Generative Adversarial Networks in Medical Imaging: Applications, Challenges, and Emerging Trends (2020–2025)

Taoyu Chen

School of Date Science and Artificial Intelligence, Chang'an University, Middle Section of South Second Ring Road, Xi'an 710064, China
`2023900902@chd.edu.cn`

Abstract. Generative Adversarial Networks (GANs) have developed into an important class of generative models in medical imaging, providing distinct properties for problems that are difficult because of limited data and costly annotation. In 2020 - 2025, GANs have also shifted from basic augmentation to more advanced tasks such as, image generation, cross-domain translations, super-resolution reconstructions, as well as data generation within domain-specific applications of medical imaging, for example, ophthalmology and neuroimaging. Although there has been ongoing progress in GAN implementations, challenges remain that are inherent to GANs, such as training instabilities, and mode collapse in training, and clinically meaningful evaluation measures have not yet evolved. Additionally, diffusion models have become increasingly popular in the generative modeling landscape and have led to some studies that highlight the benefits and limitations of GANs. This review aims to summarize the development and application of GANs in medical imaging, highlighting relevant technical and practice problems and trends including, hybrid GAN-diffusion architectures, clinical empirically-based proof methods, ethics, and explainability. Our analysis concludes that there will continue to be ongoing GAN development alongside novel generative models; however, GANs will remain a relevant and practical approach in the bounded environment of medical imaging computing.

Keywords: Generative Adversarial Networks, Medical Imaging, Image Synthesis.

1 Introduction

Medical imaging plays a great role in the diagnosis, monitoring, and treatment of diseases. However, deep learning methods are still confronted with the main challenge in this domain: the availability of high-quality, annotated medical datasets is rather limited. Medical image acquisition is expensive, time-consuming, and strictly regulated by privacy policies; expert annotation requires significant clinical effort. For this reason, using GANs for generating realistic synthetic data is an important research direction to expand datasets and improve model performance [1].

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_86

In previous works, GANs had been mainly used as data augmentation tools with relatively simple architectures, such as DCGANs and conditional GANs. The goal at that time was to increase the size of the training samples for various tasks, like classification and segmentation, due to the scarcity of data [2]. Recently, in the last five years (2020–2025), the role of GANs has significantly expanded. Recent works have begun to employ more advanced architectures, including StyleGAN and CycleGAN, which allow for complex tasks, including cross-modality image translation and anatomically consistent synthesis [3, 4]. Therefore, increasing the quality of the data, rather than its quantity, and enabling new functionalities are the current emphases. Examples include super-resolution for enhancing image clarity [5], image harmonization to decrease variability due to scanners across different centers [6], and domain-specific applications in ophthalmology and neuroimaging [7, 8].

While GANs have seen considerable development, several persistent issues remain including training instability, mode collapse, and limited evaluation frameworks that affect model reliability [9]. Further, the emergence of diffusion models as strong generative competitors has spurred renewed attention to the future position of GANs in medical imaging [10]. Given these rapid developments, a focused review of progress over the last five years was necessary to place in context how GANs have evolved and where future opportunities exist.

The purpose of this paper is to give an overview of recent advances in GAN-based medical imaging within the period from 2020 to 2025, with particular emphasis on key applications, technical developments, existing limitations, and emerging trends. This work is organized as follows: Chapter 2 discusses major application areas of GANs in medical imaging, namely synthesis, augmentation, reconstruction, and harmonization. Chapter 3 discusses the main technical challenges and comparatively studies GANs against diffusion-based generative models. Chapter 4 elaborates on ongoing research directions, such as hybrid GAN–diffusion architectures, clinical validation pathways, and explainability. Chapter 5 summarizes the results and outlines some possible future developments in the field.

2 Applications of GANs in Medical Imaging

2.1 Image Synthesis and Modality Translation

Image synthesis can be considered one of the first and most pioneering areas where GANs have contributed in the medical image processing domain. The main idea behind this use case is to create a very realistic image based on a certain medical condition or image type. In the last five years, a number of state-of-the-art GAN networks have been successfully utilized for the synthesis of a variety of medical images [3]. For example, cGANs have been utilized for the synthesis of brain MRI images based on certain tumor properties [1], while a number of GAN models, including those based on the StyleGAN model, have been successful in creating realistic fundus images based on the progression stages of a certain medical condition, namely diabetic retinopathy [7].

