

Systematic Review of Artificial Intelligence in Acute Respiratory Distress Syndrome for COVID-19 Lung Patients: A Biomedical Imaging Perspective

Jasjit S. Suri , *Fellow, IEEE*, Sushant Agarwal , *Member, IEEE*, Suneet K. Gupta , *Member, IEEE*, Anudeep Puvvula , Klaudija Viskovic, Neha Suri, Azra Alizad, Ayman El-Baz, Luca Saba , Mostafa Fatemi , *Life Fellow, IEEE*, and Subbaram Naidu, *Life Fellow, IEEE*

Abstract—SARS-CoV-2 has infected over ~165 million people worldwide causing Acute Respiratory Distress Syndrome (ARDS) and has killed ~3.4 million people. Artificial Intelligence (AI) has shown to benefit in the biomedical image such as X-ray/Computed Tomography in diagnosis of ARDS, but there are limited AI-based systematic reviews (aiSR). The purpose of this study is to understand the Risk-of-Bias (RoB) in a non-randomized AI

trial for handling ARDS using novel AtheroPoint-AI-Bias (AP(ai)Bias). Our hypothesis for acceptance of a study to be in low RoB must have a mean score of 80% in a study. Using the PRISMA model, 42 best AI studies were analyzed to understand the RoB. Using the AP(ai)Bias paradigm, the top 19 studies were then chosen using the raw-cutoff of 1.9. This was obtained using the intersection of the cumulative plot of “mean score vs. study” and score distribution. Finally, these studies were benchmarked against ROBINS-I and PROBAST paradigm. Our observation showed that AP(ai)Bias, ROBINS-I, and PROBAST had only 32%, 16%, and 26% studies, respectively in low-moderate RoB (cutoff > 2.5), however none of them met the RoB hypothesis. Further, the aiSR analysis recommends six primary and six secondary recommendations for the non-randomized AI for ARDS. The primary recommendations for improvement in AI-based ARDS design inclusive of (i) comorbidity, (ii) inter-and intra-observer variability studies, (iii) large data size, (iv) clinical validation, (v) granularity of COVID-19 risk, and (vi) cross-modality scientific validation. The AI is an important component for diagnosis of ARDS and the recommendations must be followed to lower the RoB.

Manuscript received February 21, 2021; revised May 24, 2021 and July 2, 2021; accepted August 6, 2021. Date of publication August 11, 2021; date of current version November 5, 2021. (Corresponding author: Jasjit S. Suri.)

Jasjit S. Suri is with the Stroke Diagnosis and Monitoring Division, AtheroPoint™ LLC, Roseville, CA 95661 USA (e-mail: jasjit.suri@atheropoint.com).

Sushant Agarwal is with the Advanced Knowledge Engineering Centre, GBTI, Roseville, CA 95661 USA, and also with the Department of Computer Science Engineering, Pranveer Singh Institute of Technology (PSIT), Kanpur 209305, India (e-mail: sushant.ag09@gmail.com).

Suneet K. Gupta is with the Department of Computer Science Engineering, Bennett University, Noida 524101, India (e-mail: suneet.gupta@bennett.edu.in).

Anudeep Puvvula is with the Stroke Diagnosis and Monitoring Division, AtheroPoint™ LLC, Roseville, CA 95661 USA, and also with the Annu's Hospitals for Skin and Diabetes, Nellore 524101, India (e-mail: dr.anudeepuvvula@gmail.com).

Klaudija Viskovic is with the University Hospital for Infectious Diseases, 10000 Zagreb, Croatia (e-mail: klaudija.viskovic@bfm.hr).

Neha Suri is with the Mira Loma High School, Sacramento, CA 95821 USA (e-mail: nsuri574@gmail.com).

Azra Alizad is with the Department of Radiology, Mayo Clinic College of Medicine and Science, Rochester, MN 55905 USA (e-mail: azra.alizad@gmail.com).

Ayman El-Baz is with the Department of Bioengineering, University of Louisville, Louisville, KY 40292 USA (e-mail: ayman.elbaz@louisville.edu).

Luca Saba is with the Department of Radiology, Azienda Ospedaliero Universitaria (AOU), 09124 Cagliari, Italy (e-mail: lucasabamd@gmail.com).

Mostafa Fatemi is with the Department of Physiology and Biomedical Engineering, Mayo Clinic College of Medicine and Science, Rochester, MN 55905 USA (e-mail: fatemi.mostafa@mayo.edu).

Subbaram Naidu is with the Electrical Engineering Department, University of Minnesota, Duluth, MN 55812 USA (e-mail: naiduds@isu.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/JBHI.2021.3103839>, provided by the authors.

Digital Object Identifier 10.1109/JBHI.2021.3103839

Index Terms—COVID-19, CT, ARDS, AI, Risk-of-Bias.

I. INTRODUCTION

COVID-19 or Coronavirus is a disease that was declared a “public health emergency of international concern” or “pandemic” by the International Health Regulations Emergency Committee of the World Health Organization (WHO) on January 30, 2020. As of 20th May 2021, the WHO statistics showed more than **165** million people have been infected causing Acute Respiratory Distress Syndrome (ARDS), and nearly **3.4** million have lost their lives due to this virus [1]. There is a dire necessity to flatten the pandemic curve and prevent this severe illness during the “long-COVID-19” (beyond the COVID-19 era). The SARS-CoV-2 virus directly affects the human lungs, travels through the respiratory system and into the body [2]. However, the mutated ribonucleic acid (RNA) present in the virus makes it difficult to treat the infected patient. As per the Journal of the American College of Cardiology (JACC), cardiac troponin may help determine the risk of myocarditis [3], signaling a positive COVID-19 diagnosis [4]. Imaging, therefore, also plays a vital

role in predicting and validating the severity of the infection [5]; however, the patient’s vital clinical information further improves its ability to predict the severity better and lowers the mortality rate [6].

Artificial Intelligence (AI) has been helpful in combating such diseases because of its ability to model extensive and non-linear covariates against COVID-19 deaths in a big data framework. In previous pandemics, models were developed to flatten their mortality curves. For example, the Zika epidemic [7], Influenza type A, the H1N1 pandemic [8], and the Chikungunya epidemic [9] had shown a correlation between their data streaming using telemedicine and the pandemic curve. More recently, in China, similar telemedicine models have been adapted [10]. Therefore, we firmly believe that predicting the severity of COVID-19 using AI through computational models will be significantly useful to address the lack of software “verification, scientific and clinical validation” (discussed in section VII.B) capabilities worldwide. Note that AI-based COVID-19 severity is determined by either (i) classifying the COVID-19 pneumonia patient scans against controls or other kinds of pneumonia or (ii) locating the diseased region in the lung scan(s). Ground Glass Opacities (GGO) can be used to validate the COVID-19 severity. This study is focused on (a) ARDS that deals with the lung gas exchange disorder caused by the SARS-CoV-2, and (b) imaging of infected lungs using Computed Tomography (CT) and Chest X-rays (CXR).

In 2020, there have been six AI-based systematic reviews (aiSR) on ARDS [11–16]. However, they are incomplete, not well focused, and lack practical recommendations for safe and effective AI design for ARDS analysis. A detailed comparison among the six aiSR is discussed in the benchmarking section VII.A. In general, there are limited aiSR, which rank these studies, compute their mean scores, and determine the AI studies with low RoB. Further, this aiSR establishes a link between the AtheroPoint’s artificial intelligence-based Bias (AP(ai)Bias) and previous RoB paradigms such as the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) or Prediction model Risk Of Bias Assessment Tool (PROBAST) for handling ARDS via CT or CXR.

This aiSR uses Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) model for study selection. These studies are then analyzed to understand the role of AI in the detection of COVID-19 severity in radiological images and evaluate the RoB using AI attributes. It then presents the pathophysiology of COVID-19 leading to ARDS [17], followed by different AI techniques used in the lung segmentation and classification of the disease severity. Furthermore, to help understand the non-randomized trials’ outcome, we have proposed AP(ai)Bias, a novel bias estimation, that rates each of the ten AI attributes, computes cumulative and mean scores, and ranks the selected studies. Further, AP(ai)Bias is compared against qualitative paradigms such as ROBINS-I and PROBAST. Lastly, the aiSR presents several significant recommendations for lowering the RoB in the AI design for ARDS.

II. METHODOLOGY

A. Search Strategy

A detailed search was performed using PubMed, IEEE Xplore, ScienceDirect, ArXiv and Google Scholar. The keywords used for selecting studies were COVID-19, ARDS, deep learning, lung segmentation, classification, lung CT, X-ray, and

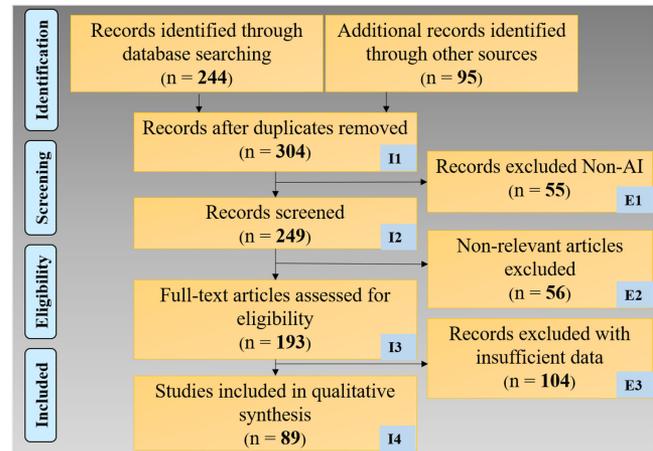


Fig. 1. Search strategy using the PRISMA model.

AI. Fig. 1 shows the PRISMA model consisting of the studies used in this review. A total of 339 studies were identified, and duplicates were removed using the feature called “Find Duplicates” in EndNote software by Clarivate Analytics [18], thus, retaining 304 records. The three exclusion criteria were (i) studies not related to AI, (ii) non-relevant articles, and (iii) having insufficient data. This excluded 55, 56, and 104 studies (marked as E1, E2 (non-AI, but COVID-19), and E3 in Fig. 1), leading to the final selection of 89 studies.

B. Hypothesis and the Acceptability Criteria

We hypothesize that “non-randomized Artificial Intelligence-based attributes can (a) detect, (b) classify, (c) estimate severity of the COVID-19 risk, and (d) meets the performance standards in lung infected ARDS patients.” Three acceptability criteria were: (i) for the AP(ai)Bias-based ranking method, the mean score must be greater than or equal to 80% for an AI-based study while taking into consideration all the AI attributes [19]. This was due to the consensus of the experienced team and five different classes for each AI attribute based on its strength (such as low, moderate, high-moderate, low-of-a-high, and high-of-a-high, ranging from 1 to 5). Similarly, for the (ii) ROBINS-I and (iii) PROBAST paradigms, our acceptability criterion must meet the score of 80% or above for AI-based studies to be in the low-RoB zone [19], [20].

