



REVIEW

## A Survey of Generative Adversarial Networks for Medical Images

Sameera V. Mohd Sagheer<sup>1,\*,#</sup>, U. Nimitha<sup>2,#</sup>, P. M. Ameer<sup>2</sup>, Muneer Parayangat<sup>3</sup>, Mohamed Abbas<sup>3</sup>  
and Krishna Prakash Arunachalam<sup>4</sup>

<sup>1</sup>Department of Biomedical Engineering, KMCT College of Engineering for Women, Kozhikode, 673601, Kerala, India

<sup>2</sup>Department of Electronics and Communication Engineering, National Institute of Technology Calicut, Kozhikode, 673601, Kerala, India

<sup>3</sup>Electrical Engineering Department, College of Engineering, King Khalid University, Abha, 61413, Saudi Arabia

<sup>4</sup>Departamento de Ciencias de la Construcción, Facultad de Ciencias de la Construcción Ordenamiento Territorial, Universidad Tecnológica Metropolitana, Santiago, 7800002, Chile

\*Corresponding Author: Sameera V. Mohd Sagheer. Email: sameeravm@gmail.com

#These authors contributed equally to this work

Received: 25 April 2025; Accepted: 08 August 2025; Published: 26 February 2026

**ABSTRACT:** Over the years, Generative Adversarial Networks (GANs) have revolutionized the medical imaging industry for applications such as image synthesis, denoising, super resolution, data augmentation, and cross-modality translation. The objective of this review is to evaluate the advances, relevances, and limitations of GANs in medical imaging. An organised literature review was conducted following the guidelines of *PRISMA* (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The literature considered included peer-reviewed papers published between 2020 and 2025 across databases including PubMed, *IEEE Xplore*, and Scopus. The studies related to applications of GAN architectures in medical imaging with reported experimental outcomes and published in English in reputable journals and conferences were considered for the review. Thesis, white papers, communication letters, and non-English articles were not included for the same. *CLAIM* based quality assessment criteria were applied to the included studies to assess the quality. The study classifies diverse GAN architectures, summarizing their clinical applications, technical performances, and their implementation hardships. Key findings reveal the increasing applications of GANs for enhancing diagnostic accuracy, reducing data scarcity through synthetic data generation, and supporting modality translation. However, concerns such as limited generalizability, lack of clinical validation, and regulatory constraints persist. This review provides a comprehensive study of the prevailing scenario of GANs in medical imaging and highlights crucial research gaps and future directions. Though GANs hold transformative capability for medical imaging, their integration into clinical use demands further validation, interpretability, and regulatory alignment.

**KEYWORDS:** Generative adversarial networks; medical images; denoising; segmentation; translation

### 1 Introduction

Medical imaging is a fundamental component of modern healthcare, offering non-invasive methods to visualize the internal structures of the human body. It supports diagnosing, planning treatment, and monitoring a range of medical conditions, utilizing common imaging techniques such as X-rays, Magnetic Resonance (*MR*), Computed Tomography (*CT*), Ultrasound, and Positron Emission Tomography (*PET*).

These imaging methods, such as *MR*, *CT*, *PET*, and ultrasound, serve various diagnostic purposes, offering detailed insights into the body's internal structures. Each modality is suited for specific clinical



applications, with *MR* excelling in soft tissue imaging, *CT* providing high-resolution bone images, *PET* detecting metabolic activity, and ultrasound enabling real-time visualization of soft tissues. An outline of the different medical imaging techniques is provided below, outlining their specific uses and advantages in healthcare.

1. **Ultrasound (US) Images:** *US* imaging is widely utilized in diagnostic fields like cardiology, obstetrics, and gynecology due to its ability to generate high-resolution images without subjecting patients to ionizing radiation [1]. The technique works by emitting high-frequency sound waves (typically between 1 and 5 MHz) from a probe into the body. As these sound waves pass through the body, they interact with different tissue boundaries. Some of the waves are reflected back to the probe. The probe captures these reflected waves and transmits the data to the ultrasound machine. The distance between the probe and the organ boundaries information is then used to create a two-dimensional image on the screen, showing the distances and intensities of the reflections [2]. Fig. 1 illustrates the formation of an *US* image. Diagnostic ultrasound typically operates at frequencies ranging from 2 to 15 MHz. Higher frequencies produce better image resolution but have lower penetration depth due to increased absorption and attenuation. For this reason, high-frequency ultrasound is used to visualize superficial structures like the thyroid, while lower frequencies are employed for imaging deeper organs. Ultrasound is a non-invasive imaging technique, making it a preferred option in many medical procedures. However, it does have limitations. They cannot be used to image bones cannot be imaged as they block or absorb the ultrasound waves [3–6].
2. **Magnetic Resonance (MR) Images:** *MR* imaging uses magnetic fields and radio waves to generate detailed images of internal body structures that are difficult to capture with other imaging modalities. The human body is composed of billions of hydrogen atoms, which align with the magnetic field when exposed to it. This alignment causes the hydrogen atoms, which are positively charged, to orient uniformly. A pulse of radio frequency energy is applied to disrupt this alignment, causing the protons to shift. The protons emit energy when they return to the initial position. The intensity of this released energy is measured and displayed on a gray scale, forming cross-sectional images of the body. *MR* images are created using complex values that correspond to the Fourier transform of the magnetization distribution [7–9]. Fig. 2 illustrates the formation of *MR* image.
3. **Computed Tomography (CT) Images:** A *CT* scan employs computer algorithms to process multiple *X*-ray images procured from various angles. The combination of these images generate cross-sectional (tomographic) images of a given region within the scanned object. This technique is particularly useful in detecting hemorrhages and other conditions that may resemble a stroke, such as tumors or subdural/extradural hematomas [10]. However, *CT* imaging relies on ionizing radiation, and the exposure from this radiation accumulates over time. To minimize the impact of ionizing radiation, Low Dose Computed Tomography (*LDCT*) images are generated as an alternative [11–14].
4. **Positron Emission Tomography (PET) Images:** *PET* is a molecular imaging method that has quickly become a vital tool for functional imaging. *PET* works by generating images of the body based on the radiation emitted by radioactive substances introduced into the body. These substances, often tagged with short-lived radioactive isotopes like Carbon-11, are created by bombarding standard chemicals with neutrons. When a positron emitted by the radioactive material interacts with an electron in the tissue, gamma rays are released, which the *PET* scanner detects. *PET* can visualize blood flow and biochemical processes. Unlike structural imaging methods, *PET* focuses on the functionality of organs and the nervous system. Despite its valuable applications, the technology is costly and not widely available [15–17]. The gamma-ray detectors are used to identify pairs of gamma photons emitted in opposite directions, which are captured by two corresponding detector elements. Once these events

are detected, the analog front-end circuitry produces an event signal. If the amplitude of the incoming gamma-ray pulse exceeds a predefined threshold, a trigger signal is generated by the analog front-end. This signal is then sent to the time-to-digital converters (*TDCs*), which convert the time interval of the detected event into a digital representation.

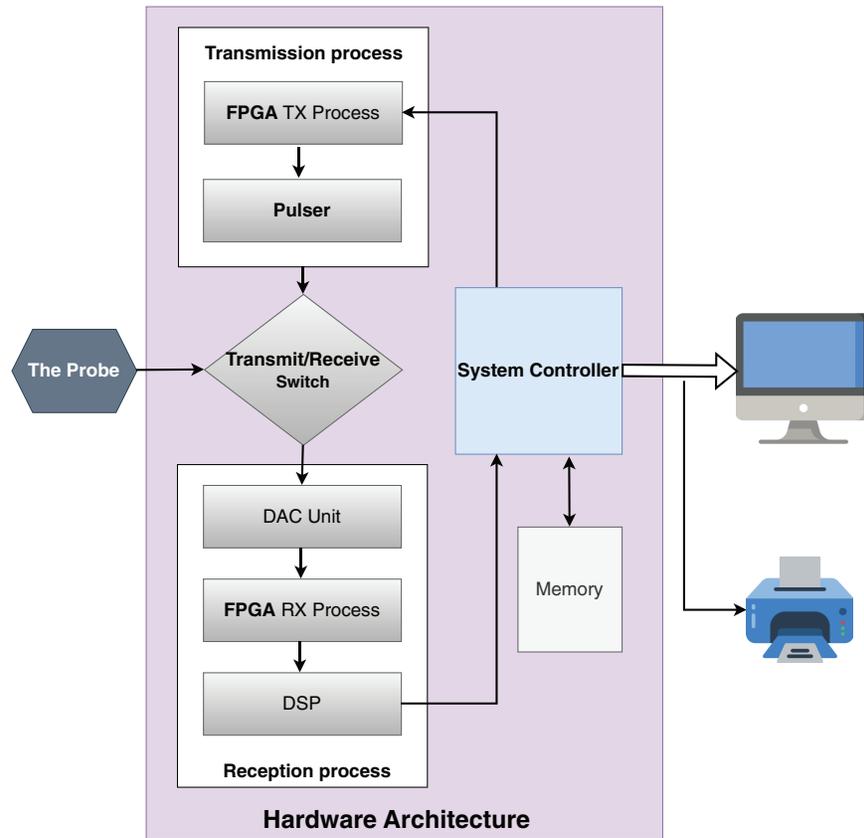


Figure 1: Formation of *US* image [2]