A very important sub-task of synthesis is modality translation, which is an image-to-image translation task. The goal here is to transfer the image from a source modality to a target modality. This has many significant implications in the clinical setting because it can potentially avoid the need for a follow-up scan, thereby cutting patient costs, saving time, as well as reducing susceptibility to ionizing radiation. A very common use case here could be the synthesis of a Computed Tomography scan from a Magnetic Resonance Imaging scan. In this regard, CycleGAN has been very helpful in creating a synthetic Computed Tomography image from a patient's MRI scan [4]. This becomes extremely important in radiotherapy because, while MRI has excellent soft tissue contrast facilitating tumor delineation, a CT scan is mandatory for the determination of a therapeutic dose. In this regard, it becomes possible to have a scan synthesized from the planning scan, thereby making it possible to conduct MRI-only radiotherapy.

More modern developments have been centered on improving the realism as well as the anatomical accuracy of the generated images [6]. The utilization of conditional information has become increasingly advanced, enabling a level of precise control over the generation task. For instance, a GAN can be conditioned on a semantic segmentation map, thereby enabling the generation of an image that has the correct anatomy, wherein structures such as organs or tumors are depicted exactly as they are segmented. Additionally, attention mechanisms have been added to GANs to enhance textural realism [9]. The attention mechanisms can be used to enable the generator to selectively converge on regions of interest, such as areas within a tumor, while maintaining the overall anatomy.

2.2 Data Augmentation & Dataset Expansion

The main use of GANs in medical imaging is in dealing with the challenge of data scarcity. In medicine, image datasets are usually small. The class distribution within a medical image dataset is imbalanced, as images from healthy subjects or those belonging to common diseases greatly outnumber those from rare diseases [1]. This has been a negative factor in deep learning, as deep learning algorithms are not very good at handling image categories with limited available images. GANs can cope with the challenge very well. GANs use the small image datasets to train systems to produce additional relevant points [2]. GANs are a very effective solution to this challenge.

For illustration purposes, suppose a convolutional neural network is being trained to classify liver lesions in a CT scan, separating benign cysts from malignant tumors. The common benign cysts would likely have thousands of example images, while malignant tumors would have only a handful of example images. If a model were to be trained on such a biased set, it would be likely to be excessively biased towards benign, potentially misdiagnosing cancer. Using a conditional GAN on a small set of malignant tumor images, one could potentially generate a large number of realistic tumor images [1]. The generated images would have the capability to mimic the textures, shapes, and sizes of actual pathological images.

The addition of such synthesized images to the training database means that the convolutional neural network is trained on more elaborate and comprehensive features of tumors. The synthesis of additional images not only results in an increase in the database but also increases the diversity of the database. The classifier, as a result, has a higher level of generalization and a higher level of sensitivity, meaning it has the capability to detect rare diseases [2]. The use of GANs has been vital in enhancing the accuracy of the model in the diagnosis of rare diseases.

2.3 Domain-Specific Applications

Apart from common tasks like data augmentation, GANs have been used for dealing with challenging areas specific to certain subdomains of medicine. This section discusses a few applications of GANs in the areas of ophthalmology, neuroimaging, and computational pathology.

In ophthalmology, one of the most important applications of GANs is the generation of high-resolution images of the retina and fundus. High-resolution images are vital for diagnosing conditions such as diabetes and glaucoma. The latest GANs have the capability to generate images that can recreate different stages of the condition, such as early-stage diabetic retinopathy and severely affected glaucoma in the advanced stages [7]. This has a huge significance in enriching the pool of images used for training AI-assisted systems to detect the most important pathological features.

In the field of neuroimaging, Generative Adversarial Networks have been employed for the diagnosis as well as prognosis of neurological and psychiatric ailments [8]. In the case of neurodegenerative ailments, specifically Alzheimer’s Disease, a prime utilization of GANs involves modeling the progression of a patient’s condition. GANs can be trained on MRI scans to forecast future advancements of a patient’s brain atrophy, thereby generating a "snapshot" of a patient’s future brain based on initial scans. In the context of psychiatry, researchers have also demonstrated the application of GANs to build structural brain scan data for supplementing datasets concerning the diagnosis of Major Depressive Disorder (MDD).

