Improving Quality of Medical Scans using GANs

[1]Tanushree Bharti; [2]Yogam Singh; [3]Mudit Jain; [4]Ankita Kumari
[1]JRF, Poornima University, Jaipur
[2]JECRC University, Jaipur
[3]Poornima University, Jaipur
[4]Teaching Assistant, Poornima University, Jaipur

Abstract:- Improving the quality of medical images is essential for precise diagnosis and treatment planning. When low quality images are used to train the neural network model, the good accuracy cannot be achieved. Nowadays, Generative Adversarial Networks (GANs) have become a potent image enhancement tool that can provide a fresh method for raising the caliber of medical images. In order to improve medical images, this paper presents a GAN-based framework that reduces noise, increases resolution, and corrects artifacts. The suggested technique makes use of a generator network to convert low-quality images into their high-quality equivalents, and a discriminator network to assess the veracity of the improved images. To ensure robustness across various modalities, the model is trained on a diverse dataset of medical images, including MRI, CT, and X-ray scans. Our experimental results show that GAN-based method significantly improves the image quality when compared to conventional methods, as evidenced by enhanced peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) according to quantitative evaluations. This study emphasizes the value of incorporating deep learning methods into medical image processing pipelines and the potential of GANs to advance medical imaging technology so that a robust neural network model can be designed.

Keywords:- *Medical Images Quality, Convolutional Neural Networks, Generative Adversarial Networks (GANs), Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM).*

I. INTRODUCTION

Medical imaging technologies, which provide non-invasive observation of internal body components, are essential to contemporary healthcare. Examples of these technologies include CT, MRI, and X-ray. These pictures are useful diagnostic tools that help doctors find anomalies, monitor the course of diseases, and schedule treatments. However, the availability of varied and representative datasets for training is a major factor in how well machine learning models perform in medical image interpretation [1]. The part of Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in segmenting images and recognizing objects in recent years. Utilizing these networks for clinical tasks, such as classifying medical pictures, and segmenting organs and diseases, should improve medical judgment [2].

Despite even the medical imaging is crucial for diagnosis, therapy planning, obtaining representative and diverse datasets for machine learning model training is still difficult because of privacy issues and restricted access to uncommon cases. In order to enhance current datasets, this research investigates the possibility of using Generative Adversarial Networks (GANs) to generate artificial medical scans [3]. We explore the difficulties in medical imaging, the shortcomings of conventional data augmentation methods, and the requirement for intelligent data augmentation strategies. A major challenge in developing models fit for clinical use is the lack of sufficient diverse labelled training data [4]. Additionally, class disparity frequently occurs in medical data. The idea behind the GANs is that it has a generator and discriminator built in, which makes it useful for comparing human scans. It is helpful in analysis for improvement of the medical scans. The practical and research was significantly improved,suggesting that GAN-based data augmentation holds promise for medical applications [5].

II. LITERATURE REVIEW AND NEED OF THE STUDY

Medical image quality can now be improved with greater effectiveness thanks to Generative Adversarial Networks (GANs), which can handle tasks like noise reduction, resolution enhancement, and image reconstruction. Many researchers did work on the GAN, some most cited research work is taken here. Qiaoying Yang et al [6] showed that GANs could reduce noise while maintaining diagnostic accuracy, resulting in a significant improvement in low-dose CT scan quality.

Moreover, super-resolution has been achieved with GANs, improving the resolution of medical images. The work of Chunyuan Li et al. [7] demonstrated that GAN-based approaches perform better in this context than traditional methods by reconstructing high-resolution images with better preservation of fine details. Specifically, the Super-Resolution GAN (SRGAN) produced high-quality images that were more accurate and aesthetically pleasing. Moreover, GANs have been used in medical image synthesis, which makes it possible to produce high-quality images from imperfect or low-quality data.

Yibin Song et al [8] did work on Liver lesion detection and classification with novel neural network architectures. Zhu Jun-Yan et al. [9] develop a model for image to image translation and created high-quality MRI images from CT scans using GANs. Zhang, et al [10] proposed a translating

ISSN No:-2456-2165

and segmenting multimodal for medical volumes with cycle- and shape-consistency GAN.

This research paper's main goal is to find out how well GANs work in producing realistic medical scans that faithfully depict patient data that has not yet been seen. Our main aim or the motive is to overcome the shortcomings of conventional methods of data augmentation and enable the production of varied and superior medical datasets by utilizing deep learning. With differing degrees of effectiveness, balancing GANSs (BAGANs)[11] have been used to correct class imbalance in GAN data and it compares the image that is uploaded in discriminator and the generator works better at low data regimes, when the model's primary issues are overfitting and inadequate generalization.

