

Title: Curbing the AI-induced enthusiasm in diagnosing COVID-19 on chest X-Rays: the present and the near-future

Authors:

Alexandru Burlacu, Radu Crisan-Dabija, Iolanda Valentina Popa, Bogdan Artene, Vasile Birzu, Mihai Pricop, Cristina Plesoianu, Daniele Generali

Emails and Affiliations:

Alexandru Burlacu: MD, PhD, Head of Department of Interventional Cardiology - Cardiovascular Diseases Institute, and “Grigore T. Popa” University of Medicine, Iasi, Romania; alexandru.burlacu@umfiasi.ro ; alburlacu@yahoo.com ; **ORCID:** <https://orcid.org/0000-0002-3424-1588>

Radu Crisan-Dabija: MD, PhD, University of Medicine and Pharmacy “Grigore T. Popa” Iași, Pulmonology Department, Head of Clinic of Pulmonary Diseases Iasi; radu.dabija@umfiasi.ro

Iolanda Valentina Popa: MD, PhD, Institute of Gastroenterology and Hepatology, Iași, Romania, and 'Grigore T. Popa' University of Medicine, Iasi, Romania; iolivp@gmail.com

Bogdan Artene: MD, Department of Interventional Cardiology - Cardiovascular Diseases Institute, Iasi, Romania; bogdan.artene@yahoo.com

Vasile Birzu: Falcon Trading, Iasi, Romania; vasile@falcontrading.ro

Mihai Pricop: Falcon Trading, Iasi, Romania; mihai@falcontrading.ro

Cristina Plesoianu: PhD, Medical Sciences Academy, Romania; cristina.plesoianu@angionet.ro

Daniele Generali: MD, DPhil, Azienda Istituti Ospitalieri di Cremona, Italy, and Dipartimento Universitario Clinico di Scienze Mediche, Chirurgiche e della Salute Università degli Studi di Trieste, Italy; dgenerali@units.it

Correspondence:

Alexandru Burlacu MD, PhD, Head of Department of Interventional Cardiology - Cardiovascular Diseases Institute, Bd. Carol 1 No 50, Iasi, 700503, Iasi Romania. Phone: + 40744488580 email: alexandru.burlacu@umfiasi.ro ; alburlacu@yahoo.com ; **ORCID:** <https://orcid.org/0000-0002-3424-1588>

Keywords: COVID-19; SARS-CoV-2; artificial intelligence; chest X-ray; diagnosis; prediction; Machine learning.

Abstract (289 words)

In the current context of COVID-19 pandemic, a rapid and accessible screening tool based on image processing of chest X-rays (CXRs) using machine learning (ML) approaches would be much needed. Initially, we intended to create and validate an ML software solution able to discriminate on the basis of the CXR between SARS-CoV-2-induced bronchopneumonia and other bronchopneumonia etiologies.

A systematic search of PubMed, Scopus and arXiv databases using the following search terms [“artificial intelligence” OR “deep learning” OR “neural networks”], AND [“COVID-19” OR “SARS-CoV-2”] AND [“chest X-ray” OR “CXR” OR “X-ray”] found 14 recent studies. Most of them declared to be able to confidently identify COVID-19 based on CXRs using deep neural networks. Firstly, weaknesses of artificial intelligence (AI) solutions were analyzed, tackling the issues with datasets (from both medical and technical points of view) and the vulnerability of used algorithms. Then, arguments were provided for why our study design is stronger and more realistic than the previously quoted papers, balancing the possible false expectations with facts.

The authors consider that the potential of AI use in COVID-19 diagnosis on CXR is real. However, scientific community should be careful in interpreting statements, results and conclusions regarding AI use in imaging. It is therefore necessary to adopt standards for research and publication of data, because it seems that in the recent months scientific reality suffered manipulations and distortions. Also, a call for responsible approaches to the imaging methods in COVID-19 is raised. It seems mandatory to follow some rigorous approaches in order to provide with adequate results in daily routine. In addition, the authors intended to raise public awareness about the quality of AI protocols and algorithms and to encourage public sharing of as many CXR images with common quality standards.

