Leveraging GANs for data scarcity of COVID-19: Beyond the hype

Hazrat Ali
College of Science and Engineering
Hamad Bin Khalifa University,
Qatar Foundation, Doha, Qatar.
haali2@hbku.edu.qa

Christer Grönlund
Department of Radiation Sciences
Umeå University,
Umeå, Sweden.
christer.gronlund@umu.se

Zubair Shah
College of Science and Engineering
Hamad Bin Khalifa University,
Qatar Foundation, Doha, Qatar.
zshah@hbku.edu.qa

Abstract

Artificial Intelligence (AI)-based models can help in diagnosing COVID-19 from lung CT scans and X-ray images; however, these models require large amounts of data for training and validation. Many researchers studied Generative Adversarial Networks (GANs) for producing synthetic lung CT scans and X-Ray images to improve the performance of AI-based models. It is not well explored how good GAN-based methods performed to generate reliable synthetic data. This work analyzes 43 published studies that reported GANs for synthetic data generation. Many of these studies suffered data bias, lack of reproducibility, and lack of feedback from the radiologists or other domain experts. A common issue in these studies is the unavailability of the source code, hindering reproducibility. The included studies reported rescaling of the input images to train the existing GANs architecture without providing clinical insights on how the rescaling was motivated. Finally, even though GAN-based methods have the potential for data augmentation and improving the training of AI-based models, these methods fall short in terms of their use in clinical practice. This paper highlights research hotspots in countering the data scarcity problem, identifies various issues as well as potentials, and provides recommendations to guide future research. These recommendations might be useful to improve acceptability for the GAN-based approaches for data augmentation as GANs for data augmentation are increasingly becoming popular in the AI and medical imaging research community.

1. Introduction

The healthcare systems worldwide faced an unprecedented challenge with the COVID-19 pandemic. As a result, the capacity of coping with faster testing and providing care was put to test. When COVID-19 spread wildly, researchers were pushed to find quick ways for developing AI techniques to aid in combating the pandemic through early diagnosis. The most promising Artificial Intelligence (AI) techniques mostly fall in the deep learning category – models that consists of multiple layers of neural network (for example, convolutional neural networks (CNNs) and their variants). However, the potential of deep learning models to learn from the data rely on very large data. For image-based diagnosis and analysis, these models required a large amount of lung radiology images data. Since the data availability was meager, researchers rushed to use Generative Adversarial Networks (GANs) to generate synthetic Computed Tomography (CT) scans or X-Ray images that may capture the characteristics of real data with COVID-19 signs. It is well understood that image data augmentation is the most common application where GANs have found promising use due to their ability to generate realistic-looking images. Consequently, many studies reported the use of GANs to combat the data scarcity problem in training AI models for COVID-19 diagnosis. In this work, we identify 43 studies [1–43]. Typically, GANs serve as a submodule of the entire framework used explicitly as a data augmentation method, while the diagnosis of COVID-19 is made by using appropriate AI methods, for example, convolutional neural networks (ResNet, VGG16, etc) [44, 45].

Problem statement: While many reviews have been published on the role of AI methods in COVID-19 diagnosis [44–47], these reviews did not explicitly cover the shortcomings and risk of bias of the studies or did not explicitly cover GAN methods. One particularly relevant work is the review by Robert et al. [48]; however, it has a broad scope as it covers various machine learning methods for COVID-19. Besides, the focus in [48] is mainly the performance of AI methods for the diagnosis of COVID-19 and does not identify the limitations of the data augmentation methods explicitly. Another relevant study is the review by Cruz et al., [49] however, the focus of their work is on reviewing X-ray datasets for COVID-19 and not on reviewing the models. With the growing number of studies on GANs for synthetic data for COVID-19, it is important to review the reported methods and analyze them, particularly from the perspectives of model generalization, data representation, and clinical translation. Furthermore, to enhance the research and developments on GAN-based methods, it is critical to understand the importance of the risk of data bias, the challenges associated with the diagnostic accuracy, the evaluation mechanisms, and the inclusion of radiologists or domain experts in the loop when the research community presents more developments on the topic.

This paper aims to appreciate the early attempts for GAN-based methods to address the challenges related to COVID-19 data scarcity and diagnosis. Besides, it seeks to highlight the importance of addressing specific blights and imperfections as overcoming these can help increase the effectiveness and usability of the findings of these research developments. Unlike previous reviews, that cover a general description of the strengths of the AI methods for COVID-19 diagnosis, this work provides a more in-depth description of the various challenges that are associated with the GAN-based augmentation of data for COVID-19. While the discussion is provided from the perspective of COVID-19 disease, the findings would be useful for medical AI and medical data augmentation applications.