III. RESEARCH METHODOLOGY

➢ *CheXpert Dataset*

We used 224,316 frontal and lateral chest radiographs of 65,240 people from Stanford Hospital from the publicly available CheXpert dataset. Fourteen common chest conditions, including edema, consolidation, cardiomegaly, pleural effusion, and pneumothorax, are labeled in the dataset. Natural language processing (NLP) techniques were used to extract the labels from radiology reports. The dataset has key features as follows:

- Uncertainty Labels: One special feature of CheXpert is its ability to provide uncertainty labels for conditions that, even for human radiologists, may be unclear or challenging to classify. With the option to label data as "positive," "negative," or "uncertain," models can be trained to handle uncertainty in medical diagnosis [12].

- Diversity: The dataset is a valuable resource for creating generalizable models because it contains X-rays from a diverse patient population that spans a wide range of ages, genders, and clinical conditions.

- Training and Validation: CheXpert consists of two sets: a training set and a validation set. Skilled radiologists manually annotate the validation set. This configuration makes it possible to evaluate models robustly.

➢ *Conventional Methods of Data Augmentation*

Medical picture collections have been extensively enhanced by the application of conventional data augmentation techniques including rotation, flipping, and cropping. These methods might not be able to capture the intricate variances found in actual medical scans. Furthermore, their efficacy in producing realistic and diverse data is limited as they fail to take into consideration the underlying anatomical structures and diseases.