Wordcount (2213 words)

I. BACKGROUND

Due to its accessibility and availability, chest X-ray (CXR) imaging is well established and widely used as a screening and diagnostic tool. In the current context of COVID-19 pandemic, there is a growing demand for imaging and diagnostic services. As the workload is rapidly increasing, it can get overwhelming for radiologists to meet the reporting requirements. As the manual analysis of CXR can be time-consuming, (combined with a growing COVID-19 incidence), a machine learning (ML) approach for the screening process would be much needed.

ML has seen some major breakthroughs over the last few years, drawing a lot of attention from the healthcare system. The recent interest of medical image processing domain in exploring deep learning (DL) has led to development of numerous Artificial Intelligence (AI) diagnostic solutions, part of them (at least 15 AI platforms involved in medicine, including 11 in the imaging field) already being approved by the US Food and Drug Administration board in 2018 (1).

Recent studies suggest neural networks (NN) have surpassed even human performance in some visual and auditory recognition tasks, which may portend its applications in medicine and healthcare, especially in medical imaging (2). Despite this huge enthusiasm, there are some uncertainties that should curb this upsurge.

The “*black-box*” issue, firstly spotlighted by Dean Pomerleau in 1991, has become exponentially harder and more urgent (3). It draws attention to the ML methods to find those worthwhile signals in the vast amount of data available, but also to the validity of the results and the trust in DL (3). Despite remarkable performance of AI, the “*black-box*” problem might be worse than anticipated, as DL models can be surprisingly easy to mislead, and also tend to make overconfident predictions (3, 4).

II. AIMS AND METHODS

Initially, we aimed to create and validate a software solution able to discriminate on the basis of the CXR between bronchopneumonia lesions induced by SARS-CoV-2 and other different bronchopneumonia lesions (e.g. influenza), using AI/ML models for scanning chest radiographs in the emergency department.

At the end of March 2020, we published and registered our protocol with ClinicalTrials.gov Identifier: NCT04313946. We envisioned a multi-centric study (*“Artificial Intelligence Algorithms for Discriminating Between COVID-19 and Influenza Pneumonitis Using Chest X-Rays - AI-COVID-Xr”*) enrolling patients from Italy, United Kingdom and Romania.

For this purpose, the electronic databases of PubMed, Scopus and arXiv were systematically searched for relevant articles from the inception until late March 2020. The search terms used were [“artificial intelligence” OR “deep learning” OR “neural networks”], AND [“COVID-19” OR “SARS-CoV-2”] AND [“chest X-ray” OR “CXR” OR “X-ray”]. After removing the duplicates, screening titles and abstracts, and assessing eligibility of the selected full texts, 14 valid articles were found. Journal articles published with full text or abstracts in English were eligible for inclusion. However, only in the first two weeks of April at least 12 studies were found declaring that deep neural networks techniques can confidently identify COVID-19 based on CXRs (see Table 1, and references (5, 6)).

Therefore, this paper aims to analyze closer the weaknesses of these AI solutions, trying to cast a glance inside the *“black box”*. Moreover, arguments are also provided supporting our study is realistic with evidence-based findings, avoiding any possible false expectations.

III. RESULTS AND DISCUSSIONS

RECENT ATTEMPTS, ALLEGED SOLUTIONS: MUCH ADO FOR ... (ALMOST) NOTHING?

Most approaches for COVID-19 pneumonia detection on CXRs are based on DL and transfer learning (TL) models. DL and TL represent state-of-the-art performance in image processing and computer vision. Both should integrate the feature detection component and the classification component. TL is a process of adapting a pre-trained model for solving a new problem. The feature detection part is left unchanged and the upper layers corresponding to the classification part are optimized.

Testing for biases

Firstly, so far there are no attempts examining whether any of these algorithms, shown in Table 1, claiming to diagnose COVID-19 on chest radiographs *display systematic bias*. In a serious context like the COVID-19 pandemic, when decision-making is vitally important, testing for biases is mandatory. This can be done using different toolboxes (FairML, Lime, Grad-CAM) (7, 8).