2. Methodology

We performed a literature search to retrieve relevant published studies that reported the use of GANs for generating lung CT scan or X-Ray image data. We retrieved the studies from Pubmed, Scopus, IEEEXplore, and Google Scholar. We found a total of 348 studies. In the first phase of study selection, we removed 81 duplicates and then performed title and abstract screening and removed 208 studies. We performed the full-text reading in the second phase and excluded 16 studies that did not fulfill the inclusion criteria (as identified in supplementary material). Finally, we included 43 studies in this mini-review. We cover only those studies that reported GANs for synthesis (data augmentation) purposes. Appendix 1 provides the search terms used for retrieval of the studies. Appendix 2 provides the criteria for inclusion and exclusion of the studies.

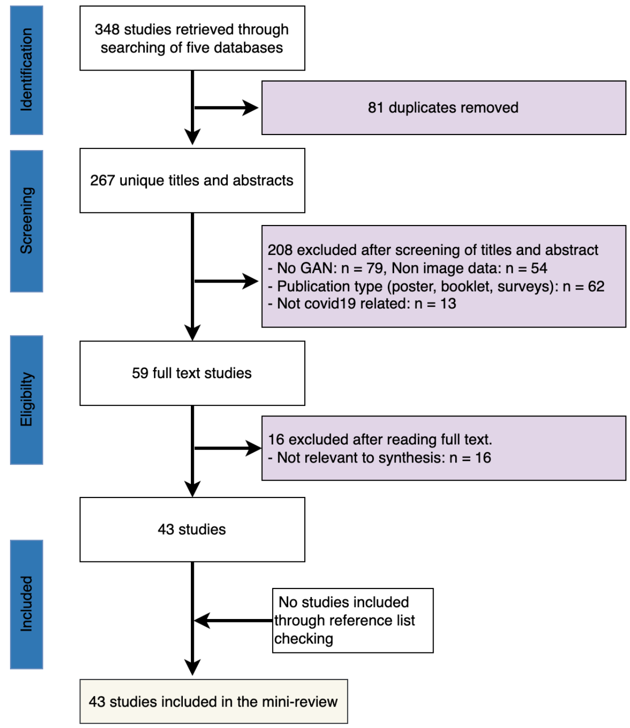

Figure 1. PRISMA flowchart for the selection of studies.

3. Results and Discussion

Fig. 1 shows the PRISMA flowchart for the selection of studies and Fig. 2 shows a summary of the demographics of the included studies. The key findings of our analysis can be categorized into the challenges related to:

1. Data proportion, such as the dataset size used for model training, the underlying bias in the data or the model training, and the associated data leakage problems.

2. The quality of the data such as the image resolution and the data modality.

3. The applications in COVID-19 such as the lack of demonstration of diagnostic performance.

4. The evaluation mechanism such as qualitative evaluation by radiologists or the metrics used for quantitative evaluation.

5. The potential clinical translations such as code availability and reproducibility.

Fig. 3 provides a summary of the challenges identified in this analysis. We discuss these in the following text.