2.4 Image Reconstruction & Super-Resolution

Generative Adversarial Networks are very important in image reconstruction as well as super-resolution, and they can be utilized for enhancing image quality in medical images, as well as extracting details from low-resolution data. In medicine, it has long been a trade-off between image resolution, scan time, and radiation exposure in imaging techniques such as PET and CT. A low-resolution image scan takes shorter time, but lacks detailed information vital for diagnosis. Conversely, acquiring a high-resolution scan takes longer and may involve a higher radiation dose.

The challenge can be overcome effectively using super-resolution generative adversarial networks (SRGANs) [5]. The approach involves training a model aimed at learning how to map a low-resolution input to a high-resolution output. The model consists of a generator, responsible for recovering high-frequency information removed during acquisition, and a discriminator trained to detect generated high-resolution images as opposed to actual ones. Both networks work in conjunction, causing the generated image to be both mathematically correct and perceptually realistic.

Most importantly, the benefits of applying GANs as a basis for super-resolution algorithms go beyond improving visual quality. The restored high-resolution images can be very useful as input sources for computational processing. For instance, the accuracy level of segmentation algorithms can be greatly improved on the basis of super-resolution techniques, especially concerning the definition of tumor edges [10]. Similarly, the accuracy of classification systems can be improved because textures are restored. GANs can therefore be considered an effective preprocessing tool in medical imaging.

2.5 Data Harmonization and Cross-Center Adaptation

In addition to generating data or improving resolution, GANs can also address the serious challenge of data heterogeneity among sites and devices. Differences in scanner parameters, acquisition protocols, and post-processing pipelines can generate different imaging characteristics that hinder model generalizability. To mitigate this, GAN-based harmonization methods have been developed to align images across multiple centers or domains [4]. CycleGAN and StarGAN architectures, for instance, can learn mappings across modalities and harmonize PET or MRI images while preserving anatomical identity. These approaches have performed remarkably well in reducing scanner bias and improving downstream analytic reliability [6]. While issues of clinical utility and synthetic artifacts must be carefully considered, this represents an important advance toward developing robust and generalizable AI models for clinical applications.

3 Challenges and Comparative Analyses

3.1 Training Stability and Evaluation Issues

Despite the success achieved by Generative Adversarial Networks (GANs), the training of GANs remains a very challenging task. The issue associated with the adversarial training process is maintaining a trade-off between the generator and the discriminator. If a disturbance occurs, causing an unbalance, it leads to a training instability, consequently a common failure referred to as "modal collapse" [11]. In this scenario, the generator finds a set of distinct data examples that can always mislead the discriminator, as a result, it only focuses on generating those limited variations instead of learning how to represent the whole distribution of data. This is a serious limitation, especially within the medical environment. For example, a GAN model tasked with generating a set of lung nodules may end up generating only nodules of a single form, missing the actual scope of pathological variability. The use of such unsustainable images within a dataset can result in datasets becoming significantly biased, hence incapable of enhancing the model's generalization capability.

The next significant drawback is the lack of unified, relevant metrics for the quality of generated images. A generated image's quality in medicine is a difficult task to measure. The common metrics used for the quality assessment of images, like Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), have several limitations. For example, the metrics measure the difference on a simple pixel level, which does not impact quality directly. Metrics specifically proposed for GANs, such as Frechet Inception Distance (FID), have gained popularity. The approach, however, requires pre-trained models on natural images. The features used are often not representatives of the actual structures in medical images. This hampers quality assessment, a labor-intensive procedure conducted by human radiologists ("visual Turing tests") [12], meaning it's not scalable for experimentation.

The aforementioned challenges have been under close scrutiny, and resolving them has become a vital agenda item on the research agenda. The remedies include a multi-faceted approach. These include the development of stronger loss functions, such as the Wasserstein loss [13], which has a smooth gradient. In addition, it includes the adoption of balanced train schedules, such as dual-scale update rules [14], to achieve a balance between the generators and discriminators. More importantly, there has been a strong emphasis on the development of clinically relevant assessment frameworks. This can be achieved from a perspective from which assessment criteria shift from pixels or features to task assessment. In this case, it involves evaluating quality based on the direct effects of generated data on the performance of the associated task, such as tumor segmentation or classification.

