

Wasserstein GAN based Chest X-Ray Dataset Augmentation for Deep Learning Models: COVID-19 Detection Use-Case

Item Type	Conference Contribution
Authors	Hussain, B. Zahid;Andleeb, Ifrah;Ansari, Mohammad Samar;Joshi, Amit Mahesh;Kanwal, Nadia
Citation	Hussain, B. Z., Andleeb, I., Ansari, M. S., Joshi, A. M., & Kanwal, N. (2022). Wasserstein GAN based chest X-Ray dataset augmentation for deep learning models: COVID-19 detection use-case. 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 2058-2061. 10.1109/EMBC48229.2022.9871519.
DOI	10.1109/EMBC48229.2022.9871519
Publisher	IEEE
Download date	2026-05-10 18:13:52
Item License	https://creativecommons.org/licenses/by-nc-nd/4.0/
Link to Item	http://hdl.handle.net/10034/627209

Wasserstein GAN based Chest X-Ray Dataset Augmentation for Deep Learning Models: COVID-19 Detection Use-Case

B. Zahid Hussain[#] *Member, IEEE*, Ifrah Andleeb[#] *Member, IEEE*, Mohammad Samar Ansari *Senior Member, IEEE*, Amit Mahesh Joshi *Senior Member, IEEE*, and Nadia Kanwal *Senior Member, IEEE*

Abstract—The novel coronavirus infection (COVID-19) is still continuing to be a concern for the entire globe. Since early detection of COVID-19 is of particular importance, there have been multiple research efforts to supplement the current standard RT-PCR tests. Several deep learning models, with varying effectiveness, using Chest X-Ray images for such diagnosis have also been proposed. While some of the models are quite promising, there still remains a dearth of training data for such deep learning models. The present paper attempts to provide a viable solution to the problem of data deficiency in COVID-19 CXR images. We show that the use of a Wasserstein Generative Adversarial Network (WGAN) could lead to an effective and lightweight solution. It is demonstrated that the WGAN generated images are at par with the original images using inference tests on an already proposed COVID-19 detection model.

Index Terms—COVID-19, Deep Learning, Data Augmentation, Generative Adversarial Network (GAN), Wasserstein GAN.

I. INTRODUCTION

Since the advent of COVID-19 (late 2019), the world has witnessed drastic public health, economical, and political conundrums. Early detection of infection has been the primary objective of response action(s), with the Reverse Transcription Polymerase Chain Reaction (RT-PCR) test being the workhorse of the detection pipeline, and a variety of artificial intelligence models being proposed to enable fast and easier (than RT-PCR) detection [1]. The recent work in [1] presented a lightweight deep learning model for COVID-19 detection from Chest X-Ray (CXR) images, while also including a concise summary of similar existing works until that time. Apart from being faster and less expensive than RT-PCR tests, such CXR-based tests can also be used to monitor disease progression with multiple X-Ray scans spread over days/weeks, which RT-PCR tests are incapable of.

While the recently proposed deep learning (DL) models for COVID-19 detection (and other similar infection/tumor detection models) do indeed offer a faster and lightweight solution, as compared to laboratory tests such as RT-PCR, the

accuracy of such models is inherently dependent on the data that is available to train these models. For applications where annotated datasets are readily available, or can be generated relatively easily, such as object detection, the research has progressed by leaps and bounds [2]. It is envisaged that medical diagnostic deep learning models will also benefit greatly with the availability of large high quality datasets. We specifically look into the use-case of Chest X-Ray (CXR) images for COVID-19 diagnosis. CXR images were found to be beneficial in the follow up of critically ill COVID-19 patients in another study [3], wherein the radiographic findings on chest x-ray in COVID-19 pneumonia patients have been found consistent with the radiographic findings detected on chest CT scans and on CXR scans in previous reports. The presence of symptoms correlated significantly with abnormal chest X-Ray findings [3]. It can be well concluded that CXR may be helpful as an aiding tool in the diagnosis and follow up in COVID-19 pneumonia [3].