III. PATHOPHYSIOLOGY OF ARDS

The number of publications and cases discussing ARDS infected by SARS-CoV-2 has been increasing over time [21]. The first set of cases in Wuhan, China, reported having COVID-19 patients hospitalized for lower respiratory tract (LRT) complaints [22]. They also indicated that the symptoms of COVID-19 are incredibly diverse, ranging from minimal LRT symptoms to significant hypoxia due to ARDS [22]. Further, Huang *et al.* [23] reported that the time gap between the onset of minimal LRT symptoms to ARDS was as short as nine days, suggesting that minimal LRT symptoms could progress rapidly in these patients. Recent studies support that in COVID-19 patients, ARDS has higher rates than extrapulmonary complications [24]–[26]. The pathophysiology of ARDS in COVID-19 patients begins with SARS-CoV-2 entering into the lung by aerosol transmission [27], as outlined in Fig. 2 (note, numbers 1-12

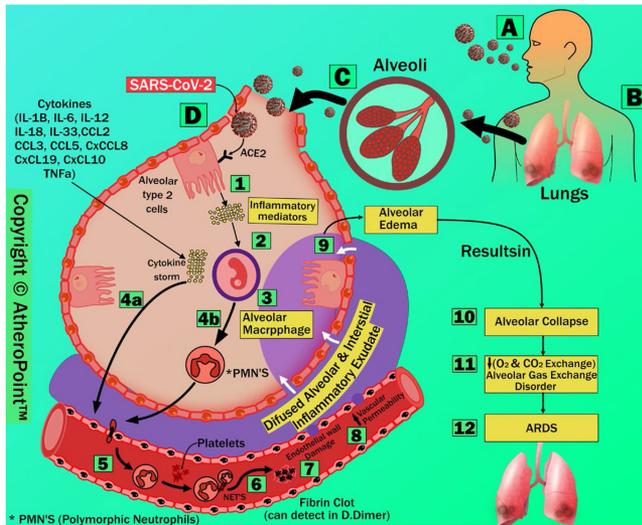


Fig. 2. Stages of ARDS (Courtesy of AtheroPoint, CA, USA).

correspond to letter **D1-D12**). The attachment of SARS-CoV-2 to the host cells occurs via the anchoring of its virion spike proteins (S1 and S2) to the angiotensin-converting enzyme 2 (ACE2) receptor on the surface of the type 2 pneumocytes on the pulmonary alveolar epithelium (**D1**). This causes respiratory symptoms to present as the earliest clinical presentation of COVID-19 [28].

Due to infection, the inflammatory process begins and leads to inflammatory mediators' production [29], [30] (**D2**). Moreover, these inflammatory mediators stimulate alveolar macrophages in producing polymorphonuclear neutrophils (PMNS) and cytokines such as IL-1, IL-6, and TNF- α (**D3**, **D4a**, and **D4b**). Additionally, cytokine hyperproduction causes a cytokine storm. The sequence of steps in the systemic inflammatory response, cytokine storm, and multiple organ failure plays a critical role in ARDS development [31].

Previous coronaviruses had also observed the same process of cytokine storm and ARDS development [32]. The produced PMNS recruits platelets and forms Neutrophil extracellular traps (NET) that causes endothelial dysfunction [33]. NET's are web-like structures of deoxyribonucleic acid (DNA) and proteins created after PMNS activation and infection. The key enzymes that help develop NET's are neutrophil elastase, type 4 peptidyl arginine deiminase, and gasdermin-D [34]. Although NET's have a beneficial role in the host defense against pathogens, they also have a detrimental role. They are created by facilitating micro thrombosis, resulting in permanent organ damage to the lung, heart, and kidney [35]. Additionally, cytokine storm causes endothelial dysfunction and creates gaps between cells, increasing vascular permeability [22], [36] (**D5**, **D7**, and **D8**). Increased vascular permeability causes fluid leakage, resulting in increased alveolar space and diffuse inflammatory alveolar exudate, causing alveolar edema [37–39] (**D8** and **D9**), typically seen in CT lung scans. Alveolar edema then leads to increased alveolar surface tension, a critical feature of diffuse alveolar atelectasis (i.e., damage or collapse) [40], [41] (**D10** and **D11**). Following the alveolar collapse, ventilation-to-perfusion mismatch, ARDS occurs, that results in an impairment of carbon dioxide excretion due to the increased alveolar dead space [42] (**D12**).

IV. ARTIFICIAL INTELLIGENCE BUILDING BLOCKS FOR ARDS: LUNG SEGMENTATION AND CLASSIFICATION

For a comprehensive RoB analysis, it is customary to investigate the basic building blocks of the ARDS pipeline. As shown in Fig. 3, the two major components are lung segmentation and COVID-19 severity classification. We will briefly study AI models' statistical distribution and AI architectures for these two components.

A. Statistical Distribution of AI Models

Even though PRISMA selected 89 studies, only 42 were AI-based studies [43]–[84]. AI-based image classification occurred in 85% [43–50], [52], [53], [55], [56], [58–62], [64–70], [73–80], [83] of the selected studies, while lung segmentation with or without classification was 3% [51], and 12% [57], [63], [71], [72], [84], respectively (Fig. 4(a)). In terms of risk granularity, AI-based on binary classification (BC) showed **46%**, while multiclass (MC) paradigm for classification showed **39%**, and hybrid (combination of segmentation and classification) were **15%** (Fig. 4(b)). Most of the studies used 2-D (82%) [43–52], [54], [57–62], [64–69], [71–80], while others were used 3-D (12%) [53], [63], [70], [83], [84], and 6% [55] used both 2-D and 3-D (Fig. 5(a)). The total images used were **51%** in 2-D, only **1%** in 3-D, and **48%** used both (Fig. 5(b)).

The landscape for AI models for image classification consisted mainly of machine learning (ML), deep learning (DL), transfer learning (TL), recurrent learning (RL), [4], and an amalgamation of these learning paradigms [85] (Fig. 6). The granular division of the AI-based classification paradigm consisted of DL (**50%**) [43], [46], [47], [49], [50], [52], [53], [55], [56], [59–61], [64–66], [69], [78–80], [83], DL and TL (**12%**) [67], [68], [75–77], TL (**10%**) [44], [46], [58], [74], DL and ML (**5%**) [48], [73], ML (**5%**) [45], [62], and RL combined with DL (**3%**) [70]. Note only **six** studies (**15%**) [51], [57], [63], [71], [72], [84] were focused on lung segmentation.

B. Lung Segmentation Architectures for COVID-19 ARDS

Lung segmentation is a process of extraction of the lung region in CXR projections or in CT-slices [86]. In the 2-D paradigm, (i) Rajaraman *et al.* [72] implemented UNet architecture with Gaussian dropout in CXR (Fig. 7). As the shape resembles the letter U, the contraction path (also known as encoder) is the left arm of U and captures the context of the image using traditional convolutions and max-pooling layers. The expanding path (also known as decoder) is the right arm of U that enables precise localization using transposed convolutions. The depth of the UNet is the number of layers in the UNet architecture and is responsible for the performance, while having a tradeoff between accuracy and computational cost [72]. (ii) Oh *et al.* [57] implemented Fully connected (FC)-DenseNet103 transfer learning model (architecture shown in Fig. 8 and results shown in Figure A(left), see Supplemental D) offered the advantage of saving time by using the pre-trained weights. In the 3-D paradigm, Wang *et al.* [63] implemented DenseNet121-FPN (Figure A(right), see Supplemental D), the only study that performed 3-D lung segmentation using CT volumes (Fig. 9). In FC-DenseNet models, 2-D or 3-D, the feature maps created by the preceding dense block are up-sampled to prevent a large number of computations and parameters. Note that, all

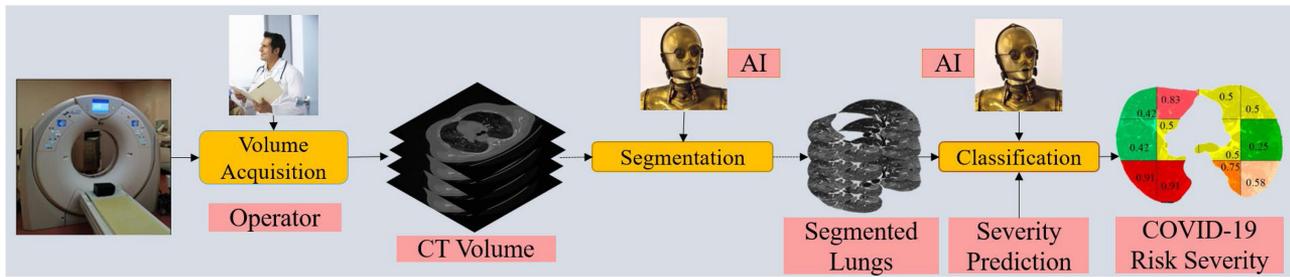


Fig. 3. AI-based ARDS pipeline in lung CT for COVID-19 severity prediction (Courtesy of AtheroPoint, Roseville, CA, USA).

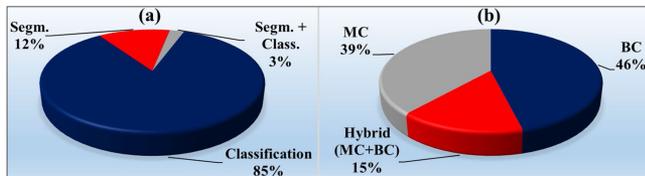


Fig. 4. (a) #studies for lung segmentation vs. classification; (b) #studies for BC and MC frameworks.

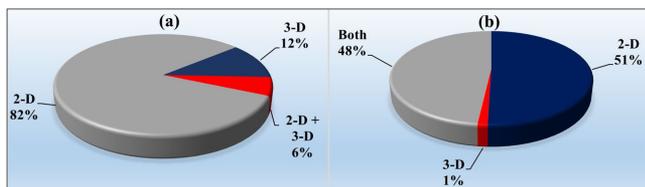


Fig. 5. (a) #studies in 2-D vs. 3-D; (b) #images in 2-D vs. 3-D.

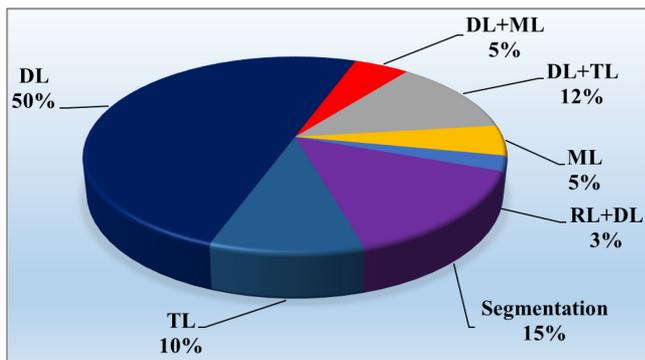


Fig. 6. Four kinds of AI models (with and without segmentation) in the selected 42 AI-based studies.

the above architectures adapted a skip connection between the down-sampling path and up-sampling paths for transferring weights. 2-D CXR is preferred over 3-D CT volume imaging due to lower costs. The in-depth comparison between 2-D and 3-D segmentation is shown in Table I.