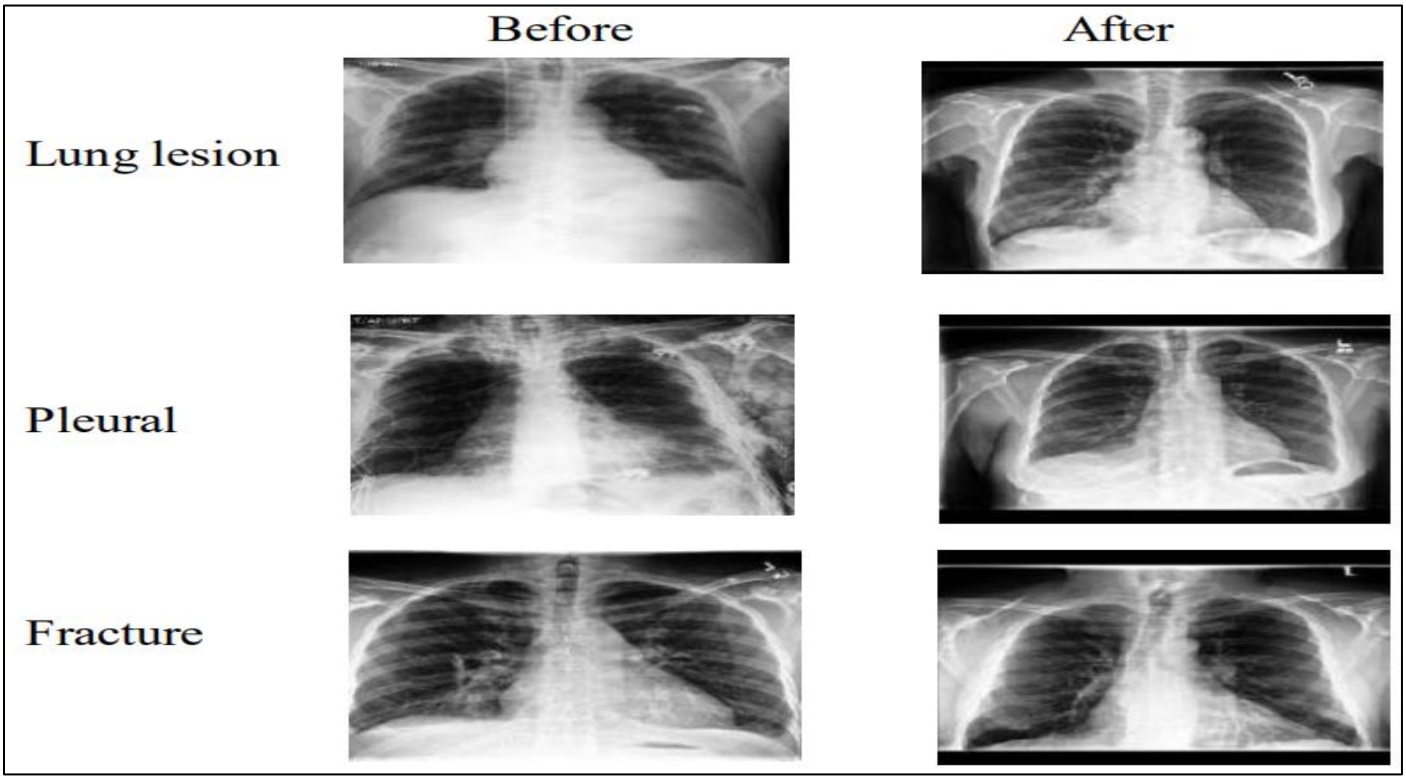

Fig 1 Data Augmentation by Generative Adversarial Network

➢ *Configuration for an Experiment*

We use a 14-way classification challenge to train DenseNet- 121, where each input image may represent more than one pathology.

We use randomly selected portions of the training dataset (1%, 10%, 50%, and 100%) to train each of the three primary experiments of the generative adversarial network Testing, validation, and training use a 90% 5% 5% split.

The example of the result including 5% of the similar to the parent data in the dataset takes places a robust results on the categorization of thoracic X-rays the construction by professor's research. We pre-trained DenseNet-121 of Torch XRay Vision on ImageNet using transfer of the all experiment of all the experiments. Figure 1 shows the augmented images by GAN.

➢ *Difficulties with Medical Imaging*

Accurate diagnosis and treatment of medical imaging require addressing a number of obstacles presented by the field. Among these difficulties are: Scan distortion, various variables, including noise, artifacts, and motion artifacts, can cause distortion in medical scans. Healthcare practitioners may find it difficult to establish precise diagnoses as a result of these aberrations, which can seriously impair the images' quality and interpretability.

➢ *The Requirement for Sensible Data Enrichment*

The drawbacks of conventional augmentation approaches may be solved by intelligent data augmentation techniques, such as those based on deep learning. These methods improve the robustness and diversity of the training dataset by producing synthetic scans that closely mimic real patient data by understanding the underlying distribution of medical pictures [13].

Generative Adversarial Networks (GANs) rank fourth.

The GAN class of deep learning models is made up of a discriminator and a generator neural network. While the generator generates synthetic data, in the generative adversarial network the part of the discriminator will havethe image of the scans, the generator produces higher- fidelity synthetic pictures by producing more realistic samples. Figure 2 shows the process of generative adversarial network for generating images.

