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

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

AI for COVID-19 Detection from Radiographs: Incisive Analysis of State of the Art Techniques, Key Challenges and Future Directions

R. Karthik a,*, R. Menaka a, M. Hariharten b, G.S. Kathiresanc

a Centre for Cyber Physical Systems, Vellore Institute of Technology, Chennai, India
b School of Computing Sciences and Engineering, Vellore Institute of Technology, Chennai, India
c School of Electronics Engineering, Vellore Institute of Technology, Chennai, India

Abstract

Background and objective: In recent years, Artificial Intelligence has had an evident impact on the way research addresses challenges in different domains. It has proven to be a huge asset, especially in the medical field, allowing for time-efficient and reliable solutions. This research aims to spotlight the impact of deep learning and machine learning models in the detection of COVID-19 from medical images. This is achieved by conducting a review of the state-of-the-art approaches proposed by the recent works in this field.

Methods: The main focus of this study is the recent developments of classification and segmentation approaches to image-based COVID-19 detection. The study reviews 140 research papers published in different academic research databases. These papers have been screened and filtered based on specified criteria, to acquire insights prudent to image-based COVID-19 detection.

Results: The methods discussed in this review include different types of imaging modality, predominantly X-rays and CT scans. These modalities are used for classification and segmentation tasks as well. This review seeks to categorize and discuss the different deep learning and machine learning architectures employed for these tasks, based on the imaging modality utilized. It also hints at other possible deep learning and machine learning architectures that can be proposed for better results towards COVID-19 detection. Along with that, a detailed overview of the emerging trends and breakthroughs in Artificial Intelligence-based COVID-19 detection has been discussed as well.

Conclusion: This work concludes by stipulating the technical and non-technical challenges faced by researchers and illustrates the advantages of image-based COVID-19 detection with Artificial Intelligence techniques.

© 2021 AGBM. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

The novel coronavirus, termed COVID-19, is recently declared as an infectious pandemic by the World Health Organization (WHO). This has created a state of alarm all over the world. While nations look for ways to contain its spread, COVID-19 remains prevalent with total global cases of over 30.6 million, claiming 950,000 lives [1]. With its high rate of spread, most health organizations and hospitals are under-prepared to handle the inflow of cases. The SARS-CoV2 virus is reported to spread through small droplets and potentially aerosol, with an incubation period of 2–14 days. COVID-19 positive patients can display symptoms like fever, dry cough, body aches, shortness of breath, loss of taste and smell, sore throat, and diarrhea [2]. They might be asymptomatic carriers as well. With such easily misinterpreted symptoms and the risk of the consequences of a wrong diagnosis, the proper identification of the virus infection is one of the topmost priorities for medical organizations.

Artificial Intelligence provides help in tackling the virus outbreak by assisting in clinical decision-making and predictive analysis. The use of Deep learning techniques has been at the forefront of accurate prediction and diagnosis of virus-infected cases. Deep learning can be used to classify infected and normal patients or perform segmentation of the infectious regions from Chest X-rays and Computed

* Corresponding author.

E-mail addresses: r.karthik@vit.ac.in (R. Karthik), menaka.c@vit.ac.in (R. Menaka), hariharan.m2016@vitstudent.ac.in (M. Hariharten), gugani.kathiresan2018@vitstudent.ac.in (G.S. Kathiresan).

https://doi.org/10.1016/j.irbm.2021.07.002
1950-0318 © 2021 AGBM. Published by Elsevier Masson SAS. All rights reserved.
Tomography (CT) scans. These AI diagnostics models could ease the load on healthcare workers allowing them to dedicate time to the care of patients and the development of a vaccine.

Early detection of the presence of infection is critical for providing treatment to save lives. Reports show that symptoms can start from a simple cold and can develop into life-threatening pneumonia. Though every individual is equally likely to be infected, the severity of its effects varies based on the age group. The severity is high for elder age groups and people with pre-existing co-morbidities like heart conditions, obesity, diabetes, and respiratory illnesses like asthma. Since the virus spreads through sneeze or cough droplets in the air, minimizing contact between individuals helps to curb the virus’s primary mode of transmission. To avoid this spread through close contact, countries have imposed lockdowns, social distancing, and quarantine restrictions.

The current major mode of diagnosis is the RT-PCR (Reverse Transcription-Polymerase Chain Reaction) testing. It detected the presence of the virus based on swab testing or blood samples. RT-PCR can give results in a period of a few hours to 2 days with an accuracy interval of 81 - 96%. But these tests cannot identify the severity of the infection and the accuracy of the results depends on the strength of the virus strain. And with the large amounts of samples flowing in, the basic differentiation between the coronavirus infections and other infections becomes a significant step towards accurate diagnosis.

The use of X-rays and CT scans for the identification of pneumonia proves to be a very useful method of diagnosis. Since the virus attacks the respiratory tracts of humans, medical imaging can be used as a diagnostic tool for detecting COVID-19 infection. X-ray and CT scan machines are already available at most hospitals, allowing them to assist in the diagnosis without the need to procure testing kits. These imaging tests highlight opacities and abscesses which are the identifiable patterns that characterize the viral infection.

Chest X-ray images for positive cases show bilateral diffuse patchy opacities with some bibasilar sparing [3], which can help identify and analyze pneumonia, lung aggravation, and developed lymph hubs. CT scans for COVID-19 cases tend to bring out patterned distributions of opacities (interlobular septal thickening superimposed on ground-glass opacities) [4]. The main goal of identifying these patterns and their density is to perform accurate diagnosis, determine the severity of the disease, and guide prognosis.

The use of Artificial Intelligence (AI) techniques for identifying cases of infection and its radiological properties from medical imaging like Chest X-ray and CT scans have proven to be effective for accurate diagnosis [5–33]. Many approaches can be taken in Machine Learning and Deep Learning to address COVID-19 detection and segmentation tasks. AI-guided medical imaging analysis offers huge potential to serve as a primary diagnostic tool for COVID-19 detection [34–41]. Learning deep features that can distinctly capture COVID-19 radiological patterns in chest X-ray and CT forms the first step in the diagnosis. The potential of Machine Learning-based predictive mechanisms can be realized in the prognostic analysis [42]. Hence many studies harness algorithms like SVMs, Random Forests, to provide the necessary insight(s) into the prediction and diagnosis of the coronavirus infection. These automation systems can ease the strain on healthcare workers and can take over the filtering process of COVID-19 diagnosis. The early diagnosis of infection can save valuable time, allowing treatment to be administered in the preliminary stages of the infection avoiding the consequences of risk.

The consequences of misdiagnosis pose a severe risk, which can even be life-threatening for the infected patient. Automated systems face many challenges due to the velocity and volume of data being huge. With the large inflow of cases, the cleaning and processing of data become a huge obstacle especially with the need for high-resolution images. These systems must also ensure the prudence of data, as cases are influenced by factors like age, gender, and pre-existing conditions. The addition of these cases can potentially impact the quality of the results.

Through this work, we aim to achieve the following objectives.

- To discuss the latest trends in the application of AI techniques for COVID-19 detection from medical images.
- to compare and analyze the performance of the recent works in COVID-19 detection and to recognize areas of improvement.
- to highlight the gaps and drawbacks of the reviewed works and identify the potentia l scope for future research.

2. Research search strategy

In this study, we have referred to research publication databases like ScienceDirectTM, IEEEExploreTM, and Google ScholarTM to seek and obtain other studies and publications related to COVID-19 diagnosis from medical imaging. Many articles were sourced from Springer and Wiley publications as well. These articles were screened in the context of Artificial intelligence methods which specified the use of medical imaging. We examined multiple publications and screened the papers before the actual collection, ensuring no redundancy and unreliability in the collection. In the preliminary screening, publications were filtered based on the title and journal publication. The subsequent screening was based on the abstract, image modality, and methods applied. Finally, an in-depth analysis of the details and implementation aspects of the method employed enabled an informed final screening. These steps resulted in a shortlist of 140 publications to be included in this review. In this review, we believe to have considered the key contributions of all AI-based research publications reported till 2020 for COVID-19 diagnosis from medical imaging. The search process used the keywords like ‘COVID-19’, ‘X-ray’, ‘Computed Tomography’, ‘Artificial intelligence’, ‘Deep learning’, ‘Machine Learning’, ‘Imaging’, ‘Classification’, ‘Segmentation’, ‘Deep Features’, ‘Neural Network’, and ‘Computer-aided diagnostic support system’ to download the relevant publications. A period of 27 days was involved to download these 148 publications and perform the screening. Table 1 provides details about the inclusion and exclusion criteria. Fig. 1 presents the month-wise split-up of the obtained articles based on the modalities used. The search strategy applied and the manuscript filtering process is illustrated in Fig. 2. The comprehensive assessment of the obtained publications involved a search for specific details, is illustrated in Table 2 and 3 for not able studies. To the best of our knowledge, this is the first review report consolidating the findings of different AI approaches for COVID-19 detection and segmentation based on the chest X-ray/CT imaging modalities.

The review is organized as follows. In Section 2. We present the details of the search strategy employed to collect and filter the research articles from different sources. Section 3 highlights the significance of the modalities of medical imaging used for COVID-19 diagnosis. Section 4 presents the consolidated review of the methods employed by the different publications, categorized into classification and segmentation. Section 5 presents the various datasets that the research works have employed. It also includes the discussion, evaluation metrics used by the works, challenges, and future scope, and the limitations of this study. Finally, Section 6 concludes the research work describing the significance of image-based AI methods in COVID-19 diagnosis. The quality assessment form utilized to review each publication is presented in the appendix.
Table 1
Inclusion and exclusion criteria.

| Factor                | Inclusion                                                                 | Exclusion                                                                                           |
|-----------------------|---------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| Dataset and research  | • Studies dealing with CT, X-ray, or combined imaging modalities.        | • Studies involving animal data.                                                                     |
| outcome               | • Studies using real-time dataset samples obtained from hospitals/clinical labs. | • Studies that did not involve COVID-19 related experiments.                                          |
|                       | • Studies detecting the presence of COVID-19, region of the infection or providing localization details. | • Studies not involving imaging modalities like X-ray or CT                                           |
|                       | • Studies describing insights into CT/X-ray manifestation of COVID-19 through automation. | • Studies describing treatment protocols/medical condition of the patient pre and post COVID-19.   |
| Study design and      | • Studies involving AI techniques for required modality                  | • Studies explaining the scientific working of the imaging modalities employed in the research articles. |
| Methodology           | • Studies involving Deep Learning or Machine learning techniques         | • Studies that are related to biochemical research on COVID-19 infections                            |
|                       | • Studies involving Classification of COVID-19 images                     | • Studies that are not related to AI methods.                                                        |
|                       | • Studies involving segmentation of COVID-19 infection region             | • Studies that consider other infections using the required modality.                               |

Fig. 1. Month-wise distribution of published papers split based on imaging modalities.

3. Radiological imaging for Covid-19 detection

Medical imaging plays a crucial role in identifying the diagnosis of COVID-19 infected patients and in determining the prognosis as well [43]. This section describes the common COVID-19 findings observed in X-ray and CT.

3.1. Chest X-ray manifestation of COVID-19

Chest X-ray-based COVID-19 diagnosis has shown promising scope and serves as a reliable and quick tool in terms of availability and cost [44]. Hospitals can perform Chest X-rays as a swift screening method for patient admission [45,46]. While many medical organizations employ CT scans as well, the radiation dose and the sanitization process set an obstacle to efficient diagnosis [47]. In these conditions, Chest X-rays become a better alternative with their portability and availability. Due to these advantages, Chest X-rays are used as the primary COVID-19 investigation modality in the United Kingdom (UK) [48].

The observable features seen in COVID-19 affected Chest X-rays are commonly manifested as a spectrum of pure ground glass, mixed ground glass opacities to consolidation in bilateral peripheral middle and lower lung zones [49]. These alveolar opacities are often found with superimposed atelectasis, appearing predominantly in the lower lobe [50]. As the disease progresses beyond the early stage, Chest X-rays can display multiple patchy opacities, which eventually, become confluent and severe cases may appear as a “whited out lung” [43].

Chest X-ray diagnosis provides an 89.0% sensitivity (95% CI, 85.5%-91.8%) as explained by Schiaffino et al. [51]. This draws a significant comparison with RT-PCR tests for which Gatti et al. [52] had reported a sensitivity of 61% (95%CI 55–67%). In another study by Stevens [48], Chest X-ray sensitivity for detecting COVID-19 was reported to be 85% compared to 93% for the initial RT-PCR. Over the course of the prognosis, Stephanie et al. found the sensitivity of the Chest X-ray increased over time [53]. It was also found that serial Chest X-ray imaging has the potential of reaching the diagnostic accuracy of chest CT [52]. Experiments by Cozzi et al. [54] revealed that
the diagnostic performance of Chest X-rays had an overall accuracy (76 %–86 %), with an 89% sensitivity. A 66% sensitivity for more experienced radiologists (66%), and 41% for the less experienced radiologists [54].