C. Lung Classification Architectures for COVID-19 ARDS

The pipeline for ARDS diagnosis consists of the classification of lung scans based on COVID-19 risk severity. The AI model using ground truth with two classes leads to binary class (BC) framework, while models using multiple ground truths yielded

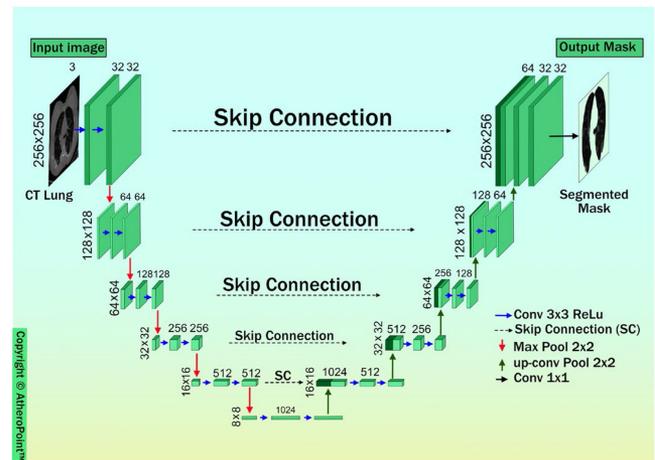


Fig. 7. UNet architecture for lung segmentation (Courtesy of AtheroPoint™, Roseville, CA, USA).

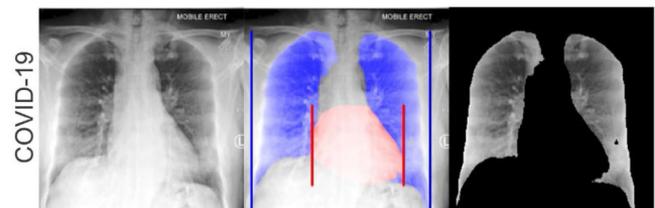


Fig. 8. Lung segmentation using DenseNet [57] using CXR (Reproduced with permission).

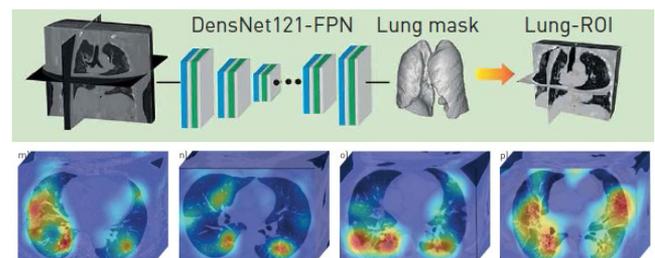


Fig. 9. (a) 3-D segmentation of the CT lung images and (b) heat map using a DL system [63] (Reproduced with permission).

multiclass (MC). Studies that adapted both segmentation and classification (using either BC or MC) were categorized into hybrid class (HC). The statistical distribution between the three types of classes was 46%, 39%, and 15%, respectively. Most of

TABLE I
2-D vs. 3-D LUNG SEGMENTATION IN COVID-19 ARDS

SN	Attributes	2-D ARDS Segmentation	3-D ARDS Segmentation
1	Acquisition	2-D CXR and 2-D CT slices	3-D Volume
2	Modality	CXR and CT slice	3-D CT or PET
3	Frame Processing	Single frame processing at a time	Spatially oriented joint processing
4	Statistical Processing	Pixel processing	Voxel processing
5	Segmented Coverage	Limited	Complete
6	Architecture Type	FC-DenseNet103 [57], UNet [72]	DenseNet121-FPN [63]
7	Arch. Complexity	Less complex (fewer #layers)	More complex (more #layers)
8	Parameters Adapted	~11.6 million	~8 million
9	Convolution Filters	2-D Convolution filter	3-D Convolution filter
10	Processing Speed	Faster, due to low volume	Slower, due to the large volume
11	Memory Size	Consumes low memory	Consumes high memory

the studies that conducted BC used CXR [44], [75], [76], [87] while MC (such as DenseNet [52] and truncated Inception Net for CXR [46], COVNet (a modified Resnet-50 architecture) [83] and COVIDNet-CT [79] for CT, and VGG-16 for Ultrasound [74]) adapted transfer learning (TL) architecture ranging from 5-fold to 10-fold cross-validation paradigms (Figure B, see Supplemental D). The best accuracies of 96% and 99% in BC and MC were obtained by Brunese *et al.* [44] and Gunraj *et al.* [79]. A unique observation was seen in Ozturk *et al.* [59] for validating the COVID-19 output that consisted of the superimposition of colored heat maps on grayscale CXR using GRAD-CAM [59]. In HC, three studies had used 2-D CXR [57], [59], [72] while two studies had used 3-D CT [53], [63]. Wu *et al.* [64] implemented a new HC *multi-view* fusion CT technique using the modified Resnet-50 on a patient size of 495, demonstrating accuracy of 83.33% and AUC~0.905 ($p < 0.001$). This was 6% better compared to the *single view* paradigm. Oh *et al.* [57] also showed that better classification performance could be achieved when trained using FC-DenseNet103 on segmented infected lung region in 2-D CXR. It is however noteworthy to explore factors that affects the performance of AI models which we discuss next.

V. PERFORMANCE EVALUATION OF AI TECHNIQUES

A. Image Quality Assessment: CT vs. CXR

The image quality plays an important role when it comes to computer-aided diagnosis (CAD) performance [86], [88], [89]. The CT image quality can also be fuzzy or degraded due to the radiation dosage [90]. Further, due to high COVID-19 severity, there is a fluid leakage or alveolar edema, which causes lung-scans to show hyperintensity distribution (due to high consolidation and ground-glass opacity). Digital CXR is preferred over conventional CXR due to high-resolution imaging [92]. Several methods were designed to assess signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) in CAD-based static and motion imagery [93]. Image registration is essential during the image quality assessment [94], [95]. However, the effect of an image quality assessment on AI for COVID-19 has not been well researched. The AI studies used in this *aiSR* did not demonstrate AI performance due to image degradation. The lack of image quality testing or the effect of image quality assessment on the AI performance will lower the overall score estimation during the comprehensive evaluation.

B. AI-Based Performance Parameters

Comparison of the AI models in the studies was based on accuracy, sensitivity (SEN), specificity (SPE), F1-score, precision, and recall metrics. The missing data was calculated based on the studies existing data, such as SEN, SPE, patient size, and sometimes using the true positive rate. The mean accuracy of the AI models extracted was $93.05 \pm 5.7\%$. The best accuracy was 98% [59], while the sensitivity and specificity had a mean of $92.01 \pm 10.25\%$ and $89.80 \pm 15.42\%$, respectively. All the studies combined achieved a mean F1-score and precision of $93.72 \pm 6.36\%$ and $93.51 \pm 4.45\%$, respectively. Thus these performance metrics show high values, so certainly, these attributes will have strong contributions to the overall score when computing $AP(\mathbf{ai})\text{Bias}$. Note that the accuracies computed above considered the augmentation technique during data preparation process (also shown in recent studies [96–98]).

VI. THREE PARADIGMS FOR THE RISK-OF-BIAS ESTIMATION

Our current novel ranking strategy, $AP(\mathbf{ai})\text{Bias}$ helps to identify AI-based studies that are comprehensive, complete, error-free, safe, and effective. On the contrary, ROBINS-I [99] and PROBAST [100] helps in identification of factors that contribute to RoB. This section is therefore focused on three kinds of paradigms for RoB in non-randomized AI for ARDS.

A. $AP(\mathbf{ai})\text{Bias}$: Ranking Paradigm for RoB

We have considered 10 attributes for AI evaluation in $AP(\mathbf{ai})\text{Bias}$ model that includes (i) segmentation and classification, (ii) cross-validation (CV) protocol, (iii) inter- & intra-observer variability (IIV), (iv) benchmarking, (v) scientific validation (SV) and clinical evaluation (CE), (vi) learning paradigm (LP), (vii) data preparation (DP), (viii) data size, (ix) performance evaluation (PE), and (x) innovation. Note that DP attribute was further subdivided into (vii.a) augmentation protocol [96], [97], [101], (vii.b) region-of-volume selection, given the CT volume, (vii.c) format conversion declaration, (vii.d) manual tracings for ground truth or goal standard to generate binary shapes, and (vii.e) baseline characteristics demonstration. Each attribute (i-x) can earn up to a maximum of five points as presented in the section “Hypothesis and the Acceptability Criteria”. The value obtained in each attribute is pro-rated based on the threshold adapted for the attributes. These values of each attribute are then added, leading to an accumulated score.

TABLE II
ROBINS-I METHOD

	Low Risk	Moderate Risk	High Risk	Unclear Risk				
Study	C1	C2	C3	C4	C5	C6	C7	
[64]	L	M	L	L	M	L	L	
[46]	H	H	M	M	H	L	H	
[58]	H	H	H	M	H	M	H	
[59]	M	L	L	L	H	L	M	
[49]	H	H	M	M	H	M	H	
[66]	H	H	L	M	H	M	H	
[43]	H	H	L	M	H	M	H	
[47]	L	L	L	M	L	U	U	
[71]	L	L	L	M	M	L	L	
[75]	H	H	H	M	H	M	H	
[62]	H	H	L	L	H	H	H	
[48]	H	H	H	U	U	H	H	
[50]	L	L	H	L	L	U	M	
[45]	L	L	L	U	L	U	M	
[73]	L	H	L	M	L	M	M	
[77]	M	H	H	M	H	M	M	
[80]	L	L	M	M	L	M	M	
[79]	L	H	H	U	L	L	M	
[81]	M	L	M	M	L	L	H	
Parameters	(i) Bias due to confounding		(ii) Bias in selection of participants into the study	(iii) Bias in classification of interventions	(iv) Bias due to deviations from intended interventions	(v) Bias due to missing data	(vi) Bias in measurement of outcomes	(vii) Bias in selection of the reported result
	Pre-intervention		At-intervention	Post-intervention				

TABLE III
PROBAST METHOD

	Low Risk	Moderate Risk	High Risk	Unclear Risk
ID	Participants	Predictors	Outcomes	Analysis
[64]	L	L	L	L
[46]	H	M	M	M
[58]	H	H	M	H
[59]	M	L	M	U
[49]	H	U	M	H
[43]	H	M	M	H
[66]	H	M	M	U
[47]	L	L	M	U
[71]	L	M	M	L
[75]	H	H	M	U
[62]	H	L	M	H
[48]	H	H	M	H
[50]	L	M	L	U
[45]	L	M	L	U
[73]	M	L	M	M
[77]	M	H	M	U
[80]	L	M	L	M
[79]	M	M	L	M
[81]	M	L	L	M

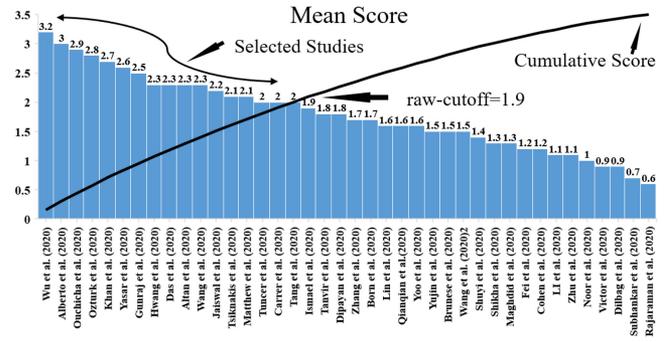


Fig. 10. Selected 19 studies out of 42 based on the raw-cutoff of a mean score greater than equal to 1.9.