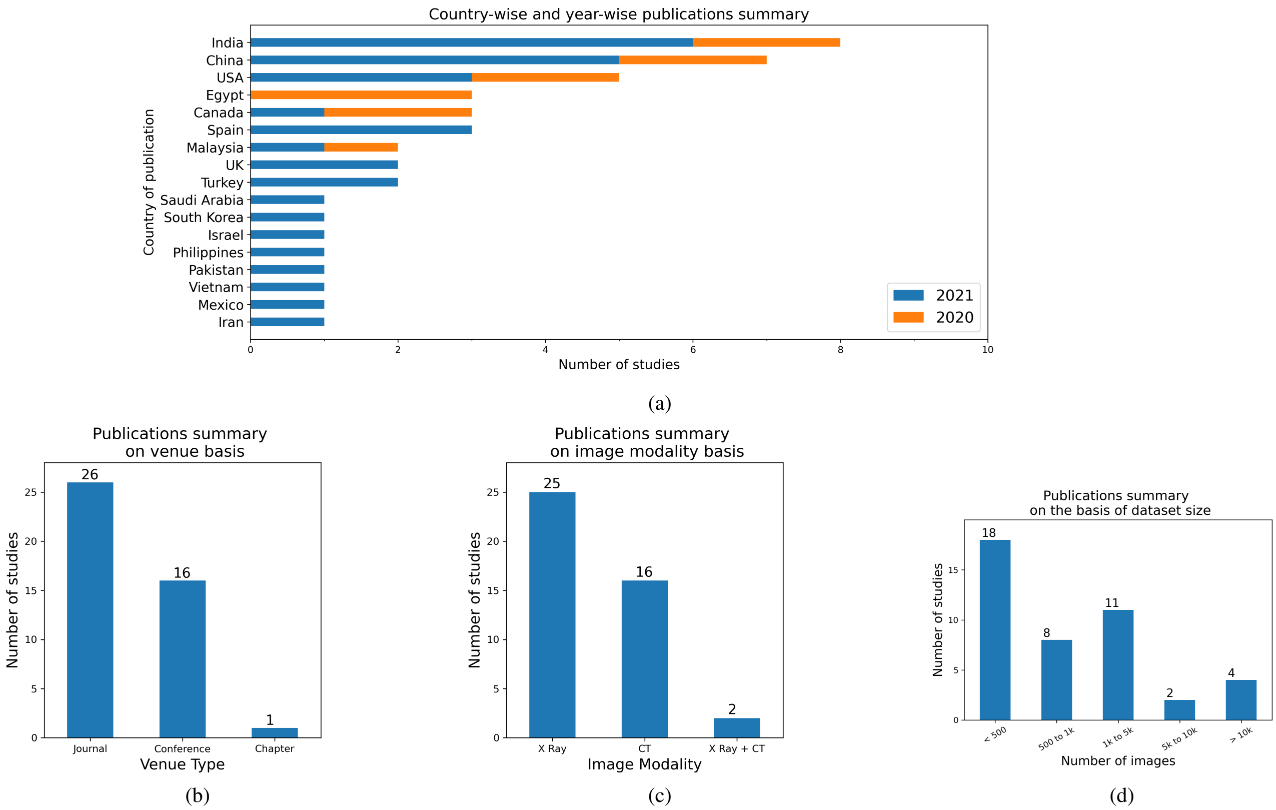

Figure 2. Demographics of the studies. (a) Summary of country-wise and year-wise number of studies. (b) Summary of number of the studies on the basis of venue. (c) Summary of number of studies on the basis of using different modalities of training data. (d) Summary of studies on the basis of the number of images used for training.

3.1. Dataset size

Better representation of the data can be obtained from large data. However, in the included studies, only four studies [17], [36], [39], [43] reported using more than ten thousand images, and only two studies [12], [37] reported using more than 5000 images. In more than half of the studies [1–7, 14–16, 20–29, 31, 34, 35, 38, 40, 42], the number of images reported for training was less than 1000. So, this training data size raises concerns about the generalization of GANs training. Similarly, in many studies, we found a lack of information on the number of patients (individuals) from whom the images have been acquired. Besides, we found that the number of individuals, even if reported in the studies, was too small compared to what would be recommended for training a GAN model. Only three studies [4], [7], [20] reported using data of more than 500 individuals. In addition, a common challenge in the crowd-source data is mapping the number of images to the number of individuals. Furthermore, the inclusion/exclusion of an individual and the methods and criteria for recruiting individuals to record data may affect the data demographics, bias, and imbalance, consequently influencing the model training.

3.2. Risk of bias

Bias in data refers to the imbalanced representation of different groups in the data. For example, there can be gender bias, age bias, or sample bias, etc. If the original data used for training GANs has a bias, then there is a high risk that the bias will be carried to the synthetic data. It is clear that the crowd-sourced datasets available publicly lacked any such specifications, or the studies fail to report to cater for the biases in the data. Eventually, the bias in data will result in a bias in the training of the AI-based diagnosis model too. For example, Garcia et al. [49] reported in their analysis of studies reporting X-ray images datasets to diagnose COVID-19 that the models trained with chest X-ray datasets were prone to high bias. The potential risk of bias in training AI models is a critical bottleneck in the generalization ability and robustness of the models. For example, at the beginning of the pandemic, most of the positive cases were for adults only. On the other hand, a pneu-