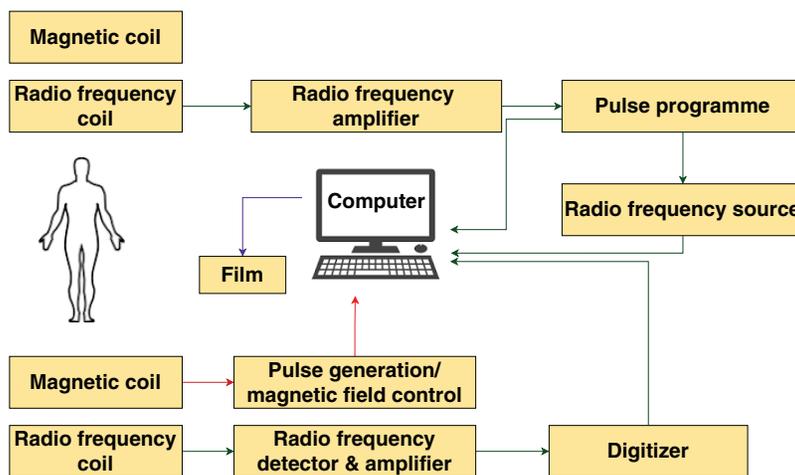


Figure 2: Formation of *MR* image [8]

Each imaging modality provides unique understanding, helping clinicians in making better decisions. The complexity and quantity of medical imaging data emphasizes for advanced computational tools to support analysis and interpretation. Generative Adversarial Networks (GANs) are well-suited to the unique expectations of medical imaging due to their ability to generate highly realistic images unlike the traditional discriminative models that only classify patterns. This capability is especially important in medical imaging, where annotated datasets are often limited, imbalanced, or expensive to obtain thus improving model robustness and performance.

Moreover, GANs excel at image to image translation tasks or enhancing image quality thus making them ideal for applications where clarity and detail are critical for diagnosis. These strengths collectively make GANs a powerful tool for advancing medical imaging. The following section gives a detailed introduction to GANs.

### 1.1 Overview of Generative Adversarial Network (GAN)

Recent advancements in computing power and big data analysis have significantly boosted the development of Artificial Intelligence (AI)—systems that mimic human cognitive abilities, such as learning, problem-solving, and decision-making [18–22]. AI can process and analyze large datasets efficiently. Machine Learning (ML), a subset of AI, learns from data by identifying patterns and features [23]. Two fundamental types of machine learning are supervised learning and unsupervised learning. While supervised learning [24,25] requires labeled data for training, unsupervised learning discovers patterns in unlabeled data, making it more applicable in scenarios where labeling is infeasible. Among these, supervised learning is the most widely utilized and successful approach. In supervised learning, algorithms are provided with a data set comprising pairs of input and output examples. The algorithm learns to map each input with its corresponding output, effectively associating input examples to output examples. A widely used form of supervised learning is classification. Once trained, supervised learning algorithms can achieve accuracy levels that exceed human performance, making them essential in various products and services. Despite these advancements, the learning process has limitations compared to human abilities. Current supervised learning approaches typically require millions of training examples [26]. To address these challenges, researchers are increasingly focusing on unsupervised learning, to reduce dependence on extensive human supervision and decrease the number of training examples needed. In general, the purpose of unsupervised learning is to extract meaningful information from a data set containing unlabeled input examples. Unlike supervised learning, unsupervised learning seeks to uncover useful patterns from unlabeled data [27,28]. Two well-known applications of unsupervised learning are clustering and dimensionality reduction.

A significant approach in unsupervised learning is generative modeling. Generative modeling aims to approximate the true data distribution  $p_{data}(x)$  with a model distribution  $p_{model}(x)$ . This is achieved by designing a function  $p_{model}(x; y)$  with parameters  $y$  and optimizing these parameters to make  $p_{model}(x)$  as similar to  $p_{data}(x)$  as possible. A common technique for generative modeling is maximum likelihood estimation, which minimizes the Kullback-Leibler (KL) divergence between  $p_{data}(x)$  and  $p_{model}(x)$ . Traditionally, explicit density models with simple probabilistic forms were used for such tasks. However, with advances in machine learning, more complex models have been developed to handle high-dimensional data. Diffusion models have recently gained attention in the field of computer vision for their impressive performance in generative tasks. These models consist of a two-phase process: a forward noise injection phase and a reverse reconstruction phase. During the forward phase, the original data is corrupted by adding Gaussian noise. The reverse phase consists of a neural network trained to reconstruct the original data by denoising it. Although these models can produce high quality and diverse outputs, they are computationally intensive and often suffer from slow generation times [29]. To address these limitations, Goodfellow et al. [30]

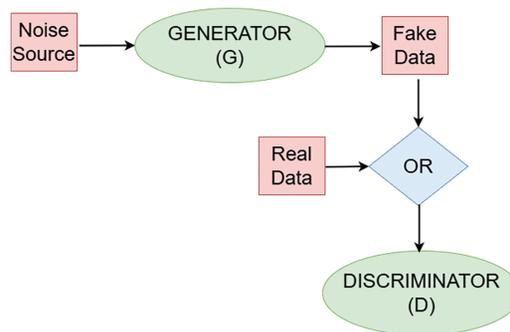
introduced Generative Adversarial Networks (GANs) [31], a novel generative model that has become a prominent tool for tasks like image denoising, translation, segmentation, and reconstruction. Generative Adversarial Networks (GANs) [32–34] enables unsupervised learning by generating new data samples from an existing data distribution.

Formally, a GAN consists of two models: a generator  $G(z; W_g)$  and a discriminator  $D(x; W_d)$ . The generator maps noise vectors  $z \sim p_z(z)$  to the data space, creating synthetic samples  $G(z)$ . The discriminator outputs a probability  $D(x) \in [0, 1]$  representing the likelihood that  $(x)$  came from the real data distribution  $p_{\text{data}}(x)$  rather than  $G(z)$ . The two models are trained simultaneously in a minimax game and the equation as follows [30]:

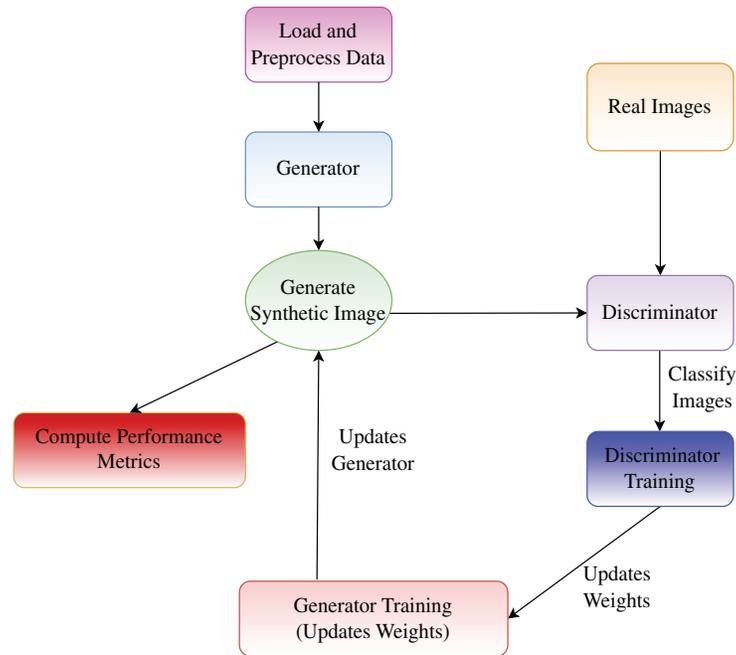
$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (1)$$

This adversarial process drives the generator to improve its outputs such that they are indistinguishable from real samples, while the discriminator becomes better at detection.

In general, the two neural networks: a generator and a discriminator, are designed to compete with each other [35]. Fig. 3 illustrates the structure of a GAN. The generator and discriminator architectures generally consist of multi-layer convolutional or fully connected layers. The generator learns the statistical properties of the training data and generates new images, while the discriminator evaluates and distinguishes between real and synthetic images [30]. Both networks serve as mappings between data domains [36]. The generator, without direct access to the real dataset, aims to create convincing synthetic images to deceive the discriminator. If the discriminator makes an incorrect classification, an error signal is generated to refine the generator's output, progressively enhancing the quality of generated images. The generator transforms a latent space into the data space, while the discriminator maps image data to a probability score, indicating whether an image is real or synthetic. If the discriminator identifies an image as real, it outputs a value close to 1, whereas for a synthetic image, it outputs a value near 0. Fig. 4 shows the training process of a GAN.