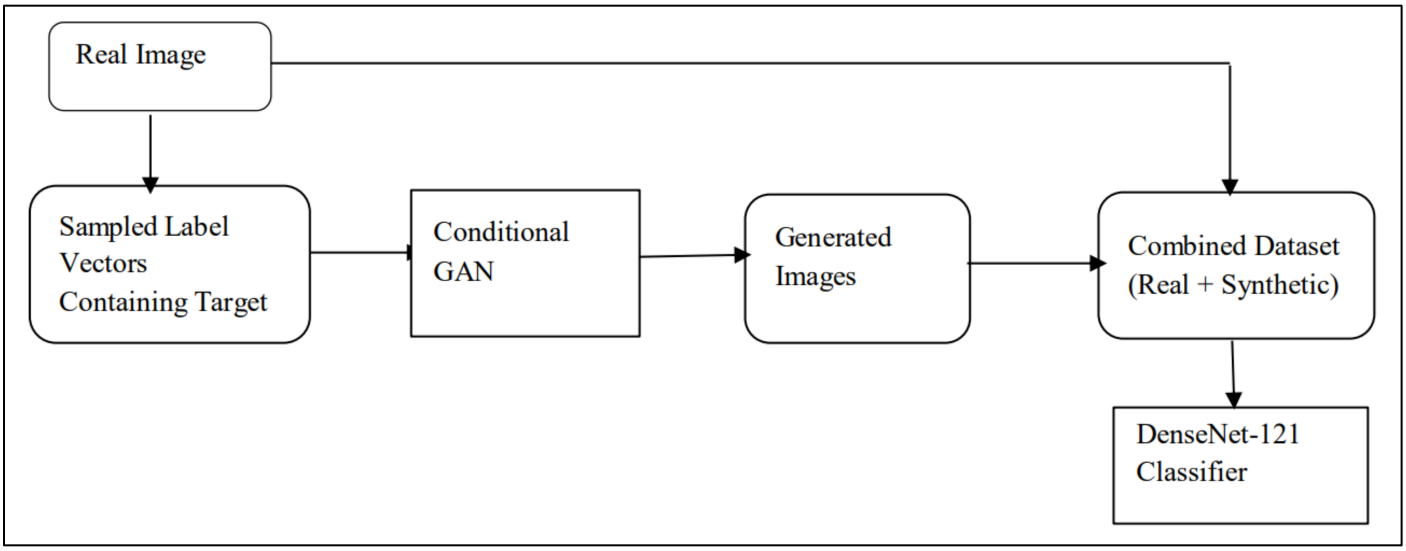

Fig 2 Process of Generative Adversarial Network

IV. IMPLEMENTATION

➢ *Deep Convolutional GAN (DCGAN)*

One essential tool for picture-generating problems is architecture. To generate visually accurate and high-resolution images, it uses convolutional layers to extract hierarchical characteristics from the input data1.

- Generator Network: This generator learns to produce visuals that mimic actual data by using random noise as input. Usually, it is made up of convolutional layers, which are followed by upsampling layers such as transposed or nearest-neighbor convolutions. The generator's output is an image.

- Discriminator Network: The discriminator is a convolutional neural network (CNN) that learns to distinguish between real images from the dataset and fake ones from the generator. It accepts an image as input. It produces a likelihood score that indicates the authenticity of the provided image.

- Adversarial Training: The discriminator and generator in a min-max game are trained simultaneously. As the discriminator endeavors to precisely distinguish between genuine and counterfeit images, the generator aims to produce visuals that are indistinguishable from real photos.To enhance both networks' performance, the parameters are optimized throughout the training phase.

- Convolutional Layers: DCGANs employ convolutional layers for the generator and discriminator rather than fully linked layers. Convolutional layers work effectively for applications like picture production because they can capture spatial patterns in images.

- Batch Normalization: To stabilize training and quicken convergence, batch normalization is frequently used in both the discriminator and generator networks. It makes the decreasing the internal covariate shift issue through the activations of each layer.

- Avoiding Pooling Layers: DCGANs usually steer clear of the generator and discriminator pooling layers. Rather, they employ fractional-strided convolutions (transposed convolutions) for upsampling in the generator and stride convolutions for downsampling in the discriminator. This improves the preservation of spatial information.

- Activation Functions: The generator network uses standard activation functions like ReLU (Rectified Linear Unit), with the exception of the output layer, which uses tanh, an appropriate activation function, to guarantee that the generated images' pixel values fall between [-1, 1]. Leaky ReLU is frequently utilized in the discriminator to avoid the dead neuron issue. DCGANs have shown remarkable success in producinglifelike images in a variety of domains such as bedrooms, faces, and landscapes. They have also served as an inspiration for a great deal of development and innovation in the generative modeling space.

➢ *Pix2pix Network*

This is a conditional GAN architecture that discovers how to translate an input picture to an output picture. It has proven effective in a variety of applications, including the creation of medical images from semantic labeling and image-to- image translation [14].

The main work or the duty is the image to image transition or Pix2Pix, is a conditional generative adversarial network (GAN) designed especially for image-to-image translation applications. This is how Pix2Pix functions and what sets it apart:

- Conditional GAN Framework: Pix2Pix expands the capabilities of the GAN framework by including a conditional setting. In a traditional GAN, the generator generates images from random noise as input, while the discriminator looks for differences between genuine and fake images. The generator in Pix2Pix gains the capability to translate images between distinct domains because both the discriminator and generator are conditioned on input images [15].

- Translation of Images to Images: Pix2Pix is especially made for jobs in which the input and "Image-to-Image Translation with Conditional Adversarial Networks," or Pix2Pix for short, is a sort of conditional generative adversarial network (GAN) in which the output images are paired. This covers tasks such as mapping satellite imagery to maps, creating realistic graphics from sketches, transforming daytime views into nighttime sceneries, and more. The network gains the ability to translate input images from one domain into equivalent output images in another.

- Generator and Discriminator Networks: In Pix2Pix, an encoder-decoder architecture is commonly used as the generator. The input image is encoded by the encoder into a latent representation, which is then decoded by the decoder to produce the output image. The part of the Convolutional neural networks are used in the discriminator to discriminate between actual input-output image pairings and fictitious ones produced by the generator [16].

- Aviation Training: Comparable to Unlike previous GANs, Pix2Pix trains the discriminator and generator at the same time via adversarial training. The discriminator seeks to discern between pairings of actual and fake images, while the generator seeks to make output images that are identical to real photos.