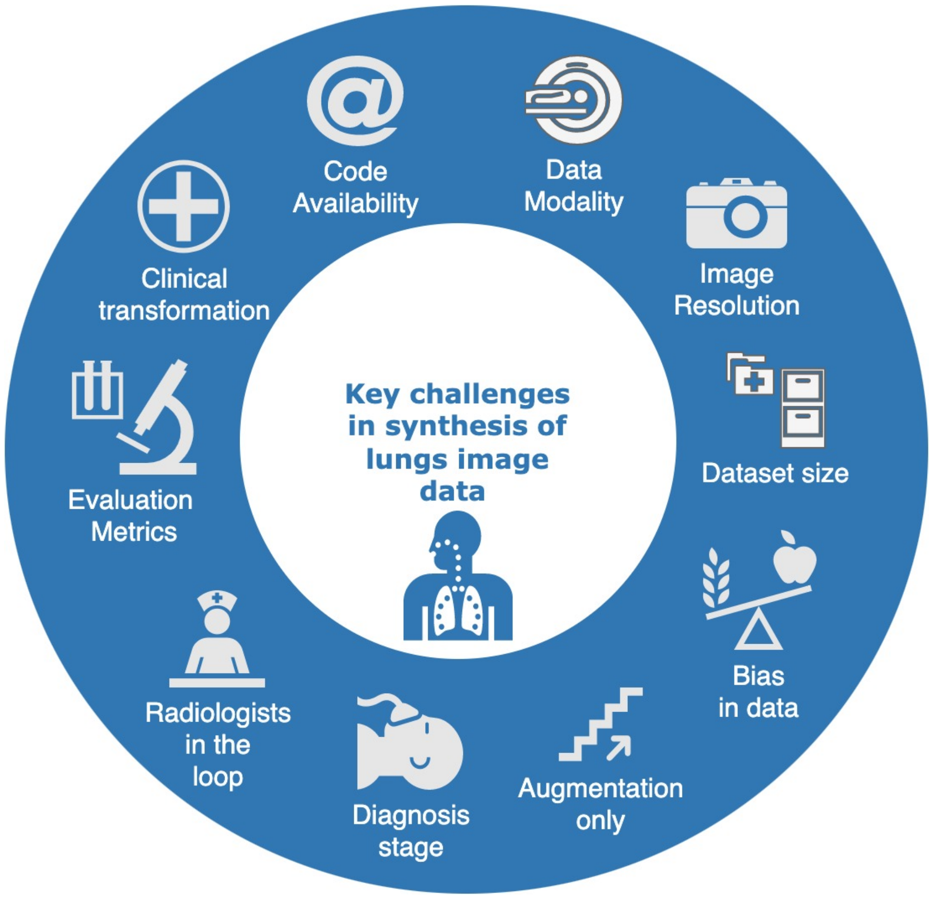

Figure 3. Key challenges identified in the studies reporting GAN-based data augmentation for COVID-19.

monia dataset may have more samples for lung X-ray images of children. A GANs model trained on such data will carry forward a similar bias. Suppose we use the pneumonia dataset with more young people/pediatrics samples and the COVID-19 dataset with more examples for adults. In that case, the training may end up in a model that has learned to classify lung images of adults versus children instead of classifying COVID-19 versus pneumonia. Some studies [4], [22], [29], [32] used the RSNA pneumonia dataset available from Kaggle[1]. This particular dataset, as reported by Kermany et al., [50] has image data for young people between the age of one to five. Ablation studies and adversarial tests are important to report to avoid such biases. So, one may test the model for extreme cases such as samples of adults with pneumonia or samples of young people that were COVID-19 positive.

3.3. Data leakage

Some studies reported results for multiple datasets. However, no cross-verification mechanism was reported to avoid data leakage. Hence, reporting results on two different public datasets does not necessarily imply validation/testing on independent samples as these datasets may have borrowed from each other. For example, [11], [18], [33] reported the use of three different datasets of chest X-ray images including [44, 51]. However, the data in Tawsif et al., [44] borrows images from Cohen et al., [51]. Amongst these, [18] reported deletion of duplicate images though details were missing on how duplicate images were identified). [11] reported that they used one dataset to train the model and then used the trained model to label images of the other dataset but did not identify that the two datasets have overlapping entries. In summary, prevention of data leakage is not guaranteed by the studies, and the studies lack explicit details on addressing the confounders. The majority of these studies (36 out of 43 studies) reported using a publicly available dataset for CT and X-ray images. For similar reasons as discussed above, many of the images in these data were collected through crowdsourcing as an early attempt to facilitate research on AI for COVID-19. Consequently, the data collection (more often through crowdsourcing) was mostly uncontrolled and not well-curated.

[1]https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data

Hence, they lacked many details such as demographics or guarantee to prevent data leakage. Only a few studies [7], [12], [14], [15], [20], [31], [34] reported using privately collected data where the researchers had more control over the data collection mechanism.

3.4. Image resolution

A common pre-processing step used in the included studies was rescaling the image data to a resolution 512×512 or 224×224 pixels. The motivation for resizing the images was mostly model-driven and not driven by the clinical objective. At least, the studies provided no insights on how or why they chose this resolution. In some cases, the rescaling seems quite an aggressive downsizing of the input X-ray images from 2170×1953 pixels to 512×512. Similarly, for volumetric CT scans, the studies had to restrict themselves to 2D images (individual scans) to train the model. Insights on such choices are either missing or very limited.

3.5. Data modality

In the 43 studies, we found that almost 60% of the studies (25 used X-ray images versus 16 used CT) reported using X-ray image data to train GANs models for data augmentation to achieve diagnosis with AI models. However, it is well understood that CT scan has higher sensitivity than X-ray images for COVID-19 diagnosis [49].

3.6. Augmentation only

Some studies reported using GAN-based methods to generate synthetic lung CT scans or X-ray images [1], [14], [21], [22], [24], [25], [34], [35], [42]. For example, [14] covers segmentation only). [25] reports the synthesis of volumetric CT scans using a 3D conditional GAN however, the study did not analyze the generated synthetic data for its ability for diagnosing COVID-19, meaning the studies did not explore the effectiveness of the generated synthetic data.

3.7. Diagnosis stage

A common trend in the studies was the staging of the disease. Most of the image data shared during the earlier days of the pandemic comprised cases of severe COVID-19 with a noticeable impact on the lungs. Consequently, the GANs trained would generate synthetic images of the extreme stage where the presence of the infection was already established. Hence, GAN-based methods that can support early diagnosis of the infection are limited.

3.8. Radiologists in the loop

In the included studies, none (but one study [22]) reported an evaluation/rating of the synthesized data by presenting it to radiologists. So, there remains an unclear interpretation of the qualitative assessment of the synthesized CT/X-Ray images. [22] reported qualitative analysis of synthesized X-ray images by a radiologist that the radiologist could distinguish from real X-ray images for COVID-19 positive cases. Besides, the qualitative analysis suggested that the synthesized images fall short of diagnosis quality.