Contribution: We propose a lightweight DL model employing a Wasserstein Generative Adversarial Network (WGAN) to synthetically generate high quality COVID-19 positive CXR images. This is becoming increasingly important as high-quality, abundant, and varying training data, without sacrificing Ground Glass Opacities (GGO), is required for effective training of potential high-performance deep learning models for COVID-19 detection. As is shown later in the paper, the quality of the generated images using the proposed approach is better than existing approaches, making the generated dataset amenable for DL applications.

The contents of the paper are organized as follows. Section II contains a summary of pertinent research. In Section III, the proposed approach and methodology are described. Section IV evaluates the proposed system performance and results. Finally, Section VI contains concluding remarks.

II. RELATED WORK

It is well known that medical image datasets are both time-consuming to generate and annotate. Additionally, the unavailability of labeled, extensive, generalized datasets is one of the most notable problems with DL in medical imaging. Hence, GANs have been used to produce synthetic samples which are very close to the real images in appearance, which are inherently labelled, and help to improve the models.

[#]Equal Contribution

B.Z. Hussain and I. Andleeb are with the Z.H. College of Engineering & Technology, Aligarh Muslim University, India. M.S. Ansari is with the Faculty of Science and Engineering, University of Chester, United Kingdom. A.M. Joshi is with the Malaviya National Institute of Technology Jaipur, India. N. Kanwal is with Keele University, United Kingdom (Emails: m.ansari@chester.ac.uk, amjoshi.ece@mmit.ac.in, n.kanwal@keele.ac.uk)

Sanfort et al. [4] evaluated use of CycleGAN for data augmentation in CT segmentation tasks. For kidney and liver the in-distribution volume measurement errors were exceptional (3% and 4% respectively), demonstrating the potential of GAN-based data augmentation for medical applications. Sundaram et al. [5] compared the GAN-based augmentation with the traditional augmentation models to synthetically expand the CheXpert dataset of chest radiographs. Radford et al. [6] demonstrated the natural general image representations by using the learned features. They proved DCGANs to yield remarkable results for evaluating the quality of unsupervised learning algorithms and to apply them as a feature extractor on supervised datasets and evaluate the performance of linear models trained on Imagenet-1K. Giraldi et al. [7] utilize Super-Resolution Generative Adversarial Network (SRGAN) to obtain high resolution digital periapical radiographs. The results of SRGAN models using transfer learning were found to be better on average achieving the desired image quality with more defined edge details and fewer blur effects. Lei et al. [8] presented an unpaired DL scheme for the reconstruction of MRI images. Taking note of the problem where high-quality training labels are inadequate, Wasserstein GANs (WGANs) are being used for the unpaired training. Their findings on abdominal and knee MRI datasets authenticates the effectiveness of Wasserstein distance based adversarial networks and also confirms the practical reconstruction of MRIs as a feasible alternative to ordinary traditional methods.

Summarily, although the potential of GANs has been demonstrated for some medical datasets, WGANs have not yet been employed for the generation of a COVID-19 CXR dataset. This paper demonstrates that WGANs are indeed very viable candidates for medical data generation.

III. PROPOSED METHODOLOGY

The performance of deep learning models is dependent on the quality and quantity of data presented to them during the training. With a reasonably good dataset, a DL model can be made to learn the inherent characteristics of the data without explicitly programmed. Here, we aim to present a WGAN-based technique which aims to synthetically generate training images for better development of deep learning models in diagnosing COVID-19 [1], [9]. This approach is useful for generating high quality Chest X-Ray (CXR) images without compromising the differentiating factor like ground glass opacities. The following sections describe the suggested operational framework, design of the proposed network, and the solution.

A. Dataset

An open-access benchmark dataset *viz.* COVID-19 Radiography Dataset¹, comprising of 3615 COVID-19 images, is used to facilitate the training and evaluating the proposed WGAN model. The respective sources from where the different COVID-19 positive CXR images are collected are also highlighted (in green) in Fig. 1.