To put into perspective, as supported by various authors [55–57], the exigency of Chest X-rays in the diagnosis of COVID-19 is bolstered by certain factors:

(1) the lack of a reliable, time-efficient diagnosis;
(2) possible obstacles faced in the availability of CT scanning machines and difficulties in the fast sanitization of CT rooms;
(3) ease of Chest X-ray performance in isolated rooms in the Emergency Room.

Another factor to consider would be the very high pre-test probability associated with the disease spectrum. This may be skewed due to factors like high disease severity and reduction in the presence of non-COVID-19 pneumonia. These factors in turn resulted in better bedside performance of diagnosing COVID-19 from Chest X-ray [58].

3.2. Chest CT COVID-19 patterns

The use of chest CT imaging as a potential tool to infer COVID-19 has been demonstrated by several clinical studies. The main factors that favor the use of chest CT modality in COVID-19 diagnosis are the high sensitivity and low miss rate [59–64]. CT has registered 97% sensitivity in a study with 1014 patients in Wuhan, China [59]. CT features were able to identify 75% of the false-negative cases that were not reported by PCR test. In a similar experiment, it was reported that a sensitivity of 96.07% for CT-based screening [60]. The predominant patterns observed in these studies include Ground-Glass Opacities, consolidation, and septal thickening. CT scans with bilateral lung involvement are also indicative of COVID-19, which is reflected in a study conducted on 34 subjects [61]. Other common CT manifestations are bilateral patchy shadowing, which was observed in 86.2% of confirmed cases in the experiment by Guan et al. [62]. The highest sensitivities of 98% and 97% for CT-based COVID-19 detection were registered in the clinical studies in Shanghai, China [62] and Italy [63].

The most prevalent manifestations of COVID-19 on CT are the peripheral and posterior ground-glass opacities with or without consolidation [59,65–69]. In patients showing severe symptoms, septal thickening in the interlobular region and air bronchogram were observed [60,67–69]. Aside from these features, crazy paving pattern and reverse halo sign as CT features for detecting COVID-19 [60,66,70]. These observations form the preliminary evidence for assessing CT manifestations of COVID-19.

4. AI techniques for COVID-19 detection and infection delineation

This organized review outlines the various AI techniques employed in recent literature for the detection of COVID-19 from medical imaging. As an overview, the majority of the literature on AI for COVID-19 detection is in the deep learning domain, although few works

---

**Fig. 2. Overview of the search strategy and manuscript filtering process.**
have explored the ML methods. The fully automatic deep learning techniques learn the feature extraction directly from the image data. CNNs for deep feature representation and classification have shown excellent results in medical image processing and have worked so well in the COVID-19 detection task. The knowledge of the prominent features, patterns learnt from data efficiently facilitate diagnostic assistance for the clinicians.

Deep neural networks are a learning algorithm that stack multiple neuronal nodes in a layer-wise fashion. They are gradient-based learners that tune their parameters in response to reducing the error made by the model towards classification/segmentation. This demands precisely setting up the model training with stratified-class sampling, controlling learning rate over the epochs, and performing a hyper-parameter search. The deep learning paradigm has made rapid strides in healthcare automation by offering enormous design possibilities that can be customized specifically to the task and dataset at hand. With the computational power of Graphics Processing Unit (GPU) and distributed computing models, these deep learning architectures can be trained and tested in a relatively quick time. In the scope of COVID-19 detection, some studies have explored a range of CNN methods, ML classifiers on deep features, capsule networks, RCNN, etc. This section reviews the different state-of-the-art AI-based methods applied for COVID-19 detection.

4.1. COVID-19 classification from radiographic imaging

In this section, various research methods that perform COVID-19 classification are extensively reviewed. These works have utilized the two major imaging modalities (chest X-ray/CT) for the COVID-19 detection task. The key insights revealed in these works are discussed in detail.

4.1.1. Chest X-ray based classification

Chest X-ray images are one of the most accessible and widely used imaging modalities in the AI literature for finding COVID-19. The X-ray based detection methods can be broadly categorized into one of these categories: 1) Transfer learning approaches [71–106], 2) customized deep architectures [107–115], 3) capsule networks and sequential CNN [118–124], 4) semi-supervised, GAN approaches [125–132], 5) deep feature extraction and image processing techniques [133–138,140–154,209–211], 6) CAD methods [155–157] and optimization algorithms [116,117,125,139].

Transfer learning models use the knowledge learnt from previous experience, through modification or addition of customized layers to fit the dataset. This arena of research has been predominant in the research on CNN-based COVID-19 findings. Within this category, the methods can be further sub-classified into VGG networks, Residual networks, Inception, Xception CNNs, and a mixture of several architectures.

Many works have analyzed the effectiveness of a VGG16-based deep Learning model for the identification of pneumonia and COVID-19. For instance, Civit-Masot al. proposed a VGG-16 model with a pre-processing stage that performed final classification using the confidence parameter obtained after the training [71]. Several other works, including Pandit al. [72], Manapure et al. [73], Do et al. [74], Vaid et al. [75], Heidari et al. [76] modified the VGG16 model with a customized fully connected layers for classification. In another study, Makris et al. performed an architectural fine-tuning of several CNN architectures via transfer learning and projected that VGG16 and VGG19 models had shown the best results [77]. Alazab et al. also applied the VGG16 model to identify COVID-19 from augmented chest X-ray data [78]. The data augmentation was achieved through geometric image transformations. In the line of researches utilizing VGG-based transfer learning, Panwar et al. proposed a fine-tuning of the pre-trained VGG-16 with custom layers and hyper-parameters [79]. Similarly, with the classic approach, Zhu et al. used a VGG16 model trained off the ImageNet dataset, and correlation analysis between AI-predicted versus radiologist scores to optimize predictions [80]. In another study by Dansana et al. took the binary classification approach for COVID-19 pneumonia [81]. The proposed architecture used a fine-tuned VGG-16 and InceptionNet-V2 architecture, after performing pre-processing to generate the vectorized feature maps. Using a different combination, Karar et al. used pre-trained CNN architectures in an 11 model cascaded configuration, including VGG and ResNet [82]. Barbano et al. applied a transfer learning mechanism using multiple standard CNN architectures (VGNet, ResNet) pre-trained on the chest X-ray pathology datasets. These pre-trained feature extraction networks are then fine-tuned on the COVID-19 datasets [83].

Residual learning is a common design paradigm in most of the CNN works, owing to their ability to prevent the vanishing gradient problem. Misra et al. presented a multi-channel pre-trained ResNet architecture to facilitate the diagnosis of COVID-19 chest X-rays [84]. Three ResNet-based models were then retrained to classify X-rays on a one-against-all basis. Comparing multiple architectures, Jain et al. employed a four-phase model that included pre-processing, augmentation, and two-stage transfer learning model [85]. These two stages were based on different ResNet architectures; the first stage extracted the viral pneumonia features from other pneumonia and the second stage for COVID extraction from other viral pneumonia. In another type of combination, Rahimzadeh et al. proposed a concatenation-based configuration of transfer learned models [86]. The ResNet50V2 and Xception models extracted deep features, which were then concatenated, resulting in better classification from both the feature vectors. Another similar work by Benbrahim et al. applied the pre-trained ResNet50 and InceptionV3 transfer learning architectures with logistic regression for COVID-19 identification [87]. El-Rashidy al. extended a similar ResNet50 CNN architecture with fully connected layers for X-ray classification [88]. The CNN was trained as a part of a remote patient monitoring IoT framework for data acquisition, learning, and cloud storage. Another customization proposed by Punia et al. used pre-trained ResNet34 and ResNet50 and included an ADADELTA based learning strategy for the classification of a constructed X-ray dataset [89]. Since airspace opacities in the X-ray have been prominently associated with COVID-19, Azemin et al. employed a Resnet based CNN to train for the task of finding airspace opacities in chest X-rays [90].

Several other works have compared the performance of multiple transfer learning CNNs. For instance, Minae et al. presented results for four different architectures, including ResNet18, ResNet50, SqueezeNet, and DenseNet-121 on a custom constructed dataset [91]. Many works are focused on the inclusion of Inception and Xception networks in performance comparison. Sethi et al. proposed a “recommendation network” which used four different pre-trained architectures, namely Xception, ResNet50, MobileNet, and Inception V3 [92]. Similarly, Khan et al. used four pre-trained deep-learning models (DenseNet121, ResNet50, VGG16, and VGG19) for the diagnosis of X-ray images as COVID-19 or normal [93]. Apostolopoulos et al., employed transfer learned VGG19, MobileNet v2, Inception, Xception, and ResNet v2 models for classifying COVID-19 X-rays [94]. DenseNet was a common inclusion in many transfer learning works. To detect COVID-19 manifestations on X-ray, Blain et al. developed a DenseNet based model for finding alveolar and interstitial opacities in the U-net segmented.
X-ray lung regions [95]. Similar to that, Elasnnaoui et al. performed automatic multi-class classification with transfer learning on seven standard CNN models, including the VGG-net, Densenet, Inception-Net, ResNet, and MobileNet [96], Rahaman et al. provide a performance comparison of several transfer learning approaches with standard deep learning models [97]. In a similar approach, Boudrioua et al. modified the DenseNet121, NASNetLarge, and NASNetMobile with fully connected layers for X-ray classification [98]. Albahl et al. performed transfer learning with three architectures - NASNetLarge Inception NetV3, Inception ResNet V2 on an augmented chest X-ray dataset [99].

In the line of research works that train with augmented data, Phankokkruad developed CNN models through transfer learning from Xception, VGG16, and Inception-ResNet-V2 on a large X-ray dataset [100]. Modified ResNet, VGG16, Xception, and Inception networks have also been used in the work by Ko et al. for COVID-19 classification [101]. Xception networks for COVID-19 classification have been explored in several works. Singh et al. [102], Khan et al. [103], Lujan et al. [104] have designed transfer learning models based on Xception network architecture for accurately finding COVID-19 from chest X-rays. Comparably, Horry et al. presented a multi-modal classification model that employed a classification pipeline that took augmented input data and was tested on 8 different transfer learning architectures [105]. Transfer learning from other architectures, like the DarkNet model, was also tried in the literature. In the case of Ozturk et al., the authors utilized a customized DarkNet model with fewer layers and filters, gradually increasing them based on experimental results [106].

Customized CNN architectures that are different from existing CNN designs are tailor-made to handle the specific task and the dataset at hand. They more often generalize well to the real world and are highly optimized for the specific use case. For instance, Abbas et al. presented a cascaded 4-stage deep CNN for the classification [107]. It identified image irregularities in its class boundaries using a class decomposition mechanism. With an external classifier, Yoo et al. presented a composition of three binary decision trees, each trained by a CNN model [108]. Das et al. used a customized deep CNN model to extract low-level features which were then classified on an Xception network [109]. In a feature engineering approach, Turkoglu performed relief feature selection on deep features from pre-trained AlexNet CNN for the classification of COVID-19 X-rays [110]. Haque et al. presented a work describing multiple linearly stacked CNN architectures with convolutional and pooling layers [111].

The need for efficient and reliable imaging-based COVID-19 diagnosis has driven many authors to pursue novel approaches to address a solution. For instance, Karthik et al. proposed a novel distinctive convolutional filter learning paradigm for characteristically identifying patterns associated with COVID-19 and other pneumonias on the X-ray [112]. In another approach, Ouchicha et al. presented residual CNN with dual convolutional pathways, to extract 16 feature maps from X-rays [113]. In another novel approach, Abdani et al. presented a compact architecture based on a custom 14 layer convolutional network that employed spatial pyramid pooling for multi-scale classification [114]. With the transfer learning approach, Das et al. proposed a truncated architecture to reduce model complexity and to reduce over-fitting [115]. The truncated InceptionV3 based architecture was pre-trained on the ImageNet database and used an adaptive learning rate protocol for training. In a similar work to optimize pre-trained InceptionV3, Bridge et al. introduced a Generalized Extreme Value distribution-based activation function, to integrate with the Inception model. This led to improvements in the classification performance over models with traditional activation methods on unbalanced datasets [116]. The GreyWolf Optimizer (GWO) algorithm was used by Goel et al. to optimize the CNN feature extraction and classification components [117].

