GAN-enabled Chest X-ray Image Synthesis for Medical Imaging

Sakshi Kharat, Dr. Bhakti Palkar, Mansi Kambli

Department of Computer Engineering, K. J. Somaiya School of Engineering,
Mumbai-400086, Maharashtra, India

Abstract: This study leverages Generative Adversarial Networks (GANs) to generate high-quality synthetic chest X-ray images, critical for advancing disease diagnosis, treatment planning, and patient monitoring. The literature underscores GANs' significant role in augmenting datasets, training deep learning models, and enhancing the interpretability of medical imaging. The GAN architecture comprises a generator that transforms random noise into realistic images and a discriminator that differentiates between real and synthetic images. Our research addresses the challenges and opportunities of deploying GANs in medical imaging, with a focus on robust training methodologies and ethical considerations. After 15 training epochs, our GAN model achieved a generator loss of 5.4734, a discriminator loss of 0.0360, an SSIM of 0.1494, and an MSE of 0.1807. These metrics demonstrate the model's capability to produce realistic synthetic images, with a precision rate of 99.27% and a recall rate of 49.82%, indicating the potential of GANs to enhance medical imaging datasets for clinical applications.

Keywords: Generative Adversarial Networks (GANs), Generator, Discriminator, Medical Imaging

1. Introduction

Medical imaging plays an important role in modern healthcare, offering critical insights for disease diagnosis, surgical intervention guidance, and treatment outcome monitoring. However, conventional methods for generating medical imaging data, such as traditional GAN models, often fall short in producing high-resolution, clinically relevant images, particularly in specialized areas like pediatric radiology. These limitations are compounded by challenges like a lack of annotated datasets, privacy concerns, and the need for diverse images to represent various medical conditions. Generative Adversarial Networks (GANs) represent a groundbreaking approach in artificial intelligence, capable of generating realistic and diverse data. GANs consist of two components: a generator that converts random noise into realistic images and a discriminator that distinguishes between real and synthetic images. This dual network structure allows for the generation of high-fidelity synthetic data that can augment existing medical image datasets, thereby improving the training and validity of machine learning models.

This study focuses on using an advanced GAN architecture specifically designed to create high-quality synthetic images of chest X-rays, with a particular emphasis on pediatric cases. Unlike traditional GAN models, our approach incorporates several innovative features that enhance both image quality and model performance. We utilize advanced upsampling techniques through Conv2DTranspose layers, allowing for the generation of high-resolution images that capture critical details often seen in pediatric radiology. Additionally, our training regimen includes progressive resizing and an adaptive learning rate, which contribute to the stability and efficiency of the training process.

Traditional GANs often struggle with generating high-resolution images that accurately capture the nuances of medical data, particularly in specialized fields like pediatric radiology. Our advanced architecture not only addresses these challenges but also ensures that the generated images maintain the variability and detail necessary for effective clinical use. Our model is particularly beneficial for pediatric patients, as it aims to address the scarcity of annotated chest X-ray images in this demographic, which can hinder accurate diagnosis and treatment planning. By generating synthetic images that closely mimic the variability and characteristics of real pediatric chest X-rays, our GAN not only improves the availability of training data for machine learning models but also enhances the diagnostic capability of medical practitioners. This research ultimately aims to advance healthcare technology by addressing data scarcity and enhancing dataset diversity, supporting more effective diagnostic processes and treatment planning for pediatric patients.

The remainder of this paper is organized as follows: Section 2 discusses related work and background on GANs in medical imaging. Section 3 details the methodology, including dataset preparation, model architecture, training procedures, and evaluation metrics. Section 4 presents the results. Finally, Section 5 concludes the study and outlines potential directions for future research

2. Related Work

The survey focuses on research involving Generative Adversarial Networks (GANs) in medical image generation and data augmentation, emphasizing the importance of realistic medical image generation for diagnostic reliability and training.

In one such study [1], the authors successfully used GANs to generate 1000 synthetic COVID-19 chest X-rays with low Fréchet Inception Distance (FID) scores, indicating high fidelity to real data. Their findings demonstrate GANs' ability to successfully augment training datasets without introducing considerable noise. Ng and Hargreaves also compared the performance of Deep Convolutional GANs (DCGANs) and Wasserstein GANs with Gradient Penalty (WGAN-GP), noting differences in stability and training efficiency across epochs. The authors [2] investigated the use of Generative Adversarial Networks (GANs) for data augmentation to improve the efficacy of Convolutional Neural Networks (CNNs) in diagnosing lung diseases from chest radiographs, especially when there are class imbalances in the dataset. The study trains DenseNet-121 on the Stanford CheXpert dataset using three augmentation techniques: GANs-based, standard, and no augmentation, across varying dataset sizes (1%, 10%, 50%, 100%).