3.9. Evaluation Metrics

One common mechanism for quantitative evaluation of synthetic CT scan or X-ray images was the use of metrics such as Structure Similarity Index Measure (SSIM), Frechet Inception Distance (FID), and Peak Signal-to-Noise Ratio (PSNR). However, the evaluation with SSIM and FID without input from radiologists can be overly optimistic as these metrics are primarily derived from computer vision literature, and their suitability for medical images might be limited. Besides, the lack of evaluation by radiologists might hinder the acceptability of these models being translated into clinical practice and hence lose the very purpose of using AI to aid in combating the pandemic.

3.10. Code availability

To advance the developments in data augmentation, there is a dire need to provide reproducible softwares or code. We found the lack of reproducibility analysis as a common trend in the studies. More specifically, only three studies [14], [24], [39] provided links to publicly available Github repositories for their code. Among these, one of the links [14] was found broken.

3.11. Can the studies be translated to clinical applications?

We believe with the existing shortcomings in the studies, GAN-based studies are not yet ready to be translated to clinical practice. Nevertheless, despite the shortcomings, the potential impact of using GAN-based methods to improve the training of AI models for COVID-19 diagnosis cannot be denied. Therefore, future translation to clinical applications is not futile.

4. Suggestions

In this section, we provide suggestions on handling the GAN-based data augmentation for COVID-19. We believe that these suggestions will provide a roadmap to the research community and improve future study designs.

4.1. Large data

Since much more data is becoming available for lung CT scan and X-Ray images; it is highly recommended that any current and future research on GANs for COVID-19 is based on much larger datasets than those reported in the earlier studies surveyed in this work. In addition, the reporting of data demographics should be encouraged as this will help mitigate challenges such as data leakage or risk of bias in data.

4.2. Reproducibility

We urge the research community to publish their codes for reproducibility to promote usability and utility. The transparency in using the models will be a key to increasing their acceptability and help the community gain better insights into the model. In addition, it will also enable reproducibility as well as future developments on top of existing works.

4.3. Adversarial test

A useful way to evaluate the model is to report aggressive adversarial test strategies, for example, by generating GAN-based data from lung CT scans before 2020 and then using the generating data to evaluate the AI model for COVID-19 diagnosis. Any positive detection will provide an opportunity to investigate the model performance as COVID-19 was only spreading in late 2020.

4.4. Evaluation methods

The research community (including the authors, the reviewers, and the readers) should consider evaluating their work against the Radiological Society of North America Checklist (RSNA CLAIM) [52], assessment tools like PROBAST [53], or similar metrics, that provide a comprehensive checklist for AI models from the perspective of data, model training, and evaluation metrics. A lack of compliance with the guidelines such as those in CLAIM [52] or validation of the studies using assessment tools such as PROBAST [53], will hamper the translation of these findings into clinical applications.

4.5. Early feedback from radiologists

The research community should consider incorporating the feedback and input from doctors and clinicians starting from the earlier phase of the study design. Doing so will help better understand the key challenges in handling lung CT scans or X-ray images data, interpret the subtle information in the data, and increase the opportunities for clinical translation of the developed methods.

5. Limitations

Our analysis includes studies from four databases, namely Pubmed, Scopus, IEEE Xplore, and Google Scholar. So, studies that are not indexed in these databases have been left out. We did not include pre-prints and unpublished literature in this analysis as they are not peer-reviewed, but might contain good research outcomes. The scope of our work is limited to images-based studies only. The analysis does not directly compare the evaluation metrics as the included studies differ in terms of the dataset size, choice of GAN architecture, the timeline of the study. The research and developments on methods for COVID-19 occur extremely fast. Hence, it is possible that as this work is being drafted, several additional studies might be published which are not covered in this analysis.

6. Conclusion

This work provided a critical analysis of the shortcomings of GAN-based methods for data augmentation in applications related to AI-based diagnosis of COVID-19. We identified many areas where the research on using GAN-based methods for data augmentation in COVID-19 could be improved. We believe that the findings of this analysis of studies reporting GAN-based synthetic data complement findings of a previous study by [48] who reported the many pitfalls in the use of machine learning methods for COVID-19, and an earlier study by [49] that identified the risk of bias in the reported AI models. The analysis in this work will help the readers to understand the limitations of published studies and design better studies in the future to overcome the shortcomings.