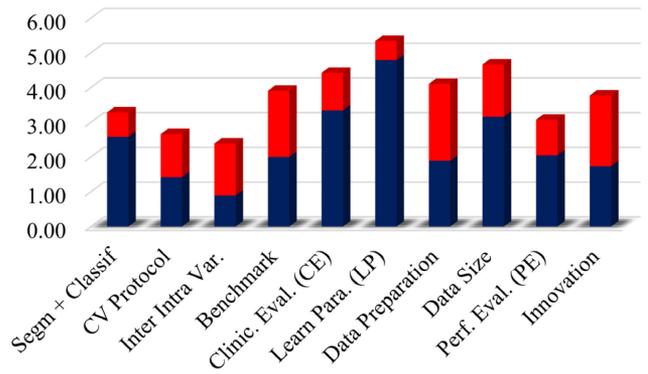


Fig. 11. Mean score using ten attributes on top 19 studies.

Finally, these 42 scores (corresponding to 42 studies) were then ranked in decreasing order (low-bias to high-bias). The raw-cutoff (1.9) for the selection of AI studies was determined based on the intersection of the “cumulative plot of mean score vs. studies” and the “descending order of distribution of the scores of the studies” (Fig. 10). Table V shows the AP(ai)Bias model-based ranking of the best 19 AI-based studies taken from a pool of 42 studies. According to the ranks, the color was assigned, where green is given to the group of the high-rank (low-bias), yellow for mid-rank (moderate-bias), and red for the low-rank (high-bias). Note that the maximum rank a study can obtain is 50 (10 attributes multiplied by a maximum of 5 points for each attribute). Using the above strategy and applying to all the 10 attributes to 19 studies, the top three contenders came out to be Wu *et al.* [64], Alberto *et al.* [71], and Ouchicha *et al.* [58].

B. Interpretation of the AP(ai)Bias Strategy

To interpret these results, we analyze the mean scores for each of the ten attributes over all the 19 studies (see Table V: C1 to C10). The mean and the standard deviation of these attributes are shown in Fig. 11, which can also be used for computing the percentage contribution of these attributes. (i) Learning paradigm (C6) was the highest scorer amongst all the attributes, since 86% of the studies were supervised. (ii) Inter- & intra-observer variability (C3, label as IIV) showed the lowest mean value amongst all the studies. This demonstrated that the evaluation of AI techniques was undermined and inconclusive. (iii) Data size (C8) attained a mean value of 2.7, meaning that most of the data size was in the range of 100 to 500 subjects. (iv) Four

TABLE IV
BENCHMARKING TABLE FOR SYSTEMATIC REVIEWS FOR COVID-19 ARDS

SN	Authors	Nagendran <i>et al.</i> [13]	Roberts <i>et al.</i> [14]	Wynants <i>et al.</i> [15]	Bao <i>et al.</i> [12]	Albahari <i>et al.</i> [11]	Montazeri <i>et al.</i> [16]	Suri <i>et al.</i>
Dates	Attributes	February 2020	October 2020	April 2020	June 2020	June 2020	April 2021	Proposed 2021
1	ROBINS	x	x	x	x	x	x	✓
2	PROBAST	✓	✓	✓	x	x	✓	✓
3	AP(ai)Bias Ranking	x	x	x	x	x	x	✓
4	AP(ai)Bias interpretation	x	x	x	x	x	x	✓
5	PRISMA	91	45	27	13	11	44	89
6	Study Considered	91	45	27	13	11	44	42
7	References	41	84	71	30	109	80	89
8	Patho section	x	x	x	x	x	x	✓
9.1	Recommendations – R1	Transparency	Input data	CRF	Multiethnic	MCDA	Fine-tuning	Data Size
9.2	Recommendations – R2	High Quality	Robust training	PTC	-	AHP	Data augm	Clinical Validate
9.3	Recommendations – R3	-	Reproducibility	VTEM	-	-	-	IIV
9.4	Recommendations – R4	-	High Q doc	-	-	-	-	Comorbidity
9.5	Recommendations – R5	-	Peer review	-	-	-	-	Risk Granularity
9.6	Recommendations – R6	-	-	-	-	-	-	Reproducible
10	Segmentation	x	x	x	x	x	✓	✓
11	AI arch layout	x	x	x	x	x	✓	✓
12	Performance of AI studies	✓	✓	✓	✓	✓	✓	✓

*CRF: Conventional risk factor consists of age, body temperature, and (respiratory) signs and symptoms; for prognostic models, age, sex, C reactive protein, lactic dehydrogenase, lymphocyte count, and potentially features derived from CT scoring; PTC: Point the challenges in design; VTEM: validate the existing models; IIV: Inter- and intra-observer.

TABLE V
AP(AI)BIAS METHODOLOGY BASED RANKING OF THE AI STUDIES FOR ARDS

SN	Author (2020)	C1 Segm. & Class.	C2 CV Protocol	C3 IIV	C4 Bench	C5 CE	C6 LP	C7 DP	C8 Data Size	C9 PE	C10 Innov	C11 Absolute Score	C12 Mean Score	C13 Study Rank
1	Wu <i>et al.</i> [64]	3	2	1	4	3	4	5	3	3	4	32	3.2	1
2	Alberto <i>et al.</i> [71]	1	2	4	1	4	3	5	5	2	3	30	3	2
3	Ouchicha <i>et al.</i> [58]	3	3	0	5	4	4	0	2	3	5	29	2.9	3
4	Ozturk <i>et al.</i> [59]	4	3	0	5	3	5	0	2	2	4	28	2.8	4
5	Khan <i>et al.</i> [49]	3	3	0	4	4.5	5	3	2	2	0	26.5	2.7	5
6	Yasar <i>et al.</i> [66]	2	2	0	3	4	5	4	3	3	0	26	2.6	6
7	Gunraj <i>et al.</i> [79]	3	0	0	2	3	5	0	5	3	4	25	2.5	7
8	Li <i>et al.</i> [81]	3	3	1	0	3	5	1	3	1	4	24	2.4	8
9	Wang <i>et al.</i> [80]	3	0	3	2	3	5	0	5	2	0	23	2.3	9
10	Altan <i>et al.</i> [43]	3	0	0	2	3	5	5	2	3	0	23	2.3	10
11	Das <i>et al.</i> [46]	3	0	0	5	4	5	0	2	3	1	23	2.3	11
12	Hwang <i>et al.</i> [47]	2	1	4	2	4	5	0	5	0	0	23	2.3	12
13	Jaiswal <i>et al.</i> [75]	2	0	0	0	3	5	4	5	3	0	22	2.2	13
14	Tsiknakis <i>et al.</i> [77]	3	3	0	0	2.5	5	1	5	1	0	20.5	2.1	14
15	Matthew <i>et al.</i> [37]	3	0	1	0	2.5	5	5	4	0	0	20.5	2.1	15
16	Tuncer <i>et al.</i> [62]	2	1	0	0	0	5	3	1	3	5	20	2	16
17	Tang <i>et al.</i> [73]	2	2	0	0	4.5	5	0	2	1	3	19.5	2	17
18	Ismael <i>et al.</i> [48]	2	2	3	0	4.5	5	0	1	2	0	19.5	2	18
19	Carrer <i>et al.</i> [45]	2	0	0	3	4	5	0	3	2	0	19	1.9	19

*Segm: Segmentation; class: Classification; IIV: Intra- and inter-variability; Innov: Innovation; Bench: Benchmarking; PE: Performance evaluation; CE: Clinical Evaluation; LP: Learning Paradigm; DP: Data Preparation

attributes scored mediocre mean values, namely segmentation and classification (C1), benchmarking (C4), scientific validation and clinical evaluation (C5), and performance evaluation (C9), contributing to be 44%, 26%, 45.5%, and 30%, respectively. This clearly shows the segmentation and classification were binary. Only 26% of the studies implemented benchmarking. Amongst all the studies, 45.5% did the clinical evaluation with the radiologist or used a pre-evaluated cohort for the training. The attribute performance evaluation (C9) scored 31.5%, indicating that the accuracy of the AI models was around 93%. Our observation showed that only 6 studies ([49], [58], [59], [64], [66], [71]) (Table V) were in the low-bias pool (bias-cutoff greater than 2.5, C12, Table V) ((6/19) * 100 ~ 32%), that passed the hypothesis.

