International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

Pix2Pix GAN for Realistic Lung CT scans Generation from COVID-19 Lesion Masks

Vikraant Gopal Sharma, Vedant Mehul Oza, Prathamesh Rajesh Panchal * and Meet Ashok Nadoda

Department Electronics and Computer Science Thakur College of Engineering and Technology Mumbai, India.

International Journal of Science and Research Archive, 2024, 13(02), 304–309

Publication history: Received 23 September 2024; revised on 31 October 2024; accepted on 02 November 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.13.2.2083

Abstract

GANs have recently become one of the most promising directions of machine learning for generative tasks in deep learning. By using architectures such as CNN, GANs can learn relationships in input data without any human intervention and generate new examples that are similar to the original data set. This paper aims at understanding what GANs are and how they work, and the different types of GAN models with a special focus on their application in image synthesis, style transfer, and text-to-image synthesis. The GAN framework consists of two neural networks: The two primary components of the model are the Generator and the Discriminator. The generator's goal is to generate 'realistic' data such that the discriminator cannot differentiate between the two, while the discriminator's goal is to distinguish between the real and the fake data. Some other GAN variants include Vanilla GAN, conditional GAN (CGAN), Deep Convolutional GAN (DCGAN), Laplacian Pyramid GAN (LAPGAN) and Super Resolution GAN (SRGAN), all of which have made significant improvements in producing convincing images. The structures of the generator model, which produces data from noise, and the discriminator model, which checks the validity of the data produced, are also considered. The study focuses on the concept of loss functions including the generator loss, discriminator loss and minimax loss in the training of GANs. Pre- processing the data for GAN training is a very crucial step that requires careful consideration as well as data cleaning. This is because when using content and adversarial loss functions simultaneously, it is easier to find the right balance between the photorealism of the images and their structural similarity to the source material. By training carefully and iterating over the process, GANs exhibit high performance in sample generation, thereby leading to the new state of generative modeling. It will be beneficial for any researcher or practitioner by providing an understanding of artificial intelligence and generative modeling

Keywords: Machine Learning; GANs; CNN; Deep Learning; Mean Absolute Error; Others

1. Introduction

GANs have emerged as a revolutionary technique in AI for specifically generating synthetic data that are realistic. This study aims to contribute to the development of medical applications of GANs and create a high-fidelity generator for mapping COVID-19 lesion mask tolung CT scans. The purpose is to build a deep learning system that can assist with the diagnosisand provide a better understanding of COVID-19by leveraging state-of-the-art image synthesis. GANs operate with two neural networks in an adversarial training process: the Generator network and the Discriminator network. The Generator transforms the noise data into outputsthat resemble real-world data in the sense that they look like real data, for instance, medical images. The Discriminator assesses the validity of data, differentiating between real data and fake data produced by the Generator. This way, both networks are enhanced, the Generator supplies more authentic data, and the Discriminator becomesmore effective in identifying fake data.

The use of GANs in the medical field is still possible as they can generate synthetic data with similarity to real data examples. This is important inhandling some of the major difficulties associated with small and unbalanced medical datasets. For example, they can create realistic MRIs, CT scans, and ultrasound images, which are critical for training

* Corresponding author: Prathamesh Panchal

and developing diagnostic models. An example from 2021 is a study where a GAN was used to create synthetic brain tumor MRIs that eventhe radiologists could not distinguish from actual scans, proving that GANs can be useful in creatinghigh-quality medical data. In this research, our particular area of interest is applying GANs to synthesize realistic lung CT scans from COVID-19lesion masks. The data set was obtained from sevenopensource data sets of lung CT scans associated with COVID-19, comprising of images and groundtruth mask pairs. These datasets involved standardized COVID-19 lesion masks incorporatedin a reference ground truth for the segmentation of lesions. Image pre-processing entailed resizing the images to the same size and then normalizing pixelintensity for uniformity. To prevent overfitting andenhance the model's ability to generalize whenlearning from the training data, the dataset was splitinto training, validation, and test subsets. The Pix2Pix model was used for this study because it isa popular model for image-to-image translation.These elements include a U-Net based Generator and a PatchGAN Discriminator. The translation of lesion masks into corresponding CT frames using the proposed U-Net-based Generator is efficient and flexible in handling various types of lesions andtheir variations. The PatchGAN