The foundational paper [3] introduces a dual-model framework where a generative model (G) mimics real data, guided by a discriminative model (D) that distinguishes between real and generated samples. This approach, using backpropagation without complex inference techniques, demonstrates effective sample generation in experiments.

In their studies, researchers investigated GANs for the generation of realistic medical images and annotations. The [4] study examined DCGAN, LAPGAN, pix2pix, CycleGAN, and UNIT, demonstrating their application in image reconstruction and synthesis. The [5] study evaluated multiple GAN architectures for synthesizing medical images spanning cardiac cine-MRI, liver CT, and RGB retina modalities, evaluating model effectiveness using FID scores and segmentation accuracy, and highlighting differences in performance between architectures. In their study [6], researchers employed GANs to synthesize MRI scans of brain tumors, combining DCGAN and WGAN variants with style transfer techniques to increase realism. Results from datasets such as BraTS 2020 revealed high SSIM scores of up to 0.83, addressing data scarcity and class imbalance in medical imaging. Another work [11] examined synthetic brain MR image generation, emphasizing GANs' function in augmenting datasets for deep learning model training.

In the study [8], researchers highlighted the use of GANs for data augmentation and anonymization, emphasizing its ability to increase the quantity of training datasets while maintaining patient privacy. This approach allows for the sharing of medical data while preserving sensitive information. In their study [9], researchers addressed the significant challenges associated with data sharing in medical imaging by using GANs to generate synthetic medical images. The thorough evaluation conducted in their study indicated that these synthetic images are of high quality and can effectively replace real medical images for training purposes. This approach not only preserves patient privacy, but also facilitates wider sharing and collaboration among researchers and institutions, thereby overcoming traditional barriers to data sharing in the healthcare sector.

In their review [10], researchers examined various GAN architectures and their applications in medical image synthesis, covering models like DCGANs, CycleGANs, and conditional GANs. They highlighted advancements in training techniques, such as adversarial loss functions and multi-scale discriminators, which improved the realism and utility of generated images. The review also explored applications in radiology, oncology, and neurology, demonstrating GANs' broad impact on diagnostic accuracy and medical research.

In their study [18], researchers proposed an innovative data augmentation pipeline that leverages GANs to create synthetic labeled datasets of 3D echocardiography images. This study showcased the versatility of GANs in generating high-quality 3D medical images, which are crucial for detailed anatomical and functional analysis. The pipeline incorporated advanced GAN architectures capable of capturing the complex structures of echocardiographic images and generating realistic synthetic counterparts.

Research [15] demonstrated that GANs can effectively synthesize realistic medical visuals that closely resemble real patient data. This approach augments small datasets, enhancing diagnostic model training and performance. The study emphasized GANs' potential to improve medical imaging, particularly in domains where data scarcity is a significant challenge. Similarly, research in neuroimaging [16] demonstrated GANs' ability to generate high-fidelity synthetic brain images, which is required for brain mapping and neurological research. These synthetic images can overcome data acquisition limitations, allowing for more thorough studies and advancing neuroimaging research. Research [7] and [17] showcased the use of dual GANs and advanced architectures for generating synthetic medical images, improving the quality and applicability across different imaging modalities. Another study [14] focused on generating realistic X-ray images, crucial for enhancing radiographic training datasets.

Research [19] introduced SRTransGAN, combining GANs with transformers for high-quality image super-resolution by leveraging transformers' ability to capture long-range dependencies and GANs' generative power. Another study [20] extended GAN applications to 3D imaging, focusing on generating realistic 3D bone images. These images preserved intricate bone details, proving useful for orthopedic research, surgical planning, and educational purposes. Research [24] focused on enhancing MRI image resolution using a perceptual conditional GAN (cGAN), improving visual quality and perceptual relevance through specialized loss functions. Similarly, [23] presented a fine-tuned GAN model for medical image super-resolution, emphasizing tailored training for enhancing resolution and quality across MRI and CT scans, benefiting medical diagnostic capabilities.

Research [21] explored combining image data with clinical information to generate comprehensive synthetic datasets. Research [22] provided insights into the practical applications and limitations of GAN-generated medical images. Research [12] and [13] highlighted specialized GAN applications for generating specific medical imaging modalities, expanding the utility of GANs in medical image synthesis. Additionally, [25] focused on generating synthetic 3-D PET images using TGANs, crucial for oncological studies by incorporating realistic tumor masks. These applications demonstrate GANs' potential to enhance diagnostic processes and advance medical research in imaging.