¹[Online] Available at: <https://www.kaggle.com/preetviradiya/covid19-radiography-dataset>

B. Dataset handling

The dataset consists of shuffled 3615 CXR images of COVID-19 positive cases, for training the proposed model. To mitigate the classification accuracy degradation due to deformations [10], the actual image size (299×299) is taken without any resizing. This size is chosen by taking into account the training time for the model while still maintaining a good model efficacy.

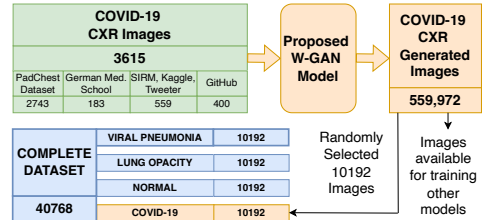


Fig. 1. Dataset: the original dataset, and the generated dataset with WGANs

C. Methodology and Approach

1) *Experimental Design:* We propose to supplement the existing CXR images datasets by generating high quality CXR images. Previous GAN based synthetic image generation approaches have the downside of low-quality of the generated images while keeping the training time fairly moderate. One major stumbling block with these studies are the differentiating factor for COVID-19 CXR images i.e the Ground Glass Opacities (GGO). A low quality generated image may not manifest a fairly discernable GGO and hence failing the model in diagnosing COVID-19 efficiently. Since the end goal of all these approaches is implementation for real-world applications, there is the need to use viable training images without compromising the efficacy of diagnosis in any form. Therefore we propose a WGAN based generative model which exhibits favorable performance gain.

2) *Image generation using WGAN:* It is well known that GANs can be used to generate new data instances which opens a whole new world for research with constrained training data. However, conventional GANs typically are unable to generate high quality images while maintaining a reasonable model size and training time. Wasserstein GANs are an extension of GANs that seek to diminish mode collapse with the increased contrivance of generating meaningful and debuggable hyper-parameter searching trajectories. WGANs make the training process more stable and make it less sensitive to model architecture and hyper-parameters. Just like any typical GAN, WGAN hosts two competing and adversarial neural networks namely the Generator and the Discriminator which are the competing elements.

The performance of the WGAN is susceptible to a particular type of failure know as ‘mode collapse’ wherein the Generator is caught in the cyclic generation of only a few ‘plausible’ outputs. Research has suggested that if the Discriminator is too good, then Generator training can fail due to vanishing gradients. In effect, an optimal Discriminator

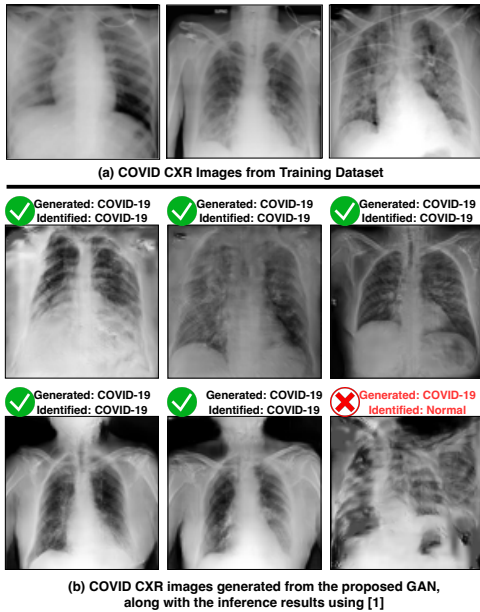


Fig. 3. Examples of (a) actual images in training dataset; (b) COVID CXR images generated by proposed WGAN, and the inference results using the earlier proposed lightweight COVID detection model in [1]

Fig. 4 presents the Accuracy plots for the training and testing phases of the model on the augmented(WGAN) dataset, which depict that after 50 epoch in the training cycle, the lightweight model of [1] witnesses high training and validation accuracy values which are very close to each other.