References

[1] Y. Jiang, H. Chen, M. Loew, and H. Ko, "Covid-19 ct image synthesis with a conditional generative adversarial network," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 441–452, 2020. 1, 3, 5

[2] Autee P, Bagwe S, Shah V, and Srivastava K, "StackNet-DenVIS: a multi-layer perceptron stacked ensembling approach for COVID-19 detection using X-ray images.," *Physical and engineering sciences in medicine*, vol. 43, no. 4, pp. 1399–1414, 2020. 1, 3

[3] Loey M, Manogaran G, and Khalifa NEM, "A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images.," *Neural computing & applications*, pp. 1–13, 2020. 1, 3

[4] Karakanis S and Leontidis G, "Lightweight deep learning models for detecting COVID-19 from chest X-ray images.," *Computers in biology and medicine*, vol. 130, p. 104181, 2021. 1, 3, 4

[5] Li Z, Zhang J, Li B, Gu X, and Luo X, "COVID-19 diagnosis on CT scan images using a generative adversarial network and concatenated feature pyramid network with an attention mechanism.," *Medical physics*, vol. 48, no. 8, pp. 4334–4349, 2021. 1, 3

[6] Rasheed J, Hameed AA, Djeddi C, Jamil A, and Al-Turjman F, "A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images.," *Interdisciplinary sciences, computational life sciences*, vol. 13, no. 1, pp. 103–117, 2021. 1, 3

[7] Morís DI, de Moura Ramos JJ, Buján JN, and Hortas MO, "Data augmentation approaches using cycle-consistent adversarial networks for improving COVID-19 screening in

portable chest X-ray images.," *Expert systems with applications*, vol. 185, p. 115681, 2021. 1, 3, 5

[8] Singh RK, Pandey R, and Babu RN, "COVIDScreen: explainable deep learning framework for differential diagnosis of COVID-19 using chest X-rays.," *Neural computing & applications*, pp. 1–22, 2021. 1

[9] Goel T, Murugan R, Mirjalili S, and Chakrabartty DK, "Automatic Screening of COVID-19 Using an Optimized Generative Adversarial Network.," *Cognitive computation*, pp. 1–16, 2021. 1

[10] Zhang L, Shen B, Barnawi A, Xi S, Kumar N, and Wu Y, "FedDPGAN: Federated Differentially Private Generative Adversarial Networks Framework for the Detection of COVID-19 Pneumonia.," *Information systems frontiers : a journal of research and innovation*, pp. 1–13, 2021. 1

[11] Y. Karbhari, A. Basu, Z. W. Geem, G.-T. Han, and R. Sarkar, "Generation of synthetic chest X-ray images and detection of COVID-19: A deep learning based approach," *Diagnostics*, vol. 11, no. 5, p. 895, 2021. 1, 4

[12] Amin J, Sharif M, Gul N, Kadry S, and Chakraborty C, "Quantum Machine Learning Architecture for COVID-19 Classification Based on Synthetic Data Generation Using Conditional Adversarial Neural Network.," *Cognitive computation*, pp. 1–12, 2021. 1, 3, 5

[13] N. Hernandez-Cruz, D. Cato, and J. Favela, "Neural style transfer as data augmentation for improving covid-19 diagnosis classification," *SN Computer Science*, vol. 2, no. 5, p. 410, 2021. 1

[14] Zhang J, Yu L, Chen D, Pan W, Shi C, Niu Y, Yao X, Xu X, and Cheng Y, "Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT images.," *Biomedical signal processing and control*, vol. 69, p. 102901, 2021. 1, 3, 5

[15] H. Jiang, S. Tang, W. Liu, and Y. Zhang, "Deep learning for covid-19 chest ct (computed tomography) image analysis: A lesson from lung cancer," *Computational and Structural Biotechnology Journal*, vol. 19, pp. 1391–1399, 2021. 1, 3, 5

[16] P. Mann, S. Jain, S. Mittal, and A. Bhat, "Generation of COVID-19 Chest CT Scan Images using Generative Adversarial Networks," in *2021 International Conference on Intelligent Technologies (CONIT)*, pp. 1–5, 2021. Journal Abbreviation: 2021 International Conference on Intelligent Technologies (CONIT). 1, 3

[17] T. M. Quan, H. M. Thanh, T. D. Huy, N. D. T. Chanh, N. T. P. Anh, P. H. Vu, N. H. Nam, T. Q. Tuong, V. M. Dien, B. Van Giang, B. H. Trung, and S. Q. H. Truong, "XPGAN: X-Ray Projected Generative Adversarial Network For Improving Covid-19 Image Classification," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 1509–1513, 2021. Journal Abbreviation: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). 1, 3

[18] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection," *IEEE Access*, vol. 8, pp. 91916–91923, 2020. 1, 4

[19] Z. Liang, J. X. Huang, J. Li, and S. Chan, "Enhancing Automated COVID-19 Chest X-ray Diagnosis by Image-to-Image GAN Translation," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1068–1071, 2020. Journal Abbreviation: 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 1

[20] D. I. Morís, J. de Moura, J. Novo, and M. Ortega, "Cycle Generative Adversarial Network Approaches to Produce Novel Portable Chest X-Rays Images for Covid-19 Diagnosis," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1060–1064, 2021. Journal Abbreviation: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 1, 3, 5