Despite the advancements made by traditional GANs and data augmentation techniques in medical image synthesis, they encounter significant challenges in generating high-resolution pediatric chest X-rays. These methods often struggle to preserve the fine details critical for accurate diagnosis, particularly in the context of small datasets and class imbalances, which can result in training instability and the production of low-quality images.

Our research directly addresses these challenges by concentrating on pediatric radiology and employing advanced upsampling techniques alongside a customized training regimen. This approach not only enhances existing datasets but also ensures that the generated images adhere to the rigorous standards necessary for clinical application in pediatric radiology. By implementing a progressive training methodology and fine-tuning model architecture for stability, our method significantly improves image fidelity, providing a robust tool to enhance diagnostic accuracy in this specialized domain.

3. Methodology

3.1 Image Dataset

In this study, 138 chest X-rays from pediatric patients aged one to five years at Guangzhou Women and Children's Medical Center were utilized. Each X-ray was converted to grayscale, resized to 128x128 pixels, and normalized to the range [-1, 1]. The images were exported in JPEG format and were part of routine clinical care.

This dataset is sourced from the following citation: Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2

Examples of the chest X-ray images used in the study are shown below:

Figure. 1 These images show anterior-posterior views of pediatric patients' chest X-rays, resized to 128x128 pixels

3.2 Software and Hardware

The study employed Python as the primary programming language, utilizing TensorFlow and Keras libraries for implementing the Generative Adversarial Network (GAN) model. NumPy was used for efficient numerical computations, and OpenCV handled image processing tasks. The T4 GPU accelerated the training process significantly by offloading intensive computations. Google Colab's environment facilitated seamless development and execution of the GAN model.

3.3 Generator Network

In our research study, we employed a generator to translate random noise vectors into high-fidelity, 128x128 pixel Chest X-ray images. This pivotal generator architecture comprises several essential layers meticulously designed for optimal performance. Initially, a dense layer transforms the input noise vector into a higher-dimensional space, yielding an output shape that is further processed by a reshape layer into a 3D tensor with dimensions 32x32x256. Subsequently, two Conv2DTranspose layers play a crucial role in upsampling the tensor, each producing output with shapes 64x64x128 and 128x128x128, respectively. Both Conv2DTranspose layers employ Leaky ReLU activation and utilize a 4x4 kernel size to achieve their upsampling functionality. Lastly, a Conv2D layer generates the final image, outputting a tensor of shape 128x128x1 with a 4x4 kernel size and utilizing a Tanh activation function to scale pixel values to the range [-1, 1]. This meticulously crafted architecture underscores our commitment to producing realistic chest X-ray images for advanced medical research and analysis

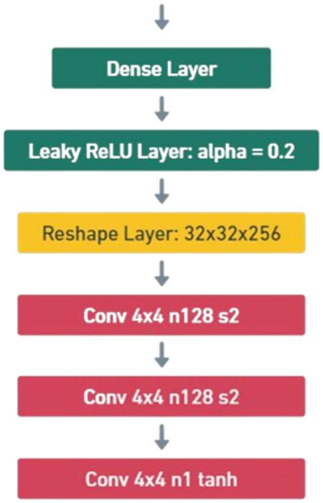

Figure. 2 Leaky ReLU-based Generator Architecture with Conv2DTranspose Upsampling

3.4 Discriminator Network

Our discriminator architecture is tailored for the task of discriminating between real and generated images in a GAN. It starts with an input layer matching the image dimensions, followed by a series of convolutional layers. The initial layer consists of 64 filters with a 3x3 kernel size and 'same' padding, preserving spatial information. Each convolutional layer is followed by a LeakyReLU activation function with a slope of 0.2, which helps the model learn complex patterns by introducing non-linearity. Subsequent layers follow a similar structure, with 128 filters in the next two convolutional layers. These layers use a 3x3 kernel size and a stride of 2 for downsampling, allowing the model to capture more abstract features as the spatial dimensions decrease. The final convolutional layer employs 256 filters with a 3x3 kernel size and a stride of 2. After the convolutional layers, the feature maps are flattened into a 1D vector and passed through a dropout layer with a rate of 0.4 to prevent overfitting. Finally, a dense layer with a single neuron and a sigmoid activation function produces the final output, representing the probability of the input image being real. This architecture is designed to effectively learn discriminative features from the input images, enabling the discriminator to make accurate distinctions between real and generated samples during the adversarial training process.