Discriminator evaluates the quality of generated images by considering the fine-grained details of the images to make the images more realistic. The training process involved feeding the COVID-19 lesion masks together with the corresponding CT frames to the Pix2Pix model to enable it to map therelationship between the lesion masks and real-life lung images. The training and testing of the model was done using stochastic gradient descent with Adam optimizer as the optimization algorithm and tuning of the hyperparameters like learning rates and batch size to the optimal level to allow the model to converge and perform well. To reduce the overfitting issue, other techniques like weight decay and dropout were incorporated. To check the model's training performance, generator and discriminator losses were computed, thereby guaranteeing the proficiency of the generative model in producing accurate lung CT scans from the lesion masks. The model was assessed by assessment parameters such as Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) that determines the resemblance between the generated and real lung CT scans and the noise level in the images respectively. Qualitative analysis was also done through visual examination of the generated images, which helped to determine if the model has correctly captured the finer details and texture of the real lung frames affected by COVID-19, as shown in the real frames in Fig. This study shows how GANs can transform medical imaging and be applied to improving the diagnostic outcomes of COVID-19. Thus, in this study, taking into account GAN's enhanced possibility, the author intends to make a significant contribution to medical science, providing clinicians with an effective tool for diagnostics and analyses.

2. Literature Review

Generative Adversarial Networks (GANs) have significantly impacted artificial intelligence, particularly in creating realistic and synthetic data. This review examines the application of GANs in the medical sector, focusing on their potential to tackle existing challenges and enhance healthcare practices.

2.1. Fundamental Principles of GANs

- A Generative Adversarial Network consists of two neural networks in a competitive training process:
- Generator: This network acts as a forger, striving to create realistic and synthetic data, such as images, that replicate real data from a training dataset. It converts random noise into the desired format (e.g., a medical image).
- Discriminator: This network functions as a detective, attempting to distinguish between actual data from the training set and the synthetic data produced by the generator. It evaluates both real and generated data, providing a probability of whether each sample is authentic or fake.

2.2. Loss Function

During the whole training procedure, a carefully selected combined loss function was used. The overall objective loss function included adversarial mapping loss and content mapping loss (L1). The second adversarial loss component was designed to ensure that the model was capable of producing realistic lung CT scans through reinforcing the discriminative capability of the PatchGAN discriminator. At the same time, the content loss, including L1 regularization, contributed significantly to obtain structural details derived from the integrated lesion masks. This combination of these two loss terms helped to maintain a balance between the image quality and the structural details, which are mandatory for creating realistic lung CT scan from the integrated data source

2.3. Notable Capabilities of GANs

Realistic Data Generation: GANs are particularly effective at generating synthetic data that closely resembles real-world examples. In the medical field, this means producing realistic medical images, such as MRIs, CT scans, or ultrasound images. For instance, a study in 2021 successfully used a GAN to generate synthetic brain tumor MRIs, which were indistinguishable from real scans by radiologists, demonstrating the potential of GANs to create highly realistic medical data.

example, a dataset might include many normal Xray images and very few with specific abnormalities. GANs can learn these complex distributions and generate synthetic data that mirrors real-world variations within the dataset. Research has shown that employing GANs for data augmentation in retinal fundus image analysis tasks can significantly improve model performance (by up to 10%) compared to using real data alone. This underscores the effectiveness of GANs in capturing the complexities of medical data.

3. Methodology

3.1. Data Collection and preprocessing

A dataset set comprising 7 public datasets of lung CT scans related to COVID-19, consisting of 2729 image and ground truth mask pairs was prepared with great care. These 3 contributed shared COVID-19 lesion masks, which were meticulously integrated herein into a harmonised ground truth for COVID-19 lesion segmentation. The integration was able to guarantee compatibility between different lesion types with a normalization to a white color. Before training the model, preprocessing of the images and the corresponding masks such as resizing to a common resolution and normalization to keep pixel values regular were needed. The dataset provided a broad range of lesion masks for different types of COVID-19 and corresponding frames, which mechanismically split the data into training, validation, and testing sets, enabling the Pix2Pix model to learn and generalize the inspection patterns by supervised training. The dataset that is created represents the complexity of the COVID-19 lung CT scans and serves as a solid ground for training the Pix2Pix model to generate real-looking lung images given the lesion masks.