C. ROBINS-I

The objective of ROBINS-I is to mimic the randomization of the non-randomized studies. It covers seven distinct attributes (domains) divided into three intervention factors (marked parameters) for bias during (a) “Pre-Intervention,” (b) “At-Intervention,” and (c) “Post-Intervention,” through which bias can be studied. Table II shows the bias due to (i) confounding factors (data size and data source); (ii) selection of participants (training and testing protocols); (iii) classification of interventions (data augmentation and use of imaging features); (iv) deviations from intended interventions (demographics, multicenter data, and comorbidity); (v) missing data (SWAB test,

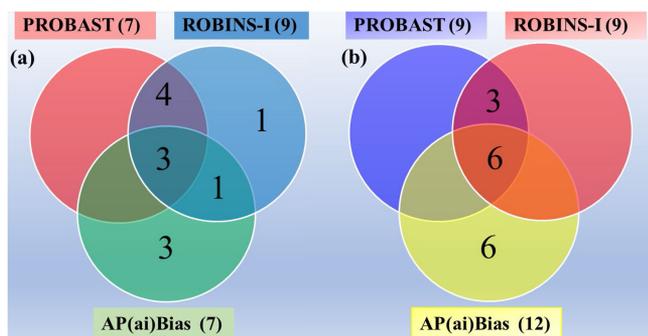


Fig. 12. Venn diagram shows the three bias paradigms using: (a) high-bias; (b) moderate-high bias scenarios.

patient follow-up); (vi) measurement of outcomes (innovation, optimization, validation by the radiologist, COVID-19 data size); and (vii) selection of the reported result (prevents it from being included in the meta-analysis). Table II (seven-column marked C1 to C7) below discusses the outcome of 19 studies (“Study” column). The three-color scheme is adopted to depict the outcome of the qualitative analysis. Red color means high-bias, indicating a severe issue in the study concerning the factors taken into consideration. Moderate-bias is depicted by yellow, indicating that the study holds good on the given set of non-randomized data, and the green means low-bias, as the study performs comparatively well on the testing parameters and the input data. We conclude from ROBINS-I (Table II), “Study column, shaded in pink color” that $\sim 73\%$ ($(14/19) * 100$) have at least one of the attributes with high-bias. Only three studies ([64], [71], [80]) passed the hypothesis test ($(3/19) * 100 \sim 16\%$). Further, we also conclude that 60% ($(79/132) * 100$) were low-moderate bias (green and yellow color) using the cell division approach.

D. PROBAST

It is a popular AI-based RoB assessment tool. It uses four attributes as shown in Table III, where (a) *participants*, were the source of the image database and whether the radiologist verified them, (b) *predictors*, demographic data availability (Yes/No), and imaging features. (c) *outcomes*, consisted of the factors that whether different datasets were combined and if the reverse transcription-polymerase chain reaction (RT-PCR) test was conducted for the cohort, and (d) *analysis*, covers the cohort size, number of COVID-19 patients, optimization techniques used, validation, and the innovation in the design. We used the same 19 studies adapted for the AP(ai)Bias analysis ranking (Table V). Using PROBAST, we conclude that $\sim 47\%$ (9 out of 19) studies were high-bias (marked as H in red color); some even had an unclear RoB. Five studies [64], [71], [79–81] out of the 19 selected studies passed the hypothesis ($(5/19) * 100 \sim 26\%$). Using the cell division approach, PROBAST showed 67% ($(51/76) * 100$) were low-moderate bias.

E. Analysis of RoB Using Venn Diagram

Following steps were adapted to generate the Venn diagram (Fig. 12). (i) Conversion of ROBINS-I (Supplemental A) and PROBAST (Supplemental C) from qualitative scheme to quantitative scheme, using the conversion scores of low-biases

to 5, moderate-bias to 3, high-bias to 1, and unclear-bias to 0. (ii) Selection of common studies between (a) ROBINS-I and PROBAST (left: [43], [48], [49], [58], [66], [75], [77], right: [43], [46], [48], [49], [58], [62], [66], [75], [77]), (b) PROBAST and AP(ai)Bias ranking (left: [48], [75], [77], right: [43], [46], [48], [62], [75], [77]). (c) ROBINS-I and AP(ai)Bias (left: [48], [62], [75], [77], right: [43], [46], [48], [62], [75], [77]) and (d) between all the three paradigms (top: [48], [75], [77], bottom: [43], [46], [48], [62], [75], [77]).

This digital count is shown in the Venn diagram, Fig. 12 for high-bias and moderate-high bias. These counts are 7, 3, 4, 3 for high-bias, and 9, 6, 6, 6 for moderate-high bias, respectively. (iii) These digital counts are then converted into percentage by normalizing it by total studies (19), shown in Fig. 12. These percentages are 39%, 17%, 23%, and 17%, and 50%, 33%, 33%, 33% for high-bias and moderate-high bias, respectively. The top two low-bias studies [64], [71] were the same in all three paradigms. It was interesting to note that reference [66] was low-bias using AP(ai)Bias, while it was in moderate-high bias for ROBINS-I and PROBAST. This is because the AP(ai)Bias takes into consideration attributes like CE, LP and DP. Note that Fig. 12 uses bias-cutoff of 2.2 and 2.5 for high-bias and moderate-high bias, respectively.

VII. DISCUSSIONS

The main contributions of the aiSR first showed the basic pipeline of the ARDS framework (Fig. 3). The study then showed the statistical distribution visually for (a) different AI models, (b) image modalities, and (c) image dimension type (2-D vs. 3-D). The crux of the aiSR was to understand the Risk-of-Bias (RoB) in a non-randomized AI trial for handling ARDS using three paradigms: AP(ai)Bias, ROBINS-I, and PROBAST. AP(ai)Bias consisted of ten main attributes and several sub-attributes, while the ROBINS-I and PROBAST consisted of seven and four attributes. These two qualitative assessments were quantified using the same strategy as in AP(ai)Bias. This framework further allowed us to show how the AP(ai)Bias based ranking strategy can be compared against ROBINS-I and PROBAST using high-RoB and low-RoB cutoff’s and pictorially represented using Venn diagram. The study further presented 12 (six primary and six secondary) recommendation for high-RoB studies for better AI designs for ARDS. Finally, the study correlated the pathophysiology of ARDS and lung damage process represented by GGO. In terms of imaging modality, CXR [46], [57], [59], [72], [80] is more economical, and it is seen that more than 50% of the studies had used them. GGO has been one of the most common manifestations in CT image volumes, but these may vary from person-to-person. It has been noted that data from the patients with rapid progression of the disease show a faster change in lung lesions, but to incorporate this in the AI framework, more data for the AI-model is required to train them. It also requires a follow-up on the patient’s condition. The evaluation of the AI models has not been done using randomized trials, so to overcome this, we have proposed the usage of novel methods for the RoB analysis using ROBINS-I and PROBAST.

A. Benchmarking: Comparative Study on aiSR

Table IV shows a comparison between the previous aiSR, where 12 types of attributes were chosen to compare the five studies [11–15]. The proposed study is in the last column,

labelled “Suri.” Note that we offer “√” in places for unique contribution in the “Suri model.” We also offer recommendations in clinical validation, inter-and intra-observer variability, comorbidity, and risk granularity.

B. Recommendations

The primary set of six-point recommendations is: (i) *Comorbidity*: This section focuses on a novel intuitive approach that can lead to new improvements to the present prevailing methods for COVID-19 diagnosis. Studies have shown that comorbidities like age, ethnicity, hypertension, diabetes, higher BMI, respiratory disorders, hyperlipidemia, and obesity lead to worsening of ARDS [102], [103]. In these studies, ARDS is more prominent among the older patients, and if given a timely prognosis, can be controlled to reduce the risk factor and enhance the efficacy related to COVID-19. For validation of this data, a randomized controlled trial is necessary where the data is clinically validated. With AI techniques and comorbidity factors, this disease’s prognosis and diagnosis can be made more accurate. (ii) *Inter- and intra-observer variability*: Our observation shows that only **12%** of the AI-based studies attempted inter-and intra-observer variability study analysis. This does not assure that the AI results are robust. An example of comprehensive IIV can be seen in [104], [105]. Similarly, inter-and intra-operator variability can be computed, ensuring further reliability in clinical settings. (iii) *Data Size*: Even though **50%** of the AI studies for ARDS had a total number of subjects <500, this could be improved by targeting 1,000 and above. Typically, clinical trials adapted in meta-analysis range to many thousands (>15,000). This also requires conducting “power-analysis” for the AI system [106–108]. (iv) *Scientific Validation and Clinical Evaluation*: It requires the engineering AI-based design to be validated by the clinical community in terms of reliability, accuracy, and reproducibility. Therefore, the medical, scientific, and engineering community needs to collaborate more closely. They could also incur more costs for system design, should the medical community participate for a longer duration [109], [110]. (v) *Risk Granularity*: Risk assessment in several medicine fields is binary; however, such a strategy poses a challenge during drug prescription and better patient care. For this reason, new strategies have evolved recently where a multiclass framework provides a stronger granularity of risk leading to better control of medications and monitoring [111], [112]. This requires multiple classes in the ground truths design for COVID-19 severity. This means a careful examination of the CT lung images by the radiologist in conjunction with the pulmonologist. The second alternative is to stratify the risks of the AI systems’ output in continuous values between 0 and 1, thereby partitioning the output into multiple classes. Note that the classes’ thresholds are based on the baseline characteristics of the input cohort’s demographics.

Note that the classes’ thresholds are based on the baseline characteristics of the input cohort’s demographics. (vi) *Scientific validation using cross-modality fusion*: Scientific validation is crucial for ensuring the AI system’s correct functioning. One way could be looking at COVID-19 using two different angles, such as imaging of the lung using CT and positron emission tomography (PET). With the advanced technology of combined PET/CT, one can visualize the PET images’ metabolic distribution corresponding to the spatial CT [113]. Fig. 13 below shows the PET image showing the functional metabolic distribution

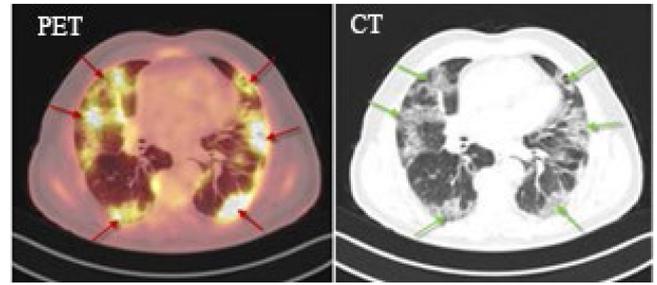


Fig. 13. PET-CT pair of a COVID-19 lung ([113]).

of COVID-19. The selected studies did not consider PET/CT fusion as part of a systematic review.

The secondary set of recommendations include (i) solid model design (training and prediction) [86], [89], [114], (ii) reproducibility [104], (iii) process of peer-review, (iv) high-quality documentation, (v) multi-ethnic and multi-regional data collection, and (vi) multiple ground-truth events for clinical validation [11]. Since all the three paradigms of hypothesis (section II.B) were invalid, we thus conclude that one requires the above concrete recommendations for improving AI design for meeting the requirements of the hypothesis keeping in mind for optimal AI performance. Note that in our study the AI solutions did not consider the socio-economic causes during the design. Recently a study [115] was conducted elaborating the need for inclusion of socio-economic conditions for COVID-19. This could possibly be adapted as an extension to the current work. The role of socio-economic conditions during the ML design was attempted by our group for neonatal deaths in different counties of Bangladesh [116]. A similar approach can also be adapted for COVID-19 based on geography, social-economic conditions.