[21] J. -A. Rodríguez-de-la-Cruz, H. -G. Acosta-Mesa, and E. Mezura-Montes, "Evolution of Generative Adversarial Networks Using PSO for Synthesis of COVID-19 Chest X-ray Images," in *2021 IEEE Congress on Evolutionary Computation (CEC)*, pp. 2226–2233, 2021. Journal Abbreviation: 2021 IEEE Congress on Evolutionary Computation (CEC). 1, 3, 5

[22] S. Menon, J. Galita, D. Chapman, A. Gangopadhyay, J. Mangalagiri, P. Nguyen, Y. Yesha, Y. Yesha, B. Saboury, and M. Morris, "Generating Realistic COVID-19 x-rays with a Mean Teacher + Transfer Learning GAN," in *2020 IEEE International Conference on Big Data (Big Data)*, pp. 1216–1225, 2020. Journal Abbreviation: 2020 IEEE International Conference on Big Data (Big Data). 1, 3, 4, 5

[23] S. Dong and Z. Zhang, "Joint Optimization of CycleGAN and CNN Classifier for COVID-19 Detection and Biomarker Localization," in *2020 IEEE International Conference on Progress in Informatics and Computing (PIC)*, pp. 112–118, 2020. Journal Abbreviation: 2020 IEEE International Conference on Progress in Informatics and Computing (PIC). 1, 3

[24] Z. Yang, L. Zhao, S. Wu, and C. Y. -C. Chen, "Lung Lesion Localization of COVID-19 From Chest CT Image: A Novel Weakly Supervised Learning Method," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 6, pp. 1864–1872, 2021. 1, 3, 5

[25] J. Mangalagiri, D. Chapman, A. Gangopadhyay, Y. Yesha, J. Galita, S. Menon, Y. Yesha, B. Saboury, M. Morris, and P. Nguyen, "Toward Generating Synthetic CT Volumes using a 3D-Conditional Generative Adversarial Network," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 858–862, 2020. Journal Abbreviation: 2020 International Conference on Computational Science and Computational Intelligence (CSCI). 1, 3, 5

[26] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and M. Guizani, "DL-CRC: Deep Learning-Based Chest Radiograph Classification for COVID-19 Detection: A Novel Ap-

proach," *IEEE Access*, vol. 8, pp. 171575–171589, 2020. 1, 3

[27] M. Loey, Smar, F. ache, and N. E. M Khalifa, "Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning," *Symmetry*, vol. 12, no. 4, p. 651, 2020. Publisher: Multidisciplinary Digital Publishing Institute. 1, 3

[28] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and S. H. N. Taha, "The Detection of COVID-19 in CT Medical Images: A Deep Learning Approach," *Big Data Analytics and Artificial Intelligence against COVID-19: Innovation Vision and Approach*, pp. 73–90, 2020. Publisher: Springer. 1, 3

[29] H. Zunair and A. B. Hamza, "Synthesis of COVID-19 chest X-rays using unpaired image-to-image translation," *Social Network Analysis and Mining*, vol. 11, no. 1, pp. 1–12, 2021. Publisher: Springer. 1, 3, 4

[30] J. S. Sachdev, N. Bhatnagar, and R. Bhatnagar, "Deep Learning Models Using Auxiliary Classifier GAN for Covid-19 Detection–A Comparative Study," *The International Conference on Artificial Intelligence and Computer Vision*, pp. 12–23, 2021. Publisher: Springer. 1

[31] D. I. Morís, J. de Moura, J. Novo, and M. Ortega, "Portable Chest X-ray Synthetic Image Generation for the COVID-19 Screening," *Engineering Proceedings*, vol. 7, no. 1, p. 6, 2021. Publisher: Multidisciplinary Digital Publishing Institute. 1, 3, 5

[32] Y. Sanajalwe, M. Anbar, and S. Al-E'Mari, "Covid-19 automatic detection using deep learning," *Computer Systems Science and Engineering*, vol. 39, no. 1, pp. 15–35, 2021. Place: ["National Advanced IPv6 Centre of Excellence (Nav6), Universiti Sains Malaysia, 11800 USM, Penang, Malaysia", "Computer Science Department, Northern Border University (NBU), 9280 NBU, Ar'ar, Saudi Arabia"]. 1, 4

[33] A. Al-Shargabi, J. Alshobaili, A. Alabdulatif, and N. Alrobah, "Covid-cgan: Efficient deep learning approach for covid-19 detection based on cxr images using conditional gans," *Applied Sciences (Switzerland)*, vol. 11, no. 16, 2021. 1, 4

[34] S. Shivadekar, J. Mangalagiri, P. Nguyen, D. Chapman, M. Halem, and R. Gite, "An intelligent parallel distributed streaming framework for near real-time science sensors and high-resolution medical images," in *50th International Conference on Parallel Processing Workshop*, (United States), 2021. 1, 3, 5

[35] J. Toutouh and U.-M. O'Reilly, "Signal propagation in a gradient-based and evolutionary learning system," in *GECCO 2021 - Proceedings of the 2021 Genetic and Evolutionary Computation Conference*, (United States), pp. 377–385, 2021. 1, 3, 5