Issues with the datasets

a) All approaches used only public data from covid-chestxray-dataset (<https://github.com/ieee8023/covid-chestxray-dataset>) and kaggle competitions datasets related to pneumonia. Data owners warn users to *not claim diagnostic performance without a clinical study*.

b) CXRs from these public datasets *are not evaluated from a medically point of view*. Each radiograph projection must be correctly identified as PA (postero-anterior) and AP (antero-posterior). Although most films are PA (with the patient in an erect position), in the context of a major outbreak, for severe cases, AP projection will be used for patients confined in bed, in a supine position. This differentiation is important because, for example, in a supine film, blood will flow more to lungs' apices than when standing. Omitting to differentiate PA from AP projections may result in a misdiagnosis of pulmonary congestion. Moreover, patients lying in bed are more likely to present obvious opacities such as electrocardiogram electrodes, chest drains, pacemakers or foreign bodies, that can alter program's output.

Another important quality assessment is checking for *signs of patient's rotation* (ideally should be minimal). Severe rotation can make diagnosis difficult by altering the normal cardiomedial contour. Checking for a good X-Rays *penetration* is also a standard part of assessing quality. Penetration is the degree to which X-rays have passed through the body and is affected by both the duration of exposure and the power of the radiation beam. An under-penetrated film looks diffusely light and anatomical structures, like the vertebrae behind the heart, are eclipsed. Over-penetration makes radiographs look diffusely dark and lung markings are poorly delineated.

c) Datasets used *are highly imbalanced*. If the training data contains 75% positive and 25% negative cases, the algorithm will immediately notice that predicting "positive" all the time improves its success rate (4, 5, 9-13). The training process is just the optimization of a loss

function, without any data interpretation. Whenever the algorithm finds a shortcut to maximize the success rate, it will use it immediately. Adjustments can be made in the training process for unbalanced data, but they are no longer effective for very high ratios like the ones presented in Table 1. To date, only 200 COVID-19 CXR images are publicly available worldwide, while there are thousands of images depicting normal or bacterial pneumonia radiographs. After study completion, our team will contribute with hundreds of COVID-19 CXRs from more medical centres: Suceava and Iasi (Romania), Cremona (Italy), London (UK).

d) There is no knowledge about *the X-ray machines* on which chest radiographs were performed. The digital X-ray machine artifacts differ from those that appear on X-ray film. Even though traditional film X-ray machines are more accessible, they have a higher frequency of artifacts resulted from user error, especially in the preparation and development of the images (14).

e) There is *no metadata about patients'* background, travel history, country of origin or other useful knowledge. Data is crucial if it is paired with enough context to create meaning. *Without context data might become noise* (15). From a radiologist point of view, there is a high diagnosis suspicion of COVID-19 if the presence of multiple, bilateral and peripheral patchy ground glass opacities is associated with epidemiological context. Although these typical radiographic characteristics are present only 70% of time (16, 17) the epidemiological context can add more certainty to diagnosis.

f) Data is *not equally distributed* on age groups, race, comorbidities and other patient characteristics. This can generate knowledge gaps and confounders in the data that inherently lead to biased predictions.

g) Artificial Intelligence is *not clever enough without qualitative data to rely on*. Data quality can induce discriminatory properties in classifiers, while unmeasured predictive features generate discrimination (18). A cliché frequently used in ML is “*Garbage in, garbage out*”, since clean and accurate data goes a long way toward determining the quality of the results.

Issues with algorithms

a) Most of the studies on DL-based COVID-19 detection have utilized convolutional NN (CNNs) for image classification. CNNs are powerful image processing techniques but have some downsides in the context of CXR-based pneumonia detection. Different models are either

encoded for image classification (VGG, DenseNet, GoogleNet, ResNet, MobileNet) or lesions localization, but not both. However, an accurate CXR diagnosis method needs to have embedded both the lesions localization and diagnostic classification components.