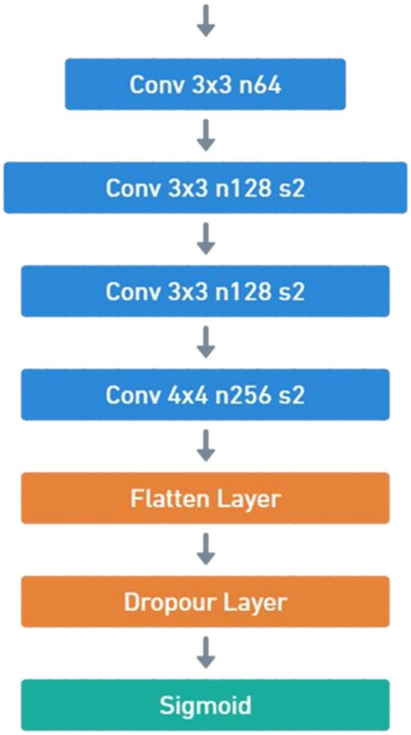

Figure. 3 Detailed Flowchart of the Discriminator Architecture in GAN

3. 5 Training Procedure

The training approach for our GAN model alternates between training the discriminator and the generator to provide balanced learning. Initially, real chest X-ray images were normalized to the range [-1, 1], which aligned with the generator's output scale. For each training step, random noise vectors of dimension 100 were fed into the generator, allowing it to generate synthetic X-rays. This process improves the generator's ability to create diverse and high-quality synthetic images. During training, the discriminator is trained using a combination of real and fake images. A batch of real images is randomly selected from the dataset, and an equal number of fake images is generated by the generator. The discriminator is trained to distinguish between real and fake images by minimizing the binary cross-entropy loss. Subsequently, the generator is trained to maximize the probability of the discriminator misclassifying fake images as real by updating its weights to minimize the binary cross-entropy loss while the discriminator's weights are frozen.

For evaluating the quality of the generated images, Structural Similarity Index (SSIM) and Mean Squared Error (MSE) were calculated for each batch of images and averaged across epochs. SSIM compares the similarity of generated and real images, with values closer to 1 indicating higher resemblance, whereas MSE measures the average squared difference, with lower values indicating higher quality. These metrics provided a comprehensive evaluation of the generated images quality. The GAN model was trained for 15 epochs, each with 4000 steps, and both the discriminator and generator were iteratively improved. At the end of each epoch, the models' performance was evaluated using the SSIM and MSE metrics. Furthermore, the model weights were saved at specific intervals for checkpointing and restoration. This extensive training resulted in a generator capable of producing synthetic X-rays that are quite similar to real ones, as indicated by high SSIM and low MSE values.

3.5.1 Theoretical Justification

The theoretical foundation of our approach lies in adversarial training, where the generator G and discriminator D are optimized simultaneously through a minimax game. The generator's objective is to minimize its loss L_G, which is defined as the negative log probability of the discriminator classifying generated images as real:

$$\mathcal{L}_{\mathcal{G}} = -E_{z\sim p_z(z)}\left[\log D\left(G(z)\right)\right] \quad (1)$$

Here, z is sampled from a latent $p_z(z)$. This loss function encourages the generator to produce images that are indistinguishable from real data. Conversely, the discriminator aims to maximize its loss L_D, which combines the log likelihoods of classifying real and generated images:

$$L_D = -E_{x\sim pdata(x)}\left[\log\left(D(x)\right)\right] - E_{z\sim pz(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (2)$$

The theoretical principles behind adversarial training ensure that the generator and discriminator improve in tandem, with the generator creating increasingly realistic images and the discriminator enhancing its ability to distinguish between real and synthetic images. This approach has been proven effective in various generative tasks, particularly in image synthesis. Our empirical results, including a high Structural Similarity Index (SSIM) and a low Mean Squared Error (MSE), validate the theoretical soundness of our model.

3. 6 Evaluation Metrics

To assess the performance of our GAN model, we utilized several evaluation metrics

3.6.1 Generator Loss

The Generator Loss L_G quantifies the generator's effectiveness in producing realistic images that can deceive the discriminator. It is calculated using binary cross-entropy:

$$\mathcal{L}_{\mathcal{G}} = -E_{z\sim p_z(z)}\left[\log D\left(G(z)\right)\right] \quad (3)$$

A lower Generator Loss indicates that the generator is producing high-fidelity images. This metric is crucial for monitoring the quality of the generated images and guiding the generator's training process.