Several works validate the effectiveness of the Capsule network. To study the performance of the Capsule net configurations, Asfah et al. proposed the COVID-CAPS model that was pre-trained on an external X-ray dataset [118]. To further understand its contribution, Toraman et al. presented a capsule network-based model that included 5 custom convolutional layers to provide better feature maps [119]. Dhaya et al. also experimented on a capsule model that was optimized with weighted cross-entropy loss [120]. Similar to the capsule net, Chowdhury et al. proposed a parallel-dilated CNN-based COVID-19 detection system, from enhanced chest X-ray images [121].

Apart from the complex architectures, many researches have employed relatively simpler architectures to achieve an acceptable degree of model fitting. In the early dated works, sequential CNN for COVID-19 classification from X-ray was common. For instance, Bakhrani et al. [122], Haque et al. [123], and Salman et al. [124] presented a sequential CNN for classifying COVID-19 infected X-rays.

To include the advantages of data augmentation and network optimization, Ucar et al. proposed a Deep Bajes-SqueezeNet based COVIDDiagnosis-Net [125]. The network was based on the pre-trained SqueezeNet architecture that utilized Bayesian optimization, and also employed offline augmentation for a chest X-ray dataset. In another study, Zebin et al. have used multiple pre-trained convolutional backbones as a feature extractor with a CycleGAN to increase sample count [126]. Rajaraman et al. proposed a weakly labeled data augmentation approach to solve the dearth of COVID-19 chest X-rays for the infection detection task [127]. A CNN trained on pneumonia dataset was used to generate labels for augmenting COVID-19 samples. To improve the results with augmentation on a constructed dataset, Marques et al. employed customized EfficientNetB4-based layers for binary and multi-class classification experiments [128].

Data generation with GAN has been an important approach in the COVID-19 AI literature. For instance, Waheed et al. proposed a synthetic dataset generation method using an Auxiliary Classifier GAN-based model, used with VGG-16 for classification [129]. Sedik et al. proposed to train a convolutional GAN model for data augmentation that enhances the learnability of the CNN and Conv-LSTM discriminator models for COVID-19 classification [130]. Loey et al. presented an approach that trains a GAN for COVID-19 X-ray data augmentation [131]. The augmented X-ray dataset was classified on a CNN that additionally predicts candidate bounding box regions with high inference probability. Zulkifli et al. presented DC-GAN in the augmentation of chest X-ray data followed by a lightweight CNN to perform classification on the enhanced dataset [132]. The lightweight model design was based on multi-scale spatial pyramid pooling and DenseNet layers.

Some studies focus on rigorous feature extraction for COVID-19 from Chest X-rays. In the case of Abraham et al., the authors utilized multi-CNN feature extractors and the combined features used as input to a Bayesian classifier [133]. The multi-CNN was made up of 10 pre-trained architectures for feature extraction, and the feature selection was done by a Correlation-based Feature Selection (CFS) algorithm with Subset Size Forward Selection (SSFS) to determine the optimal feature subsets. A study by Tuncer et al. proposed a novel feature generation model with feature extraction by using Residual exemplars local binary patterns, and iterative Relief for feature selection [134]. This was verified with 5 different classifiers, being Decision Trees, Linear Discriminant, kNN, and SVM. Altan et al. presented a hybrid model which employed 2D curvelet transform, with a Chaotic Salp Swarm Algorithm (CSSA) optimized feature matrix, which was based on the EfficientNet-B0 model [135]. Noise in the X-ray data can be compensated through pre-processing. Togrça et al. proposed a pre-processing technique that used Fuzzy color to remove noise, which was then combined with the original dataset using a stacking algorithm [136]. This stacked dataset was used with MobileNet and SqueezeNet models and a SVM classifier. In a similar attempt, Vinod et al. discussed the use of a pre-trained CNN model and Decision Tree for the fast classification of pre-processed COVID-19 X-ray and CT images [137].
To investigate the use of other classification algorithms, Islam et al. developed a combination of CNN for complex feature extraction, and LSTM as the classifier for the proposed architecture [138]. To combine the results of pre-processing and deep features, Nour et al. presented a custom CNN-based feature extraction model coupled with Bayesian optimization, used with offline augmented data [139]. The features were then given to various ML-based classifiers like SVM, kNN, and Decision trees. Comparably, Bharati et al. proposed a hybrid VGG, data augmentation, and spatial transformer network (STN) pre-processed model, termed as VDSNet [140]. Alqudah et al. applied the features extracted from the Advanced Optical Coherence Tomography network (AOCT-net) with ML classifiers as a hybrid transfer learning approach [141]. Pereira et al. introduced a hierarchical classification schema for pre-trained CNNs and feature extractors [142]. The model considered different feature extraction methods for Early Fusion, Data re-sampling, and a multi-classification approach. To draw a comparison between neural network architectures, Varella-Santos et al. utilized feed-forward and CNN models to classify custom databases to develop a baseline model for X-ray and CT-based COVID-19 diagnosis [143]. A feature extraction method was also considered, namely Texture Features, which included GLCM and BLP methods for the same. In another study, Apostolopoulos et al. used an end-to-end trained MobileNet v2 architecture, with a Relief feature selection algorithm and SVM to classify COVID-19 X-rays images [144]. Lopez et al. presented a customized CNN approach that used histogram contrast equalized X-ray images were used to train the model [145]. In the work by Sahlo et al. the InceptionNet CNN was used to extract deep features, which were refined with the fractional-order marine predators' algorithm for COVID-19 classification [146]. Novitasari et al. presented a COVID-19 detection framework that uses CNN for feature extraction, PCA for feature selection, and SVM for classification [147]. Searching for optimal network configurations is an important factor in the design aspect of the CNN towards enhancing performance. For instance, Singh et al. proposed a hyper-parameter tuned CNN that performs softmax classification for COVID-19 on chest X-rays. The model was fine-tuned with Multi-objective adaptive differential evolution [148]. Sekeroglu et al. experimented with customized convolutional network architectures for multiple classification criteria, such as COVID-19/normal, COVID-19/pneumonia [149]. In the study conducted by Woźniak et al., a probabilistic neural network was used to identify lung carcinomas from X-ray scans [209]. The local variance around the pixels was used to generate possible lung nodules, which were then discriminated using a probabilistic neural network. Ke et al. presented a heuristic classification technique to identify degenerated tissues and lung diseases from X-ray images [210]. Employing a bio-inspired approach, Woźniak et al. proposed a feature-specific configuration of models that search for lung disease features in X-ray images [211]. This configuration is based on heuristic models that perform a search with a dedicated fitness function.

Some works have also employed object detection CNNs for COVID-19 detection. Shibly et al. used a Faster RCNN based architecture with inherited features from a VGG based network [150]. Saiz et al. applied an SSD model with VGG-16 backbone to generate candidate infection region proposals on X-ray and to perform COVID-19 instance classification [151]. Some studies have proposed ensemble-based approaches for COVID-19 classification. Chandra et al. presented a majority voting-based classifier that consisted of an ensemble of 5 different classification algorithms [152]. The model used GLCM and HOG-based feature extractors and supervised classifiers, namely, SVM, ANN, kNN, Naïve Bayes, and Decision Trees. Ahsan et al. proposed a fused MLP-CNN model to simultaneously fit mixed numerical data and imaging data [153]. Rajaraman et al. used a configuration that combined results of iteratively pruned classification models in a stacking ensemble [154].

The use of commercial AI software for testing has helped study the models already trained to identify X-ray features. For instance, Hwang et al. utilized a deep learning-based CAD system for assisting radiologists in diagnosing COVID-19 [155]. The CAD deep learning algorithm was trained to distinguish between normal and X-rays affected by 4 major thoracic diseases, which was exploited by the radiologists for COVID-19 analysis. In the same arena of research in CAD-assisted learning, Mohammed et al. proposed a multi-stage CAD framework for classifying X-rays [156]. The particle swarm intelligence was applied to get lung region segmentation, from which GLCM texture features are extracted and classified on a SVM ensemble. In related work, Murphy et al. explored the CAD commercial software pre-trained for detection of tuberculosis from X-ray for the COVID-19 identification task [157]. The AI model was orchestrated as a U-net for lung segmentation followed by patch-based ensemble networks for infection classification. The software methods provide a verifiable way to comprehend the stage-wise processing of the image towards the classification.

4.1.2. CT based classification

CT-based approaches have mainly employed a variety of feature extraction schemes and ensembling methods to perform COVID-19 prediction. In contrary to chest X-ray literature, only some works have taken the transfer learning approach for CT image classification. Of these works, Pathak, et al. utilized deep transfer learning on ResNet32 with custom layers, for classifying COVID-19 positive and negative CT images [158]. Employing a dual-sampling strategy, Ouyang et al. used the 3D ResNet34 architecture was used as the backbone for the classification [159]. In another study, Silva et al. utilized a pre-trained EfficientNet based architecture [160]. The model also employed a 5-fold evaluation that made use of “Random”, “slices” and “Voting” based scenarios. Hu et al. proposed a strategy for a weakly supervised model that performs CT classification with VGG-16 networks [161]. Xu et al. presented a location attention model based on the ResNet18 model to classify 3d segmented CT images as having COVID-19 pneumonia or not [162]. The effectiveness of residual networks is also captured in the work by Ardakani et al., which presented a comparison of 10 standard CNN architectures [163]. ResNet-101 had resulted in the highest overall accuracy. Perumal et al. proposed a transfer learning-based model along with Haralick features obtained from the enhanced input images [164]. Using the classic approach, Jaiswal et al. applied the ImageNet pre-trained CNN for extracting deep features [165]. Mishra et al. explored various Deep CNN-based approaches for detecting the presence of COVID19 from chest CT images [166]. A decision fusion-based approach was also proposed, which combined predictions from multiple individual models, to produce a final prediction. Tabrizchi et al. presented an approach to use standard ML and DL algorithms for identifying COVID-19 [167]. While MLP, CNN, SVM were single point classifiers, ensemble methods like Adaboost and gradient boosted trees were also used in the detection task.

For CT-based COVID-19 detection, there are several studies with feature extraction as their motivation. For instance, Yan et al. presented a CNN with a multi-scale spatial pyramid (MSSP) decomposition architecture, which learned feature representations of multi-scale inputs that did not require large-scale training data [168]. In combination with a classifier, Shaban et al. proposed a hybrid feature selection methodology and introduced the Enhanced kNN [169]. The methodology included both wrapper and filter feature selection methods while the enhanced classifier added new heuristics to a traditional kNN classifier. Han et al. applied the deep 3d multi-instance learning model to extract features at an instance level. An attention-based pooling of such instance labels is done to derive patient-level classification [170]. Similarly, Li et al. utilized a self-supervised approach, which used a modified Rubik's cube Pro model to extract 3D features and was
also used as the backbone for the classification network [171]. Redesigning the previously proposed pre-trained COVID-Net architecture, Wang et al. made modifications in terms of the network architecture and cosine annealing learning strategy [172]. They further proposed a joint learning scheme to solve the data heterogeneity problem to improve the performance of the model. A similar study by Öztürk et al. utilized a 2-stage classification model with a SVM classifier [173]. The data is subjected to shallow augmentation and was subjected to different feature extraction algorithms, which was then over-sampling using SMOTE algorithm. Hasan et al. used Q-deformed entropy-based texture features and deep CNN features to train a Bi-LSTM classifier for COVID-19 detection from CT slices [174]. The combined feature set was refined through the statistical ANOVA method.

Some of the works propose to include parameter-tuning configurations that build on the classic CNNs. Pathak et al. proposed a deep bidirectional classification model which employs a LSTM network [175]. The Bi-directional LSTM network makes use of a mixture density network, in which the hyper-parameters were fine-tuned using a memetic adaptive differential evolution algorithm. In another attempt, King et al. used an unsupervised clustering-based approach to identify COVID-19 features from X-ray images [176]. They employed a Self-Organizing Feature Map to cluster instances of infections by independently considering each region of the image. Similarly to draw a comparison of these networks, Singh et al. proposed a deep CNN architecture for COVID-19 classification using multi-objective differential evolution. It is a class of genetic algorithms that offer multiple stages of mutation, crossover, and selection in optimizing the search for hyperparameters [177].

Generating candidate regions/bounding boxes for training at a patch level can reinforce classification decisions and also be effectively used in localization. Butt et al. presented a classification model which used pre-processing to extract effective pulmonary regions from CTs [178]. A 3D CNN model was used to identify multiple candidate image cubes. In another study, Ahuja et al. introduced a three-phase detection model including: data augmentation, COVID-19 detection, and abnormality localization [179]. Xiao et al. proposed a multi-instance learning framework with ResNet CNN [180]. In the work by Bai et al., CT slices with existent lung abnormalities were modeled on the EfficientNet B4 architecture [181]. The samples were classified into either COVID-19 or other pneumonia classes. The CT slice probabilities for a patient are pooled through a fully connected neural network to generate a patient-level prediction.