Most of the models referred in Table 1 are trained using the whole image, without any interpretation or lesion location information, having only the classification label (e.g. positive/negative). Classification models do not encode the position and orientation of objects ignoring spatial relations between them. Consequently, CNNs of this type are not invariant to rotations and transformations of input data and will need very large datasets to capture all the possible transformations in order to achieve very good performance. However, only small datasets are yet available for COVID-19 classification (19).

In the present context of COVID-19 imagistic detection, integrating models like Mask-RCNN (20) that take into account the localization of lesions might prove more useful and accurate. Also, the RSNA dataset contains annotations, bounding box on each of the pneumonia opacities (<https://www.rsna.org/en/education/ai-resources-and-training/ai-image-challenge/RSNA-Pneumonia-Detection-Challenge-2018>).

b) Most of the models are based only on DL and TL. Even though DL algorithms are very powerful they need to be fed with much larger and qualitative datasets for the network to learn correctly the distinctive diagnostic features. For now, DL may not be the solution and the widespread concept that DL is the only ML method worth investigating could be untrue. Other baseline models with similar performance but with more explainability might prove to be a wiser choice in a context of small-scale datasets available (21).

c) Most images in the referred papers were rescaled to a resolution of 224x224 before being used for the DL models training. The information might be present in the shrunken images, but it would be more reliable if the images could be processed at their (higher) initial resolution.

A NOVEL APPROACH OF AI-ALGORITHMS IN COVID-19 RECOGNITION ON X-RAYS

To our knowledge, no clinical study for COVID-19 CXR diagnosis was conducted before. For more accurate results, we propose rigorous criteria to be followed in the context of a prospective clinical trial.

Firstly, care must be taken for performing good quality radiographs and a more rigorous selection of proper images, equally distributed on age groups. A prospective study allows for better compliance monitoring on the quality of this step.

All images must be subjected to the preprocessing stage. Data normalization to the same range of grays, resolution and zoom is required. All images must have the same mean gray value. Sharpening, histogram equalization (HGE) and Perona-Malik diffusion may be performed for edge enhancement, contrast enhancement and noise reduction on entire CXR images (22). HGE can attain almost equally distributed intensities through blending gray-levels with low frequencies into one and expanding frequent intensities over a higher range of gray levels (5).

Segmentation is another essential step in AI-based COVID-19 image processing and analysis. It outlines the CXR regions of interest (lungs, diseased regions and lesions) in order to feed the ML algorithm with localization information and more meaningful data. Many publications consider segmentation as an essential stage in analyzing COVID-19 images. However, works involving segmentation in COVID-19 are only conducted on CT imaging studies (23).

Segmentation of X-ray images is difficult due to the overlapping of the ribs and clavicles with lesions of interest. There are no methods developed, to date, for segmenting X-ray images of COVID-19 patients. Still, there are several models proposed for CXR segmentation in pneumonia (24-26).

Data enrichment is a pivotal step for bringing context into data. Context is used by the algorithms to discriminate the noise from meaningful data. Table 2 shows main history, clinical and biological variables to be employed in the analysis.

For ML models training and testing, we propose a step-up approach from simpler algorithmic methods to more complex, while successively integrating both the localization and the classification modules in a composite framework.

Considering the serious context created by the COVID-19 pandemic and the perilous consequences that can be attained by erroneous models that move from paper to product, the use of debiasing methods in algorithm development should be considered. Debiasing techniques have been proposed for medical imaging ML models like unlearning dataset membership (27). Other methods use adversarial learning to generate a debiased input data (28). Also, a fairness

check study was carried out on SOTA models that are trained on large public datasets to predict diagnosis from CXR images (18).