3.6.2 Discriminator Loss

The Discriminator Loss $L_D$measures the discriminator's ability to distinguish between real and generated images. It is also computed using binary cross-entropy:

$$(\mathcal{L}_{\mathrm{D}} = -\frac{1}{m}\sum_{i=1}^{m}\left[y_{\text{true}}^{(i)}\log\left(y_{\text{pred}}^{(i)}\right) + \left(1 - y_{\text{true}}^{(i)}\right)\log\left(1 - y_{\text{pred}}^{(i)}\right)\right]) \quad (4)$$

Here, m is the batch size, $y_{\text{true}}^{(i)}$ represents the true label of the $i - th$ image, and $y_{\text{pred}}^{(i)}$ is the discriminator's prediction. Lower values of Discriminator Loss indicate improved classification accuracy and better discriminator performance.

3.6.3 Structural Similarity Index (SSIM) and Mean Squared Error (MSE)

In our study, we use two fundamental metrics to evaluate generated image quality in Generative Adversarial Networks (GANs): the Structural Similarity Index (SSIM) and Mean Squared Error. SSIM provides a perceptual measure by quantifying the similarity of two X-rays in terms of contrast, luminance, and structure. We can define the SSIM as:

$$\mathrm{SSIM}\left(x, G(z)\right) = \frac{\left(2\mu_x\mu_{G(z)} + C_1\right)\left(2\sigma_{xG(z)} + C_2\right)}{\left(\mu_x^2 + \mu_{G(z)}^2 + C_1\right)\left(\sigma_x^2 + \sigma_{G(z)}^2 + C_2\right)} \quad (5)$$

Where, μx and $\mu_{G(z)}$ are the mean values, x, and $\sigma_{G(z)}$are the standard deviations, $\sigma_{xG(z)}$ is the covariance, and C_1 and C_2 are constants for stability. SSIM values closer to 1 indicate higher similarity between generated and real images.

MSE, on the other hand, measures the average squared difference between corresponding pixel intensities in generated and real images, resulting in an overall image fidelity measure. This can be computed as:

$$\mathrm{MSE}\left(x, G(z)\right) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - G(z)_i\right)^2 \quad (6)$$

Where n is the number of pixels. Lower MSE values reflect better image quality and fidelity.

3.6.4 Precision and Recall

In our study, precision and recall are critical metrics to evaluate the performance of our Generative Adversarial Network (GAN) in producing realistic chest X-rays. Precision is measured as the ratio of correctly generated X-rays to the total number of generated X-rays, demonstrating the reliability of the generated images. In contrast, recall measures the ratio of correctly generated X-rays to the total number of actual X-rays, indicating the GAN's ability to capture all relevant features. We used these measures additionally, to other evaluation metrics to comprehensively asses the quality of generated images.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (7)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (8)$$

3.6.5 Real Accuracy and Fake Accuracy

In this study, real accuracy measures how accurately the discriminator identifies real images, while fake accuracy assesses its ability to detect generated (fake) images. These metrics are essential for evaluating the discriminator's effectiveness in distinguishing between real and fake images, crucial for generating high-quality and realistic images in GANs.

$$\text{Real Acc} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (9)$$

$$\text{Fake Acc} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (10)$$

4. Result

4.1 Inference of Losses, Average SSIM, MSE, Precision, and Recall

Throughout the 15 epochs of training, the generator loss stabilized at 5.4734, indicating that it produced challenging outputs for the discriminator, as compared to 2.5321 in Epoch 1. The discriminator loss decreased from 0.1619 in Epoch 1 to 0.0360 in Epoch 15, indicating high classification accuracy.

Table 1. Evaluation Metrics and Performance per Epoch

Epoch	**Gen/ Disc**	**SSIM/ MSE**	**R / F**	**P / R**
1	2.53/0.16	0.14/0.19	0.76/0.95	0.95/0.55
2	5.12/0.10	0.15/0.18	0.93/0.99	0.99/0.51
3	2.39/0.63	0.15/0.18	0.98/0.94	0.94/0.49
4	6.20/0.05	0.15/0.17	0.99/0.99	0.99/0.50
5	6.77/0.12	0.15/0.17	1.00/0.99	0.99/0.49
6	4.75/0.01	0.15/0.17	1.00/1.00	1.00/0.50
7	8.76/0.008	0.15/0.18	1.00/1.00	1.00/0.50
8	5.98/0.008	0.15/0.17	0.97/1.00	1.00/0.50
9	6.82/0.009	0.15/0.17	0.99/0.95	0.95/0.49
10	7.67/0.20	0.15/0.17	1.00/0.99	0.99/0.49
11	6.9/0.0002	0.14/0.18	1.00/1.00	1.00/0.50
12	5.47/0.036	0.15/0.17	1.00/1.00	1.00/0.50
13	9.8/0.0022	0.15/0.17	1.00/0.99	0.99/0.49
14	11.9/0.002	0.15/0.17	1.00/0.99	0.99/0.49
15	5.47/0.036	0.15/0.18	1.00/0.99	0.99/0.49