4.2. COVID-19 regions localization

The literature on COVID-19 detection presents several ways to visualize how the manifested infection features are extracted and exploited in classification. Localization techniques, like the GRAD-CAM highlight such hotspot regions in the input image. They make the results seem more interpretable and useful for the clinician. To this end, several works have utilized the Grad-CAM algorithm for localization. For instance, Mahmud et al. proposed a custom CovXNet architecture that used GRAD-CAM for localization [182]. The model was pre-trained on non-COVID-19 X-rays and employs a stacking algorithm for its optimization by generalizing the extracted features. Brunese et al. proposed a three-fold method: first detects healthy or generic disease; second discriminates between generic diseases and COVID-19; and third highlights the areas of the COVID-19 disease in the X-ray [183]. The first two methods employed pre-trained VGG-16 models, and GRAD-CAM to enable visualization. Similarly, Ezzat et al. presented a transfer learned DenseNet121 classification model which employed GSAoptimization, and GRAD-CAM to visualize the localization regions [184]. Similar work on DenseNet-based classification was carried out by Oh et al., in which GRAD-CAM was used to aid better interpretability of the results [185].

ScoreCAM is an improved weighing model over the Grad-CAM. For instance, Fan et al. proposed a DenseNet model for X-ray classification and the score-CAM method for visualizing the infected regions [186]. Exploiting the concept of attention maps, Tsiknakis et al. presented an interpretable AI framework assessed by expert radiologists [187]. The validation of attention maps focuses on the diagnostically-relevant image regions.

Such class activation mapping for salient region identification is used for CT scans too. In the case of CT scans, Sharma et al. discussed the performance of a commercial classification software on a constructed CT image dataset [188]. A ResNet-based architecture of the Microsoft Azure software was used, along with GRAD-CAM for localization. The linear combination of weight and activation graph is used in the infection localization analysis. In another work utilizing the DenseNet CNN, Yang et al. trained on the dataset of CT samples with COVID-19 manifestations [189]. The CAM feature activations were shown to highlight regions of GGO/consolidation. Considering 2 imaging modalities, Panwar et al. took a transfer learning-based approach on the VGG19 model for binary classification and used GRAD-CAM to provide visualization [190].

Region localization can also be achieved through training a model to identify candidate regions from a model prediction. Wang et al. presented a two-stage feature pyramid CNN model that localizes the infection location with residual attention networks [191]. U-net architectures are a common feature in localization tasks. Hurt et al. proposed a U-net based CNN annotated for the task of COVID-19 localization [192]. The predicted pneumonia probability map is used to infer the regions of interest for the radiologist inspection.

4.3. COVID-19 infection segmentation

This section presents a detailed review of various segmentation approaches to detect COVID-19 from X-ray/CT modalities. The key insights and methods described in these works are discussed elaborately.

4.3.1. X-ray based methods

The majority of AI works for COVID-19 identification from X-ray modality have carried out the classification task. There are few works on segmenting COVID-19 affected regions in X-ray. This is mainly because X-ray features are not primarily used in the localization and quantification of COVID-19 in the clinical setting compared to that of CT. CT manifestations of COVID-19 are relatively well explored and their features are predominantly studied for the identification of COVID-19 affected regions. Still, X-rays are a very useful modality for analyzing pneumonia of any nature and that has driven some works to apply it for COVID-19 infection segmentation. Most of these are optimization algorithms. For instance, Abdel-Basset et al. proposed a meta-heuristic algorithm that integrates Slime Mould Algorithm (SMA) with the whale optimization algorithm in order to maximize Kapur's entropy [193]. The model extracts the regions of interest in the X-ray image by applying thresholding schemes. The extracted areas in the image denote ground-glass or consolidative pulmonary opacities. These X-ray observations are some of the manifestations associated with COVID-19. The performance of integrated SMA has
been evaluated on Chest X-rays and compared with five algorithms: Lshade algorithm, WOA, FireFly algorithm FFA, HHA, and salp swarm. In another study by Abdel-Basset et al., a hybrid detection model based on an improved marine predators algorithm (IMPA) was presented, along with a ranking-based diversity reduction (RDR) strategy for X-Ray image segmentation [194]. As described in Section 4.2, several works have demonstrated attentional localization of COVID-19 on X-ray, nonetheless only the aforementioned works were explicitly trained for COVID-19 segmentation.

4.3.2. CT based methods
There are a diversity of AI methods that attempt accurate delineation of COVID-19 infected regions from chest CT. These methods broadly deal with attention mechanisms, CNN, and ML optimizations.

CNN for medical image segmentation has been employed to solve a range of lung abnormalities segmentations, including Ground glass opacities and consolidation. The established proficiency of CNN in segmentation has driven many researches to come up with novel deep architectures for the COVID-19 dataset. For instance, in the work by Wang et al., the authors proposed an adaptive self-ensembling framework that uses the student-teacher model to perform COVID-19 segmentation [195]. The exponential moving average of the student is used as a teacher model to guide the CNN. Additionally, a dice loss formulation was proposed for learning noise-robust features. Similarly, Fan et al. presented a segmentation network with a parallel partial decoder to extract the high-level features [196]. The semi-supervised network used implicit reverse attention and explicit edge-attention to model the boundaries and enhance the representations of the images, suitable for limited datasets. In another study, Ni et al. used a deep learning algorithm consisting of lesion detection, segmentation, and location-based processing [197]. The quantitative determination performance was made with radiological residents. To estimate the location of COVID-19 infected tissues in the CT scan, Hassanatabar et al. trained a feedforward CNN that linearly stacks several convolutional layers [198]. U-net based architectures are traditionally used as a baseline for measuring model performance on any dataset. For the COVID-19 data, Trivizakis et al. trained a U-net segmentation model to perform infection delineation [199]. The lung segmented CT slices were segmented on a custom encoder-decoder architecture and classified to COVID-19/pneumonia.

Another line of research has been in the image thresholding algorithms to learn COVID-19 segmentation. For instance, as a way of improving the Marine Predators Algorithm (MPA) for thresholding-based segmentation, Elaziz et al. proposed an improved moth-flame optimized MPA for segmenting COVID-19 regions from CT [200]. Mishra et al. implemented a Kapur/Otsu image thresholding based pneumonia segmentation, which is optimized with Cuckoo-Search-Algorithm (CSA) [201].

CAD-assisted diagnosis has provided meaningful interpretations of infected hotspot regions in the clinical diagnosis of COVID-19. Grassi et al. presented research in which radiologists carried out a computer-aided analysis of lung abnormalities (lung parenchyma volume, GGO, and consolidations) with three commercial AI software 1) Thoracic VCAR, 2) Myrian, Intrasense and 3) InferRead, InferVision [202]. The computer-aided pneumonia quantification was made towards assessing COVID-19 traces on the CT scan.

4.4. COVID-19 risk assessment and prognosis

Risk analysis of COVID-19 takes a critical role in early medication and to decide the course of the follow-up treatment. Some of the works have explored ways to provide an indication of the severity of the virus infection to aid clinical prognosis. For instance, Cohen et al. trained a DenseNet model on a regression task to predict the extent of lung involvement and degree of opacity in the COVID-19 affected chest X-rays [203]. The architecture consisted of CNN layers for feature extraction followed by fully connected layers for performing target predictions. In another work to assess the pulmonary disease severity of COVID-19, Li et al. developed a convolutional Siamese network algorithm that learns from chest X-rays [204]. The DenseNet121 was modeled as a Siamese network and was pre-trained with a weakly labeled CheXpert dataset. The CNN learning was transferred onto a smaller COVID-19 training dataset for pulmonary risk assessment. In the line of works that exploit clinical factors, Iwendi et al. developed a random forest classifier on patient health records, symptoms, age, gender to predict the influence of COVID-19 [205]. The node split criteria enable finding the critical features that affect COVID-19 prediction. Ng et al. also presented a multi-variable logistic regression-based risk prediction model that considered factors like patient sex, age, symptoms, blood test results, and CXR findings [206]. Similarly, Liang et al. proposed a deep learning-based survival model that can predict the risk of COVID-19 patients developing critical illness based on clinical characteristics at admission [207]. The study constructed a three-layer feed-forward neural network for survival modeling which was then combined with a Deep learning survival Cox model. Zhang et al. presented a risk factor analysis using Kaplan-Meier curves that take CT segmented lung lesion areas and the clinical parameters as input to categorize patients into high/low-risk groups [208]. The CT segmentation was performed to identify five types of lesions, including consolidation (CL), ground-glass opacity (GGO), pulmonary and pleural effusion. Research in severity assessment and criticality prediction are the next steps in the automation for administering COVID-19 treatment protocols.

5. Discussion

This work carries out a comprehensive study of the various AI approaches in the literature for detecting COVID-19 from different imaging modalities. To the best of our knowledge, this is the first work presenting a modality-wise review of AI methods specifically for COVID-19 classification and segmentation.

5.1. Datasets for COVID-19 detection and infection delineation

In this section, the different CT and X-ray datasets used in the research literature for COVID-19 classification and segmentation have been described. For the classification task, a large number of researches have utilized one or more of the open-source datasets while some of them have trained on private datasets collected at hospitals and clinical labs. Table 4 presents a summary of the various openly available datasets on the web for COVID-19 detection. In the infection segmentation task, most of the datasets used in the works are held privately, while some of them are publicly accessible. Table 5 gives a compilation of such freely usable segmentation datasets.
Table 2
A comprehensive view of some notable methods for COVID-19 classification.

| S.no | Source                     | Modality | Methodology            | Implementation libraries | Validation                      | Target classes (pneumonia/COVID-19/normal) |
|------|----------------------------|----------|------------------------|--------------------------|---------------------------------|------------------------------------------|
| 1    | Khan et al. [103]          | X-ray    | CNN                    | Keras, Tensorflow        | 4-fold cross-validation         | COVID-19/Normal/Bacterial Pneumonia/COVID-19/Normal |
| 2    | Karthik et al. [112]       | X-ray    | CNN                    | PyTorch                  | 5-fold cross-validation         | Pneumonia/COVID-19/Normal Viral pneumonia/Bacterial pneumonia/COVID-19/Normal |
| 3    | Goel et al. [117]          | X-ray    | CNN                    | MATLAB 2020a             | 10-fold cross-validation        | COVID-19/Normal/Pneumonia                 |
| 4    | Chowdhury et al. [121]     | X-ray    | CNN                    | Keras, Tensorflow        | 80:10:10% cross-validation      | COVID-19/Normal/Viral Pneumonia          |
| 5    | Ucar et al. [125]          | X-ray    | CNN                    | MATLAB                   | 3687 training, 462 validation, 459 testing | COVID-19/Normal/Pneumonia                 |
| 6    | Marques et al. [128]       | X-ray    | CNN                    | Keras                    | 10-fold cross-validation        | Normal/Pneumonia/COVID-19                 |
| 7    | Waheed et al. [129]        | X-ray    | CNN                    | Keras                    | 932 training, 192 testing       | COVID-19/Normal                          |
| 8    | Abraham et al. [133]       | X-ray    | CNN                    | MATLAB 2020a             | 10-fold cross-validation        | COVID-19/Non-COVID-19                    |
| 9    | Altan et al. [135]         | X-ray    | CNN                    | MATLAB 2019b             | 5-fold cross-validation         | COVID-19/Normal/Viral Pneumonia          |
| 10   | Toğacar et al. [136]       | X-ray    | CNN                    | MATLAB 2019b             | 70:30% cross-validation         | COVID-19/Non-COVID-19                    |
| 11   | Islam et al. [138]         | X-ray    | CNN                    | Keras, TensorFlow        | 5-fold cross-validation         | Normal/COVID-19/COVID-19                 |
| 12   | Nour et al. [139]          | X-ray    | CNN                    | MATLAB (2019a)           | 70:30% cross-validation         | COVID-19/Normal/Viral Pneumonia          |
| 13   | Shibly et al. [150]        | X-ray    | CNN                    | TensorFlow               | 10-fold cross-validation        | COVID-19/Non-COVID-19                   |
| 14   | Chandra et al. [152]       | X-ray    | Vote based Classifier | MATLAB R2018a            | 10-fold cross-validation setup  | Pneumonia/COVID-19                       |
| 15   | Ouyang. [159]              | CT       | CNN                    | PyTorch                  | 5-fold cross-validation        | COVID-19/Community-Acquired Pneumonia (CAP) |
| 16   | Wang et al. [172]          | CT       | CNN                    | PyTorch                  | 4-fold cross-validation        | COVID-19/Non-COVID-19                    |
| 17   | Pathak et al. [175]        | CT       | LSTM                   | MATLAB 2018b             | 20-fold cross-validation       | COVID-19 (+)/COVID-19 (-)/Pneumonia      |
| 18   | King et al. [176]          | X-ray    | SOFM                   | Python, OpenCV           | 80:20% testing                 | Normal/COVID-19                         |
| 19   | Singh et al. [177]         | CT       | CNN                    | MATLAB 2019a             | 20-fold cross-validation       | COVID-19 (+)/COVID-19 (-)                |
| 20   | Ahuja et al. [179]         | CT       | CNN                    | MATLAB 2019a             | 70:30% cross-validation        | COVID-19 (+)/COVID-19 (-)/Bacterial/COVID-19/Normal |
| 21   | Oh et al. [185]            | X-ray    | CNN                    | MATLAB 2015a             | 70:10:20% testing              | Normal/Bacterial/TB/COVID-19/Normal      |