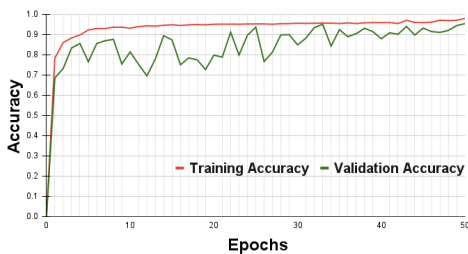


Fig. 4. Accuracy plots for WGAN images on the model in [1]

Table IV contains a comparison of the performance of the Non-GAN based Augmented data and GAN based Augmented data on our lightweight CNN model [1]. It can be observed that the parameters like accuracy, precision, specificity at sensitivity and AUC score has increased from 93.04% to 95.34%, 92.81% to 94.2%, 99.82% to 99.91% and 98.38% to 99.40% respectively. This shows that our WGAN based generated augmented image dataset obtained better results than the Non-GAN based image dataset. Table II presents a comparison of the performance of the proposed WGAN model with similar and recent counterparts.

V. CONCLUSION

Recognizing the need for augmentation for medical image datasets, this paper presented a WGAN-based COVID-19

TABLE I
COMPARISON OF THE PERFORMANCE OF THE NON-GAN AND WGAN BASED AUGMENTED DATA ON [1]

Dataset	Non-GAN based Augmented Data	WGAN based Augmented Data
Class	COVID-19 Normal Lung-Opacity Viral Pneumonia	COVID-19 GAN Normal Lung-Opacity Viral Pneumonia
Accuracy (%)	93.04	95.34
Precision (%)	92.81	94.2
Specificity at Sensitivity (%)	99.82	99.91
AUC Score (%)	98.38	99.4

TABLE II
COMPARISON OF PERFORMANCE OF WGAN MODEL WITH ITS COUNTERPARTS

Model	Accuracy (%)	Specificity (%)
WGAN	95.34	99.91
DCGAN [11]	80.0	81.0
IAGAN [11]	80.0	84.0
CovidGAN [12]	95.0	97.0

image generator for use in training of high-performance DL models for COVID-19 detection. The quality of the generated images was assessed using the generated images for inference using a recently proposed lightweight COVID-19 detection model, and the pre-trained model (on the original dataset) performed exceptionally well on the generated images as well – thereby leading credibility to the generated images.

REFERENCES

- [1] B. Z. Hussain, I. Andleeb, M. S. Ansari, and N. Kanwal, "Lightweight deep learning model for automated covid-19 diagnosis from cxr images," in *2021 IEEE Int. Conf. on Computing (ICOCO)*, 2021, pp. 218–223.
- [2] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, "A survey of modern deep learning based object detection models," *Digital Signal Processing*, p. 103514, 2022.
- [3] M. K. . Y. K. Liqa A. Rousan, Eyhab Elobeid, "Chest x-ray findings and temporal lung changes in patients with COVID-19 pneumonia," *BMC Pulmonary Medicine*, 09 2020.
- [4] V. Sandfort, K. Yan, P. Pickhardt, and R. Summers, "Data augmentation using generative adversarial networks (cyclegan) to improve generalizability in ct segmentation tasks," *Scientific Reports*, vol. 9, 11 2019.
- [5] S. Sundaram and N. Hulkund, "Gan-based data augmentation for chest x-ray classification," 2021.
- [6] S. C. Alec Radford, Luke Metz, "Unsupervised representation learning with deep convolutional generative adversarial networks," 11 2015.
- [7] M. B. Moran, M. D. Faria, G. A. Giraldi, L. F. Bastos, and A. Conci, "Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs," *Computers in biology and medicine*, vol. 129, 2021.
- [8] K. Lei, M. Mardani, J. Pauly, and S. Vasawanala, "Wasserstein gans for mr imaging: from paired to unpaired training," 10 2019.
- [9] G. H. Yann LeCun, Y. Bengio, "Deep learning," *Nature*, 2015.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [11] S. Motamed, P. Rogalla, and F. Khalvati, "Data augmentation using generative adversarial networks (gans) for gan-based detection of pneumonia and covid-19 in chest x-ray images," *Informatics in Medicine Unlocked*, vol. 27, p. 100779, 2021.
- [12] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. Pinheiro, "Covidgan: Data augmentation using auxiliary classifier gan for improved covid-19 detection," *IEEE Access*, vol. PP, 05 2020.