images using artificial intelligence. Comput Math Methods Med, 2020
55. Alghamdi H, Amoudi G, Elhag S, Saeedi K, Nasser J (2021) Deep learning approaches for detecting COVID-19 from chest x-ray images: a survey. IEEE Access
56. de Sousa OL, Magalhães DM, Vieira PDA, Silva R (2021) Deep learning in image analysis for COVID-19 diagnosis: a survey. IEEE Lat Am Trans 19(6):925–936
57. Bhattacharya S, Maddikunta PKR, Pham Q, Gadekallu TR, Chowdhary CL, Alazab M, Piran MJ (2021) Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey, vol 65
58. Kumar A, Gupta PK, Srivastava A (2020) A review of modern technologies for tackling COVID-19 pandemic. Diabetes Metab Syndr: Clinical Research & Reviews 14(4):569–573
59. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, He K, Shi Y, Shen D (2020) Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. IEEE Rev Biomed Eng 14:4–15
60. Swapnarekha H, Behera HS, Nayak J, Naik B (2020) Role of intelligent computing in COVID-19 prognosis: a state-of-the-art review. Chaos, Solitons & Fractals 138:109947
61. Albahri OS, Zaidan AA, Albahri AS, Zaidan BB, Abdulkareem KH, Al-Qaysi ZT, Alamoodi AH, Aleesa AM, Chyad MA, Alesa RM, Kim LC (2020) Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects, Journal of Infection and Public Health
62. Sufian A, Ghosh A, Sadiq AS, Smarandache F (2020) A survey on deep transfer learning to edge computing for mitigating the COVID-19 pandemic. J Syst Architect 108:101830
63. Mohamadou Y, Halidou A, Kapen PT (2020) A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. Appl Intell 50(11):3913–3925
64. Shuja J, Alanazi E, Alasmary W, Alashaikh A (2020) COVID-19 datasets: a survey and future challenges. Development 11:12
65. Latif S, Usman M, Manzoor S, Iqbal W, Qadir J, Tyson G, Castro I, Razi A, Boulos MNK, Weller A, Crowcroft J (2020) Leveraging data science to combat COVID-19: a comprehensive review, IEEE Transactions on Artificial Intelligence
66. Hussein HA, Abdulazeez AM (2021) COVID-19 pandemic datasets based on machine learning clustering algorithms: a review. PalArch's J Archaeol Egypt/ Egyptol 18(4):2672–2700
67. Chandra TB, Verma K, Singh BK, Jain D, Netam SS (2021) Coronavirus disease (COVID-19) detection in chest x-ray images using majority voting based classifier ensemble. Expert Syst Appl 165:113909
68. Singh AK, Kumar A, Mahmud M, Kaiser MS, Kishore A (2021) COVID-19 infection detection from chest x-ray images using hybrid social group optimization and support vector classifier. Cognit Compu, 1–13
69. Aradhya VNM, Mahmud M, Guru DS, Agarwal B, Kaiser MS (2021) One-shot cluster-based approach for the detection of COVID–19 from chest x–ray images. Cognit Compu, 1–9
70. Yasar H, Ceylan M (2021) A novel comparative study for detection of COVID-19 on ct lung images using texture analysis, machine learning, and deep learning methods. Multimed Tools Appl 80(4):5423–5447
71. Ismael AM, Şengür A (2021) Deep learning approaches for COVID-19 detection based on chest x-ray images. Expert Syst Appl 164:114054
72. Sharifrazi D, Alizadehsani R, Roshanzamir M, Joloudari JH, Shoeibi A, Jafari M, Hussain S, Sani ZA, Hasanzadeh F, Khozeimeh F, Khosravi A (2021) Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using x-ray images. Biomed Signal Process Control 68:102622
73. Toğaçar M, Ergen B, Cömert Z (2020) 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 121:103805
74. Elkorany AS, Elsharkawy ZF (2021) Covidetection-net: a tailored COVID-19 detection from chest radiography images using deep learning. Optik 231:166405