**Figure 3:** Block diagram of GAN



**Figure 4:** Flowchart of GAN training process

**Table 1:** Comparison of the advantages and limitations of specific GAN variants across different medical imaging tasks

	<i>Vanilla GAN</i> [37]	<i>DCGAN</i> [38]	<i>cGAN</i> [39]	<i>WGAN</i> [40]	<i>CycleGAN</i> [41]
<b>Training stability</b>	Low-prone to mode collapse and convergence issues	Moderate-more stable due to architectural constraints (e.g., use of CNNs)	Moderate-improved with conditioning but can still be unstable	High-uses Wasserstein loss for better convergence	High-designed for unpaired image to image translation
<b>Data requirement</b>	Moderate-requires substantial labeled data	Moderate-unlabeled data sufficient for many tasks	High-requires paired data for training	Moderate-flexible with labeled/unlabeled data	Low-does not require paired data
<b>Performance metrics</b>	Variable-depends heavily on task and tuning	Good for image synthesis; less suitable for segmentation	High accuracy in segmentation and paired synthesis tasks	Superior generative quality and stable training	Excellent in unpaired synthesis (e.g., MR to CT), not suitable for segmentation

Over the years, GANs have undergone substantial development, starting from the original adversarial learning concept, although initial versions struggled with fully capturing the data distribution [42]. A comparison of some popular GAN models is mentioned in Table 1. The original GAN, commonly referred to as the Vanilla GAN [37], utilizes random noise as input to the generator, which synthesizes photorealistic images. The discriminator differentiates between real and generated (fake) images. In the absence of ground truth images, the generator is trained solely using adversarial loss, while the discriminator is optimized using classification loss. However, the Vanilla GAN is limited in its ability to generate images across diverse classes effectively. Training Vanilla GANs is inherently unstable and often results in generators yielding outputs

that lack coherence structure. A set of architectural constraints was proposed to overcome the un-stability issue and evaluated for Convolutional GANs, termed Deep Convolutional GANs (DCGAN) [38]. Trained discriminators demonstrate competitive performance in image classification tasks compared to other unsupervised methods. Visualization of learned filters reveals that specific filters specialize in generating distinct objects. Compared to Vanilla GAN, DCGAN replaces pooling layers with strided convolutions in the discriminator and fractional-strided convolutions in the generator. Additionally, it incorporates batch normalization in both the generator and discriminator. The emergence of DCGAN brought improvements in image fidelity [38], and WGAN later addressed challenges related to mode collapse and training instability [40]. Vanilla GAN and DCGAN exhibit several limitations, one of which is the inability to control the generated outputs. For example, while a GAN can train a generator to produce images of digits (0–9) from random noise, practical applications often require generating a specific image. This limitation can be addressed by incorporating an additional input to guide the generation process. Previously, the generative model was  $p_g(x)$ . Now, it is designed to produce  $p_g(x|c)$ , where  $c$  is a conditional input used to control the generation process. This input  $c$  can be a string of codes representing the desired output or intent. GAN models often encounter challenges like mode collapse, which hinder their ability to provide meaningful learning curves that are crucial for debugging and hyper-parameter tuning. This issue can be addressed by using Wasserstein GAN (WGAN) [40], which incorporates the Wasserstein distance, also known as the Earth Mover's distance. In the WGAN architecture, the discriminator is replaced with a critic, which evaluates the Wasserstein distance rather than performing binary classification.

GANs generate images by sampling from a probability distribution, while CycleGANs [41] perform image-to-image translation between two domains. CycleGAN establishes a mapping  $G: X \rightarrow Y$  such that the output  $\hat{y} = G(x)$ ,  $x \in X$ , is indistinguishable from real images  $y \in Y$  by an adversary trained to differentiate  $\hat{y}$  from  $y$ . Unlike conventional GANs, CycleGAN employs two generators and two discriminators to enable bidirectional translation between the domains. To enhance translation quality, CycleGAN incorporates cycle consistency loss in addition to adversarial loss, ensuring that a translated image can be mapped back to its original domain. CycleGAN enabled transformations between image domains without the need for paired datasets [41], and PGAN introduced a stepwise training method to generate increasingly detailed images [43]. Subsequent models like SAGAN concentrated on identifying important image areas and capturing long-range dependencies [41]. More recent innovations include RANDGAN, which improves segmentation for anomaly detection and outperforms earlier GAN frameworks in the medical context [44]. DGGAN focuses on generating anonymized brain vascular imagery using MRA patches [45], and ED-GAN integrates VAEs with GANs to enhance generative performance [46]. These advancements collectively demonstrate the substantial progress of GANs, broadening their applicability in areas such as medical image synthesis, anomaly detection, and complex data representation.

With the variety of GAN models available for different applications, the base model or its specific variants can be selected based on the task. For instance, DCGAN is suitable for image generation, SRGAN for super-resolution, U-Net-based GANs for segmentation, and CycleGAN for image translation.

GANs [47–50] have demonstrated their potential in various fields, including image generation, style transfer, and data augmentation. Their ability to generate high-fidelity synthetic images without extensive labeled datasets makes them valuable for medical imaging. The following are the applications of GANs [51] in medical imaging:

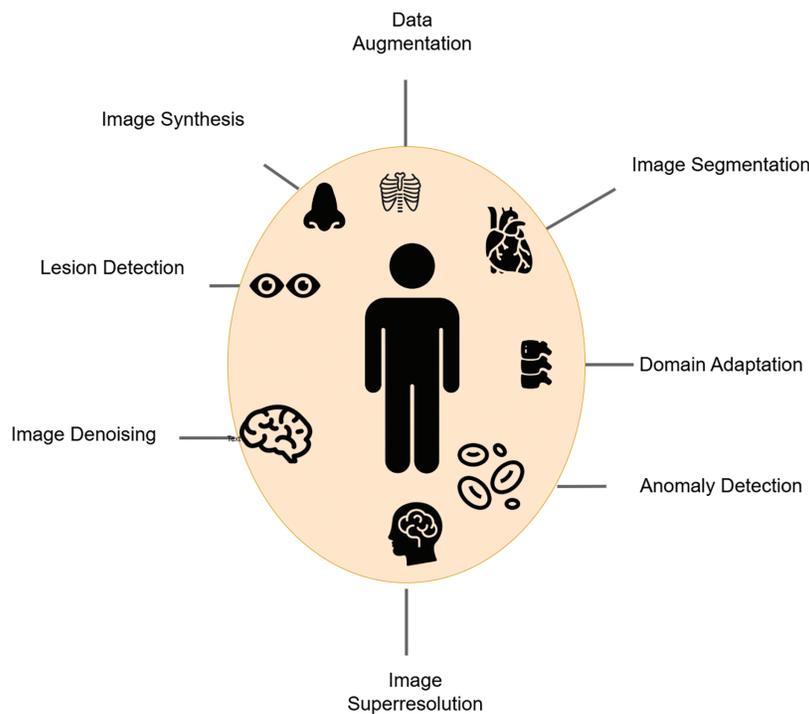
- Image Denoising: Medical images, such as low-dose CT scans or accelerated MRI, often have lower quality due to noise or reduced resolution. GANs enhance these images by reducing noise and increasing resolution [39,52–54].

- **Image Segmentation:** Segmentation is a crucial task in medical imaging, as it helps in determining and extracting areas of interest from medical images which has lately been well achieved using GANs [55–57].
- **Image Super Resolution:** Super-Resolution enhances the resolution thereby improving the clarity and detail of anatomical structures. GANs consist of the generator and the discriminator that work in opposition to refine the super resolution process [58,59].
- **Image Translation:** Translation involves converting images from one modality to another. GANs have demonstrated great ability in medical image translation by generating realistic images while preserving anatomical accuracy [60–63].
- **Image Reconstruction:** GANs enable the generation of medical images from limited data. For instance, they can generate 3D images from 2D slices or convert MR [64–67] into CT images, facilitating multi-modality analysis.
- **Data Augmentation:** GANs can create synthetic medical images that capture realistic variations, helping to balance datasets and improve model training. This approach is particularly useful when dealing with rare diseases or underrepresented classes [68–72].
- **Anomaly Detection:** By learning the distribution of normal anatomical structures, GANs can identify deviations from this norm, aiding in the detection of anomalies [73,74] and for early disease detection and screening [75–78].
- **Domain Adaptation:** Variations in medical images resulting from differences in imaging devices, acquisition protocols, or healthcare institutions can affect model performance. GANs facilitate domain adaptation, allowing models trained on one dataset to generalize effectively to other datasets [79–81].

The use of GANs in medical imaging represents a significant improvement in healthcare. GANs offer distinct advantages in addressing the challenges of medical imaging, primarily due to their capacity to produce highly realistic images, even with limited supervision which is in contrast to conventional models that focus solely on classification or detection tasks. This is important in the medical domain, where labeled datasets are often scarce, imbalanced, or costly to acquire. In addition, GANs are effective in tasks that require one form of medical image to be transformed into another, such as improving the resolution of scans, converting between imaging modalities. These applications are vital for accurate clinical assessment.

Ongoing research and technological progress continue to expand the scope and impact of GANs in healthcare [82]. For instance, *CycleGAN* is employed for domain adaptation, facilitating the translation of images across different modalities [41], while *pix2pix* supports tasks like resolution improvement and denoising through image-to-image translation [83]. A further example is *UNITGAN*, which allows cross-modal image integration by establishing a shared latent space between modalities. In contrast, *ProGAN* contributes to the generation of high-resolution images [43]. These GAN models have played a key role in enhancing both the diversity and quality of medical images, which in turn supports improved healthcare analysis and outcomes. Fig. 5 summarizes the application of GAN in medical imaging.

Despite their potential, GANs present certain challenges. Training GANs can be computationally demanding, and achieving stable training requires careful model design and parameter tuning. The interpretability of GAN-based models is another concern, particularly in high-stakes medical applications where model transparency is essential. Ethical considerations are also crucial, especially regarding patient privacy and the potential misuse of synthetic medical images.