In Epoch 15, the average Mean Squared Error was 0.1807, slightly increasing from 0.1780 in Epoch 5, indicating consistent pixel-wise error between generated and real images. The Average Structural Similarity Index improved to 0.1494 from 0.1409 in Epoch 1, showing enhanced perceptual quality in the generated images throughout training. In Epoch 15, the discriminator showed 99.28% precision and 49.82% recall, indicating high precision but moderate recall in classifications. This reflects a balance between confidence in positive predictions and correct identification of all positives, with precision improving from 95.65% and a slight recall decrease compared to Epoch 1 (55.46%).

By Epoch 15, the discriminator achieved 100% accuracy in identifying real images, a significant improvement from 76.81% in Epoch 1. Fake image accuracy reached 99.28% in Epoch 15, up from 95.65% in Epoch 1, showcasing the discriminator's effective ability to distinguish between real and generated images and marking substantial progress during training.

4.2 Observations on Image Quality Progression After Training

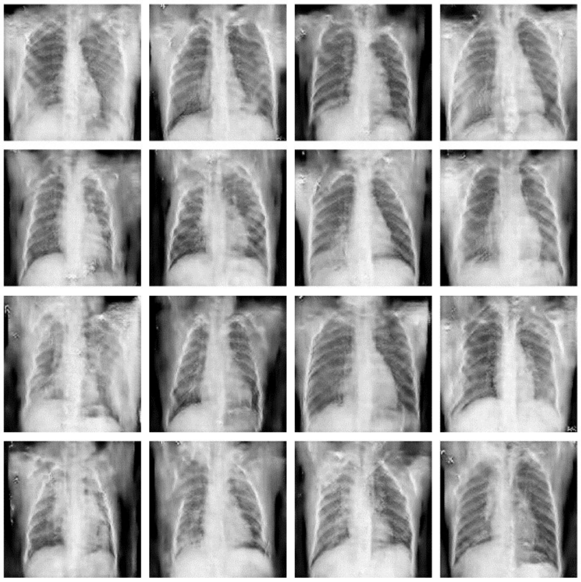

Figure. 4 Generated Chest X-ray Images after 1 Epoch

After training, we observed notable improvements in image quality. In Figure 4, representing Epoch 1, images varied in clarity and structure, often appearing blurred or noisy, with noticeable artifacts and less defined lung structures.

By Figure 5, depicting Epoch 15, images became more uniform and consistent, showing reduced artifacts and noise. These images closely resembled realistic chest X-rays, with improved perceptual quality and a more accurate depiction of the ribcage and lung fields.

In Figure 6, we observe clear trends in generator and discriminator performance over epochs. The generator loss decreases, indicating improved realism in generated data, while discriminator loss diminishes, reflecting better discrimination between real and fake data. The MSE plot trends downward, showing closer pixel value alignment between generated and real images, and the SSIM plot rises towards 1, indicating enhanced structural similarity.

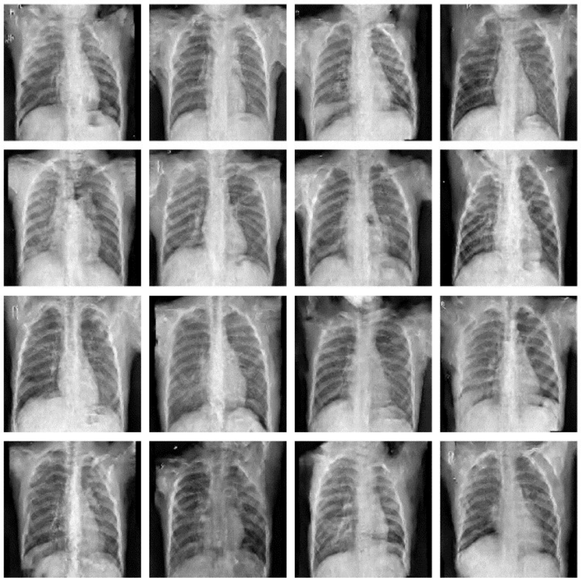

Figure. 5 Generated Chest X-ray Images after 15 Epoch

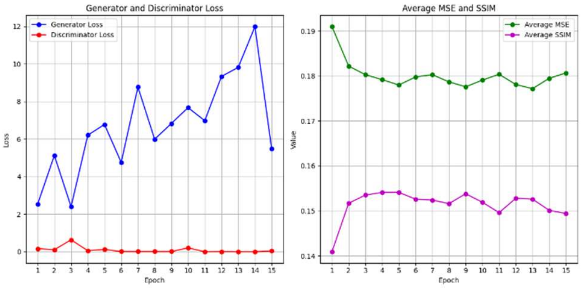

Figure. 6 Performance Metrics Evolution Over 15, which includes Generator and Discriminator Loss, MSE and SSIM