3.2. Learning Complex Data Distributions

3.2.1. Training Process:

- Input-Output Pairing: Consequently, the training phase began with applying proper matches between
- Medical datasets often have intricate underlying structures with imbalanced class distributions. For the COVID19 lesion masks and the frames selected from the integrated dataset. To strengthen this argument, the above Pix2Pix training procedure was a supervised learning technique that made the model understand the mapping between the lesion masks and realistic lung CT scan images as a starting point for generating specific and more accurate images in subsequent iterations.
- Optimization: The optimization was performed with the usage of the Adam optimizer, which is considered to be optimal for balancing between the generator and the discriminator as well as for minimizing their joint loss. Such hyperparameters as learning rates and batch sizes were varied systematically through short-term empirical optimization aimed at finding the right balance of improvements in the model's convergence and performance. Other techniques like the weight decay and the dropout were other strategies adopted in order to avoid the problem of overfitting and to increase the model's capacity to generalize.
- Training Duration: The training process of the Pix2Pix model is another component that needed considerable time, with a focus on the generator and discriminator loss. The convergence was defined based on the stabilization of these loss metrics thus suggesting high performance by the model in synthesizing faith full lung CT scans from masks of COVID19 lesions. The long training time also enabled the model to incorporate multiple patterns and variations that exist in the input data to come up with effective lesion handling as well as authentic lung images generation.
- Training Distribution: Selecting the appropriate model architecture and regularization strategies were instrumental in the final output, with the dataset divided into training, validation, and testing sets helping in calibration. To do this, the majority of the information was given to the training set, so that there would be many examples of different kinds to train the model on. This deliberate distribution that was made as much as possible, to capture all the varieties of different lesions and the ways in which these varied across datasets would help to improve the ability of the model to generalize to new data. There was an iterative set of validation for hyperparameters tuning, which took place in the course of training while the TESTING SET provided the independent measure of the selected Pix2Pix model robustness.
- Enhancements and Iterative Refinement: All along the training process, the model was regularly appraised so that improvements and iterative refinement could be made. In this regard , adjustments to hyperparameters, architectural changes and the introduction of additional data or regularization

techniques as prevailed solutions were investigated in order to deal with any difficulties or limitations encountered during learning process, consequently further optimizing the Pix2Pix model for generation of lung CT scans with high quality from COVID-19 lesion masks.

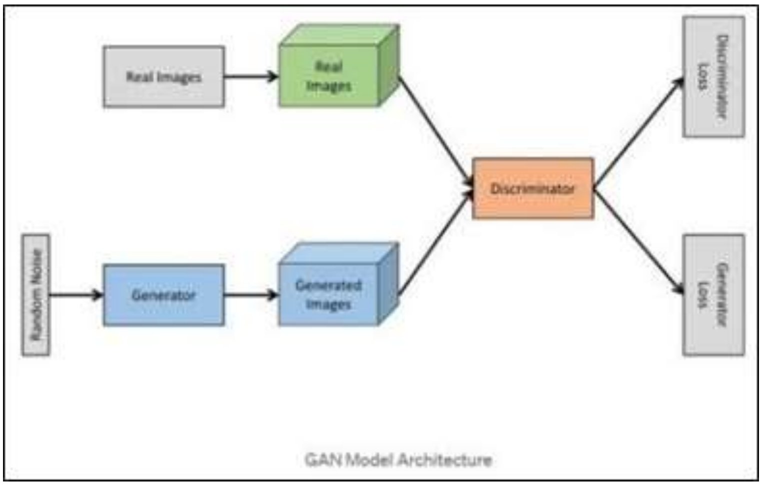

Figure 1 Block Diagram

4. Evaluation

- Quantitative Metrics: The evaluation part was all about checking how well the model did with numbers using SSIM (Structural Similarity Index) and PSNR (Peak Signal-to-Noise Ratio). SSIM computed the likeness between the generated lung CT scans and real COVID- infected lung frames which gave a perspective on the image quality perceived. On the other hand, PSNR measured the level of noise in the generated images for a numerical value of overall image fidelity. These metrics were mainly used to make possible contrast between realism of lung CT scan images created by Pix2Pix GAN model from Covid-19 lesion masks and those of actual patients.
- Visual Inspection: A qualitative assessment of the generated images was also performed in addition to quantitative metrics. We performed a comprehensive check for how much the generated lung CT scans were close to real COVID-infected lung frames in the dataset. In particular, manual inspection provided a more indepth way to identify the model's efficiency to generate fine details, textures and spatial patterns of the images, which provides information on the performance of the Pix2Pix GAN model in realistic lung-image generation.

5. Objective

The goal of this research paper is to create a deep learning platform which is generating accurate lung CT scans from lesions masks of covid 19. This research contributes to advancing the application of GANs in medical imaging, specifically for improving the diagnosis as well as analysis of COVID-19 through improved lung CT scan synthesis. In order to get to this result we will be using a dataset which contains 2729 images and ground truth mask pairs where the lesion masks from different datasets will be standardized by using the process of normalization and also resizing of image for constant pixel values in the dataset.