**Figure 5:** Applications of GAN in medical imaging

## 1.2 Methodology

This research adopts a Systematic Literature Review (SLR) approach, guided by the methodologies introduced by Kitchenham et al. [84]. The study design follows the structured process outlined in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework. This ensures a clear, consistent, and reproducible process for gathering and analyzing data. All key literature included in this review was selected in accordance with PRISMA's widely accepted standards for conducting systematic reviews. The following steps were included in the search strategy.

1. **Databases Searched:** The review targets scholarly, peer-reviewed articles sourced from well-established databases, including ScienceDirect, SpringerLink, the ACM Digital Library, IEEE Xplore, PubMed, Scopus and Web of Science. At the outset of this study, a total of 500 articles were initially selected for review. The articles were sourced from literature published between 2020 and 2025, with the final selection process carried out between *November 2024* and *January 2025*. Following a thorough evaluation, 167 articles were shortlisted based on their relevance to applications, challenges, and recent advancements. The deliberate focus on recently published literature reflects our commitment to providing a forward-looking and state-of-the-art analysis.
2. **Search Terms and Keyword Strategy:** To capture the full range of applications of GANs in medical imaging, we developed a structured list of keywords representing both the core concept (GANs) and specific application domains in medical imaging. Table 2 indicates the keywords used for selecting the data and the number of selected papers under each category.
3. **Boolean Operators and Search String Construction:** Boolean operators were applied to systematically combine the two groups of keywords. The operator *OR* was used within each group to capture synonyms and variations, while the operator *AND* was used to combine the technology-related terms with the application-related terms. The following is an example of the search string used:

(“Generative Adversarial Networks” or “GAN” or “GANs”) AND (“Medical Imaging” or “Medical Image Denoising” or “Medical Image Super Resolution” or “Medical Image Segmentation” or “Medical Image Translation” or “Medical Image Reconstruction” or “Medical Data Augmentation”).

This logic ensures that the search retrieves studies that discuss GANs in any of the specified application domains within medical imaging.

4. **Search Execution and Documentation:** The search was performed using both keyword and *MeSH* term combinations (where applicable, such as in PubMed). Filters were applied to include only peer-reviewed journal articles and conference papers published in English. Search results were exported to a reference management tool (e.g., Zotero or EndNote), and duplicates were removed prior to screening.
5. **Screening Process:** Following retrieval, titles and abstracts were screened independently by two reviewers. Full texts were then assessed for eligibility based on predefined inclusion and exclusion criteria. [Table 3](#) indicates the inclusion and exclusion criteria used for selecting the articles.
6. **Quality Assessment Method:** To assess the methodological quality and reporting transparency of included studies, the *CLAIM* (Checklist for Artificial Intelligence in Medical Imaging) guideline was used. This checklist evaluates critical domains relevant to *AI* studies, including dataset characteristics, model evaluation procedures, validation methodology, and reproducibility. [Table 4](#) adapted from the *CLAIM*, indicates the criterion used to evaluate the methodological quality of included studies. Only studies meeting key *CLAIM* criteria were retained for final synthesis to ensure reliability of the review findings. Each study was assessed independently by two reviewers. Disagreements were resolved through discussion or by consulting a third reviewer.

**Table 2:** Keywords used for article selection

Keyword	Paper count
GANs in medical imaging	145
GANs in medical image denoising	25
GANs in medical image super resolution	27
GANs in medical image segmentation	23
GANs in medical image translation	28
GANs in medical image reconstruction	30
GANs in medical data augmentation	10

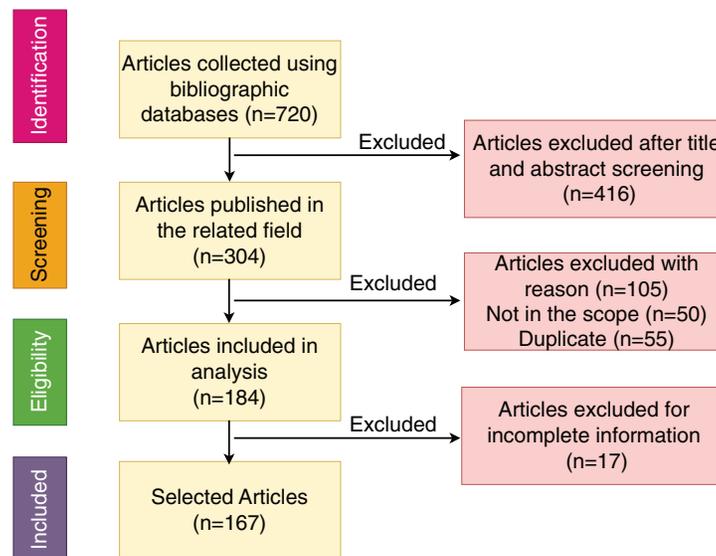
**Table 3:** Inclusion and exclusion criteria for selected articles

	Inclusion criteria	Exclusion criteria
Types of study	Original and review article	Thesis, white papers, communication letters and editorials
Language	Research published in English language	Duplicate and non-English articles
Publication Year (For result analysis and application)	Articles published in 2020–2025	–
Source	Publications in peer reviewed journals and conferences	–
Settings	GANs in medical imaging	Not related to medical imaging
Intervention	ML and GAN	Traditional methods

**Table 4:** CLAIM based quality assessment criteria applied to included studies

S. No.	CLAIM Domain	Assessment Criteria
1	Study design & objective	Clear research objective; rationale for using GANs clearly stated
2	Dataset description	Data source and acquisition protocols described; dataset size, diversity, and labeling provided
3	Data splitting & Preprocessing	Clear description of training/validation/test splits; preprocessing steps explained
4	Model architecture	GAN architecture detailed (e.g., type of GAN, loss functions, training parameters)
5	Evaluation metrics	Appropriate and consistent metrics used (e.g., PSNR, SSIM, Dice, AUC)
6	Validation strategy	Cross-validation or external validation employed to avoid overfitting
7	Ground truth & annotation	Reference standards and annotation protocols clearly described
8	Reproducibility & code availability	Whether code, models, or datasets were publicly shared
9	Clinical relevance & interpretation	Results discussed in the context of clinical utility, risks, or limitations
10	Bias and limitations disclosure	Potential sources of bias and study limitations explicitly acknowledged

This approach ensured a comprehensive, reproducible, and methodologically sound search process in alignment with PRISMA standards. The PRISMA flow diagram of the article selection procedure is shown in Fig. 6. PRISMA checklists can be found in the supplementary files.



**Figure 6:** PRISMA flow diagram of the article selection procedure

The remaining sections of this paper explore in detail how GANs are applied within the field of medical imaging. Section 2 discusses the various applications of GANs in medical images. Section 4 details the different challenges of GANs for clinical use under regulatory frameworks. Section 5 gives an insight to the other generative models present. Finally, the Section 6 concludes the paper providing an insight into the limitations and future of GANs in the medical field.

## 2 Medical Applications

Medical imaging application uses GAN in two separate ways; one as generator which examines the underlying data distribution and generates new (synthetic) images. The discriminator section can classify normal and abnormal images. An overview of usage of GANs in various medical applications is presented in Tables 5–7. The following subsection describes the medical imaging applications of GAN for denoising, segmentation, super resolution, translation, reconstruction and data augmentation.

**Table 5:** Summary of methods & applications (Arranged by Year)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
1.	Ali et al. [85]	2024	Medical image processing	–		Overview of GAN for medical image processing.
2.	Tu et al. [86]	2024	Image translation	summer2winter, facade dataset	FID and EGF	Introduces the unpaired image-to-image translation method using a diffusion adversarial network (UNDAN). The method captures seasonal features well and enhances clarity but occasionally misinterprets details or leaves inconsistencies.
3.	Fan et al. [66]	2023	Image fusion	AANLIB	SSIM, RMSE, MI and Gradient Evaluation	U-Patch GAN, built on a GAN framework to achieve self-supervised fusion of multimodal brain images, aiming to improve the quality of the fusion process. Effectively fuses various modal images while preserving rich texture and functional information, however, but there is potential to develop more effective mechanisms to improve the generator structure.
4.	Wang et al. [87]	2023	Image translation	Cityscapes, Facades, CUHK Face Sketch, Day-Night, and Oil-Chinese Paintings	FCN-score, semantic segmentation metrics, FID, and LPIPS	Introduces an unsupervised approach to image-to-image translation using long-short cycle-consistent adversarial networks. Enhanced image quality by reducing information loss, lacks from computational overhead without proportional performance gains.
5.	Zhou et al. [48]	2023	Medical image fusion	–	–	Survey of Generative Adversarial Networks (GANs).