Table 3
A comprehensive summary of few notable COVID-19 segmentation methods.

| S.no | Source                      | Modality | Methodology                                   | Testing dataset size |
|------|-----------------------------|----------|-----------------------------------------------|----------------------|
| 1    | Abdel-Basset et al. [194]   | X-ray    | Improved Marine predators algorithm           | 9 X-ray images       |
| 2    | Wang et al. [195]           | CT       | Adaptive self-ensembling CNN                 | 130 CT scans         |
| 3    | Fan et al. [196]            | CT       | Attention CNN                                | 50 CT images         |
| 4    | Hassanatbar et al. [198]    | CT       | Marine predators algorithm                    | 104 CT images        |
| 5    | Elaiz et al. [200]          | CT       | Transfer learning CNN                         | 939 CT slice images  |

5.2. Evaluation metrics for COVID-19 classification and segmentation

Different evaluation metrics were commonly used by researchers to measure the performance of the proposed models and techniques. For studies taking the classification approach, Accuracy (ACC), Specificity (SPE), Sensitivity (SEN), Precision (PRE) F1 Score (F1) are frequently used.

5.2.1. Accuracy (ACC)

Accuracy is the fraction of correctly identified predictions, as given by Eq. (1). It measures the overall performance of the model on the test set.

\[
ACC = \frac{TP + TN}{TP + TN + FP + FN}
\]

where \(TP = \) true positive, \(TN = \) true negative, \(FP = \) false positive, \(FN = \) false negative.
Table 4
Publicly available Imaging Datasets for the COVID-19 detection.

| S. No | Modality | Description | No of Samples | Link to Dataset |
|-------|----------|-------------|---------------|----------------|
| 1     | X-ray    | COVID-chest X-ray-dataset (Accessed 26 Oct 2020) | 584 COVID-19 | https://github.com/ieeep023/covid-chestxray-dataset |
| 2     | CT       | SARS-COV-2 CT-scan, A large dataset of real patients CT scans for SARS Cov-2 identification | 1252 COVID-19, 1230 normal | https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset |
| 3     | X-ray    | Kaggle’s COVID-19 Radiography Database | 219 - COVID-19, 1341- normal, 1345 viral pneumonia | https://www.kaggle.com/tawsifurrahman/covid19-radiography-database |
| 4     | X-ray    | Covid-19 X rays, Dadario AMV. | 78 COVID-19 | https://www.kaggle.com/dsv/1019469 |
| 5     | X-ray    | COVID19 Dataset | 13,975 COVID-19 | https://arxiv.org/abs/2003.05871 |
| 6     | X-ray    | Chest x-ray images (pneumonia) | 4265 COVID-19 | https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia |
| 7     | CT       | COVID-CT-Dataset: a CT scan dataset about COVID-19 | 1575 Normal | http://arxiv.org/abs/2003.13865 |
| 8     | X-ray    | COVID-19 chest X-ray (Accessed 30 Oct 2020) | 55 COVID-19 | https://github.com/agchung/Figure1-COVID-chestxray-dataset |
| 9     | X-ray    | Radiopaedia (Accessed 30 Oct 2020) | 98 COVID-19 samples | https://radiopaedia.org/articles/covid-19-37lang=en |
| 10    | CT       | The cancer imaging archive (TCIA)(Accessed 26 Oct 2020) | 650 3d CT scans | https://wiki.cancerimagingarchive.net/display/Public/CT+Images+in+COVID-19/3d/COVID-19/2020 |
| 11    | X-ray    | E.H.C. Muhammad, et al. COVID-19 radiology database. Can AI help screen viral COVID-19 pneumonia?, 2020 | 23 COVID-19 (+), 1485 Viral pneumonia, 1579 Normal | https://arxiv.org/abs/2003.13145 |
| 12    | X-ray    | Italian society of medical and interventional radiology. | 115 COVID-19 (+) | https://www.sirm.org/category/senza-categoria/covid-19/ |
| 13    | CT       | COVID-19 and common pneumonia chest CT dataset | 412 Non-COVID-19 pneumonia, 412 COVID-19 pneumonia | https://data.mendeley.com/datasets/3y5sgckg6/1 |
| 14    | X-ray    | Actualmed COVID-19 chest X-ray dataset, 2020 | 239 COVID-19 | https://github.com/agchung/Actualmed-COVID-chestxray-dataset |
| 15    | X-ray    | Praveen. Corona Hack: Chest X-Ray Dataset. | 58 COVID-19 | https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset |
| 16    | X-ray    | TWITTER COVID-19 CXR DATASET | 134 COVID-19 | https://twitter.com/ChestImaging/ |

Table 5
Open access COVID-19 Segmentation datasets.

| S. No | Modality | Description | No of Samples | Link to Dataset |
|-------|----------|-------------|---------------|----------------|
| 1     | CT       | COVID-19 CT Segmentation Dataset (Apr 2020) | 100 COVID-19 2D CT slices | https://medicalsegregation.com/covid19/ |
| 2     | CT       | COVID-19 CT Lung and Infection Segmentation Dataset (Apr 2020) | 20 COVID-19 3D CT scans | https://zenodo.org/record/3757476#XS4FS8RgI8 |
| 3     | CT       | MosMed COVID-19 CT Scans | 1100 studies with 50 COVID-19 annotated CT scans | https://mosmed.ai/en/ |

5.2.2. Specificity (SPE)
Specificity measures the proportion of negative class samples that were correctly identified, as presented in Eq. (2).

\[
SPE = \frac{T_N}{T_N + F_P}
\]  

(2)

5.2.3. Sensitivity (SEN) / Recall (REC)
Sensitivity/Recall measures the proportion of positive class samples that were correctly identified, which is given in Eq. (3).

\[
SEN = \frac{T_P}{T_P + F_N}
\]  

(3)

5.2.4. Precision (PRE)
Precision gives the rate of the truly classified positive images among the classes, which is expressed in Eq. (4).

\[
PRE = \frac{T_P}{T_P + F_P}
\]  

(4)

5.2.5. F1 score (F1)
F1 is a joint score of Precision and Recall, expressed as a harmonic mean of these metrics. It is given in Eq. (5).

\[
F1 = \frac{2 \times PRE \times REC}{REC + PRE}
\]  

(5)
For segmentation-based studies, along with Accuracy, F1-score, Precision, and Sensitivity, other evaluation metrics were utilized too. These are Jaccard Similarity Index (JI), Peak Signal to Noise Ratio (PSNR), Structured Similarity Index Metric (SSIM), Universal Quality Index (UQI), Dice similarity coefficient (DSC), Relative Volume Error (RVE), and Hausdorff Distance (HD_{95}).

5.2.6. Jaccard similarity index (JI)

IOU or Jaccard Index is determined as the area of overlap between the prediction and actual ground truth, divided by the area of union between the prediction and ground truth. Eq. (6) illustrates this.

\[
JI = \frac{|I_G \cap I_R|}{|I_G \cup I_R|}
\]  

(6)

where \(I_G\) = Ground Truth Image, and \(I_R\) = Predicted Segmentation map.

5.2.7. Peak signal to noise ratio (PSNR)

PSNR is used to measure the similarity of the segmented image's quality compared with the original image, expressed in Eq. (7).

\[
PSNR = 10 \log_{10}(\frac{255^2}{MSE})
\]  

(7)

5.2.8. Structured similarity index metric (SSIM)

SSIM is used to measure the similarity, contrast distortion, and brightness between the original and the segmented feature. This is given in Eq. (8).

\[
SSIM(O, S) = \frac{(2\mu_O \mu_S + a)(2\sigma_O \sigma_S + b)}{\mu_O^2 + \mu_S^2 + \sigma_O^2 + \sigma_S^2 + b}
\]  

(8)

where \(\mu_O\) is the average intensities of the original image, while \(\mu_S\) indicates the average intensities of the segmented image. \(\sigma_O, \sigma_S\) are the standard deviations of the original and segmented images, respectively. \(\sigma_{OS}\) is the covariance between the two images. \(a, b\) is equal to 0.001 and 0.003, respectively.

5.2.9. Universal quality index (UQI)

UQI is an indicator similar to SSIM. As mentioned in Eq. (9), the index is used in measuring the quality of the segmented image based on the similarity structure between the two images rather than the error rate.

\[
UQI(O, S) = \frac{(4\mu_O \mu_S \sigma_{OS})}{(\mu_O^2 + \mu_S^2)(\sigma_O^2 + \sigma_S^2)}
\]  

(9)

5.2.10. Dice similarity coefficient (DSC)

DSC measures the overlapping volume between two segmentations, as seen in Eq. (10).

\[
DSC = \frac{2 | A \cap B |}{| A | + | B |}
\]  

(10)

where A and B denote the voxel sets for segmentation and ground truth respectively.

5.2.11. Relative volume error (RVE)

RVE, given in Eq. (11), is used to compare automatic segmentation results to the manually traced ground truth.

\[
RVE = \frac{\text{abs}(|A| - |B|)}{|B|}
\]  

(11)

5.2.12. Hausdorff distance (HD)

HD, as illustrated in Eq. (12) and (13), measures the maximum distance separating the volumes of two surface points and illustrates the outliers obtained in the segmentation. HD_{95} is similar to HD, but instead takes the 95th percentile instead of the maximum value in Eq. (12).

\[
HD(S_A, S_B) = \max_{i \in S_A} \min_{j \in S_B} ||i - j||_2
\]  

(12)

\[
HD(S_A, S_B) = \max (HD'(S_A, S_B), HD'(S_B, S_A))
\]  

(13)

5.3. Analysis of existing works

This section discusses salient performance trends of the different AI methods for COVID-19 detection and segmentation. These trends include the algorithms and systems that were employed in notable recent studies and an analysis of their performance in terms of various metrics.

Table 6 provides a consolidation of the state-of-the-art works for the classification of COVID-19 from X-ray and CT. The table highlights the top research papers from different model types. We developed a taxonomy of the classification methods and bucketed the research works into the following categories: Deep Transfer Learning, Ensemble CNN, Capsule Networks, Feature Selection techniques, Semi-supervised/GAN models, RCNN, Optimization Algorithms, CAD, Hybrid CNN, Sequential CNN, Evolutionary Algorithms. From Table 6,
Table 6

Analysis of results presented by few notable COVID-19 classification methods.