IV. CONCLUSIONS

The authors consider the potential of AI use in COVID-19 diagnosis on CXR is real. However, scientific community should be careful in interpreting statements, results and conclusions regarding AI use in imaging. It is therefore necessary to adopt standard criteria for research and publication of data, as it seems that in the recent months scientific reality suffered manipulations and distortions. Also, a call for responsible approaches to the imaging methods in COVID-19 is raised. It seems mandatory to follow some rigorous approaches in order to provide with adequate results in daily routine. In addition, the authors intended to raise public awareness about the quality of AI protocols and algorithms and to encourage public sharing of as many CXR images with common quality standards.

Registration

ClinicalTrials.gov Identifier: NCT04313946 (<https://clinicaltrials.gov/ct2/show/NCT04313946>)

Data availability

Data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding Statement

This study was funded by the Romanian Academy of Medical Sciences and European Regional Development Fund, MySMIS 107124: Funding Contract 2/Axa 1/31.07.2017/ 107124 SMIS

Table 1. Reported papers using Artificial Intelligence in COVID-19 chest x-ray diagnosis.

Number	Studies	Algorithm	Image rescale	No. of x-rays	Classification	Performance
1	Hall L.O. et al. (9)	Deep transfer learning (DTL) on pre-trained VGG16; 13 convolutional layers	224x224	122 COVID-19 4273 Viral and bacterial pneumonia	COVID-19/ Other pneumonia	ACC=96.1%
2	Narin A. et al. (29)	DTL using 3 pre-trained deep CNNs (DCNNs)	224x224	50 COVID-19 50 normal CXR	COVID-19/ Normal	ResNet50: ACC=98%
3	Chowdhury M.E.H. et al.(10)	DTL using 4 pre-trained DCNNs	224x224 or 227x227	190 COVID-19 1341 Normal 1345 Viral Pneumonia	COVID-19/ Normal/ Other viral pneumonia	SqueezeNet: 3-class ACC=98.3%
4	Zhang J. et al. (11)	Stochastic gradient descent with pre-trained CNN on ImageNet; 18 convolutional layers	512 x 512	100 COVID-19 1431 Other pneumonia	COVID-19/ non COVID-19	AUC=95.18%
5	Apostolopoulos Bessiana T. (30)	I.D., DTL using 4 pre-trained DCNNs	200 x 266	224 COVID-19 700 Common pneumonia 504 Normal	COVID-19/ Pneumonia/ Normal	3-class ACC=93.48%
6	Afshar P. et al.(19)	Capsule Networks; 4 convolutional layers and 3 Capsule layers	N.A.	76 COVID-19 8066 Normal 5526 Other pneumonia	COVID-19/ non-COVID-19	ACC=95.7%
7	Wang L., Wong A.(12)	Generative synthesis developed DCNN that uses lightweight residual projection-expansion-projection-extension design pattern	N.A.	76 COVID-19 8066 Normal 5526 Other pneumonia	COVID-19/ Normal/ Pneumonia	3-class ACC=92.4%
8	Farooq M., Hafeez A. (13)	DTL on pre-trained ResNet50; cyclical and discriminative learning rates; progressive image resizing	N.A.	68 COVID-19 1203 Normal 931 Bacterial pneumonia 660 Other	COVID-19/ Normal/ Bacterial/ Viral pneumonia	4-class ACC=96,23%

				viral pneumonia			
9	Li X., Zhu D. (31)	DTL on pre-trained DenseNet-121; 1 fully connected layer for extracting general CXR features and 8 fully connected disease-specific layers	N.A.	185 COVID-19 185 Normal 185 Other pneumonia	COVID-19/ Normal/ Pneumonia	3-class ACC=88.9%	
10	Hemdan E.E. et al.(32)	Comparing 7 DCNNs	224x224	25 COVID-19 25 Normal	COVID-19/ Normal	VGG19 and DenseNet201: ACCs=90%	
11	Khalifa N. E. M. et al. (33)	Generative adversarial networks and DTL using 4 DCNNs models	N.A.	624 Normal and pneumonia	Normal/ Pneumonia	Resnet18: ACC=99%	
12	Maghdid H. S. et al. (6)	CNN with 1 convolutional layer Pre-trained AlexNet	N.A.	85 COVID-19 85 Normal	COVID-19/ Normal	CNN ACC=94.1% AlexNet: ACC=98%	
13	Karim M. R. et al. (5)	Neural ensemble method combining several deep architectures	N.A.	190 COVID-19 8614 Other pneumonia 8066 Normal	COVID-19/ Pneumonia/ Normal	PPV=89.61%	
14	Ghoshal B., Tucker A. (4)	Uncertainty estimation and interpretability with drop-weights based Bayesian CNN	224 x 224	68 COVID-19 1583 Normal 2786 Bacterial pneumonia 1506 Viral pneumonia	COVID-19/ Normal/ Bacterial pneumonia/ Viral pneumonia	ACC=89.82%	