[36] E. Acar, E. Sahin, and I. Yilmaz, "Improving effectiveness of different deep learning-based models for detecting COVID-19 from computed tomography (CT) images," *Neural Computing and Applications*, 2021. 1, 3

[37] S. Sheykhivand, Z. Mousavi, S. Mojtahedi, T. Yousefi Rezaii, A. Farzamnia, S. Meshgini, and I. Saad, "Developing an efficient deep neural network for automatic detection of COVID-19 using chest X-ray images," *Alexandria Engineering Journal*, vol. 60, no. 3, pp. 2885–2903, 2021. Place: ["Biomedical Engineering Department, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran", "Department of Mechanical Engineering, Faculty of Mechanical Engineering, University of Tabriz, Tabriz, Iran", "Department of Biomedical Engineering, Faculty of Engineering, Işık University, Istanbul, Turkey", "Faculty of Engineering, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia"]. 1, 3

[38] A. Rangarajan and H. Ramachandran, "A preliminary analysis of AI based smartphone application for diagnosis of COVID-19 using chest X-ray images," *Expert Systems with Applications*, vol. 183, 2021. 1, 3

[39] H. Li, Y. Hu, S. Li, W. Lin, P. Liu, R. Higashita, and J. Liu, "CT Scan Synthesis for Promoting Computer-Aided Diagnosis Capacity of COVID-19," in *Intelligent Computing Theories and Application*, vol. 12464 of *Lecture Notes in Computer Science*, Springer, 2020. 1, 3, 5

[40] M. Zulkifley, S. Abdani, and N. Zulkifley, "COVID-19 screening using a lightweight convolutional neural network with generative adversarial network data augmentation," *Symmetry*, vol. 12, no. 9, 2020. 1, 3

[41] A. Ambita, E. Boquio, and P. Naval, Jr., "COViT-GAN: Vision Transformer for COVID-19 Detection in CT Scan Images with Self-Attention GAN for Data Augmentation," in *International Conference on Artificial Neural Networks*, vol. 12892, Springer, 2021. 1

[42] J. Toutouh, M. Esteban, and S. Nesmachnow, "Parallel/Distributed Generative Adversarial Neural Networks for Data Augmentation of COVID-19 Training Images," in *Latin American High Performance Computing Conference*, vol. 1327 of *Communications in Computer and Information Science*, (United States), Springer, 2021. 1, 3, 5

[43] A. Bar-El, D. Cohen, N. Cahan, and H. Greenspan, "Improved CycleGAN with application to COVID-19 classification," in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, vol. 11596, 2021. 1, 3

[44] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, and N. Al Emadi, "Can AI help in screening viral and COVID-19 pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020. Publisher: IEEE. 1, 2, 4

[45] A. Abd-Alrazaq, M. Alajlani, D. Alhuwail, J. Schneider, S. Al-Kuwari, Z. Shah, M. Hamdi, and M. Househ, "Artificial intelligence in the fight against COVID-19: scoping review," *Journal of medical Internet research*, vol. 22, no. 12, 2020. Publisher: JMIR Publications Inc., Toronto, Canada. 1, 2

[46] L. Wang, Y. Zhang, D. Wang, X. Tong, T. Liu, S. Zhang, J. Huang, L. Zhang, L. Chen, and H. Fan, "Artificial intelligence for COVID-19: a systematic review," *Frontiers in medicine*, p. 1457, 2021. Publisher: Frontiers. 2

[47] A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe, and X. Cheng, "Artificial intelligence and machine learning to fight COVID-19," *Physiological genomics*, vol. 52, no. 4, pp. 200–202, 2020. Publisher: American Physiological Society Bethesda, MD. 2

[48] M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I. Aviles-Rivero, C. Etmann, C. McCague, and L. Beer, "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans," *Nature Machine Intelligence*, vol. 3, no. 3, pp. 199–217, 2021. Publisher: Nature Publishing Group. 2, 6

[49] B. G. Santa Cruz, M. N. Bossa, J. Sölter, and A. D. Husch, "Public Covid-19 X-ray datasets and their impact on model bias–A systematic review of a significant problem," *Medical image analysis*, vol. 74, p. 102225, 2021. Publisher: Elsevier. 2, 3, 5, 6

[50] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, and F. Yan, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018. Publisher: Elsevier. 4

[51] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future," *arXiv preprint arXiv:2006.11988*, 2020. 4

[52] J. Mongan, L. Moy, and C. E. Kahn Jr, "Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers," *Radiology: Artificial Intelligence*, vol. 2, no. 2, p. e200029, 2020. Publisher: Radiological Society of North America. 6

[53] R. F. Wolff, K. G. Moons, R. D. Riley, P. F. Whiting, M. Westwood, G. S. Collins, J. B. Reitsma, J. Kleijnen, S. Mallett, and PROBAST Group†, "PROBAST: a tool to assess the risk of bias and applicability of prediction model studies," *Annals of internal medicine*, vol. 170, no. 1, pp. 51–58, 2019. Publisher: American College of Physicians. 6