- L1 Loss: Pix2Pix augments the objective function with a pixel-wise L1 loss component in addition to the adversarial loss. This loss encourages the generated images to more closely approximate the ground truth images at the pixel level. Together, adversarial loss and L1 loss yield visually pleasing results while maintaining minute details.

- Design Modifications: To enable information flow at various sizes, Pix2Pix frequently employs a U-Net design for the generator, which incorporates skip connections between the encoder and decoder. This architecture aids in the capture of both local and global elements, improving performance in jobs involving the translation of images.

Pix2Pix is a popular tool for many computer vision applications, including semantic segmentation, style transfer, image colorization, and more. It is a useful tool for many image alteration applications because of its capacity to learn mappings across various visual domains from paired training data

➢ *StarGAN*

With just one model, StarGAN is a flexible GAN architecture that can translate images between different domains. It makes it possible to create a variety of medical images with various features and qualities. StarGAN, or "Star Generative Adversarial Network," is a generative adversarial network (GAN) architecture designed for multi-domain image-to- image translation. This is how StarGAN functions and what makes it unique:

- Multi-Domain Image Translation: StarGAN can handle many domains inside a single model, in contrast to standard image-to-image translation techniques that call for distinct models for each translation assignment. It supports several target domains at once and has the ability to convert images across them. It may, for instance, convert pictures of human faces into many facial styles, representing various ages, genders, and races [17].

- Single Generator and Discriminator: For all domain translations, StarGAN has a single generator and discriminator architecture. This indicates that the accountable generator is the same for creating images in every target domain, and the discriminator gains the ability to discern between images produced by the generator that are false and genuine images from any domain.

- Conditional GAN Framework: To enable multi-domain translation, StarGAN expands the conditional GAN framework. The target domain labels, which indicate the desired domain for the output image, and the source domain labels, which indicate the domain of the input image. With this conditioning, the network is able to learn how to produce images in various target domains according to the label of the input domain.

- Adversarial and Cycle-Consistency Losses: It offers details on how well the model performs across dataset regimens and augmentation techniques as indicated by the ROC- AUC score. The little data alignment we can the improvement in the image of the scan by the generator. This is observed across all disorders, with an AUC increase of 0.07 for fractures and 0.03 for lung lesions and pleural abnormalities that the generated images in the target domain are identical between GAN augmentation and no augmentation to genuine images. It also features a cycle-consistency loss, motivated by CycleGAN, which aims to translate back to the source domain such that the reconstructed images are almost identical to the original input images. By doing this, the produced photos' artifacts are lessened and uniformity is preserved [18].

The fifth feature is the Domain Adaptation Module that StarGAN offers. It gains the capacity to alter the generated images according to the characteristics of the target domain. The quality and realism of the translated images are improved by this module, which adjusts global and local characteristics to match the target domain.

StarGAN has addressed a wide range of image translation problems, including face attribute change, style transfer, and domain adaptation. Its capacity to manage several domains with a single model accounts for its effectiveness and adaptability for a broad spectrum of computer vision and image processing applications.

➢ *Teaching the Generator Model*

Training the generator model includes optimizing the network parameters to minimize the discrepancy between the generated and real images. This process requires appropriate loss functions and training strategies in addition to a large and diverse dataset of medical scans to ensure convergence and stability.

GANs can capture the hidden underlying properties of the training dataset, including illnesses, textures, and anatomical structures. By selecting samples from the learnt latent space, the generator model can be trained to produce previously unobserved medical scans. These produced scans can be utilized for a number of tasks, such as creating pathological instances to train reliable diagnostic models, domain adaption, and data augmentation.

It should be emphasized that traditional augmentation improves performance when compared to no augmentation, however not as much as GAN augmentation. Similarly, our results for the second- smallest data regimen (10%) demonstrate that the conditions are improved by GAN augmentation. These positive results demonstrate the superior performance of models trained with GAN augmentation. The learning curves for training without augmentation and with GAN augmentation are seen in Figure 3. It turns out that GAN-based overfitting is mitigated in the 1% low-data regimen by augmentation.

In the most recent epoch, the difference between the training and validation losses of the GAN enhanced model is 0.03, whereas the non-augmented model's difference is 0.06. In the 10% regimen, both models show comparable overfitting, the variation in the difference in the generative adversarial network would be the interval of 3 to 4.

This is in line with our AUC findings, which show that while adding synthetic photos helps reduce overfitting for extremely small data batches, it may not always be helpful as dataset sizes grow.