Table 2. Data enrichment with context variables from COVID-19 patients.

Category	Variables	Reference
Patient's history	Age	(34)
	Travel history	
	Country and city of origin	
	Comorbidities	
Clinical data	Fever, cough, fatigability	Fever and cough are the most frequent symptoms (35, 36)
	Blood pressure, Respiratory rate, dyspnea, oxygen saturation	(37)
Biological data	Leucocyte, lymphocyte, neutrophils and platelet counts	Lymphopenia is common (35)
	Alanine aminotransferase, aspartate aminotransferase	Liver function tests may be elevated (35, 38)
	D-dimer	Commonly elevated (38)
	CRP, ESR	Elevated and appear to trend upward with disease progression (25, 39)
	LDH, ferritin, troponin, CK, CKMB	(35, 36, 38)

CRP: C reactive protein; ESR: erythrocyte sedimentation rate; LDH: lactate dehydrogenase; CK: creatine kinase; CKMB: creatine kinase-MB.

References

1. Administration USFD. Artificial Intelligence and Machine Learning in Software as a Medical Device 2020 [Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>].
2. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. Korean J Radiol. 2017;18(4):570-84.
3. Castelvechi D. Can we open the black box of AI? Nature. 2016;538(7623):20-3.
4. Ghoshal B, Tucker A. Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection. arXiv preprint arXiv:200310769. 2020.
5. Karim M, Döhmen T, Rebholz-Schuhmann D, Decker S, Cochez M, Beyan O. DeepCOVIDExplainer: Explainable COVID-19 Predictions Based on Chest X-ray Images. arXiv preprint arXiv:200404582. 2020.
6. Maghdid HS, Asaad AT, Ghafoor KZ, Sadiq AS, Khan MK. Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms. arXiv preprint arXiv:200400038. 2020.
7. Adebayo JA. FairML : ToolBox for diagnosing bias in predictive modeling: Massachusetts Institute of Technology; 2016.
8. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, editors. Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision; 2017.
9. Hall LO, Paul R, Goldgof DB, Goldgof GM. Finding COVID-19 from Chest X-rays using Deep Learning on a Small Dataset. arXiv preprint arXiv:200402060. 2020.
10. Chowdhury ME, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, et al. Can AI help in screening Viral and COVID-19 pneumonia? arXiv preprint arXiv:200313145. 2020.
11. Zhang J, Xie Y, Li Y, Shen C, Xia Y. COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection. arXiv preprint arXiv:200312338. 2020.
12. Wang L, Wong A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images. arXiv preprint arXiv:200309871. 2020.
13. Farooq M, Hafeez A. COVID-ResNet: A Deep Learning Framework for Screening of COVID19 from Radiographs. arXiv preprint arXiv:200314395. 2020.
14. Nuth EK, Armbrust LJ, Roush JK, Biller DS. Identification and effects of common errors and artifacts on the perceived quality of radiographs. Journal of the American Veterinary Medical Association. 2014;244(8):961-7.
15. Divvala SK, Hoiem D, Hays JH, Efros AA, Hebert M, editors. An empirical study of context in object detection. 2009 IEEE Conference on Computer Vision and Pattern Recognition; 2009 20-25 June 2009.