In Figure 7, the graph shows precision-recall and real-fake accuracy metrics across training epochs. Initially, precision is higher than recall, indicating early strength in identifying positives but missing some. As training progresses, recall improves, capturing more positives. For real and fake accuracy, the model initially better identifies real cases.

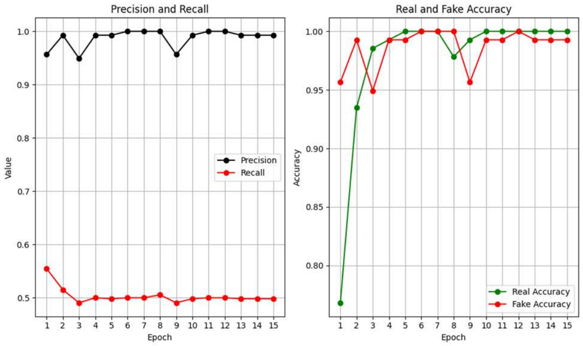

Figure. 7 Performance Metrics Evolution Over 15 Epochs, which includes Precision and Recall, Real Accuracy and Fake Accuracy

4.3 Comparison with State-of-the-Art Methods

To evaluate the effectiveness of our GAN-based approach for generating 2D chest X-ray images, we compare it with several state-of-the-art methods. Our approach prioritizes high image fidelity and noise reduction, focusing on metrics such as SSIM and MSE, which are crucial for diagnostic accuracy. Existing methods have used various GAN architectures and techniques. Some studies have employed 3D GANs for generating 3D medical images, focusing on segmentation and feature preservation. These approaches are resource-intensive and emphasize different aspects of image quality compared to our 2D-focused methodology.

Other research has developed hybrid synthetic datasets by combining medical images with non-image clinical data. While this method offers privacy benefits and potential for secondary analysis, it has shown slightly lower performance compared to private datasets, illustrating the impact of dataset size and structure on results.

Several advanced GAN models have been used for image super-resolution and perceptual quality, achieving high detail and visual quality. However, our approach stands out by emphasizing quantitative metrics to ensure structural integrity and suitability for clinical use, particularly in pediatric radiology.

Additional studies have used GANs with conditional discriminators to generate specific medical images, with promising results but limitations due to single dataset use. Comparisons of different GAN models for synthetic medical image generation have highlighted the importance of high-quality synthesis and privacy preservation, addressing challenges related to fidelity and feature consistency.

This comparative analysis demonstrates the effectiveness of our method in maintaining high image quality and addressing specific needs in medical imaging, thereby justifying its contribution to the field.

5. Future Work

The current study, based on a dataset of 138 X-ray images, lays the groundwork for further exploration in several areas. Experimenting with different batch sizes and extending the number of training epochs could potentially result in the generation of even more realistic lung images. Additionally, the integration of more sophisticated GAN architectures, such as StyleGAN or BigGAN, could significantly enhance the quality of the generated images. These advancements could contribute to the development of more effective diagnostic tools and a deeper understanding of lung pathologies.

6. Conclusion

The training of our GAN model for lung image generation has been a rigorous and resource-intensive process, demanding substantial computational power and meticulous tuning. Despite these challenges, our efforts have yielded promising results. Over the course of training, the model has demonstrated a notable advancement in generating increasingly realistic lung images. Through 4000 steps, using a dataset of 138 images over 15 epochs and spanning 12 hours, 10 minutes, and 58 seconds of training time, we observed significant improvements in discriminator and generator metrics. Looking forward, our focus will be on refining the model architecture and exploring more efficient training methodologies to further enhance the quality and accuracy of the generated images. These advancements hold significant promise for advancing diagnostic capabilities in pulmonary imaging and contributing to the broader field of medical image analysis.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, methodology, software, formal analysis, investigation, resources, data curation, writing original draft preparation, writing review and editing, and visualization were performed by the first author. Supervision, project administration, and funding acquisition were handled by the second author.