75. Khan AI, Shah JL, Bhat MM (2020) Coronet: a deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed 196:105581
76. Mahmud T, Rahman MA, Fattah SA (2020) Covxnet: a multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest x-ray images with transferable multi-receptive feature optimization. Comput Biol Med 122:103869
77. Bhattacharyya A, Bhaik D, Kumar S, Thakur P, Sharma R, Pachori RB (2022) A deep learning based approach for automatic detection of COVID-19 cases using chest x-ray images. Biomed Signal Process Control 71:103182
78. Loey M, El-Sappagh S, Mirjalili S (2022) Bayesian-based optimized deep learning model to detect COVID-19 patients using chest x-ray image data, vol 142
79. Ahsan M, Based M, Haider J, Kowalski M (2021) COVID-19 detection from chest x-ray images using feature fusion and deep learning. Sensors 21(4):1480
80. Perumal V, Narayanan V, Rajasekar SJS (2021) Detection of COVID-19 using CXR and CT images using transfer learning and Haralick features. Appl Intell 51(1):341–358
81. Wang L, Linda ZQLIN, Wong A (2020) COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest x-ray images. Sci Rep 10(1):1–12
82. Hussain E (2021) Corodet: a deep learning based classification for COVID-19 detection using chest x-ray images. Chaos Solitons Fractals 142:110495
83. Narin A, Kaya C, Pamuk Z (2021) Automatic detection of coronavirus disease (COVID-19) using x-ray images and deep convolutional neural networks. Pattern Anal Applic, 1–14

84. Luz E, Silva P, Silva R, Silva L, Guimarães J, Miozzo G, Moreira G, Menotti D (2021) Towards an effective and efficient deep learning model for COVID-19 patterns detection in x-ray images. Research on Biomedical Engineering, 1–14
85. Das AK, Ghosh S, Thunder S, Dutta R, Agarwal S, Chakrabarti A (2021) Automatic COVID-19 detection from x-ray images using ensemble learning with convolutional neural network. Pattern Anal Applic, 1–14
86. Tang S, Wang C, Nie J, Kumar N, Zhang Y, Xiong Z, Barnawi A (2021) EDL-COVID: ensemble deep learning for COVID-19 cases detection from chest x-ray images. In: IEEE Transactions on Industrial Informatics
87. Zhou T, Lu H, Yang Z, Qiu S, Huo B, Dong Y (2021) The ensemble deep learning model for novel COVID-19 on CT images, vol 98
88. Jain R, Gupta M, Taneja S, Hemanth DJ (2021) Deep learning based detection and analysis of COVID-19 on chest x-ray images. Appl Intell 51(3):1690–1700
89. Degerli A, Ahishali M, Yamac M, Kiranyaz S, Chowdhury MEH, Hameed K, Hamid T, Mazhar R, Gabbouj M (2021) COVID-19 infection map generation and detection from chest x-ray images. Health Inf Sci Syst 9(1):1–16
90. Gilanie G, Bajwa UI, Waraich MM, Asghar M, Kousar R, Kashif A, Aslam RS, Qasim MM, Rafique H (2021) Coronavirus (COVID-19) detection from chest radiology images using convolutional neural networks. Biomed Signal Process Control 66:102490
91. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR (2020) Automated detection of COVID-19 cases using deep neural networks with x-ray images. Comput Biol Med 121:103792
92. Redmon J, Farhadi A (2017) Yolo9000: better, faster, stronger. Proc IEEE Conf Comput Vis Pattern Recognit, 7263–7271
93. Xu Y, Lam H, Jia G (2021) Manet: a two-stage deep learning method for classification of COVID-19 from chest x-ray images. Neurocomputing 443:96–105
94. Hertel R, Benlamri R (2022) A deep learning segmentation-classification pipeline for x-ray-based COVID-19 diagnosis. Biomed Eng Adv 3:100041