16. Bhatraju PK, Ghassemieh BJ, Nichols M, Kim R, Jerome KR, Nalla AK, et al. Covid-19 in Critically Ill Patients in the Seattle Region — Case Series. *New England Journal of Medicine*. 2020.
17. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis*. 2020;20(4):425-34.
18. Seyyed-Kalantari L, Liu G, McDermott M, Ghassemi M. CheXclusion: Fairness gaps in deep chest X-ray classifiers. *arXiv preprint arXiv:200300827*. 2020.
19. Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A. COVID-CAPS: A Capsule Network-based Framework for Identification of COVID-19 cases from X-ray Images. *arXiv preprint arXiv:200402696*. 2020.
20. He K, Gkioxari G, Dollár P, Girshick R, editors. Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV); 2017 22-29 Oct. 2017.
21. Garnelo M, Arulkumaran K, Shanahan M. Towards deep symbolic reinforcement learning. *arXiv preprint arXiv:160905518*. 2016.
22. TAHIR DN. *INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)*.
23. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, et al. Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19. *arXiv preprint arXiv:200402731*. 2020.
24. Gaál G, Maga B, Lukács A. Attention U-Net Based Adversarial Architectures for Chest X-ray Lung Segmentation. *arXiv preprint arXiv:200310304*. 2020.
25. Souza JC, Bandeira Diniz JO, Ferreira JL, França da Silva GL, Corrêa Silva A, de Paiva AC. An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. *Computer Methods and Programs in Biomedicine*. 2019;177:285-96.
26. Saad MN, Muda Z, Ashaari NS, Hamid HA, editors. Image segmentation for lung region in chest X-ray images using edge detection and morphology. 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014); 2014 28-30 Nov. 2014.
27. Ashraf A, Khan S, Bhagwat N, Chakravarty M, Taati B. Learning to unlearn: Building immunity to dataset bias in medical imaging studies. *arXiv preprint arXiv:181201716*. 2018.
28. Zhang BH, Lemoine B, Mitchell M, editors. Mitigating unwanted biases with adversarial learning. *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*; 2018.
29. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *arXiv preprint arXiv:200310849*. 2020.
30. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*. 2020:1.

31. Li X, Li C, Zhu D. COVID-MobileXpert: On-Device COVID-19 Screening using Snapshots of Chest X-Ray. arXiv e-prints [Internet]. 2020 April 01, 2020:[arXiv:2004.03042 p.]. Available from: <https://ui.adsabs.harvard.edu/abs/2020arXiv200403042L>.
32. Hemdan EE-D, Shouman MA, Karar ME. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. arXiv preprint arXiv:2003.11055. 2020.
33. Khalifa NEM, Taha MHN, Hassanien AE, Elghamrawy S. Detection of Coronavirus (COVID-19) Associated Pneumonia based on Generative Adversarial Networks and a Fine-Tuned Deep Transfer Learning Model using Chest X-ray Dataset. arXiv preprint arXiv:2004.01184. 2020.
34. Du RH, Liang LR, Yang CQ, Wang W, Cao TZ, Li M, et al. Predictors of Mortality for Patients with COVID-19 Pneumonia Caused by SARS-CoV-2: A Prospective Cohort Study. Eur Respir J. 2020.
35. Guan W-j, Ni Z-y, Hu Y, Liang W-h, Ou C-q, He J-x, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. New England Journal of Medicine. 2020.
36. Ji W, Bishnu G, Cai Z, Shen X. Analysis clinical features of COVID-19 infection in secondary epidemic area and report potential biomarkers in evaluation. medRxiv. 2020:2020.03.10.20033613.
37. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A Novel Coronavirus Emerging in China — Key Questions for Impact Assessment. New England Journal of Medicine. 2020;382(8):692-4.
38. Lippi G, Plebani M. Laboratory abnormalities in patients with COVID-2019 infection. 2020(0):20200198.
39. Miao C, Zhuang J, Jin M, Xiong H, Huang P, Zhao Q, et al. A comparative multi-centre study on the clinical and imaging features of confirmed and unconfirmed patients with COVID-19. medRxiv. 2020:2020.03.22.20040782.