C. Strength, Weakness, and its Extension

This main strength of this review is the selection of the best 19 AI studies for analyzing RoB. Most of the studies adapted DL as their base architecture due to the medical imaging source. As a result, DL was the best suited for its purpose. We observed that data augmentation was also used in **43%** of the studies where the data was lacking. For the selected set of studies, we have successfully analyzed the risk factors using two schemes, ROBINS-I and PROBAST. There was a gap in the studies that there was no linking on how they did COVID severity, clinical validation, variability study, and cross-modality [117]. Some of the studies did not mention about hardware constraints. On top of all, none of them showed any 510(K) FDA approval. There is much inconsistency in the studies, which can be fixed by adopting some initial clinical validation on the patients using cross-modality. Even RL, TL with cross-modality, and comorbidity can introduce innovation to this lifesaving study worldwide. Understandably, this may take more time for the medical community but could cut corners on patient care.

VIII. CONCLUSION

This study used the PRISMA model to select **89** studies, which was then AI-filtered to **42**. Based on the mean score ranking, the selection was further refined to **19** studies using a raw cutoff of **1.9**. Further, the three RoB paradigms were analyzed

using the Venn diagram. The percentage of studies that satisfied the “non-randomized Artificial Intelligence-based” hypothesis for ARDS-based COVID-19-infected lungs were only **32%**, **16%**, and **26%**, corresponding to AP(ai)Bias, ROBINS-I, and PROBAST, respectively. This percentage obtained was using high-bias cutoff of **2.2** and moderate-high bias cutoff of **2.5**, respectively. None of the three RoB models met the requirement of the hypothesis.

The aiSR’s overall presents a set of six-point primary and six-point secondary recommendations for improving the AI design for ARDS were (i) the inclusion of comorbidity in AI design, (ii) increase in data size, (meets the performance of all the standards in ARDS COVID-19 lung infected patients), (iii) scientific validation using cross-modality, (iv) conducting the clinical validations, (v) improved inter-and intra-observer variability studies, (vi) risk granularity for better drug prescription. The secondary set of recommendations include (i) solid model design (training and prediction), (ii) reproducibility of the proposed model, (iii) process of peer-review, (iv) high-quality documentation, (v) multiethnic and multi-regional data collection, and (vi) multiple ground truth events for clinical validation.

REFERENCES

- [1] WHO COVID-19 Dashboard, Geneva: World Health Organization, 2020. [Online]. Available: <https://covid19.who.int/>
- [2] L. Saba *et al.*, “Molecular pathways triggered by COVID-19 in different organs: ACE2 receptor-expressing cells under attack? A review,” *Eur. Rev. Med. Pharmacological Sci.*, vol. 24, no. 23, pp. 12609–12622, 2020.
- [3] R. Cau, P. P. Bassareo, L. Mannelli, J. S. Suri, and L. Saba, “Imaging in COVID-19-related myocardial injury,” *Int. J. Cardiovasc. Imag.*, vol. 37, no. 4, pp. 1349–1360, Apr. 2021, doi: [10.1007/s10554-020-02089-9](https://doi.org/10.1007/s10554-020-02089-9).
- [4] J. S. Suri *et al.*, “Integration of cardiovascular risk assessment with COVID-19 using artificial intelligence,” *Rev. Cardiovasc. Med.*, vol. 21, no. 4, pp. 541–560, 2020.
- [5] J. S. Suri *et al.*, “COVID-19 pathways for brain and heart injury in comorbidity patients: A role of medical imaging and artificial intelligence-based COVID severity classification: A review,” *Comput. Biol. Med.*, vol. 124, Sep. 2020, Art. no. 103960, doi: [10.1016/j.combiomed.2020.103960](https://doi.org/10.1016/j.combiomed.2020.103960).
- [6] Y. Sandoval, J. L. Januzzi, and A. S. Jaffe, “Cardiac troponin for the diagnosis and risk-stratification of myocardial injury in COVID-19: JACC review topic of the week,” *J. Amer. College Cardiol.*, vol. 76, 2020.
- [7] A. R. Daughton and M. J. Paul, “Identifying protective health behaviors on Twitter: Observational study of travel advisories and Zika virus,” *J. Med. Internet Res.*, vol. 21, no. 5, May 2019, Art. no. e13090, doi: [10.2196/13090](https://doi.org/10.2196/13090).
- [8] A. Signorini, A. M. Segre, and P. M. Polgreen, “The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic,” *PLoS One*, vol. 6, no. 5, 2011, Art. no. e19467.
- [9] N. Mahroum *et al.*, “Public reaction to Chikungunya outbreaks in Italy—Insights from an extensive novel data streams-based structural equation modeling analysis,” *PLoS One*, vol. 13, no. 5, 2018, Art. no. e0197337.
- [10] J. Li, Q. Xu, R. Cuomo, V. Purushothaman, and T. Mackey, “Data mining and content analysis of the Chinese social media platform Weibo during the early COVID-19 outbreak: Retrospective observational infodiversion study,” *JMIR Public Health Surveill.*, vol. 6, no. 2, Apr. 2020, Art. no. e18700, doi: [10.2196/18700](https://doi.org/10.2196/18700).
- [11] O. S. Albahri *et al.*, “Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects,” *J. Infection Public Health*, vol. 13, no. 10, pp. 1381–1396, Oct. 2020, doi: [10.1016/j.jiph.2020.06.028](https://doi.org/10.1016/j.jiph.2020.06.028).
- [12] C. Bao, X. Liu, H. Zhang, Y. Li, and J. Liu, “Coronavirus disease 2019 (COVID-19) CT findings: A systematic review and meta-analysis,” *J. Amer. College Radiol.*, vol. 17, no. 6, pp. 701–709, Jun. 2020, doi: [10.1016/j.jacr.2020.03.006](https://doi.org/10.1016/j.jacr.2020.03.006).
- [13] M. Nagendran *et al.*, “Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies,” *BMJ*, vol. 368, Mar. 2020, Art. no. m689, doi: [10.1136/bmj.m689](https://doi.org/10.1136/bmj.m689).
- [14] M. Roberts *et al.*, “Machine learning for COVID-19 detection and prognostication using chest radiographs and CT scans: A systematic methodological review,” 2020, *arXiv:06388s*.
- [15] L. Wynants *et al.*, “Prediction models for diagnosis and prognosis of COVID-19 infection: Systematic review and critical appraisal,” *BMJ*, vol. 369, Apr. 2020, Art. no. m1328, doi: [10.1136/bmj.m1328](https://doi.org/10.1136/bmj.m1328).
- [16] M. Montazeri, R. ZahediNasab, A. Farahani, H. Mohseni, and F. J. J. M. I. Ghasemian, “Machine learning models for image-based diagnosis and prognosis of COVID-19: Systematic review,” *JMIR Med. Informat.*, vol. 9, no. 4, 2021, Art. no. e25181.
- [17] L. Saba *et al.*, “Molecular pathways triggered by COVID19 in different organs: ACE2 receptor-expressing cells under attack? A review,” *Eur. Rev. Med. Pharmacol. Sci.*, vol. 24, pp. 12609–12622, 2020.
- [18] Meghan Hupe, “EndNote X9,” *J. Electron. Resources in Medical Libraries*, vol. 16, no. 3/4, pp. 117–119, 2019, doi: [10.1080/15424065.2019.1691963](https://doi.org/10.1080/15424065.2019.1691963).
- [19] N. N. Khanna *et al.*, “Effect of carotid image-based phenotypes on cardiovascular risk calculator: AECRS1.0,” *Med. Biol. Eng. Comput.*, vol. 57, no. 7, pp. 1553–1566, Jul. 2019, doi: [10.1007/s11517-019-01975-2](https://doi.org/10.1007/s11517-019-01975-2).
- [20] A. D. Jamthikar *et al.*, “Artificial intelligence framework for predictive cardiovascular and stroke risk assessment models: A narrative review of integrated approaches using carotid ultrasound,” *Comput. Biol. Med.*, vol. 126, Nov. 2020, Art. no. 104043, doi: [10.1016/j.combiomed.2020.104043](https://doi.org/10.1016/j.combiomed.2020.104043).
- [21] C. Pan *et al.*, “Lung recruitability in COVID-19-associated acute respiratory distress syndrome: A single-center observational study,” *Amer. J. Respir. Crit. Care Med.*, vol. 201, no. 10, pp. 1294–1297, 2020.
- [22] C. Huang *et al.*, “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China,” *Lancet*, vol. 395, no. 10223, pp. 497–506, Feb. 2020, doi: [10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [23] T. Singhai, “A review of the coronavirus disease-2019,” *Indian J. Pediatr.*, vol. 87, pp. 281–286, 2020.
- [24] N. Chen *et al.*, “Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study,” *Lancet*, vol. 395, no. 10223, pp. 507–513, Feb. 2020, doi: [10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7).
- [25] D. Wang *et al.*, “Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China,” *JAMA*, vol. 323, no. 11, pp. 1061–1069, 2020.
- [26] F. Zhou *et al.*, “Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study,” *Lancet*, vol. 395, no. 10229, pp. 1054–1062, Mar. 2020, doi: [10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3).
- [27] O. Dyer, “COVID-19: Airborne transmission is being underestimated, warn experts,” *BMJ*, vol. 370, 2020, Art. no. m2720.
- [28] Y. Zheng, Y. Ma, J. Zhang, and X. Xie, “COVID-19 and the cardiovascular system,” *Nat Rev Cardiol.*, vol. 17, no. 5, pp. 259–260, 2020, doi: [10.1038/s41569-020-0360-5](https://doi.org/10.1038/s41569-020-0360-5).
- [29] Z. Qian *et al.*, “Innate immune response of human alveolar type ii cells infected with severe acute respiratory syndrome–coronavirus,” *Amer. J. Respir. Cell Mol. Biol.*, vol. 48, no. 6, pp. 742–748, 2013.
- [30] Y. Fu, Y. Cheng, and Y. Wu, “Understanding SARS-CoV-2-mediated inflammatory responses: From mechanisms to potential therapeutic tools,” *Virologica Sinica*, vol. 35, no. 3, pp. 266–271, Jun. 2020, doi: [10.1007/s12250-020-00207-4](https://doi.org/10.1007/s12250-020-00207-4).
- [31] Y. Ding *et al.*, “The clinical pathology of severe acute respiratory syndrome (SARS): A report from China,” *J. Pathol.*, vol. 200, no. 3, pp. 282–289, Jul. 2003, doi: [10.1002/path.1440](https://doi.org/10.1002/path.1440).
- [32] J. Liu *et al.*, “Overlapping and discrete aspects of the pathology and pathogenesis of the emerging human pathogenic coronaviruses SARS-CoV, MERS-CoV, and 2019-nCoV,” *J. Med. Virol.*, vol. 92, no. 5, pp. 491–494, May 2020, doi: [10.1002/jmv.25709](https://doi.org/10.1002/jmv.25709).
- [33] G. Schonrich and M. J. Raftery, “Neutrophil extracellular traps go viral,” *Front. Immunol.*, vol. 7, p. 366, Sep. 2016, doi: [10.3389/fimmu.2016.00366](https://doi.org/10.3389/fimmu.2016.00366).
- [34] K. W. Chen *et al.*, “Noncanonical inflammasome signaling elicits gasdermin D-dependent neutrophil extracellular traps,” *Sci. Immunol.*, vol. 3, no. 26, 2018, Art. no. eaar6676.
- [35] V. Papayannopoulos, “Neutrophil extracellular traps in immunity and disease,” *Nat. Rev. Immunol.*, vol. 18, no. 2, pp. 134–147, Feb. 2018, doi: [10.1038/nri.2017.105](https://doi.org/10.1038/nri.2017.105).
- [36] L. Saba *et al.*, “Can COVID-19 trigger the plaque vulnerability—a Kounis syndrome warning for ‘asymptomatic subjects’,” *Cardiovasc. Diagnosis Ther.*, vol. 10, no. 5, pp. 1352–1355, Oct. 2020, doi: [10.21037/cdt-20-561](https://doi.org/10.21037/cdt-20-561).