3.2 Comparative Analysis: GANs vs. Diffusion

The rise of Diffusion models has significantly altered the state of art in the domain of Generative Modeling. Comparative evaluations conducted recently on the perceptual quality as well as quantitative assessment indicators, such as Müller-Franzes et al. in 2023 [10], indicate the superiority of latent diffusion models over GANs [15]. The quality of structural textures as well as convergence rates has been observed to be finer in Diffusion-driven models, albeit at a significantly higher computational cost. For smaller data sets, the faster inference processing along with computational constraints remains an important consideration, which favors GANs. It appears from the evaluative assessment of the two distinct variants of Generative Modeling, a complementary relationship instead of competition subsists. A novel direction could be derived based on combining Adversarial Learning from GANs along with Denoising Diffusion [16].

4 Future Trends

4.1 Hybrid Generative Architectures: GAN–Diffusion Integration

One interesting direction in generative modeling does not consist of a choice between GANs and diffusion models but involves using both: Researchers now see them as complementary instead of competing, with many exploring ways to build hybrid models that take the best from each. The idea is simple: use the stability of diffusion models and the sharp detail generation ability of GANs.

A hybrid model works by mingling the learning methods of both systems. The diffusion typically acts as the base generator: it adds and removes noise in steps, keeping the training steady and the overall structure clear to avoid Mode Collapse problems [11]. However, when used in isolation, the output from diffusion models can sometimes appear overly smooth and slightly blurry, without the sharp edges and fine textures characteristic of real photography.

To address this issue, the paper added a discriminator: this is a key element in GANs. The discriminator learns to pick up on even small flaws—fake textures or soft outlines. Its feedback tells the generator what looks wrong and pushes it to make more realistic and detailed images. In doing so, the adversarial training introduces a level of detail and realism that the pure diffusion process can often not achieve.

4.2 Clinical Validation and Real-World Deployment

Beyond the architectural innovations, the most critical future trend for GANs in medical imaging will be the drive for actual clinical validation and practical deployment. Although there has been much technical advancement, clinical translation of these models is generally limited. Among the main reasons for this gap lies the fact that most of the state-of-the-art studies evaluate performance with perceptual or statistical metrics only, such as SSIM, PSNR, or FID. These might be helpful for technical benchmarking but do not truly capture the diagnostic value of generated images. For instance, an image with a good FID score may still lack the specific, subtle pathological features a clinician relies on to make a diagnosis.

However, in order to close the gap from research to reality, the focus of validation needs to be reconsidered. Indeed, as underlined by recent reviews, few GAN-based systems have undergone prospective clinical trials or rigorous, physician-blinded evaluations, considered as the gold standard of medical evidence [1]. Thus, future work should focus on task-based validation, in which the success of a GAN is directly measured by its value to a clinical task, such as increasing diagnostic accuracy or delineating tumors more accurately. This will require much closer collaboration between AI developers and medical professionals.

4.3 Ethical Considerations and Data Privacy

With GANs increasingly capable of generating highly realistic synthetic medical images, they introduce a critical duality in terms of ethics and privacy. While they can be used as a strong tool to address privacy concerns by generating artificial data sets, such data sets capture the statistical properties of a patient population without including direct, identifiable records, therefore facilitating wider sharing for research. On the other hand, this very capability raises significant ethical questions. There is a risk that GANs could, through overfitting, inadvertently memorize and reproduce features of real patient data, potentially leading to privacy breaches or "membership inference" attacks where an adversary could deduce if a specific individual's data was in the training set.

4.4 Explainability to Build Trust

Finally, an urgent frontier of clinical adoption pertains to enhancing the explainability of GANs. These models are often criticized as being "black-box" systems; thus, it remains quite challenging for clinicians to instill confidence in their outputs. Knowing how a GAN produces certain anatomical structures or pathological details is essential to further validate the reliability of the synthetic data. This is where research is now considering the use of explainable AI. For example, recent works have leveraged attention-based visualization to expose how the model's focus aligns with clinically relevant diagnostic cues [2]. Incorporating such XAI techniques can help clinicians more thoroughly examine the generated images and assess their medical soundness. Establishing comprehensive interpretability measures will be essential for linking algorithmic progress with dependable clinical practice.