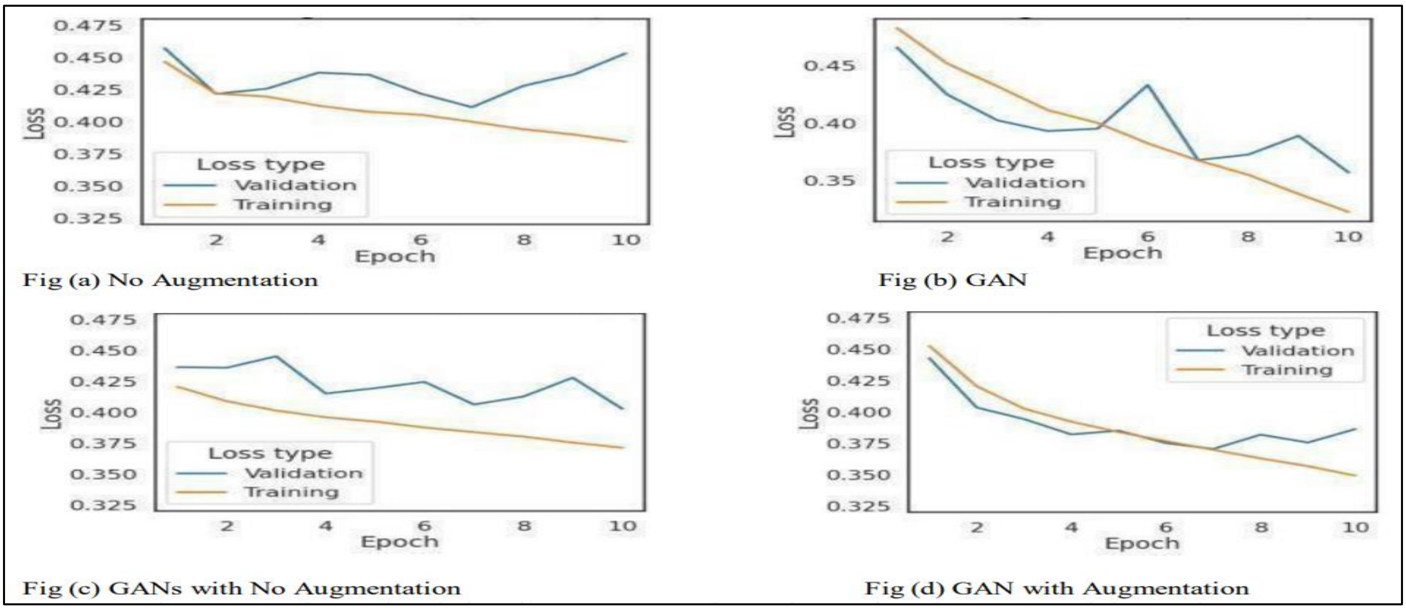

Fig 3 Testing and Updating the Curve Comparison of GAN-based Augment Training against no Augment

V. RESULTS AND DISCUSSION

Since this model performed better than the others overall, we show the image that is been compared between the generator and discriminator and the process of identifying the real image, suggesting that the network activations might be the same in both cases.

Even so, the strongest activations seem to be localized to a certain area in each X-ray for all CAMs. In a clinical setting, CAMs may offer helpful, interpretable insight regarding the areas of an X-ray that signal particular diseases due to the appearance of high activation at specific spots. Class activation map visualization presents the image that is best discriminative for a given category, where the category is determined by projecting the output layer weights onto the final convolutional feature maps. Figure 4 shows the image before GAN and after applying GAN. Table 2 shows the AUC results for not augmented, standard augmented, and GAN augmented classes across dataset regimens.

It is essential to evaluate the quality and realism of the generated medical scans to ensure that they are suitable for use in subsequent tasks. This involves both qualitative assessment by medical experts and quantitative metrics, including pixel-level similarity evaluations, to confirm the clinical relevance of the generated images.