- [37] M. A. Matthay, L. B. Ware, and G. A. Zimmerman, "The acute respiratory distress syndrome," *J. Clin. Invest.*, vol. 122, no. 8, pp. 2731–2740, Aug. 2012, doi: [10.1172/JCI160331](https://doi.org/10.1172/JCI160331).
- [38] A. Fein *et al.*, "The value of edema fluid protein measurement in patients with pulmonary edema," *Amer. J. Med.*, vol. 67, no. 1, pp. 32–38, Jul. 1979, doi: [10.1016/0002-9343\(79\)90066-4](https://doi.org/10.1016/0002-9343(79)90066-4).
- [39] L. Gattinoni *et al.*, *COVID-19 Pneumonia: Different Respiratory Treatments for Different Phenotypes?* Berlin, Germany: Springer, 2020.
- [40] P. Cardinal-Fernandez, J. A. Lorente, A. Ballen-Barragan, and G. Matute-Bello, "Acute respiratory distress syndrome and diffuse alveolar damage. New insights on a complex relationship," *Ann. Amer. Thoracic Soc.*, vol. 14, no. 6, pp. 844–850, Jun. 2017, doi: [10.1513/AnnalsATS.201609-728PS](https://doi.org/10.1513/AnnalsATS.201609-728PS).
- [41] A. L. Katzenstein, C. M. Bloor, and A. A. Leibow, "Diffuse alveolar damage—the role of oxygen, shock, and related factors. A review," *Amer. J. Pathol.*, vol. 85, no. 1, pp. 209–228, Oct. 1976. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/788524>
- [42] T. J. Nuckton *et al.*, "Pulmonary dead-space fraction as a risk factor for death in the acute respiratory distress syndrome," *New Engl. J. Med.*, vol. 346, no. 17, pp. 1281–1286, Apr. 2002, doi: [10.1056/NEJMoa012835](https://doi.org/10.1056/NEJMoa012835).
- [43] A. Altan and S. Karasu, "Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique," *Chaos Solitons Fractals*, vol. 140, Nov. 2020, Art. no. 110071, doi: [10.1016/j.chaos.2020.110071](https://doi.org/10.1016/j.chaos.2020.110071).
- [44] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays," *Comput. Meth. Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105608, doi: [10.1016/j.cmpb.2020.105608](https://doi.org/10.1016/j.cmpb.2020.105608).
- [45] L. Carrer *et al.*, "Automatic pleural line extraction and COVID-19 scoring from lung ultrasound data," *IEEE Trans. Ultrasonics, Ferroelect., Freq. Control*, vol. 67, no. 11, pp. 2207–2217, Nov. 2020.
- [46] D. Das, K. C. Santosh, and U. Pal, "Truncated inception net: COVID-19 outbreak screening using chest X-rays," *Phys. Eng. Sci. Med.*, vol. 43, no. 3, pp. 915–925, Sep. 2020, doi: [10.1007/s13246-020-00888-x](https://doi.org/10.1007/s13246-020-00888-x).
- [47] E. J. Hwang, H. Kim, S. H. Yoon, J. M. Goo, and C. M. Park, "Implementation of a deep learning-based computer-aided detection system for the interpretation of chest radiographs in patients suspected for COVID-19," *Korean J. Radiol.*, vol. 21, no. 10, pp. 1150–1160, Oct. 2020, doi: [10.3348/kjr.2020.0536](https://doi.org/10.3348/kjr.2020.0536).
- [48] A. M. Ismael and A. Sengur, "Deep learning approaches for COVID-19 detection based on chest X-ray images," *Expert Syst. Appl.*, vol. 164, Feb. 2021, Art. no. 114054, doi: [10.1016/j.eswa.2020.114054](https://doi.org/10.1016/j.eswa.2020.114054).
- [49] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images," *Comput. Meth. Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105581, doi: [10.1016/j.cmpb.2020.105581](https://doi.org/10.1016/j.cmpb.2020.105581).
- [50] M. D. Li *et al.*, "Improvement and multi-population generalizability of a deep learning-based chest radiograph severity score for COVID-19," 2020, doi: [10.1101/2020.09.15.20195453](https://doi.org/10.1101/2020.09.15.20195453).
- [51] Z. Li *et al.*, "From community-acquired pneumonia to COVID-19: A deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans," *Eur. Radiol.*, vol. 30, no. 12, pp. 6828–6837, Dec. 2020, doi: [10.1007/s00330-020-07042-x](https://doi.org/10.1007/s00330-020-07042-x).
- [52] T. Mahmud, M. A. Rahman, and S. A. Fattah, "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization," *Comput. Biol. Med.*, vol. 122, Jul. 2020, Art. no. 103869, doi: [10.1016/j.combiomed.2020.103869](https://doi.org/10.1016/j.combiomed.2020.103869).
- [53] V. Mergen *et al.*, "Deep learning for automatic quantification of lung abnormalities in COVID-19 patients: First experience and correlation with clinical parameters," *Eur. J. Radiol. Open*, vol. 7, 2020, Art. no. 100272, doi: [10.1016/j.ejro.2020.100272](https://doi.org/10.1016/j.ejro.2020.100272).
- [54] N. N. Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays," *Ing. Rech. Biomed.*, Jul. 2020, doi: [10.1016/j.irbm.2020.07.001](https://doi.org/10.1016/j.irbm.2020.07.001).
- [55] Q. Ni *et al.*, "A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images," *Eur. Radiol.*, vol. 30, no. 12, pp. 6517–6527, Dec. 2020, doi: [10.1007/s00330-020-07044-9](https://doi.org/10.1007/s00330-020-07044-9).
- [56] N. M. Noor *et al.*, "Automatic lung segmentation using control feedback system: Morphology and texture paradigm," *J. Med. Syst.*, vol. 39, no. 3, Mar. 2015, Art. no. 22, doi: [10.1007/s10916-015-0214-6](https://doi.org/10.1007/s10916-015-0214-6).
- [57] Y. Oh, S. Park, and J. C. Ye, "Deep learning COVID-19 features on CXR using limited training data sets," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2688–2700, Aug. 2020.
- [58] C. Ouchicha, O. Ammor, and M. Meknassi, "CVDNet: A novel deep learning architecture for detection of coronavirus (COVID-19) from chest X-ray images," *Chaos Solitons Fractals*, vol. 140, Nov. 2020, Art. no. 110245, doi: [10.1016/j.chaos.2020.110245](https://doi.org/10.1016/j.chaos.2020.110245).
- [59] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, Jun. 2020, Art. no. 103792, doi: [10.1016/j.combiomed.2020.103792](https://doi.org/10.1016/j.combiomed.2020.103792).
- [60] S. Roy *et al.*, "Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2676–2687, Aug. 2020.
- [61] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 39, no. 7, pp. 1379–1389, Jul. 2020, doi: [10.1007/s10096-020-03901-z](https://doi.org/10.1007/s10096-020-03901-z).
- [62] T. Tuncer, S. Dogan, and F. Ozyurt, "An automated residual exemplar local binary pattern and iterative ReliefF based COVID-19 detection method using chest X-ray image," *Chemometr. Intell. Lab. Syst.*, vol. 203, Aug. 2020, Art. no. 104054, doi: [10.1016/j.chemolab.2020.104054](https://doi.org/10.1016/j.chemolab.2020.104054).
- [63] S. Wang *et al.*, "A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis," *Eur. Respir. J.*, vol. 56, no. 2, Aug. 2020, Art. no. 2000775, doi: [10.1183/13993003.00775-2020](https://doi.org/10.1183/13993003.00775-2020).
- [64] X. Wu *et al.*, "Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study," *Eur. J. Radiol.*, vol. 128, Jul. 2020, Art. no. 109041, doi: [10.1016/j.ejrad.2020.109041](https://doi.org/10.1016/j.ejrad.2020.109041).
- [65] S. Yang *et al.*, "Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: A pilot study," *Ann. Transl. Med.*, vol. 8, no. 7, p. 450, Apr. 2020, doi: [10.21037/atm.2020.03.132](https://doi.org/10.21037/atm.2020.03.132).
- [66] H. Yasar and M. Ceylan, "A novel comparative study for detection of COVID-19 on CT lung images using texture analysis, machine learning, and deep learning methods," *Multimedia Tools Appl.*, vol. 80, no. 4, pp. 5423–5447, Oct. 2020, doi: [10.1007/s11042-020-09894-3](https://doi.org/10.1007/s11042-020-09894-3).
- [67] S. H. Yoo *et al.*, "Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging," *Front. Med. (Lausanne)*, vol. 7, 2020, Art. no. 427, doi: [10.3389/fmed.2020.00427](https://doi.org/10.3389/fmed.2020.00427).
- [68] J. Zhu, B. Shen, A. Abbasi, M. Hoshmand-Kochi, H. Li, and T. Q. Duong, "Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs," *PLoS One*, vol. 15, no. 7, 2020, Art. no. e0236621, doi: [10.1371/journal.pone.0236621](https://doi.org/10.1371/journal.pone.0236621).
- [69] F. Shan *et al.*, "Lung infection quantification of COVID-19 in CT images with deep learning," 2020, *arXiv:04655*.
- [70] S. Chaganti *et al.*, "Automated quantification of CT patterns associated with COVID-19 from chest CT," *Radiol. Artif. Intell.*, vol. 2, no. 4, Jul. 2020, Art. no. e200048, doi: [10.1148/ryai.2020200048](https://doi.org/10.1148/ryai.2020200048).
- [71] A. Signoroni *et al.*, "End-to-end learning for semiquantitative rating of COVID-19 severity on chest X-rays," 2020, *arXiv:04603*.
- [72] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio, and S. K. Antani, "Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays," *IEEE Access*, vol. 8, pp. 115041–115050, 2020.
- [73] Z. Tang *et al.*, "Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images," 2020, *arXiv:11988*.
- [74] J. Born *et al.*, "POCOVID-Net: Automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS)," Apr. 2020, *arXiv:2004.12084*.
- [75] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, pp. 1–8, Jul. 2020, doi: [10.1080/07391102.2020.1788642](https://doi.org/10.1080/07391102.2020.1788642).
- [76] H. S. Maghdid, A. T. Asaad, K. Z. Ghaffoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms," *Multimodal Image Exploit. Learn. 2021, Int. Soc. Optics Photonics*, 2021, Art. no. 117340E.
- [77] N. Tsiknakis *et al.*, "Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays," *Exp. Ther. Med.*, vol. 20, no. 2, pp. 727–735, Aug. 2020, doi: [10.3892/etm.2020.8797](https://doi.org/10.3892/etm.2020.8797).