(Continued)

Table 5 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
6.	Esmaeili et al. [75]	2023	Anomaly detection	Head-Ct, Brain Tumor MRI, Br35h-MRI, MIAS-MAMMO, MIAS-Patches MAMMO, C-NMC-Leukemia, and Retinal OCT	Precision, Recall, Specificity, FI-Score, ROC AUC, and PR AUC	Overview of the application of GANs in Alzheimer's Disease (AD), along with an exploration of cutting-edge GAN-based methods for AD in biomedical imaging. Examines the outcomes from both data and model perspectives reveals key limitations of existing techniques, but struggled to detect breast cancer lesions due to their complexity and similarity to normal tissue.
7.	Fu et al. [52]	2022	Medical image denoising			Denoising low-dose CT images (LDCT) through a two-phase deep learning approach known as the Noisy Generation-Removal Network (NGRNet).
8.	Zhang et al. [78]	2022	Anomaly detection	CIFAR-10, ILD, HAM 10000	AUC, Accuracy, Precision, Recall and FI-Score	An unsupervised approach for deep anomaly detection using an enhanced adversarial auto-encoder. The unsupervised anomaly detection method using an improved adversarial auto-encoder with convolution block chains (CCB) replacing skip connections, offering more stable training, but struggles with anomaly detection involving temporal variations.
9.	Sun et al. [88]	2022	Medical image synthesis	COPDGene, GSP	FID, MMD and IS	Introduce a new end-to-end GAN architecture designed to generate high-resolution 3D images. HA-GAN serves as plug and play module for all existing GANs; but higher memory utilization with increase in image resolutions.
10.	Xia et al. [74]	2022	Anomaly detection	-	-	Survey of GAN based anomaly detection methods.
11.	Xun et al. [89]	2022	Medical image synthesis	-	-	Presents the origins, operating principles, and extended versions of GAN, along with a review of the most recent advancements in GAN-based methods for medical image segmentation.
12.	Singh et al. [65]	2021	Medical image generation	-	-	Survey of GAN based medical image generation methods.
13.	Suganthi et al. [47]	2021	Medical image synthesis	-	-	Survey of medical image synthesis using GAN.

(Continued)

Table 5 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
14.	Ma et al. [49]	2021	Medical image enhancement	CORN-2, Fundus Multi-Disease Diagnosis, EASE	SNR, AUC, Accuracy, Sensitivity, G-mean, Dice	Two novel constraints, illumination regularization to enhance illumination uniformity and structure loss to preserve structural details, are incorporated into the objective function to improve clinical interpretability and subsequent analysis. StillGAN, carries the risk of incorrect translations, such as altering colors or introducing lesion-like artifacts, during image enhancement.
15.	Li et al. [53]	2021	Medical image denoising	JSRT LIDC-IDRI or LIDC	PSNR, SSIM	A medical image denoising technique utilizing a conditional generative adversarial network (CGAN) to address various types of unknown noise. The method handles spatially varying and unknown noises, using noise and gradient images as conditional inputs to enhance structural specificity and preserve image details. Selecting an optimal threshold to accurately derive gradient information poses a significant challenge due to the variability in image intensity and structural patterns.
16.	Goodfellow et al. [30]	2020	Overview of GAN	-	-	Overview of generative adversarial networks.
17.	Xie et al. constraint [80]	2020	Domain adaptation	Colonoscopic, REFUGE	Dice score	A GAN, referred to as MIGAN, is employed to preserve image content during cross-domain image-to-image (I2I) translation. The model preserves image content by disentangling content features from domain information and maximizing mutual information between the content features of source and translated images.

Table 6: Summary of methods &amp; applications (Arranged by Year) (Contd.)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
18.	Sun et al. [71]	2020	Medical image segmentation	BRATS17, LIVER100	Dice Score	MM-GAN, simulates the distribution of real data and generates new samples from the limited data distribution to enrich the training set. This model can translate label maps into 3D MR images without the concern of violating the underlying pathology. MM-GAN enables the translation of label maps into 3D MR images while ensuring the integrity of pathological features.

(Continued)

Table 6 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
19.	Kazemina et al. [51]	2020	Medical image analysis	–	–	Overview of <i>GANs</i> for medical image analysis.
20.	Armanious et al. [90]	2020	Image translation	Anonymized datasets	SSIM, PSNR, MSE, VIF, UQI, LPIPS	MedGAN is utilized for medical image-to-image translation, functioning at the image level in a fully end-to-end manner. Model integrates a conditional adversarial framework with novel non-adversarial losses and a CasNET generator architecture to improve global consistency and preserve high-frequency details in the results, but is unable to process 3D medical volumes effectively.
21.	Kim et al. [39]	2020	Medical imaging	LIDC-IDRI, AAPM	SSIM	Conditional Generative Adversarial Networks ( <i>CGAN</i> ) are applied to low-dose chest imaging, a widely used diagnostic tool in medical imaging. The technique offers universal applicability to various medical images without complex parameter tuning, image information loss, or high modality dependence, while enabling rapid inference through optimized convolutional operations. Using batch normalization in the generator can hinder the production of diverse and realistic images.
22.	Zhang et al. [91]	2019	Medical image denoising	4D Extended Cardiac Torso (XCAT) phantom	Qualitative Analysis	<i>SPECT</i> Image Denoising: A study assessing the application of Generative Adversarial Networks ( <i>GAN</i> ) for denoising static <i>SPECT</i> images. The proposed method can reduce <i>SPECT</i> image noise, enabling lower injection doses or shorter acquisition times while preserving image quality suitable for clinical diagnosis. The technique addresses only a single problem in low-level medical image vision, resulting in limited clinical applicability and significant residual noise after SR processing.
23.	Pan et al. [23]	2019	Survey on <i>GAN</i>	–	–	Overview on Generative Adversarial Networks ( <i>GANs</i> ).

(Continued)

Table 6 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
24.	Dar et al. [61]	2019	MRI Synthesis	MIDAS, IXI	SSIM, PSNR	MRI synthesis using a conditional generative adversarial network. The model synthesizes distinct contrasts from a single modality, with applications in multi-contrast brain MRI and improved accuracy using cross-sectional correlations. The model has the disadvantage of multi-modal imaging for large cohorts due to its high economic and time costs.
25.	Mahapatra et al. [92]	2019	Medical image analysis	EYEPACS	SSIM, PSNR, S3, P	Progressive Generative Adversarial Networks ( <i>P-GANs</i> ) based super resolution technique. The superior performance of the proposed method is evident in multiple medical image modalities like retinal vessel segmentation, micro aneurysm detection, MRI super-resolution, and cardiac LV segmentation. It struggles with minor anatomy and pathologies, limiting early disease detection.
26.	You et al. [93]	2019	Image reconstruction	Tibia, Abdominal	PSNR, SSIM, IFC	Deep learning method designed to effectively reconstruct <i>HRCT</i> images from their <i>LRCT</i> images. Achieved promising results in preserving anatomical details and reducing image noise using both supervised and semi-supervised learning approaches. The model requires significantly longer training than standard GAN methods (1–2 days) and struggles to recover subtle features.
27.	Creswell et al. [36]	2018	Study of GAN	–	–	Survey of generative adversarial networks.
28.	Yang et al. [94]	2018	Medical image denoising	2016 NIH-AAPM Mayo Clinic Low Dose CT Grand Challenge	PSNR, SSIM	A <i>CT</i> image denoising approach utilizing a Generative Adversarial Network ( <i>GAN</i> ) with Wasserstein distance and perceptual similarity. The WGAN framework reduces over-smoothing, yielding images with higher PSNR and more faithful statistical properties. The model risks losing critical features and requires re-tuning for datasets with varying noise properties.

(Continued)

Table 6 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
29.	Xue et al. [56]	2018	Medical image segmentation	MICCAI BRATS	Dice score	An end-to-end adversarial neural network, known as <i>SegAN</i> , designed for medical image segmentation tasks. The multi-scale loss function integrates CNN feature differences across multiple layers, enabling more effective training compared to classic GANs with separate losses.

Table 7: Summary of methods &amp; applications (Arranged by Year) (Contd.)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
30.	Han et al. [57]	2018	Medical image segmentation	Private dataset	Average pixel-level accuracy, Dice coefficient	Spine-GAN, a Recurrent Generative Adversarial Network, is employed for the segmentation of various spinal structures. The model integrates atrous convolution and autoencoder for fine-grained semantic representation, employs long short-term memory module (LSTM) to model spatial pathological relationships, and uses an auxiliary CNN for error correction. This model is limited to segmenting and classifying spinal structures, falling short of enabling human-like understanding of MR images.
31.	Li et al. [95]	2018	Medical image segmentation	I3A, MIVIA	Precision, Recall, SEG	Generative adversarial networks ( <i>cC-GAN</i> ) are utilized for the robust segmentation of various HEp-2 datasets. The framework leverages auxiliary classifiers to tackle HEp-2 cell segmentation, reducing over-fitting and enhancing transfer capacity in deep learning networks.
32.	Sanchez et al. [59]	2018	MR image generation	ADNI	SSIM, PSNR	An adversarial learning method for generating high-resolution MRI scans from low-resolution images. The model incorporate adversarial loss for realism and content loss to minimize differences between real and generated images, achieving promising results in high-factor downsampling for 3D medical imaging super-resolution. A limitation is the need for improved architectures and the integration of additional perceptual terms in the loss function to enhance the quality of the generated volumes.