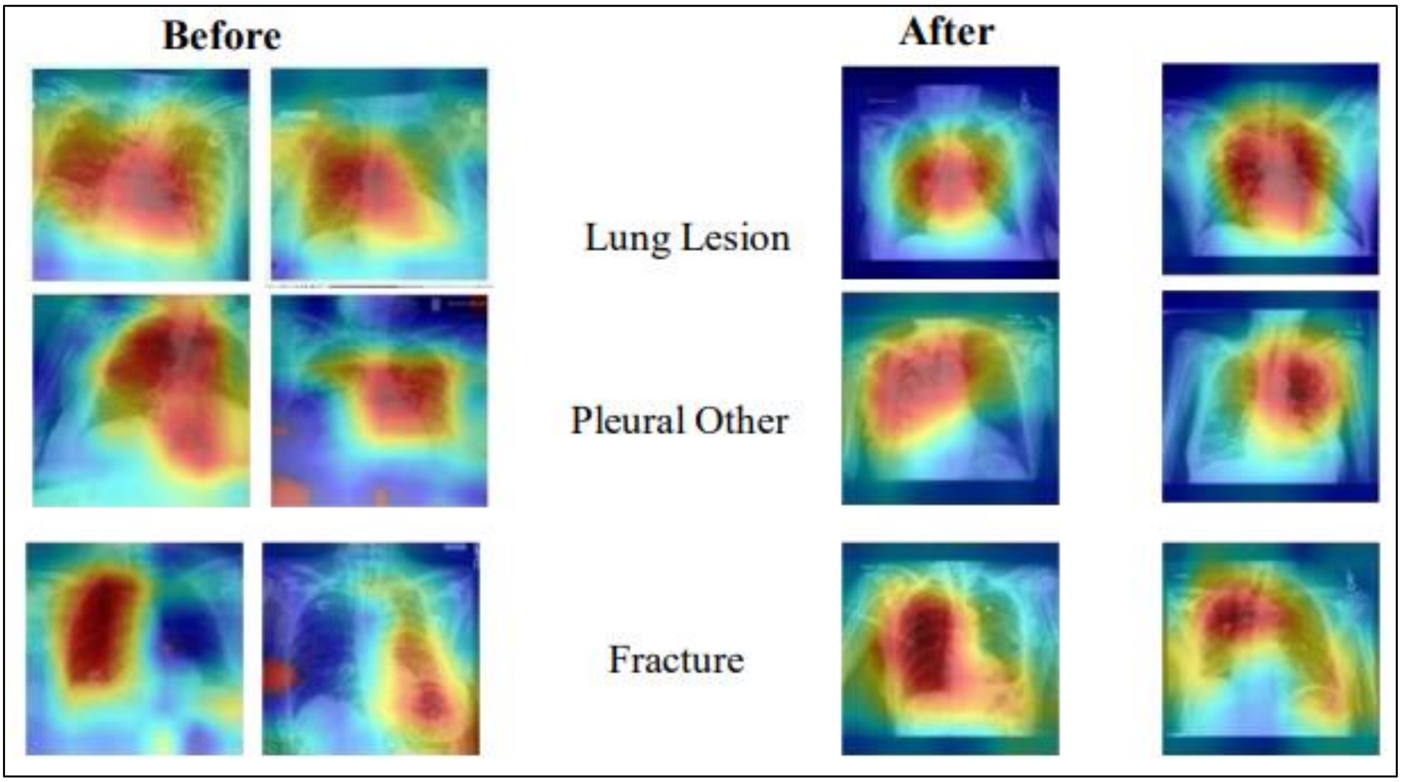

Fig 4 Images before and after GAN

Table 2: AUC Results for Augmented Classes across Dataset Regimens and Augmentation Strategies

Dataset Size	Pathology	Not the augment	Standard Augment	Generative adversarial network augment
	The Lung part	1.727	1.728	1.758
1%	Pleural	1.566	1.550	1.594
	Broken part	1.583	1.601	1.656
	The Lung part	2.809	1.796	2.852
10%	Pleural	1.632	1.655	1.670
	Broken part	1.700	1.723	1.742
	The Lung part	1.826	1.822	1.828
50%	Pleural	1.710	1.696	1.706
	Broken part	1.789	1.780	1.793
	The Lung part	1.835	2.945	1.834
100%	Pleural	1.721	1.712	1.727
	Broken part	0031.811	1.793	1.807

VI. CONCLUSION

In conclusion of this study, we compare the performance of non-augmented and standard augmented models over a range of data schedules. Our results suggest that class- imbalanced medical datasets can be effectively corrected via GAN-based data augmentation. Through a range of dataset sizes and pathologies, we show improvements in table 2 and figure 4. The discriminator and generator are the two main components of the generative adversarial network process. The discriminator compares the uploaded fake image to the original version of the image, and this process continues until the generator and discriminator have similar image connections. Eventually, we receive the comparison version of the image in the scans along with an update on a big data set, kindly note that this may not always be thecase. To sum up, GANs present a viable method for creating invisible medical scans with a variety of traits and diseases. These models can help create large-scale, high-quality medical datasets by utilizing deep learning to overcome the drawbacks of standard data augmentation methods. To address issues including data scarcity, model interpretability, and ethical concerns about the use of synthetic data in medical research and practice, more studyis required that can be taken in future work.