95. Ahmed I, Chehri A, Jeon G (2022) A sustainable deep learning-based framework for automated segmentation of COVID-19 infected regions: using u-net with an attention mechanism and boundary loss function. Electronics 11(15):2296
96. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, Springer, Cham, pp 234–241
97. Punn NS, Agarwal S (2022) Chs-net: a deep learning approach for hierarchical segmentation of COVID-19 via CT images. Neural Process Lett, 1–22
98. Voulodimos A, Protopapadakis E, Katsamenis I, Doulamis A, Doulamis N (2021) A few-shot u-net deep learning model for COVID-19 infected area segmentation in CT images. Sensors 21(6):2215
99. Cong R, Zhang Y, Yang N, Li H, Zhang X, Li R, Chen Z, Zhao Y, Kwong S (2022) Boundary guided semantic learning for real-time COVID-19 lung infection segmentation system. IEEE Trans Consum Electron
100. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M (2020) COVID-19 image data collection: prospective predictions are the future," arXiv preprint arXiv: 2006.11988
101. Bganglia (2022) COVID-chestxray-dataset, Available at https://github.com/ieee8023/covid-chestxray-dataset. Accessed 09 Sep 2022
102. Larxel (2022) COVID-19 x rays," Available at https://www.kaggle.com/datasets/andrewmvd/convid19-x-rays. Accessed 09 Sep 2022
103. Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, Islam KR, Khan MS, Iqbal A, Al-Emadi N, Reaz MBI, Islam MT (2020) Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8:132 665–132 676
104. Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Kashem SBA, Islam MT, Al Maadeed S, Zughaier SM, Khan MS, Chowdhury ME (2021) Exploring the effect of image enhancement techniques on COVID-19 detection using chest x-ray images. Comput Biol Med 132:104319
105. Rahman T (2022) COVID-19 radiography database, Available at https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database. Accessed 09 Sep 2022
106. Lindawangg (2022) Covidx dataset, Available at https://github.com/lindawangg/COVID-Net/blob/master/docs/COVIDx.md. Accessed 09 Sep 2022
107. Gunraj H, Wang L, Wong A (2020) Covidnet-ct: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest CT images. Front Med 7:1025
108. Gunraj H (2022) Covidx CT, Available at https://www.kaggle.com/datasets/c395fb339f210700ba392d81bf200f766418238c2734e5237b5dd0b6fc724fcb/versions/1. Accessed 09 Sep 2022
109. Hgunraj (2022) Covidx CT-3," Available at https://www.kaggle.com/datasets/hgunraj/covidxct. Accessed 09 Sep 2022
110. Zhao J, Zhang Y, He X, Xie P (2020) COVID-CT-dataset: a CT scan dataset about COVID-19," arXiv preprint arXiv: 2003.13865, 490
111. Jkooy (2022) COVID-CT, Available at https://github.com/UCSD-AI4H/COVID-CT. Accessed 09 Sep 2022
112. NHS (2022) National COVID-19 chest image database (nccid), Available at nhsx.github.io/covid-chest-imaging-database/. Accessed 09 September 2022
113. Bachir (2022) COVID-19 chest xray, Available at www.kaggle.com/bachrr/covid-chest-xray. Accessed 09 Sep 2022
114. Zhang K, Liu X, Shen J, Li Z, Sang Y, Wu X, Zha Y, Liang W, Wang C, Wang K, Ye L (2020) Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. Cell 181(6):1423–1433
115. CNCB-NGDC (2022) AI diagnosis, Available at ncov-ai.big.ac.cn/download?lang=en. Accessed 09 Sep 2022
116. Muhammedtalo (2022) COVID-19 x-ray image dataset, Available at github.com/muhammedtalo/COVID-19/tree/master/X-Ray. Accessed 09 Sep 2022
117. Jannisborn (2022) COVID19_ultrasound, Available at github.com/jannisborn/covid19_ultrasound/tree/master/data. Accessed 09 Sep 2022