(Continued)

Table 7 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
33.	Hu et al. [96]	2018	Medical image synthesis	ICPR	Precision, Recall, F1-score	GAN based method with a novel loss formulation to enable robust cell-level visual representation learning in an unsupervised environment. The approach relies on a novel pipeline combining cell-level visual representation learning and nuclei segmentation to distinguish cellular elements. Incorporating the gradient penalty into the network architecture necessitates computing second-order derivatives, which significantly increases training time.
34.	Nie et al. [97]	2018	Medical image synthesis	ADNI, Pelvic	PSNR, MAE, Dice Index	An adversarial learning strategy applied to improve the modelling of the FCN. The supervised deep convolutional model is using adversarial learning to estimate target images from source images, even across modalities, with an image gradient difference loss to reduce blurriness in generated outputs. A limitation is the dependency on paired source-target images from a large cohort, as acquiring multi-modal imaging at this scale is often constrained by significant economic and time costs.
35.	Bissoto et al. [98]	2018	Skin lesion synthesis	2018 ISIC Challenge	AUC	Focuses on skin lesion synthesis with GAN. A visual comparison showcasing high-resolution samples with fine-grained details, including accurately placed and sharply defined malignancy markers, resulting in visually compelling images. The model requires annotated data for image generation.
36.	Han et al. [64]	2018	MR image generation	BRATS 2016	Visual turing test, accuracy	Generates synthetic multi-sequence brain Magnetic Resonance (MR) images through the use of Generative Adversarial Networks (GANs). This research demonstrates the potential of WGAN to generate realistic multi-sequence brain MR images, enabling applications like data augmentation and enhanced physician training. A limitation of this work is its reliance solely on sagittal MR images, with coronal and transverse images yet to be generated.

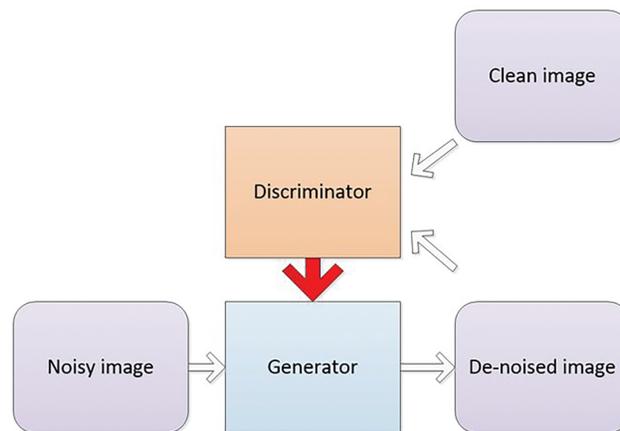
(Continued)

Table 7 (continued)

S. No.	Name	Year	Purpose	Dataset	Metrics	Remark
37.	Udrea et al. [99]	2017	Skin lesion detection	Skin lesion dataset	Accuracy	Accuracy an artificial intelligence approach for segmenting pigmented and non-pigmented lesions using GAN. The ability to achieve these results using smartphone-acquired images from laypersons highlights its value for mHealth and eHealth dermatological applications. The approach requires improved GAN architecture, a more diverse dataset, and analysis on color channels beyond RGB.

## 2.1 Denoising

Image denoising is a vital preprocessing act in analysis, as all types of medical images are susceptible to noise [100–104]. The sources of noise in medical imaging can be categorized into sensor-related, acquisition-related, and radiation-related factors. Computed Tomography (CT) is a widely utilized technique for disease diagnosis, but it carries the potential risk of radiation exposure [105,106]. Reducing radiation levels can lead to increased noise in CT images. Reconstruction of Low-Dose CT (LDCT) images offers an effective approach to address this issue. Fig. 7 shows a GAN based framework for medical image denoising. A GAN utilizing Wasserstein distance (WGAN) and perceptual similarity was applied to denoise CT images in [94]. The perceptual loss minimized noise by aligning output and ground truth features in a defined space, while the GAN shifted the noise distribution from strong to weak. Wasserstein distance was used to compare the distributions of normal-dose CT (NDCT) and LDCT data. Feature extraction was performed using Convolutional Neural Networks (CNN) based on the Visual Geometry Group (VGG). Performance metrics such as PSNR and SSIM were employed to evaluate the outputs of different networks, with WGAN–VGG achieving superior performance. Huang et al. in [107] presented a denoising method (DU–GAN) which utilized U-Net-based discriminators to assess the global and local variations in the denoised and normal images. A denoising method based on Conditional GAN was popularized by Li et al. in [53] in which the image context relationship and structural information was preserved. The method was tested on LIDC dataset and was seen to outperform the state of the art works. Zhu et al. had introduced a denoising method based on GAN in [108]. Molecular activity in human tissues was captured using Single-Photon Emission Computed Tomography (SPECT), which relies on gamma rays for image acquisition in [91]. High-noise SPECT images are input into the generator, while the discriminator assesses the generated images against real samples, specifically low-noise SPECT images. The loss, which quantifies the difference between generated and real images, is used to update both the generator and discriminator simultaneously. Both components were optimized using the Adam optimizer, with a learning rate of 0.00001 and 800 training epochs. Noise levels were assessed using the normalized standard deviation (NSD), enabling a comparison of results with and without conditional GAN denoising applied to the reconstructed SPECT images.



**Figure 7:** GAN-based framework for medical image denoising: enhancing image quality with adversarial training

## 2.2 Segmentation

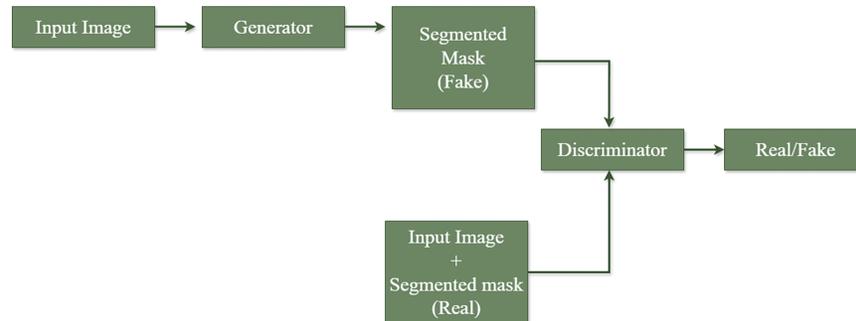
Segmentation plays a crucial role in medical image analysis. Automating the segmentation process is highly challenging due to variations in anatomical structures across different patients [55,109–111]. Skin cancer, common among individuals with fair skin, is classified into melanoma (pigmented lesions) and non-melanoma (non-pigmented lesions). Early detection of melanoma is critical. Dermoscopic images captured via smartphones can be analyzed to detect pigmented lesions. A GAN was trained using 3000 color images in [99], with the generator employing a *U-Net* architecture and the discriminator using the Adam optimizer. Training was conducted in two phases: without image rotation and with image rotation. The highest segmentation accuracy was achieved when training included image rotation. However, the proposed method lacked pre-processing stages. Future improvements in the same work were involved in enhancing the GAN architecture, utilizing larger datasets, and analyzing specific color channels beyond conventional *RGB* spaces. *X-ray* imaging, which uses high-frequency electromagnetic waves, depends on the radiological density of tissues to determine the level of absorption. Chest *X-rays* can detect infections, tuberculosis, cancer, and other chronic chest conditions. Segmenting postero-anterior chest *X-ray* images involves isolating the left and right lung fields and the heart. Since chest *X-rays* are 3D projections onto 2D images, overlapping structures complicate segmentation. The Structure Correcting Adversarial Network (SCAN) framework [99] applied adversarial techniques for semantic segmentation, comprising a segmentation network and a critic network that were trained jointly. Both the segmentation and critic networks in SCAN employed Fully Convolutional Networks (FCNs). While FCNs were effective for *RGB* images, SCAN used FCNs for grayscale chest *X-ray* images, enabling extraction of high-level data representations. Metrics such as Intersection-over-Union (*IoU*) and Dice coefficient were used to evaluate performance. The SCAN framework was trained on 247 chest *X-ray* images from the Japanese Society of Radiological Technology (JSRT) dataset, which included 154 images with lung nodules and 93 without. Additionally, 138 chest *X-ray* images from the Montgomery dataset were used, with 117 for training and 21 for testing. Approximate *IoU* values for the JSRT and Montgomery datasets were 95.1% and 93%, respectively. This method represented a successful application of convolutional neural networks for accurate segmentation results. Spinal *MR* image enabled the detection of defects in the spinal cord, vertebrae, intervertebral discs, and more. The *Spine-GAN* [57] method employed a dynamic optimization algorithm combined with a hybrid learning strategy. The *Spine-GAN* included a segmentation network with three components: an encoder, a local long short-term memory module (*LSTM*), and a decoder. An Atrous Convolutional Autoencoder (*ACAE*),

was designed for spinal image representation and pixel-level classification. The *LSTM*, a recurrent neural network (*RNN*), modelled spinal structural details by leveraging spatial correlations, thereby enhancing network performance. A Convolutional Neural Network (*CNN*) discriminator mitigated overfitting and incorporated a robust learning strategy with a flexible optimization algorithm. The discriminator accepted inputs from either the Segmentor or the ground truth. Spinal *MR* data from 253 patients, captured using 1.5 *T* equipment, was used for training and testing, comprising 5343 samples of neural foramen images, 1818 disc images, and vertebrae images. Data was split into five subgroups, with one group used for training and the others for testing. Performance metrics included pixel-level accuracy, Dice coefficient, specificity, and sensitivity, yielding an accuracy of 96.2%, specificity of 89.1%, and sensitivity of 86%. Limitations include its inapplicability to all *MR* image types and the lack of prior clinical spine knowledge in the diagnosis. Future improvements could address these limitations.