REFERENCES

[1]. Hoo-Chang Shin et al. "Synthetic data augmentation usingGAN for improved liver lesion classification." In: International Workshop on Medical Imaging Simulation and Synthesis, pages 1-11. Cham, Springer, 2018.

[2]. Eunji Choi et al. "Stargan: Unified generative adversarialnetworks for multi-domain image-to-image translation." In 2018's IEEE Conference on Computer Vision and Pattern Recognition Proceedings, pp. 8789–8797.

[3]. Jelmer M. Wolterink and colleagues "Generative adversarialnetworks for noise reduction in low-dose CT." IEEE MedicalImaging Transactions 36.12, 2017; 2536–2545.

[4]. Maayan Frid-Adar et al. "Synthetic data augmentation using GAN for improved liver lesion classification." In: International Workshop on Medical Imaging Simulation and Synthesis, pages 1-11. Cham, Springer, 2018.

[5]. Armanious, Kristina, et al. "Using adversarial networks to synthesize CT from ultrasound images." In International Conference on Computer-Assisted Intervention and Medical Image Computing, pp. 81–89. Springer, Cham (2017).

[6]. Qiaoying Yang et al. "Low-dose CT image denoising using a generative adversarial network with Wasserstein distanceand perceptual loss." 2018 IEEE Access 6: 47958–47966.

[7]. Chunyuan Li and colleagues, "Unsupervised image-to-image translation networks." In Neural Information Processing Systems Advances, pp. 700–708 in 2017.

[8]. Yibin Song and colleagues, "Liver lesion detection and classification with novel neural network architectures." In: International Conference on Computer-Assisted Intervention and Medical Image Computing, 830-838. Springer, Cham (2017).

[9]. Zhu, Jun-Yan, and associates. "Toward multimodal image-to-image translation." Pages 465–476 in Advances inNeural Information Processing Systems, 2017.

[10]. Zhang, et al., "Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network." In International Conference on Computer-Assisted Intervention and Medical Image Computing, pages 56–64. Cham, Springer, 2018.

[11]. Kavita Lal, Madan Lal Saini; A study on deep fake identification techniques using deep learning. *AIP Conf. Proc.* 15 June 2023; 2782 (1): 020155. https://doi.org/10.1063/5.0154828

[12]. Y. Singh, M. Saini and Savita, "Impact and Performance Analysis of Various Activation Functions for Classification Problems," *2023 IEEE International Conference on Contemporary Computing and Communications (InC4)*, Bangalore, India, 2023, pp. 1-7, doi: 10.1109/InC457730.2023.10263129.

[13]. Sarmah, J., Saini, M.L., Kumar, A., Chasta, V. (2024). Performance Analysis of Deep CNN, YOLO, and LeNet for Handwritten Digit Classification. In: Sharma, H., Chakravorty, A., Hussain, S., Kumari, R. (eds) Artificial Intelligence: Theory and Applications. AITA 2023. Lecture Notes in Networks and Systems, vol 844. Springer, Singapore. https://doi.org/10.1007/978-981-99-8479-4_16

[14]. M. L. Saini, A. Patnaik, Mahadev, D. C. Sati and R. Kumar, "Deepfake Detection System Using Deep Neural Networks," *2024 2nd International Conference on Computer, Communication and Control (IC4)*, Indore, India, 2024, pp. 1-5, doi: 10.1109/IC457434.2024.10486659.

[15]. P. D. S. Prasad, R. Tiwari, M. L. Saini and Savita, "Digital Image Enhancement using Conventional Neural Network," *2023 2nd International Conference for Innovation in Technology (INOCON)*, Bangalore, India, 2023, pp. 1-5, doi: 10.1109/INOCON57975.2023.10100995.

[16]. Skandarani, Youssef, Pierre-Marc Jodoin, and Alain Lalande. "Gans for medical image synthesis: An empirical study." *Journal of Imaging* 9.3 (2023): 69.

[17]. Showrov, Atif Ahmed, et al. "Generative Adversarial Networks (GANs) in Medical Imaging: Advancements, Applications and Challenges." *IEEE Access* (2024).

[18]. E. G. Kumar, M. Lal Saini, S. A. Khadar Ali and B. B. Teja, "A Clinical Support System for Prediction of Heart Disease using Ensemble Learning Techniques," *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)*, Theni, India, 2023, pp. 926-931, doi: 10.1109/ICSCNA58489.2023.10370569.