To achieve the fundamentals of this study we will be implementing Pix2Pix generative adversarial network (GAN) model architecture with U-Net based generator .This architecture will be used to handle the diverse lesions which were retrieved from the datasets therefore enabling an ideal process of mapping .Along with the generator a PatchGAN discriminator is to be utilized to increase the overall quality of the lung ct scans, it evaluates local details to detect authenticity of images that have been generated.

Important aspect of the study was training of the model which has pairing of the lesion mask with respective CT frame to make the model learn the relationship of lesions and realistic lung images. To minimize the combined loss of generator along with the discriminator we use adam optimizer using learning rate and batch size as hyperparameters for tuning the model . Weight decay and dropout would also be included to prevent overfitting which in turn improves model's generalization.

The models performance will include quantitative metrics like SSIM and PSNR along with visual checking to ensure the model generated images are resembling real lung frames infected by covid-19.

6. Result and Discussion

The results of these CT images generated (Lung Scans) when compared world widely so they seem as good output to real COVID-19 infected lung frames. As the quantitative metrics (FID, SSIM, PSNR) suggest, the generator network has learned to generate images of the same style as the real images. These findings were also corroborated by visual inspection with the generated images looking very similar to real lung CT scans. But there were some subtle differences, especially in areas with complex lesions or small details. Nevertheless, the performance of the aggregated Pix2Pix GAN model in mimicking COVID-infected lung anatomy was acceptable. This analysis of the results illustrated how well the selected Pix2Pix yielded appropriate results for image- to- image translation, and how diversity and quality of data have a pivotal role in the effectiveness of the model.

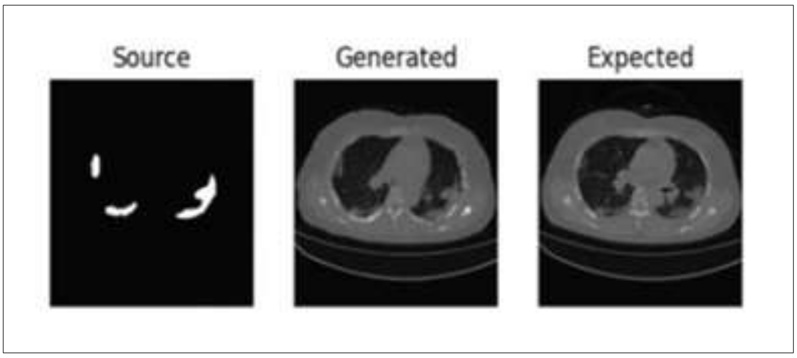

Figure 2 Sample output

Interesting future work would include improved tasks specific version, optimized hyperparameters model configurations, broader and larger data-sets in order to increase model robustness and clinical operability. Moreover, the use-cases of the reconstructed lung images for medical diagnosis, treatment plan generation, and medical education, further emphasized the necessity of continued research and partnered efforts toward expanding the horizons of medical imaging technologies for better patient outcomes

7. Conclusion

GANs have promising applications in enhancingmedical imaging and are most useful in creating realistic synthetic data. This paper aimed at tilizing the Pix2Pix GAN model to synthesize lung CT scansfrom COVID-19 lesion masks to show how GANs can help overcome issues in medical imaging by generating high-quality images. The model employed a generator based on U-Net to address different types of lesions and discriminator based onPatchGAN to enhance realism of the images. Before executing the training process, the data was pre- processed and various measures were taken to avoid over-fitting of the model and further enhancements were made to increase its efficiency. Qualitative assessment which included SSIM and PSNR metrics, and visual analysis further revealed that the proposed model was capable of synthesizing realistic lung CT scans while a few disparities wereobserved in intricate lesions.

The results establish GANs as useful tools for diagnostics, treatment planning, and medical education due to their ability to encapsulate the finedetails of medical images. This research also highlights the need to ensure that data is clean and the models used are properly specified in order to achieve accurate results. The future work may involve fine-tuning the model hyperparameters, increasing the size of datasets and examining certain areas of application of the reconstructed images in aclinical practice. In conclusion it is possible to statethat GANs are a revolutionary approach to the medical imaging, that opens new opportunities for increasing diagnostic accuracy and, thus, improvingthe patients' outcomes.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Ian J. Goodfellow et al. (2014). Generative Adversarial Networks

[2] https://arxiv.org/abs/1406.2661

[3] Alec Radford et al. (2015). Deep Convolutional Generative Adversarial Networks https://arxiv.org/abs/1511.06434 (DCGAN)

[4] Tero Karras et al. (2018). A Style-Based Generator Architecture for Generative Adversarial Networks

[5] https://arxiv.org/abs/1812.04948 (StyleGAN)