Fig. 8 shows a framework for medical image segmentation using GAN. A healthy immune system defends against foreign bodies in the body. In autoimmune disorders, the immune system attacks healthy tissues. Human epithelial type 2 (*HEp-2*) cell images can be analyzed to diagnose autoimmune disorders [95]. Due to the large variety of patterns, the available datasets were limited. *cC-GAN*, a transfer learning framework in *GAN*, is a powerful segmentation approach for *HEp-2* datasets, addressing overfitting and improving transfer capacity. The *cC-GAN* used three losses: *L1* Loss for label prediction, *GAN* Loss to help the discriminator distinguish between outputs, and Softmax Loss for classification. The generator in *cC-GAN* employed Residual *U-Net* (*RU-Net*) architecture, while an additional epoch training scheme (*AET*) was used for stable training of the generator and discriminator. Two datasets were used: the first dataset contained 252 specimens from six categories (Homogeneous, Speckled, Nucleolar, Centromere, Golgi, Nuclear membrane, and Mitotic spindle), while the second dataset contained 28 green-channel *HEp-2* images with six categories (Centromere, Homogeneous, Fine Speckled, Coarse Speckled, Nucleolar, and Cytoplasmic). The smaller size of the second dataset impacted fine-tuning, making it less effective for segmentation. Segmentation accuracy and precision were used for evaluation, with *cC-GAN* achieving 86.15% accuracy compared to Fully Convolutional ResNet (*FCRN*), which achieved 87.29%. The *FCRN*'s deeper network architecture improved accuracy but also increased the risk of overfitting due to a higher number of parameters. *cC-GAN* effectively distinguished between real and synthetic samples and classified cells. A multi-scale *L1* loss function was effective for semantic segmentation tasks [56]. Training for *SegAN* followed a conventional *GAN* min-max framework, involving a Segmentor network (*S*) and a Critic network (*C*). Both networks were trained using back propagation, optimizing the multi-scale *L1* loss. The Segmentor is a fully convolutional encoder-decoder with a kernel size of  $4 \times 4$  and a stride of 2 for downsampling. The same loss function was applied for training both the Segmentor and the Critic. The method was evaluated using 220 high-grade and 54 low-grade brain *MR* images, with Dice, Precision, and Sensitivity serving as evaluation metrics. *SegAN* compared the segmented image with the ground truth across multiple critic layers. Xue et al. in [112] had introduced a method for multi organ thorax segmentation using *U-GAN*.

The neural network architecture proposed in [113] is composed of two main components: a Segmentor and a Critic. To enhance semantic feature extraction, a novel model named *Transformer-CV-Unet* (*TCUnet*) is introduced. *TCUnet* was utilized as the generator within a *GAN* framework to perform image segmentation, improving both robustness and computational efficiency. Once the generator segmented the target images, the Critic integrated the latent features with hierarchical information from various modalities. Additionally, a hybrid adversarial mechanism incorporating a multi-phase *CV* energy functional is developed. This integrated framework, referred to as *AdvTCUnet*, leverages the strengths of both the generator and critic modules. Comprehensive evaluations on the *BraTs* 2019–2021 datasets demonstrate that the proposed method outperforms current leading approaches in brain tumor segmentation on *MRI* scans,

achieving Dice Similarity Coefficients of 0.8642 for Enhancing Tumor (ET), 0.9303 for Whole Tumor (WT), and 0.9060 for Tumor Core (TC) on *BraTs2021*. A Local Cross-Attention Unet (*LCAUnet*) was introduced by Wang et al. in [114] for the segmentation of skin lesions using edge and body features.



**Figure 8:** Framework for medical image segmentation using GAN for lesion, infection detection, etc.

### 2.3 Super Resolution

High-resolution images play a crucial role in accurate disease diagnosis. Despite advancements in medical imaging techniques, factors such as imaging conditions, equipment limitations, and environmental obstructions can result in low-resolution images [115–119]. Resolution can be improved by various methods such as enhancing the spatial resolution, interpolation, multi image super resolution methods. *MR* image resolution can be improved using *GAN* technique by taking advantage of the volumetric information in the *3DMR* image [59,120]. The proposed architecture to improve the *MRI* resolution was based on Super Resolution *GAN* (*SRGAN*). The difference between conventional *GAN* and *SRGAN* is in the convolutional filters, they are *3D* convolutional layers capable to handle volumetric information. To overcome the problem of vanishing gradient in cross entropy loss function, adversarial loss was used in the discriminator and in generator network, combination of adversarial and content loss is used. Generator network contains had six residual blocks with 32 convolutional filters of size  $3 \times 3 \times 3$ . The discriminator network had eight convolutional layers of kernel size  $3 \times 3 \times 3$ . 589 *T1*-weighted images (470 for training and 119 for testing) from the Alzheimers Disease Neuro imaging Initiative (*ADNI*) database was used and skull stripping was done for all images. Adam optimization was employed in both generator and discriminator. The generated super resolution images for upsampling factor of 2, 4 were compared against the ground truth in terms of peak signal-to-noise ratio (*PSNR*) and structural similarity index measure (*SSIM*). Both metrics yielded a better value for the upsampling factor of 2.

Histopathology is the study of presence of disease in biopsy or surgical samples using microscopes [121]. Histopathological images provides a comprehensive view of cell tissues. Different segments in a tissue are visualized by pigmenting it using different dyes. Histopathological images analysis involves segmentation, detection, feature extraction, classification, etc. A unified *GAN* framework with a newly designed loss function was introduced for the same in [96]. The loss function is derived from both *WGAN-GP* and Information Maximizing Generative Adversarial Networks (*InfoGAN*). During training, the generator (*G*) aimed to estimate the distribution  $P_g$  such that it aligns with the real data distribution  $P_r$ . This estimation was done by optimizing the Earth Mover (*EM*) distance approximation between generator (*G*) and discriminator (*D*). In the meantime the auxiliary network (*Q*) maximized the mutual information between the chosen random variables and the samples generated by *Q*. In the testing phase *Q* was capable of classifying as softmax function as the final layer in the *Q* network. Proposed unified architecture was modelled as a pipeline comprising of the following four stages-Nuclei segmentation,

Clustering/classification of cell by Q network, calculating the cell proportions and image level prediction based on cell proportions, generating interpretable cell categories by generator for different noises. Four different bone marrow datasets were examined for testing and training the network. Images from these datasets were further sub divided into cell-level images and were classified into categories such as neutrophils, myeloblasts, monocytes and lymphocytes. While training, 32 or 64 Gaussian noise variables were used and 64 categorical variable were used. While segmenting the cell images, to separate out the touched cells, morphological opening with kernel size  $7 \times 7$  was carried out. Feature extraction process could be disturbed by the color and feature contrast, this could be avoided by using non segmented histopathological images. By using such images, cells could overlap drastically. The problem of overlapping cells could be excluded by using a bounding box of size  $32 \times 32$  on non segmented image. For each cell class, precision, recall and F-Score were the segmentation evaluation parameters. Cell clustering evaluation metrics were purity, entropy and F-score. High purity value with low entropy indicated better clustering. Gradient penalty required to compute the second order derivative, so is really time consuming while training. The proposed architecture could be improvised by changing the segmentation method. It could be further improved by considering patient related data such as clinical trials, gene expression data. Anatomical segmentation of lesions and locating pathology becomes complex when anatomy or pathology is small like retinal images, cardiac MR, etc. or the image quality is too low due the acquisition process [92].

Diabetic retinopathy, glaucoma are the diseases which are diagnosed with the help of retinal fundus images. Retinal fundus image resolution are not good enough to detect microaneurysms, hemorrhages as they cover small image areas. Progressive GAN (*P-GAN*) generated a high resolution image from a low resolution input image. The proposed *P-GAN* architecture incorporated two stages-stage 1 and stage 2, with output from stage 1 being the input to the stage 2. Triplet loss function ensured quality improvement in images as they progressed from one stage to another. Low resolution version from high resolution image could be acquired by the application of Gaussian filter followed by down sampling. The proposed network could yield better quality images for high scaling factors, greater than 8. The proposed architecture with 2 stages have a set of generator (*G1*) and discriminator (*D1*) in stage 1 and another set of generator (*G2*) and discriminator (*D2*) in stage 2. The loss function for the stage 1 include Mean square error (*MSE*) and *CNN* loss terms. In contrast to stage 1, the loss function for stage 2 is the triplet loss function. Every generator block the network contains a up sampler with factor 2. Triplet loss function is a minmax function such that it minimizes the distance between output image from stage 2 and ground truth and maximizes the distance between output image from stage 2 and output image from stage 1. This dual constrains made sure that improvement was achieved both qualitatively and quantitatively. The network was trained with 5000 retinal fundus images and tested with another 1000 images. The dataset was down sampled by factors 2, 4, 8, 16, 32 and for each sampling the time taken to generate super resolved image was computed. As the up sampling factor increases, there was evident increase in the time to generate super resolved images. The proposed *P-GAN* was compared with various *SR-GAN* algorithms. *SSIM*, *PSNR* (dB), sharpness metric (*S*) [122] were the metrics for network performance evaluation and was computed by varying the scaling factor and for various noise types like Gaussian, salt and pepper noise and speckle noise. With the increase in upsampling factor, a decrease in *SSIM*, *PSNR*, *S3* metrics was estimated. Out of the three noises, low quality images with Gaussian noise, only yielded the expected results. Vessel segmentation with *P-GAN* outperform all other GAN algorithms and was compared with respect to segmentation accuracy, Specificity, area under curve (*AUC*) and sensitivity. The proposed *P-GAN* had been evaluated using cardiac MR dataset for improving the resolution and for cardiac LV segmentation.