118. D v7 Labs (2022) COVID-19 chest x-ray dataset, Available at https://darwin.v7labs.com/v7-labs/covid-19-chest-x-ray-dataset. Accessed 09 Sep 2022
119. Aysendegerli (2022) Qata-cov19 dataset, Available at www.kaggle.com/aysendegerli/qatacov19-dataset. Accessed 09 Sep 2022
120. AI AS (2022) COVID-19 CT segmentation dataset, Available at https://medicalsegmentation.com/covid19/. Accessed 09 Sep 2022
121. Ma J, Ge C, Wang Y, An X, Gao J, Yu Z, Zhang M, Liu X, Deng X, Cao S, Wei H, Mei S, Yang X, Nie Z, Li C, Tian L, Zhu Y, Zhu Q, Dong G, He J (2020) COVID-19 CT lung and infection segmentation dataset, Zenodo

122. Zenodo (2022) COVID-19 CT lung and infection segmentation dataset, Available at https://zenodo.org/record/3757476#.YxwGenbMJPa. Accessed 09 Sep 2022
123. Ma J, Wang Y, An X, Ge C, Yu Z, Chen J, Zhu Q, Dong G, He J, He Z, Nie Z, Yang X (2020) Towards efficient COVID-19 CT annotation: a benchmark for lung and infection segmentation, arXiv e-print: pp.arXiv-2004
124. Ma J (2022) COVID-19-CT-Seg-Benchmark, Available at https://gitee.com/junma11/COVID-19-CT-Seg-Benchmark/tree/master#datasets. Accessed 09 Sep 2022
125. GoogleDrive (2022) COVID-SemiSeg, Available at https://drive.google.com/file/d/1bbKAqUuk7Y1q3xsDSwP07oOXN_GL3SQM/view. Accessed 09 Sep 2022
126. BSTI (2022) COVID-19 imaging database, Available at https://www.bsti.org.uk/training-and-education/covid-19-bsti-imaging-database/. Accessed 09 Sep 2022
127. RAIOSS (2022) Coronacases, Available at coronacases.org/. Accessed 09 Sep 2022
128. ChestImaging (2022) Chest imaging, Available at https://twitter.com/chestimaging. Accessed 09 Sep 2022
129. SIRM (2022) COVID-19 database, Available at https://sirm.org/covid-19/. Accessed 09 Sep 2022

Kashfia Sailunaz completed her MSc in Computer Science at the University of Calgary, Alberta, Canada. She is currently a PhD candidate at the University of Calgary, Alberta, Canada. Her research interests include machine learning, social media analysis, and medical image analysis.

Tansel Özyer is a professor of Computer Enginnering at Ankara Medipol University, Ankara, Turkey. He completed his BSc in Computer Engineering and Information Sciences from Bilkent University, Ankara, Turkey, MSc in Computer Engineering from Middle East Technical University, Ankara, Turkey, and PhD in Computer Science from University of Calgary. He is an associate editor SNAM journal and NetMAHIB journal. He is the coordinator of ASONAM conference. His research interests include data science, social network analysis, machine learning, homeland security, health informatics, and image processing.

Jon Rokne is a professor of Computer Science at the University of Calgary, Alberta, Canada. He is an associate editor SNAM journal and NetMAHIB journal. He is the general chair of ASONAM conference. He is a co-editor of the Encyclopedia on social network analysis and mining by Springer. His research interests include computer graphics, data science, social network analysis, machine learning, health informatics, and image processing.

Reda Alhajj is a professor of Computer Science at the University of Calgary, Alberta, Canada. He is also affiliated with Istanbul Medipol University, Istanbul, Turkey, and University of Southern Denmark. He is a co-editor of the Encyclopedia on social network analysis and mining by Springer, editor in chief of SNAM journal and NetMAHIB journal. He is founding steering chair of ASONAM conference. His research interests include data science, social network analysis, machine learning, homeland security, health informatics, and image processing.