5 Conclusion

This survey has reviewed the recent progress of GANs in medical imaging through the five major application domains comprising image synthesis and modality translation, data augmentation, domain-specific medical applications, image reconstruction and super-resolution, and cross-center data harmonization. Besides, the paper analyzed major challenges currently limiting the clinical use of GAN-based systems, including training instability, limited evaluation metrics, and the rising competition from diffusion models, and highlighted emerging research directions: hybrid architectures, clinical validation strategies, and explainability techniques.

In summary, developments between 2020 and 2025 demonstrate that GANs have come a long way from simple tools for generating additional data to sophisticated components for enabling high-fidelity image generation, structural consistency, and overall improved downstream clinical task performance. However, while the recently increased visibility of diffusion models might hint at such developments, our analysis has shown that GANs maintain singular strengths, mainly regarding fast inference speed and efficiency on limited hardware, which are highly valuable in real-time or resource-constrained clinical workflows. Rather than being supplanted by them, GANs will more likely coexist with diffusion models, with each paradigm supporting different facets of medical image analysis.

Looking ahead, the future of GANs in healthcare is likely to be shaped by three major trends: hybrid GAN – diffusion architectures that combine stability with high-frequency detail; thorough, task-based, clinically focused validation to close the gap between algorithm development and real-world deployment; and greater transparency together with stronger ethical controls to ensure that synthetic medical data is trusted and safely integrated into clinical workflows. GANs continue to be one of the most dominant and effective generative paradigms in medical imaging. With ongoing innovation in model design, clinical evaluation, and interpretability, GAN-based approaches are well-positioned to drive further advances toward robust, scalable, and ethically aligned AI technologies in data-driven healthcare.

References

1. Makhlouf, A., Cherradi, B., El Hassani, A.: The use of generative adversarial networks in medical image augmentation: A systematic review. Neural Computing and Applications 35, 27873–27902 (2023)
2. Kebaili, A., Hameurlaine, A., Teyssier, R.: Deep learning approaches for data augmentation in medical imaging. Journal of Imaging 9(3), 56 (2023)
3. Skandarani, Y., Jodoin, P.-M., Lalande, A.: GANs for medical image synthesis: an empirical study. Medical Image Analysis 68, 101989 (2021)
4. Li, J., Zhang, Q., Chen, L.: Multicenter PET image harmonization using generative adversarial networks. European Journal of Nuclear Medicine and Molecular Imaging 51(8), 2234–2245 (2024)
5. Bao, Y., Lin, H.: Medical image super-resolution reconstruction using generative adversarial networks. Advances in Computational Engineering 12(2), 10833 (2024)
6. Heng, Y., Kim, M., Park, S.: Survey: Application and analysis of generative adversarial networks in medical images. Artificial Intelligence Review 58(39) (2024)
7. Dev, S., Ghose, A., Swaroop, K.: Application of generative adversarial networks for ophthalmology image domains. Eye and Vision 9, 6 (2022)
8. Huynh, K., Deshpande, G.: Applications of generative adversarial networks to structural and functional MRI-based diagnostic classification of brain disorders. Frontiers in Neuroscience 18, 1333712 (2024)
9. Singh, A., Kumar, D., Saini, R.: A survey on training challenges in generative adversarial networks for biomedical image analysis. Artificial Intelligence Review 56(12), 15201–15242 (2023)
10. Müller-Franzes, G., Henschel, L., Kainz, B.: A multimodal comparison of latent denoising diffusion and GAN-based models for medical image synthesis. Scientific Reports 13, 13921 (2023)
11. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Nets. In: Ghahramani, Z. et al. (eds.) Advances in Neural Information Processing Systems (NeurIPS), vol. 27, pp. 2672–2680 (2014)
12. Chuquicusma, M. J. M., Hussein, S. A., Burt, J., Bagci, U.: How to fool radiologists with generative adversarial networks? A visual Turing test for lung nodule simulation. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 240–244 (2018)
13. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs trained by a two-time-scale update rule converge to a local Nash equilibrium. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 6626–6637 (2017)
14. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.: Improved training of Wasserstein GANs. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 5767–5777 (2017)
15. Preechakul, K., Sun, W., Esser, P., Rombach, R.: Diffusion models beat GANs on medical image reconstruction. In: MICCAI Workshop on Deep Generative Models, pp. 15–29 (2022)
16. Lyu, Y., Zhang, F., Liu, X., Li, S.: Hybrid adversarial–diffusion models for high-fidelity medical image generation. Medical Image Analysis 89, 102969 (2023)