Acknowledgments

We would like to express our gratitude to the faculty and staff of KJ Somaiya College of Engineering, Mumbai, for providing the necessary resources and facilities and for supporting our work. Additionally, we thank Guangzhou Women and Children’s Medical Center for providing the data on pediatric patients aged one to five years, which was crucial for this study.

References

[1] M. F. Ng and C. A. Hargreaves, "*Engineering Proceedings*", vol. 31, pp. 84, 2023.

[2] S. Sundaram and N. Hulkund, "*Proc. of International Conf. On Pattern Recognition*, Vienna, Austria, pp. 900-904, 2021.

[3] I. J. Goodfellow et al., "*arXiv preprint arXiv:1406.2661*".

[4] N. K. Singh and K. Raza, "*arXiv preprint arXiv:2005.10687*".

[5] Y. Skandarani, P.-M. Jodoin, and A. Lalande, "*MDPI Electronics*", vol. 9, no. 3, pp. 69, 2020.

[6] D. Mukherkjee, P. Saha, D. Kaplun, A. Sinitca, and R. Sarkar, "*Sci. Rep.*", vol. 12, article no. 8418, 2022.

[7] J. T. Guibas, T. S. Virdi, and P. S. Li, "*arXiv:1709.01872 [cs.CV]*".

[8] Hoo-Chang Shin, Neil A Tenenholtz, Jameson K Rogers, Christopher G Schwarz, Matthew L Senjem, Jeffrey L Gunter, Katherine Andriole, Mark Michalski, "*arXiv preprint arXiv:1807.10225*".

[9] August DuMont Schütte, Jürgen Hetzel, Sergios Gatidis, Tobias Hepp, Benedikt Dietz, Stefan Bauer, Patrick Schwab, "*arXiv preprint arXiv:2012.03769*".

[10] M. Krithika alias AnbuDevi and Dr. K. Suganthi, "*School of Electronics Engineering, Vellore Institute of Technology, Chennai*".

[11] Changhee Han, Hideaki Hayashi, Leonardo Rundo, Ryosuke Araki, Wataru Shimoda, Shinichi Muramatsu, Yujiro Furukawa, Giancarlo Mauri, Hideki Nakayama, "*IEEE Transactions on Medical Imaging*", vol. 37, no. 10, pp. 2302-2312, October 2018.

[12] M. Habijan and I. Galic, "*arXiv preprint arXiv:2205.09842*".

[13] S. Divya, L. Padma Suresh, Ansamma John, "*IEEE Transactions on Biomedical Engineering*", 2021.

[14] Tahsina Muthaki, Safwan Ibne Masuk, "*IEEE Access*", vol. 9, pp. 96636-96647, 2021

[15] Yinqiu Feng, Bo Zhang, Lingxi Xiao, Yutian Yang, Tana Gegen, Zexi Chen, "*arXiv:2406.18547 [cs.CV]*".

[16] Drici Mourad, Kazeem Oluwakemi Oseni, "*arXiv:2404.08703 [cs.CV]*".

[17] John R. McNulty, Lee Kho, Alexandria L. Case, Charlie Fornaca, Drew Johnston, David Slater, Joshua M. Abzug, Sybil A. Russell, "*arXiv:2403.19107 [cs.CV]*".

[18] Cristiana Tiago, Andrew Gilbert, Ahmed S. Beela, Svein Arne Aase, Sten Roar Snare, Jurica Sprem, "*arXiv:2403.05384 [cs.CV]*".

[19] Neeraj Baghel, Shiv Ram Dubey, Satish Kumar Singh, "*arXiv:2312.01999 [cs.CV]*".

[20] Christoph Angermann, Johannes Bereiter-Payr, Kerstin Stock, Markus Haltmeier, Gerald Degenhart, "*arXiv:2310.17216 [cs.CV]*".

[21] Tomohiro Kikuchi, Shouhei Hanaoka, Takahiro Nakao, Tomomi Takenaga, Yukihiro Nomura, Harushi Mori, Takeharu Yoshikawa, "*arXiv:2308.07573 [cs.CV]*".

[22] Muhammad Danyal Malik, Danish Humair, "*arXiv:2305.18927 [cs.CV]*".

[23] Alireza Aghelan, Modjtaba Rouhani, "*arXiv:2211.00577 [cs.CV]*".

[24] S. Almahfouz Nasser, S. Shamsi, V. Bundele, B. Garg, and A. Sethi, "*arXiv preprint arXiv:2201.09314*".

[25] R. V. Bergen, J.-F. Rajotte, F. Yousefirizi, I. S. Klyuzhin, A. Rahmim, and R. T. Ng, "*arXiv:2111.01866*".