- [78] J. P. Cohen *et al.*, "Predicting COVID-19 pneumonia severity on chest X-ray with deep learning," *Cureus*, vol. 12, no. 7, Jul. 2020, Art. no. e9448, doi: [10.7759/cureus.9448](https://doi.org/10.7759/cureus.9448).
- [79] H. Gunraj, L. Wang, and A. Wong, "COVIDNet-CT: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest CT images," *Front. Med. (Lausanne)*, vol. 7, 2020, Art. no. 608525, doi: [10.3389/fmed.2020.608525](https://doi.org/10.3389/fmed.2020.608525).
- [80] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, no. 1, Nov. 11, 2020, Art. no. 19549, doi: [10.1038/s41598-020-76550-z](https://doi.org/10.1038/s41598-020-76550-z).
- [81] C. Li *et al.*, "Classification of severe and critical COVID-19 using deep learning and radiomics," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 12, pp. 3585–3594, 2020.
- [82] L. Sun *et al.*, "Adaptive feature selection guided deep forest for COVID-19 classification with chest CT," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 10, pp. 2798–2805, Oct. 2020.
- [83] L. Li *et al.*, "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.
- [84] K. Zhang *et al.*, "Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography," *Cell*, vol. 181, no. 6, pp. 1423–1433, 2020.
- [85] L. Saba *et al.*, "The present and future of deep learning in radiology," *Eur. J. Radiol.*, vol. 114, pp. 14–24, May 2019, doi: [10.1016/j.ejrad.2019.02.038](https://doi.org/10.1016/j.ejrad.2019.02.038).
- [86] A. El-Baz and J. S. Suri, *Lung Imaging and Computer Aided Diagnosis*, Boca Raton, FL, USA: CRC Press, 2011.
- [87] L. Saba *et al.*, "Six artificial intelligence paradigms for tissue characterization and classification of non-COVID-19 pneumonia against COVID-19 pneumonia in computed tomography lungs," *J. Med. Syst.*, vol. 16, no. 3, pp. 423–434, 2021.
- [88] R. M. Rangayyan and J. S. Suri, *Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer*, Bellingham, Washington, DC, USA: SPIE Publications, 2006.
- [89] A. El-Baz and J. Suri, *Lung Imaging and CADx*. Boca Raton, FL, USA: CRC Press, 2019.
- [90] L. Saba *et al.*, "Radiation dose and image quality of computed tomography of the supra-aortic arteries: A comparison between single-source and dual-source CT scanners," *J. Neuroradiol.*, vol. 45, no. 2, pp. 136–141, Mar. 2018, doi: [10.1016/j.neurad.2017.09.007](https://doi.org/10.1016/j.neurad.2017.09.007).
- [91] W. J. Veldkamp, L. J. Kroft, and J. Geleijns, "Dose and perceived image quality in chest radiography," *Eur. J. Radiol.*, vol. 72, no. 2, pp. 209–217, Nov. 2009, doi: [10.1016/j.ejrad.2009.05.039](https://doi.org/10.1016/j.ejrad.2009.05.039).
- [92] J. S. Suri, T. Danielson, Y. Guo, and R. Janer, "Fischer's fused full field digital mammography and ultrasound system (FFDMUS)," *Stud. Health Technol. Informat.*, vol. 114, pp. 177–200, 2005. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/15923774>
- [93] Y. Guo, R. Sivaramakrishna, C. C. Lu, J. S. Suri, and S. Laxminarayan, "Breast image registration techniques: A survey," *Med. Biol. Eng. Comput.*, vol. 44, no. 1/2, pp. 15–26, Mar. 2006, doi: [10.1007/s11517-005-0016-y](https://doi.org/10.1007/s11517-005-0016-y).
- [94] J. Suru, D. Wilson, and S. Laxminarayan, *Handbook of Biomedical Image Analysis*, Berlin, Germany: Springer, vol. 3, 2005.
- [95] M. Agarwal *et al.*, "Wilson disease tissue classification and characterization using seven artificial intelligence models embedded with 3D optimization paradigm on a weak training brain magnetic resonance imaging datasets: A supercomputer application," *Med. Biol. Eng. Comput.*, vol. 59, no. 3, pp. 511–533, 2021.
- [96] L. Saba *et al.*, "Ultrasound-based internal carotid artery plaque characterization using deep learning paradigm on a supercomputer: A cardiovascular disease/stroke risk assessment system," *Int. J. Cardiovasc. Imag.*, vol. 37, pp. 1511–1528, 2021.
- [97] S. S. Skandha *et al.*, "3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: Atheromatic 2.0," *Comput. Biol. Med.*, vol. 125, 2020, Art. no. 103958.
- [98] J. A. Sterne *et al.*, "ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions," *BMJ*, vol. 355, Oct. 2016, Art. no. i4919, doi: [10.1136/bmj.i4919](https://doi.org/10.1136/bmj.i4919).
- [99] R. F. Wolff *et al.*, "PROBAST: A tool to assess the risk of bias and applicability of prediction model studies," *Ann. Intern. Med.*, vol. 170, no. 1, pp. 51–58, Jan. 2019, doi: [10.7326/M18-1376](https://doi.org/10.7326/M18-1376).
- [100] S. S. Skandha *et al.*, "3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: Atheromatic 2.0," *Comput. Biol. Med.*, vol. 125, Oct. 2020, Art. no. 103958, doi: [10.1016/j.combiomed.2020.103958](https://doi.org/10.1016/j.combiomed.2020.103958).
- [101] V. Viswanathan *et al.*, "Bidirectional link between diabetes mellitus and COVID-19 leading to cardiovascular disease: A narrative review," *World J. Diabetes*, vol. 12, no. 3, pp. 215–237, 2021.
- [102] J. S. Suri *et al.*, "A narrative review on characterization of acute respiratory distress syndrome in COVID-19-infected lungs using artificial intelligence," *Comput. Biol. Med.*, vol. 130, Mar. 2021, Art. no. 104210, doi: [10.1016/j.combiomed.2021.104210](https://doi.org/10.1016/j.combiomed.2021.104210).
- [103] L. Saba *et al.*, "Intra-and inter-operator reproducibility analysis of automated cloud-based carotid intima media thickness ultrasound measurement," *J. Clin. Diagn. Res.*, vol. 12, no. 2, pp. 649–664, 2018.
- [104] L. Saba *et al.*, "Inter-observer variability analysis of automatic lung delineation in normal and disease patients," *J. Med. Syst.*, vol. 40, no. 6, Jun. 2016, Art. no. 142, doi: [10.1007/s10916-016-0504-7](https://doi.org/10.1007/s10916-016-0504-7).
- [105] A. Jamthikar *et al.*, "Cardiovascular/stroke risk predictive calculators: A comparison between statistical and machine learning models," *Cardiovasc. Diagnosis Ther.*, vol. 10, no. 4, pp. 919–938, Aug. 2020, doi: [10.21037/cdt.2020.01.07](https://doi.org/10.21037/cdt.2020.01.07).
- [106] E. J. Mascha and T. R. Vetter, "Significance, errors, power, and sample size: The blocking and tackling of statistics," *Anesth. Analg.*, vol. 126, no. 2, pp. 691–698, Feb. 2018, doi: [10.1213/ANE.0000000000002741](https://doi.org/10.1213/ANE.0000000000002741).
- [107] S. M. Smith, "Determining sample size," vol. 23, Feb. 2013, Art. no. 2017.
- [108] F. Molinari *et al.*, "Completely automated multiresolution edge snapper—A new technique for an accurate carotid ultrasound IMT measurement: Clinical validation and benchmarking on a multi-institutional database," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1211–1222, Mar. 2011. [Online]. Available: <https://ieeexplore.ieee.org/document/6026248/>
- [109] N. Ikeda *et al.*, "Automated and accurate carotid bulb detection, its verification and validation in low quality frozen frames and motion video," *Int. Angiol.*, vol. 33, no. 6, pp. 573–589, Dec. 2014. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/24658129>
- [110] A. D. Jamthikar *et al.*, "Multiclass machine learning vs. conventional calculators for stroke/CVD risk assessment using carotid plaque predictors with coronary angiography scores as gold standard: A 500 participants study," *Int. J. Cardiovasc. Imag.*, vol. 37, no. 4, pp. 1171–1187, Apr. 2021, doi: [10.1007/s10554-020-02099-7](https://doi.org/10.1007/s10554-020-02099-7).
- [111] G. S. Tandel, A. Balestrieri, T. Jujaray, N. N. Khanna, L. Saba, and J. S. Suri, "Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm," *Comput. Biol. Med.*, vol. 122, Jul. 2020, Art. no. 103804, doi: [10.1016/j.combiomed.2020.103804](https://doi.org/10.1016/j.combiomed.2020.103804).
- [112] H. Amini *et al.*, "A case of COVID-19 lung infection first detected by [18F]FDG PET-CT," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 47, no. 7, pp. 1771–1772, Jul. 2020, doi: [10.1007/s00259-020-04821-y](https://doi.org/10.1007/s00259-020-04821-y).
- [113] A. El-Baz, G. Gimel'farb, and J. S. Suri, *Stochastic Modeling for Medical Image Analysis*, Boca Raton, FL, USA: CRC Press, 2015.
- [114] E. Moss and J. Metcalf, "High tech, high risk: Tech ethics lessons for the COVID-19 pandemic response," *Patterns*, vol. 1, no. 7, 2020, Art. no. 100102.
- [115] M. Maniruzzaman *et al.*, "Risk factors of neonatal mortality and child mortality in Bangladesh," *J. Glob. Health*, vol. 8, no. 1, 2018, Art. no. 010417.
- [116] M. Agarwal *et al.*, "A novel block imaging technique using nine artificial intelligence models for COVID-19 disease classification, characterization and severity measurement in lung computed tomography scans on an Italian cohort," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 45, 2021, Art. no. 28.