| S.no | Source | AI Model | Accuracy | Sensitivity | Specificity | Precision | F1 score | AUC |
|------|--------|----------|----------|-------------|-------------|-----------|----------|-----|
| 1    | Altan et al. [135] | Swarm Optimization | 96.99 | 99.44 | 99.81 | 99.62 | 99.53 | - |
| 2    | Islam et al. [138] | Hybrid CNN | 99.4 | 99.3 | 99.2 | - | 98.9 | 99.9 |
| 3    | Ahuja et al. [179] | Deep Transfer learning – Fixed feature extractor | 99.4 | 100 | 98.6 | 99 | 99.5 | 99.65 |
| 4    | Toğacı et al. [136] | Social Mmmic Optimization | 99.34 | 99.32 | 99.37 | 99.66 | 99.49 | - |
| 5    | Ucar et al. [125] | Bayesian Optimization | 99.18 | 99.13 | - | 99.48 | 99.3 | - |
| 6    | Apostolopoulos et al. [144] | Deep Transfer learning – Training from scratch | 99.18 | 97.36 | 99.42 | 97.36 | 97.36 | - |
| 7    | Nour et al. [139] | Hybrid CNN | 98.97 | 89.39 | 99.75 | - | 96.72 | - |
| 8    | Haque et al. [111] | Ensemble CNN – Stacking | 98.3 | 100 | 96.61 | 97.62 | 98.3 | - |
| 9    | Ouyang [159] | Attention CNN | 98.1 | 99.4 | 87.3 | - | 94.7 | 98.7 |
| 10   | Ozurk et al. [110] | Deep Transfer learning – Fine-tuning | 98.08 | 95.13 | 95.3 | 98.03 | 96.51 | - |
| 11   | Karthik et al. [112] | Shuffled Residual CNN | 97.94 | 97.54 | - | 96.34 | 96.9 | 98.39 |
| 12   | Goel et al. [117] | Capsule net CNN | 97.78 | 97.75 | 96.25 | 92.88 | 95.25 | - |
| 13   | Jain et al. [65] | Deep Transfer learning – Fixed feature extractor | 97.77 | 97.14 | - | 97.14 | - | - |
| 14   | Abraham et al. [133] | Feature Selection | 97.4359 | 98.6 | - | 98.6 | 98.6 | - |
| 15   | Narayan Das et al. [109] | Deep Transfer learning – Fixed feature extractor | 97.4068 | 97.0921 | 97.2973 | - | 96.9697 | - |
| 16   | Shibly et al. [103] | Deep Transfer learning – Fixed feature extractor | 97.36 | 97.65 | 95.48 | 99 | 98.46 | - |
| 17   | Marques et al. [128] | Deep Transfer learning – Fixed feature extractor | 96.7 | 96.69 | - | 97.59 | 97.11 | - |
| 18   | Chowdhury et al. [121] | Capsule net CNN | 96.58 | 91.3 | - | 95.45 | 93.33 | - |
| 19   | Vaid et al. [75] | Deep Transfer learning – Fixed feature extractor | 96.3 | 97.1 | - | 91.7 | 94.3 | - |
| 20   | Pathak et al. [175] | Hybrid CNN | 96.1983 | 96.2295 | 96.1667 | - | 96.1667 | 96.2295 |
| 21   | Alqudah et al. [141] | Hybrid CNN | 95.2 | 93.3 | 100 | 100 | 96.53 | - |
| 22   | Wahed et al. [129] | Generative Adversarial Networks | 95 | 90 | - | 96 | 92.9 | - |
| 23   | Duran-Lopez et al. [145] | Class Activation Maps | 94.43 | 92.53 | 96.33 | 93.76 | 93.14 | 98.8 |
| 24   | Misra et al. [84] | Ensemble CNN – Stacking | 93.9 | 100 | - | 89.6 | 94.5 | - |
| 25   | Abbas et al. [107] | Deep Transfer learning – Fixed feature extractor | 93.1 | 87.09 | 100 | - | - | - |
| 26   | Sethi et al. [92] | Deep Transfer learning – Fine-tuning | 93 | 78 | - | 97 | 86 | - |
| 27   | Wozniak et al. [209] | Neural Network | 92 | 95 | 89.7 | 95.23 | 95.11 | - |
| 28   | Chandra et al. [152] | Ensemble – Majority Voting | 91.329 | 96.512 | 96.207 | 97.368 | 91.713 | 91.4 |
| 29   | Dansana et al. [81] | Ensemble CNN – Probability averaging | 91 | 94 | - | 100 | 97 | - |
| 30   | Luján-García et al. [104] | Deep Transfer learning – Fine tuning | 91 | 87 | - | 92 | 88 | 98 |
| 31   | Wang et al. [172] | Hybrid CNN | 90.83 | 85.89 | - | 95.75 | 90.87 | 96.24 |
| 32   | Singh et al. [177] | Multi-objective Differential Evolution Optimization | 90.22 | 98.4 | 89.2 | 89.8 | 93.9 | - |
| 33   | Khan et al. [103] | Deep Transfer learning – Fixed feature extractor | 89.6 | 89.92 | 96.4 | 90 | 89.8 | - |
| 34   | Pandey et al. [79] | Deep Transfer learning – Fixed feature extractor | 88.10 | 97.62 | 78.57 | 82 | 89.13 | - |
| 35   | Oh et al. [185] | Grad-CAM localization | 84.8 | 80.1 | 94.8 | 78.8 | 79.3 | - |

Table 7

Analysis of results presented by notable COVID-19 segmentation methods.

| S.no | Source | Method | PSNR | SSIM | UQI | DSC | RVE | HD | Accuracy | Sensitivity | Specificity | Precision | JI |
|------|--------|--------|------|------|-----|-----|-----|-----|----------|-------------|------------|-----------|----|
| 1    | Trivizakis et al. [159] | Transfer Learning – Fine-tuning | 99.6 | 91.1 | 92 | 98.5 | - | - | - | - | - | - | - |
| 2    | Satapathy et al. [201] | Otsu Thresholding | - | - | - | 90.32 | - | - | 97.62 | - | - | - | 83.18 |
| 3    | Wang et al. [185] | Teacher-Student | - | - | - | 80.29 | 17.72 | 18.72 | - | - | - | - | - |
| 4    | Fan et al. [196] | Attention CNN | - | - | - | 57.9 | - | - | 87 | 97.4 | 50 | - | - |
| 5    | Abdel-Basset et al. [194] | Marine Predators Optimization | 33.26 | 0.98 | 0.98 | - | - | - | - | - | - | - | - |
| 6    | Ni et al. [197] | 3D U-Net CNN | - | - | - | 85 | 97 | 61 | - | - | - | - | - |
| 7    | Hassanabadi et al. [198] | CNN | - | - | - | 83.84 | - | - | - | - | - | - | 40.00 |
| 8    | Elaziz et al. [200] | Marine Predators Thresholding | 25.43 | 0.81 | - | - | - | - | - | - | - | - | - |

although few techniques offer the highest results, the variance in performance metrics between different methodologies is very less. For instance, the studies on optimization algorithms and CNN transfer learning have reported similar results, in terms of F1-score, sensitivity, accuracy. Thus, a test of statistical significance is needed to identify which sets of algorithms are better when compared to the rest. Besides, though transfer learning-based works in overwhelming majority, concepts like attention mechanism, ensembling, hybrid multi-stage CNN are gaining huge traction in recent works.

Similarly, Table 7 summarizes the top research works that perform the infection segmentation from CT/X-ray. As discussed in Section 5.2, for segmentation, metrics including Dice metric, Jaccard Index, and PSNR yield insights into the effectiveness of these methods. The range of methods used in segmentation can be grouped into Optimization algorithms, U-net CNN, Attention CNN, Ensemble CNN, Thresholding Algorithms, and CAD techniques. It could be inferred from the results that thresholding techniques presented in Elaziz et al. [200], Satapathy et al. [201] have been able to efficiently capture the intensity range of the COVID-19 infection, in the same levels as the other CNN models. Hyper-parameter fine-tuning and optimizations of current CNN architectures have significantly led to better generalization on the COVID-19 dataset.

Table 8 captures statistical trends of the recent methodologies by grouping them based on the learning objective and the modality. The major COVID-19 AI topics reviewed in this study include classification, segmentation, region localization, and risk assessment from medical images. The radiological modalities included in the study are X-ray and CT scans. Table 8 also gives the mean and standard deviation of 498
the accuracy metric registered for these method categories. Given that some categories have a large number of papers and some have lesser papers, it is important to robustly validate the efficacy of these approaches using statistical significance tests. When comparing the approaches, several factors such as dataset source, size, complexity, optimization criterion, train-test setup are involved. At a high level, for a type of method, averaging the accuracies reported by the papers in that group directly reflects on the group’s performance as a whole. Thus, in our hypothesis testing experiments, the accuracy score is considered to represent each research paper (observation) in the method category (sample set).

The hypotheses of this review are three-fold. First, though a diverse range of deep learning models is learnt for the task for COVID-19 diagnosis and infection segmentation, the difference in their efficacy measured on a certain set of metrics only varies by a small margin. Second, there can be multiple best-performing methods, going by the statistical significance. Lastly, since the usefulness of the current methods is established for the major detection/segmentation tasks, the future research direction lies in the exploration of these models for less traversed areas such as severity assessment and weakly supervised localization.

We use the two-sample t-test on pairs of AI methods for comparing their performance. Statistical hypothesis tests are used to quantify the likelihood of metric values, assuming that the models were drawn from the same population. When this assumption (null hypothesis) is rejected, there is a significant difference in the metrics between the two groups. For a two-sample Student’s t-test, we assume that the metrics recorded by the model are independent of each other, and the two independent AI method groups are randomly sampled from two normal populations.

Let \( \bar{x}_1, \bar{x}_2 \) denote the mean accuracy of the two AI methods \( M_1, M_2 \) with \( n_1, n_2 \) number of research works respectively. Let \( s_1, s_2 \) be the standard deviation of the accuracies registered on the sample sets \( M_1, M_2 \). Then, the pooled standard deviation \( s_p \), which is a combined estimate of the overall deviation between the groups is given by Eq. (14).

\[
    s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}
\]

(14)

The combined estimate adjusts for the size of the two groups. Then, the t-statistic can be calculated as shown in Eq. (15).

\[
    t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}
\]

(15)

Then the p-value of the t-statistic needs to be computed, using the degree of freedom = \( n_1 + n_2 - 2 \). For a significance level \( \alpha = 0.05 \), if the p-value is smaller than 5%, then the null hypothesis is rejected, which would mean that the models perform differently. If not, the null hypothesis that the models’ performance is equivalent does not fail.

For each pair of methods pertaining to a task-modality bucket listed in Table 8, the T-test was applied. That is, all AI methods only within the learning objective and modality category will be compared and not across a different task or a modality. This was to ensure that there is a fair comparison of AI methods only with respect to other methods performing the same learning task on the same kind of data. Fig. 3 shows the t-statistics registered for AI methods concerning four types of task-modalities from Table 8.

From Fig. 3, these four combinations of task-modality have at least more than one kind of AI method to perform comparison. For the other task-cum-modality categories listed in Table 8, the research works are only from a single AI method. The t-statistic values computed
for a pair of AI methods as per Eq. (15) can be converted into a p-value taking into account the degrees of freedom. Fig. 4 presents the p-value tables for the four problem types. The p-value is the probability of obtaining a test accuracy value as extreme as the observed results, assuming the null hypothesis is true.

Fig. 4 validates the first hypothesis that the difference between the type of AI models that perform the same task on similar data is not statistically significant. For a significance level of 5%, the p-value is greater than this threshold in most of the pairwise method comparisons. Thus, the null hypothesis that the performance of the methods is similar cannot be rejected. As an example, for X-ray classification, the CAD performs very similarly to all other methods except the ensemble CNN. In other cases, the comparisons between the methods have not differed significantly. So, the first hypothesis has been validated with convincing evidence from the t-test.

Secondly, to identify the best-performing models, we inspect the p-value to pick clusters of methods that are highly correlated to each other. From the heatmap visualization for X-ray classification, the following highly accurate methods perform similarly - Capsule networks, Ensemble CNN, Feature selection, GAN, Hybrid CNN, and Optimization Algorithms. In the same manner, from the p-value matrix in Fig. 4 for COVID-19 CT classification, the Evolutionary Algorithms and Feature Selection form a cluster that has high mean accuracy. For CT segmentation, Thresholding techniques and U-net CNN have 28% and 17.5% better accuracy compared to the correlated cluster formed by Attention CNN, CAD, Ensemble CNN. Therefore, they are the better approach. Weak region localization is better achieved using object detection algorithms which are 7.55% more than CAM algorithms.

The future research direction in this field would be exploring the current techniques and radiological data for predicting more complex versions of classification/segmentation, such as the criticality assessment at the patient level, for different lung regions, etc. A fusion of these techniques can be developed to address different tasks in a unified framework. A radiological timeline trend of the COVID-19 progression from sets of patients can be analyzed to reveal the radiological features responsible for complications at every stage. Section 5.5 elaborately discusses the future implications of the current methods.
Fig. 4. P-value of the t-test performed between pairs of AI methods for different combinations of task and modality.

5.4. Country/continent wise analysis of COVID-19 works

To gain insights into which countries are actively engaged in the health informatics for COVID-19, two kinds of analysis were performed. In the first study, we identify continents/countries that are actively contributing to the creation of COVID-19 clinical datasets. The second analytics reveals countries that are currently working with the best AI results.

Fig. 5 presents the bar plot showing the number of unique datasets published by different regions of the world. These unique datasets ranged from manually collected datasets to aggregates of other open-source datasets. A few cases of super-aggregates were found as well. Every instance of data used from a particular region was noted and consolidated. This consolidation can be seen in Fig. 5. It can be observed that a majority of instances sourced data from the North American and Asian regions, followed by European samples as well. Within Asia, countries like India, China, Bangladesh are actively engaged in sourcing clinical data, reports from public hospitals. In North America, the USA, Canada, Mexico are leading data providers. Poland, Italy, UK, Germany are the biggest contributors from Europe. Although few in number, Russia has created high-quality segmentation databases for lung infection delineation, criticality assessment, and prognosis.

The second analysis attempts a performance comparison of published works from each country. The accuracy metrics reported in the research works were aggregated to form a country-wise boxplot distribution. The box plot shown in Fig. 6, gives the minimum, maximum, median, and quartile accuracy scores registered in the research by various countries. It can be seen that out of 25 countries reviewed, 13 countries have median accuracy of over 95%.
5.5. Challenges in COVID-19 detection and future trends

There are several challenges faced in automatic COVID-19 detection from radiological imaging. This section provides a compilation of all such influencing factors in the literature.

5.5.1. Availability of sufficient dataset

Owing to the recency of the COVID-19 pandemic, imaging databases are being constantly updated with clinical data gathered from hospitals and testing labs. There is a lack of benchmarked large-scale datasets for robustly validating the performance of the methods proposed in AI research. Many works have presented the model results on an inadequate number of test-set samples, which might not generalize to real-world data. Furthermore, for the COVID-19 segmentation task, a large number of works have used privately acquired data, making it difficult to compare with newer methods. Lack of expert annotations for CT is another major challenge faced in building segmentation datasets. To counter that, several weakly supervised approaches, data augmentation techniques have been proposed in the literature. However, the substantial need for labeled, annotated datasets is critical to develop robust learning algorithms and establish benchmarks. Besides, the varying conditions, time duration under which these images are obtained, have to be consistent with the radiological observations of when the infection gets highlighted in the CT/X-ray. X-ray samples taken during the very early stage might
not reveal the manifestation of the virus, which might not be useful in training the model. Moreover, the large heterogeneity of samples collected from different scanners has to be intensity normalized to prevent any operations bias from interfering in the training.

5.5.2. Challenges in applied deep learning

The development of AI techniques for COVID-19 identification has been largely in the transfer learning approach. The standard CNN architectures that were pre-trained for a different task had to be customized for the COVID-19 dataset through altering specific layers and extensive hyperparameter tuning. While few works have applied such optimizations, several other works directly compared the performances of the standard deep learning models. To achieve the best results on the medical imaging dataset, transfer learning models have to be trained from scratch.

Also, given the limited size of the available training dataset, it is important to design strategies that avoid model overfitting. Some works have trained GAN for data augmentation while several other types of research have applied image affine transformations. Since CNN have large discriminative power, in transfer learning from large CNN like the Inception-V3, ResNet50, suitable training optimizations (like regularization, optimal stopping) are needed to address the issue of generalizability. Moreover, large models like the Capsule-net require sophisticated computing hardware resources to train and take time to predict.

The lack of proper benchmarked COVID-19 data has prompted the researches to utilize related datasets for building the AI model. For example, several works have trained AI to detect COVID-19 related lung diseases, like viral pneumonia. Some works fit AI models to identify lung abnormalities such as airspace opacities, alveolar and interstitial opacities, which are signs of COVID-19 [90], [95]. While these X-ray features can suggest COVID-19, the vice-versa needn't be true. Knowing the applicability of these generic X-ray features towards COVID-19 detection under appropriate conditions would justify training the model to find these X-ray patterns.

5.5.3. Research on predicting risk of COVID-19, follow-up treatment, and prognosis

Research on risk assessment of the COVID-19 infected patient has been a less explored arena in the medical imaging perspective. While most of the works have used CT/X-ray images to detect the virus or propose infection region segmentation, very few works have demonstrated the severity of the infection. Using radiographs to learn a regression model of the patient criticality can help prioritize the serious cases on a first-serve basis. AI for prognosis from clinical data can guide doctors in knowing the degree to which the patient is suffering from pneumonia, mainly alert if the patient needs immediate care. With a huge inflow of cases and running out of hospital beds, such AI mechanisms would also enable grouping patients by the likely course of treatment or medication. This helps to avoid complications in critical patients by tending to them first.

In addition to performing classification, the class-activation maps generated from the intermediate layers provide a useful indication of the infection hotspots in the radiological image. Such infected regions visualization is key for understanding CNN focus areas even for the CNN trained on the classification task. While the GRAD-CAM, Score-CAM, guided backpropagation techniques are widely used in many works, newer researches on refining these algorithms in the context of X-ray/CT will enable more accurate localization.

5.5.4. Future trends

The COVID-19 detection, localization, and segmentation from imaging modality is an active research area in AI-assisted clinical diagnosis and it offers huge scope for driving innovations in the radiology domain. Currently, the X-ray/CT features for different kinds of lung abnormalities and pulmonary diseases are well documented through several clinical studies. The COVID-19 being very recent doesn’t have such well-defined prescriptions for human-level interpretations. Though GGO and consolidation are used as a generic reference in many COVID-19 studies, with data mining and AI, the micro-level latent patterns specific to COVID-19 can be precisely revealed. AI-assisted pattern mining should help the radiologist/clinician develop guidelines for analyzing the image from a biomedical perspective and decide prognosis. Also, future works should attempt to track the progression of these subtle, distinct COVID-19 X-ray/CT manifestations across time, so that the patient medical condition and the evidence from the radiograph can be correlated for drawing clear directions in medical studies. The upcoming researches would largely be along the lines of directly assisting the doctor in medication through explainable AI.

Further, there exists a great demand for generating high-resolution radiographs to solve data augmentation. Generative adversarial networks are a promising research direction for accomplishing that. Since such high-resolution COVID-19 images are generated from the same data distribution, they can facilitate building AI models, even though they might not directly come from the clinical setting. Another prominent line of research is in the patient severity assessment and prognosis which are a critical part of the clinical use case. Future AI work should majorly focus on fusion approaches that employ multiple data representations (genome sequences, medical images, variables involved in medical analysis) to regress the patient criticality score.

5.6. Limitations of this study

The following are some of the limitations of this review.

- This review has only considered AI works that have used radiological imaging data for prediction. In a different research space, several works have employed Genome sequence data, patient medical conditions, previous health records for predicting chances of COVID-19. While those factors might influence the probability of the person being infected, imaging is the principal modality that provides unbiased evidence for COVID-19, regardless of patient age or other factors.
- The review has taken considerable effort and time to compile diverse researches and draw key insights from them. In the article searching process though, a certain set of keywords have been used to curate articles. Some works might have been missed due to incomplete/full keyword matching.
- This study does not focus on describing imaging features that decide the symptoms/course of treatment for the infected patient. The review is mainly concerned with testing if an individual has COVID-19 or not.
6. Conclusion

The COVID-19 outbreak has created a sense of perturbation all over the world. In a silver lining, this agitation can present as motivation for advancements and research in Artificial Intelligence, to assist medical workers to tackle the pandemic. While the advantages it brings are quite clear, it is a fact that Artificial Intelligence models can never replace doctors and radiological experts. Nevertheless, Computer-aided techniques for the analysis of medical images have grown significantly in recent times, contributing to medical research and clinical applications. Recent studies have proven the reliability of image-based COVID-19 diagnosis, through deep learning and machine learning architectures. This research intended to review the recent achievements and progress of these architectures in the classification and segmentation of a COVID-19 infection, through the modalities undertaken. Albeit these improvements, some limitations have persevered, allowing scope for more developments as well. With the urgency this pandemic brings, mankind depends on scientific revolution for an answer. And if the medical doctors and radiological experts also play a part in the conception and building of the framework for artificial intelligence models, breakthroughs can materialize faster. Though deep learning and machine learning has yielded significant results in the medical domain, their potential is immense for other image-based classification and segmentation tasks. This potential is currently inhibited by its large consumption of time and resources, high implementation costs, etc. Classification and segmentation algorithms suffer another obstacle, being insufficient and unbalanced data, which leads to overfitting and unreliable predictions. Further advancements and novelties devised to overcome these limitations, can contribute significantly to breakthroughs in the field of biomedical image processing.

CRediT authorship contribution statement

R. Karthik: Conceptualization; Data curation, Methodology; Project administration, Resources, and Reviewing. R. Menaka: Conceptualization, Writing- Reviewing and Editing. Hariharan: Data curation, Writing- Original draft preparation, Software, Validation. Gugan S Kathiresan: Writing and Editing.

Declarations

Human and animal rights

The authors declare that the work described has not involved experimentation on humans or animals.

Informed consent and patient details

The authors declare that the work described does not involve patients or volunteers.

Funding

This work did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors’ contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.

Declaration of competing interest

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

Appendix. Evaluation form for review

| S. No | Question                                                                 | Yes | No | Other (NR/CD) |
|-------|---------------------------------------------------------------------------|-----|----|---------------|
| 1     | ✓ Is the article a bone fide and original publication, proposing a novel artificial intelligence method for COVID-19 detection based on medical imaging? |     |    |               |
| 2     | Objectives of the research<br>✓ Is the work a deep learning, or machine learning, specific approach for COVID-19 detection?<br>✓ Are the sole datasets used in the studies X-ray or CT images?<br>✓ Is there a clear explanation of the approach taken by the research?<br>✓ Are the inferences of the study having clarity and are comprehensible?<br>✓ Is the numerical reports and results presented by the study precise and significant? |     |    |               |

If the answer for question 2 is ‘no’, further scrutiny is omitted.
Secondary scrutiny and filtering: Based on datasets and methods

| S. No | Question                                                                 | Yes | No | Other (NR/CD) |
|-------|--------------------------------------------------------------------------|-----|----|---------------|
| 1     | Research Methodology                                                     |     |    |               |
|       | ✓ Is there substance to support the motivation of the study in the form of a ‘Related Works/Literature Survey’ section? |     |    |               |
|       | ✓ Are all the defined sections of the study presented clearly?          |     |    |               |
| 2     | Datasets employed                                                       |     |    |               |
|       | ✓ Is there a proper citation of the sources and the important features of the dataset used? |     |    |               |
|       | ✓ Is the dataset considered from that of a unique and categorized population? |     |    |               |
|       | ✓ Is the sample size and labels clearly mentioned for this research?     |     |    |               |
|       | ✓ Has the validation system of the methods employed been explained clearly? |     |    |               |
| 3     | Performance and implementation analysis                                  |     |    |               |
|       | ✓ Is the basic structure of the methodology employed laid out clearly?  |     |    |               |
|       | ✓ Are the implementation features required and utilized for the method provided in detail? |     |    |               |
|       | ✓ Are reviewed comparisons drawn between the proposal of the publication and the methods reported in their literature? |     |    |               |
|       | ✓ Are the metrics employed for the model validation in relation the same ideology of the problem definition? |     |    |               |
| 4     | Key findings                                                             |     |    |               |
|       | ✓ Are there limitations against the problem statement reported in the article? |     |    |               |
|       | ✓ Is there a constructive and notable conclusion presented in the paper? |     |    |               |

Tertiary scrutiny and filtering: Based on architecture and training-validation

| S. No | Question                                                                 | Yes | No | Other (NR/CD) |
|-------|--------------------------------------------------------------------------|-----|----|---------------|
| 1     | Utilized architecture                                                   |     |    |               |
|       | ✓ Is the proposed method involving deep learning architectures like CNN, FCN, FFNN, etc.? |     |    |               |
|       | ✓ Is the proposed method involving machine learning architectures like SOFM, etc.? |     |    |               |
|       | ✓ Is the proposed method involving any hybrid or ensemble based architectures? |     |    |               |
| 2     | Training-Validation methods                                              |     |    |               |
|       | ✓ Are the batch size and training details of the proposed work mentioned? |     |    |               |
|       | ✓ Is there any use of pre-processing presented with the proposed method? |     |    |               |
|       | ✓ Are all the hyper-parameters and network configurations of the proposed method listed out? |     |    |               |
|       | ✓ Has the study clearly reported all results, based on the defined performance metrics? |     |    |               |

(NR: Not Reported and CD: Cannot Determine)

Conclusion of evaluation form: Accept [ ] Reject [x] Seek Further details [ ]

Comments:

References

[1] https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200921-weekly-epi-update-6.pdf?sfvrsn=d9cf9496_6.
[2] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses.
[3] Shaikh F, Anderson M, Sohail MR, Mulero F, Awan O, Dupont-Roettger D, et al. Current landscape of imaging and the potential role for artificial intelligence in the management of COVID-19. Curr Probl Diagn Radiol 2021;50(3):430–5.
[4] Xie Y, Wang X, Yang P, Zhang S. COVID-19 complicated by acute pulmonary embolism. Radiol Card Imaging 2020;2(2):e200067.
[5] Shuja J, Alani S, Faiz A, Alshehri A, Covid-19 open source data sets: a comprehensive survey. Appl Intell 2020;1–30.
[6] Alimdadzade A, Arya S, Manandhar I, Munroe PB, Joe B, Cheng X. Artificial intelligence and machine learning to fight COVID-19. Physiol Genomics 2020;52(4):200–2.
[7] Wong HY, Lam HY, Fong AH, Leung ST, Chin TW, Lo CS, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. Radiology 2020;296(2):E72–8.
[8] Rekha Hanumanthu S. Role of intelligent computing in Covid-19 prognosis: a state-of-the-art review. Chaos Solitons Fractals 2020;109947.
[9] Lalmanawma W, Hussain J, Chahakhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. Chaos Solitons Fractals 2020;139:110059.
[10] Chamola V, Hassija V, Gupta V, Guzani M. A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact. IEEE Access 2020;8:90225–65.
[11] Pham QV, Nguyen DC, Hwang WJ, Pathirana PN. Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: a survey on the state-of-the-arts. IEEE Access 2020;8:130820–39.
[12] Emile SH, Hamid HK. Fighting COVID-19, a place for artificial intelligence. Transbound Emerg Dis 2020. https://doi.org/10.1111/tbed.13648.
[13] Dananjayan S, Raj GM. Artificial Intelligence during a pandemic: the COVID-19 example. Int J Health Plann Manage 2020.
Hwang EJ, Kim H, Yoon SH, Coo JM, Park CM. Implementation of a deep learning-based computer-aided detection system for the interpretation of chest radiographs in patients suspected for COVID-19. Korean J Radiology 2020;21.

Mohammed S, Allikainen F, Hassan Y. Automatic computer aided diagnostic for COVID-19 based on chest X-ray image and particle swarm intelligence. Int J Intell Syst 2020;35(1):63–71.

Murphy K, Smiti H, Knoops AJ, Korst MB, Samson T, Scholten ET, et al. COVID-19 on chest radiographs: a multireader evaluation of an artificial intelligence system. Radiology 2020;296(3):E166.

Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S, Shukla PK. Deep transfer learning based classification model for COVID-19 disease. IRBM 2022;43(2):87–92.

Ouyang X, Huo J, Xia L, Shan F, Liu Ji, Mo Z, et al. Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. IEEE Trans Med Imaging 2020.

Silva P, Luz E, Silva G, Moreira G, Silva R, Lucio D, et al. COVID-19 detection in CT images with deep learning: a voting-based scheme and cross-datasets analysis. Inf Med Unlocked 2020;100427.

Hu S, Gao Y, Niu Z, Jiang Y, Li L, Xiao X, et al. Weakly supervised deep learning for COVID-19 infection detection and classification from CT images. IEEE Access 2020;8:118869–83.

Xu X, Jiang X, Mao D, Du P, Li L, Yi S, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 2020.

Aradkani AA, Kanari AR, Acharaya UR, Khadem H, Mohammad A. Application of deep learning techniques to manage COVID-19 with routine clinical practice using CT images: results of 10 convolutional neural networks. Comput Biol Med 2020;103795.

Perumal V, Narayanan V, Rajasekar JS. Detection of COVID-19 using CNN and CT images using transfer learning and haralick features. Appl Intell 2020;1–18.

Jaiswal Ajayesh, Gianchandani Neha, Singh Dilbag, Kumar Vijay, Kaur Manjit. Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. J Biomech Struct Dyn 2020. https://doi.org/10.1007/s00714-020-08864-2.

Mishra AK, Das SK, Roy P. Bandypadhyay S. Identifying COVID19 from chest CT images: a deep convolutional neural networks based approach. J Healthc Eng 2020;2020;9:613.

Tabrizchi H, Nadali L, Mosavi A, Saboo-Gali A. Rapid COVID-19 diagnosis using deep learning of the computerized tomography scans. 2020. (No. 4226). EasyChair.

Yan Tao, Wong Pak Kin, Ren Han, Wang Huaiqiao, Wang Jiangtao, Li Yang. Automatic distinction between COVID-19 and common pneumonia using multi-scale convolutional neural network on chest CT scans. Chaos Solitons Fractals 2020. https://doi.org/10.1016/j.chaos.2020.110153.

Shaban WM, Babie AH, Saleh AI, Abdo-Elsad MA. A new COVID-19 patients detection strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier. Knowl-Based Syst 2020;205:106270.

Han Z, Wei B, Hong Y, Li T, Cong J, Zhu X, et al. Accurate screening of COVID-19 using attention based deep 3D multiple instance learning. IEEE Trans Med Imaging 2020.

Li Y, Dong W, Chen J, Cao S, Zhou H, Zhu Y, et al. Effective and efficient training of COVID-19 classification networks with self-supervised dual-track learning to rank. IEEE J Biomed Health Inform 2020.

Wang Z, Liu Q, Dou Q. Contrastive cross-site learning with redesigned net for COVID-19 CT classification. IEEE J Biomed Health Inform 2020.

Öztürk Ş, Özkaya U, Barstugan M. Classification of coronavirus (COVID-19) from X-ray and CT images using shrunken features. Int J Imaging Syst Technol 2021;31(1):15–15.

Hasan AM, Al-Jawad MM, Jalal HA, Shalha H, Ibrahim RW, Al-Shamsaneh AAR. Classification of COVID-19 coronavirus, pneumonia and healthy lungs in CT scans using q-deformed entropy and deep learning features. Entropy 2020;22(5):517.

Pathak Y, Shukla PK, Arya KV. Deep bidirectional classification model for COVID-19 disease infected patients. IEEE/ACM Trans Comput Biol Bioinformat 2020.

King B, Barve S, Ford A, Jha R. Unsupervised clustering of COVID-19 chest X-ray images with a self-organizing feature map. In: 2020 IEEE 63rd international symposium on circuits and systems, MWSCAS. IEEE. 2020. p. 395–8.

Singh D, Kumar V, Kaur M. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks. Eur J Clin Microbiol Infect Dis 2020;1–11.

Buit C, Gill J, Chun D, Babu BA. Deep learning system to screen coronavirus disease 2019 pneumonia. Appl Intell 2020;1.

Ahuja S, Panigrahi BK, Dey N, et al. Deep transfer learning-based automated detection of COVID-19 from lung CT scans. Appl Intell 2021;51:571–85. https://doi.org/10.1007/s10489-020-01826-w.

Xiao LS, Li P, Sun F, Zhang Y, Xu C, Zhu H, et al. Development and validation of a deep learning-based model using computed tomography imaging for predicting disease severity of coronavirus disease 2019. Front Bioeng Biotechnol 2020;8:498.

Bai HK, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. AI augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other etiology on chest CT. Radiology 2020;290(3):201491.

Mahmud T, Rahman MA, Fatmah SA, CoxNext: 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 2020;122:103869.

Brunese Luca, Mercaldo Francesco, Reginnelli Alfonso, Santone Antonella. Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Comput Methods Programs Biomed 2020. https://doi.org/10.1016/j.cmpb.2020.105068.

Ezat D, Hassaniah AE, Ella HA. An optimized deep learning architecture for the diagnosis of COVID-19 disease based on genetic search optimization. Appl Soft Comput 2020;906742.

Oh Y, Park S, Ye JC. Deep learning features on chest radiographs using transfer learning. IEEE Int Conf Med Imaging 2020.

Fan BB, Yang H. Analysis of identifying COVID-19 with deep learning model. J Phys Conf Ser 2020;1601:052010. IOP Publishing.

Tisknakis N, Trivizakis E, Vassalou EE, Papadakis GZ, Spandidos DA, Tsatsakis A, et al. Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays. Exp Ther Med 2020;20:727–35.

Sharma S. Drawing insights from COVID-19-infected patients using CT scan images and machine learning techniques: a study on 200 patients. Environ Sci Pollut Res Int 2020;1–9.

Yang S, Jiang L, Cao Z, Wang L, Cao J, Feng R, et al. Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: a pilot study. Ann Transl Med 2020;8(7).

Pamwar H, Gupta PK, Siddiqui MK, Morales-Mendez R, Bhardwaj P, Singh V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest x-ray and CT-scan images. Chaos Solitons Fractals 2020;110190.

Wang Z, Xiao Y, Li Y, et al. Automatically discriminating and localizing COVID-19 from community-acquired pneumonia on chest X-rays. Pattern Recognit 2021. https://doi.org/10.1016/j.patcog.2020.107613.

Hurt B, Kligerman S, 2020;39(8):2653–63.

Xiao Y, Kligerman S, 2020;39(8):2653–63.

Wang L, et al. Applying deep learning to detect COVID-19 pneumonia from chest X-ray images. J Biomed Inform 2020;509.

Dey N, Papadakis GZ, 2019.10160;20:79521–40.

Sarigiannis D, et al. Circuits and Systems, MWSCAS. IEEE. 2020. p. 395–8.

Mohammadi A. Detecting COVID-19 from X-ray and CT images using deep learning with an hierarchical approach. 2020;35(3):W87–9. EasyChair.

Chang K, et al. A hybrid novel Slime mould algorithm with whale optimization algorithm for tackling the image segmentation problem of chest X-ray images. Appl Soft Comput 2020;86:6642.

Chao Z, et al. Automatic COVID-19 lung infection segmentation from CT images. IEEE Trans Med Imaging 2020;39(8):2563–63.

Fan DP, Zhou T, Ji GP, Zhou Y, Chen G, Fu H, et al. Inf-Net: automatic COVID-19 lung infection segmentation from CT images. IEEE Trans Med Imaging 2020.

Ni G, Sun ZY, QL, Chen W, Yang W, Wang L, et al. A deep learning approach to characterize 2019 coronavirus disease in chest CT images. Eur Radiol 2020;1–11.

Hassantabar S, Ahmadi M, Sharifi A. Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches. Chaos Solitons Fractals 2020;140:110170.

Trivizakis E, Tisknakis N, Vassalou EE, Papadakis GZ, Spandidos DA, Sarigiamnis D, et al. Advancing COVID-19 differentiation with a robust preprocessing and integration of multi-institutional open-repository computer tomography datasets for deep learning analysis. Exp Ther Med 2020;20(5):78.

Elazziz Mohamed Abd, et al. An improved Marine Predators Algorithm with fuzzy entropy for multi-level thresholding: real world example of COVID-19 CT image segmentation. IEEE Access 2020;8:125306–30.
[201] Satapathy SC, Hemanth Dj, Kadry S, Manogaran G, Hannon NMS, Rajinikanth V. Segmentation and evaluation of COVID-19 lesion from CT scan slices - a study with Kapur/Otsu function and Cuckoo Search Algorithm. Res Square 2020. https://doi.org/10.21203/rs.3.rs-40148/v1.

[202] Grassi R, Cappabianca S, Urraro F, Feragalli B, Montanelli A, Patelli G, et al. Chest CT computerized aided quantification of PNEUMONIA lesions in COVID-19 infection: a comparison among three commercial software. Int J Environ Res Public Health 2020;17(18):6914.

[203] Cohen JP, Dao L, Roth K, Morrison P, Bengio Y, Abbasi AF, et al. Predicting COVID-19 pneumonia severity on chest X-ray with deep learning. Cureus 2020;12(7).

[204] Li MD, Arun NT, Gidwani M, Chang K, Deng F, Little BP, et al. Automated assessment and tracking of COVID-19 pulmonary disease severity on chest radiographs using convolutional Siamese neural networks. Radiol Artif Intell 2020;2(4):e200079.

[205] Iwendi C, Bashir AK, Peshkar A, Sujaatha R, Chatterjee JM, Pasupuleti S, et al. COVID-19 patient health prediction using boosted random forest algorithm. Front Public Health 2020;8:357.

[206] Ng MY, Wan EYF, Wong HYF, Leung ST, Lee JCY, Chin TWY, et al. Development and validation of risk prediction models for COVID-19 positivity in a hospital setting. Int J Infect Dis 2020.

[207] Liang W, Yao J, Chen A, Lv Q, Zanin M, Liu J, et al. Early triage of critically ill COVID-19 patients using deep learning. Nat Commun 2020;11(1):1–7.

[208] Zhang K, Liu X, Sheng J, Li Z, Sang Y, Wu X, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. Cell 2020.

[209] Woźniak M, Polap D, Capizzi G, Scuito GL, Kośmider L, Frankiewicz K. Small lung nodules detection based on local variance analysis and probabilistic neural network. Comput Methods Programs Biomed 2018;161:173–80.

[210] Ke Q, Zhang J, Wei W, Polap D, Woźniak M, Kośmider L, et al. A neuro-heuristic approach for recognition of lung diseases from X-ray images. Expert Syst Appl 2019;126:218–32.

[211] Woźniak M, Polap D. Bio-inspired methods modeled for respiratory disease detection from medical images. Swarm Evol Comput 2018;41:69–96.