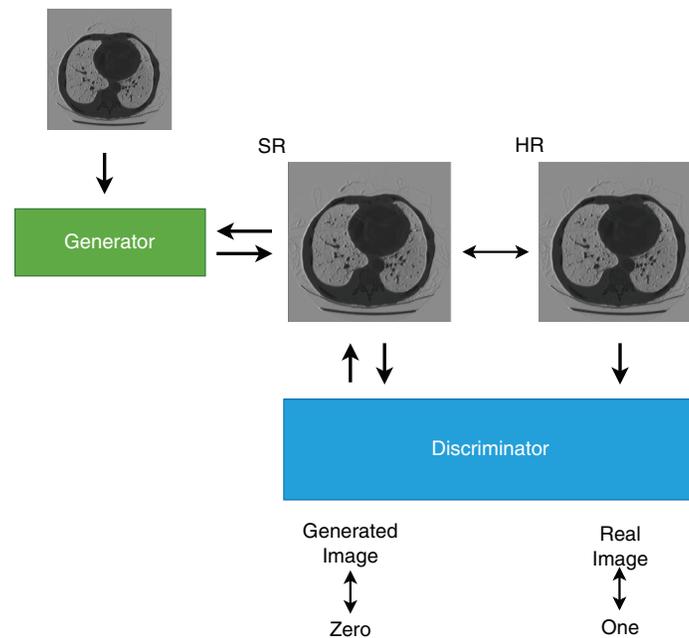
X-Ray Computed Tomography (*X-Ray CT*) uses X-Rays for screening, diagnosis, image guided surgery, etc. The quality of *X-Ray CT* images can be improved in two ways-either by using a good quality hardware or

by computationally enhancing the images. The former is not economical and also are of high radiations [123]. High radiations cause gene damages and can cause cancer. Low Dose *CT* (*LDCT*) uses lower radiation, which results in low quality *CT* images. Computational techniques can be employed to obtain High Resolution *CT* (*HRCT*) from *LDCT*. A novel residual *CNN*-based network in the *CycleGAN* framework known as Deep Cycle-Consistent Adversarial *SRCT* network (*GAN-CIRCLE*) [93] was one among the computational techniques which preserved the anatomical details in *CT* images. The proposed algorithm had the following advantages over the other conventional *GAN* architectures

- cycle consistency to ensures strong across domain consistency between *LRCT* and *HRCT*,
- exclusion of Nash equilibrium [30] problem for training *GAN*,
- omission of over fitting problem by optimizing the network,
- Inclusion of multiple cascaded layers to extract hierarchical features,
- to enhance deblurring *L1* norm is employed instead of *L2* norm.

The main challenges in recovering *HRCT* images from noisy *LRCT* images are-complex spatial variations in the images, presence of unique noise patterns, sampling and degradation makes the image blur. To handle these limitations, non linear *SR* functional blocks with residual module is included in the proposed framework, which have the ability to learn high frequency details. Adversarial learning in a cyclic manner is followed, which ensures superior quality *CT* images. The proposed network has two generators—*G* & *F*, with feature extraction network and reconstruction network. Reconstruction network has a Parallelized *CNN* with multi layer perceptron (*MLP*) to carry out nonlinear projection in the spatial domain. Parallelized *CNN* are network within network which can perform dimensionality reduction with faster computation and less information loss, able to learn the complex mapping at finer levels with better accuracy. The framework has been trained and tested using two datasets-twenty five images from Tibia dataset and 5936 images from Abdominal dataset. Performance of proposed *GAN-CIRCLE* is compared with various *GAN* using the metrics *PSNR*, *SSIM* and Information fidelity criterion (*IFC*). Enhancing the image quality acquired by portable miniature devices is of great attention recently.

A deep learning approach named *MedSRGAN* was proposed to enhance the resolution of medical images using Generative Adversarial Networks (*GANs*). Fig. 9 shows the architecture of *MedSRGAN*. At the core of this method is a specially designed generator known as the Residual Whole Map Attention Network (*RWMAN*), which effectively extracts important features across multiple channels while emphasizing critical regions within the images. To optimize performance, a composite loss function was utilized—combining content loss, adversarial loss, and feature-based adversarial loss. The *MedSRGAN* model was trained and evaluated using a dataset comprising 242 thoracic *CT* scans and 110 brain *MR* scans. Experimental results indicated that the method not only maintains fine texture details but also produces more lifelike and accurate high-resolution images. Furthermore, a Mean Opinion Score (*MOS*) assessment conducted by five experienced radiologists on *CT* image slices confirmed the model's effectiveness.



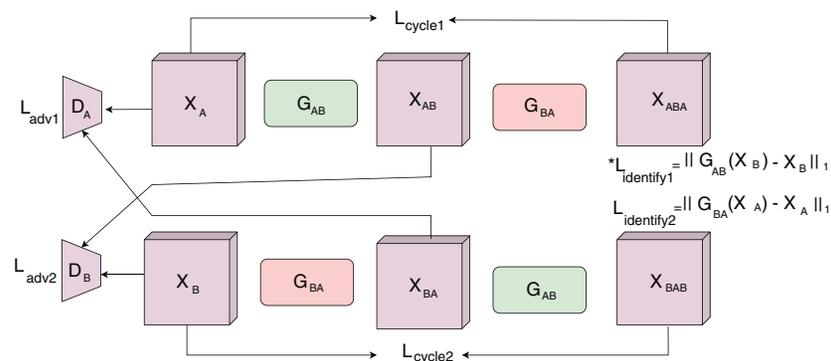
**Figure 9:** Architecture of *MedSRGAN* [124]

Degraded images are the major drawback of portable imaging devices. Spatial resolution, contrast and noise are the three issues related with ultrasound images. Portable devices produce low resolution, low contrast and high noise images, which makes the disease diagnosis inaccurate. GAN based method provides promising results for resolution enhancement with the following advantages—non linear multi level mapping between *LR* and *HR* images, adaptive feature extraction without human intervention, image quality enhancement using discriminator *D*, direct and efficient one step feed forward reconstruction procedure and easier implementation on hardware like *FPGAs*. The encoder-decoder based GAN has the problem of bottleneck which constrain the information sharing between *LR* image and *HR* image. To eliminate this problem, U-net based architecture is employed successfully. U-net model lacks in performance due to the presence of speckle noise in the *HR* image. In order to overcome the problems of encoder decoder model and U-net model, sparse skip connection UNet (*SSCU-Net*) is proposed [125]. Local patches in the low resolution images are used in the discriminator. Assuming these patches are independent, discriminator is capable of modeling high frequency details. The input image patches makes the training easier and lower the memory constraints. The proposed network uses two losses—*L1* loss for conserving the low frequency information and differential loss for high frequency information like edge sharpness. Dataset for testing and training the proposed network is done using 50 simulation pairs, 72 vivo image pairs and 40 phantom pair. Each dataset is divided into five groups—four groups for training and one group for testing. *PSNR*, *SSIM*, contrast resolution (*CR*) and mutual information (*MI*) are estimated for the performance evaluation. *PSNR*, *SSIM* measure similarity between LR and HR images, higher the *PSNR*, higher intensity similarity. *CR* estimates the ability to differentiate the intensity difference. Even though the *SSIM* for *U-net* is higher, the *U-net* images have a over smoothed appearance which implies the loss of some high frequency details. The proposed *SSCU-Net* out performed the *U-Net* and the encoder-decoder model and also the images have better resolution with preserving more edge details. Luan et al. in [126] have introduced a deep learning based adaptive matching network (*AM-Net*) alongwith a dataset generation method named Multi-mapping (*MMP*) for Ultrasound Localization Microscopy (*ULM*). A GAN based MR image super resolution technique was proposed in [127] that uses a generator mechanism incorporating multiple feature selection methods.

Jia et al. in [128] have introduced a super resolution method for retinal fundus image based on GANs which uses a vascular structure prior. This is shown to have overcome the shortcomings of its previous model *Real-ESRGAN* [129].

## 2.4 Translation

Disease diagnosis and treatment with single image modality may not be adequate in most of the cases. The images acquired cannot outline the complete anatomical details or fails to acquire the details with the desired imaging modality [85,90,130–132]. Image translation is an optimum solution, where required image is synthesized from a different image modality, without inducing much cost or risk. The challenge involved in translating one image modality to another is the presence of unrealistic data in the output image [133]. Fully supervised learning methods are among the most widely used deep learning approaches for this task [134]. However, these methods require paired low- and high-quality images for training, which is especially challenging in medical imaging, where obtaining such aligned image pairs is difficult in real-world scenarios. To address this limitation, several unsupervised learning frameworks have been developed [132,135]. Despite their potential, these approaches often face issues such as instability, noise amplification, and the occurrence of halo artifacts. A well-known solution for unpaired image-to-image translation is the Cycle-consistent Generative Adversarial Network (*CycleGAN*) [100,136]. The block diagram of (*CycleGAN*) for image-to-image translation between domains A and B is shown in Fig. 10. This architecture enables the model to learn domain-specific knowledge from representative images and transfer it to another domain without requiring paired training images. Nevertheless, most bidirectional GAN-based models are insufficiently constrained. For instance, while *CycleGAN* excels at capturing inter-domain cycle-consistency and global appearance within domains, it often struggles to preserve local details. This limitation is particularly significant in medical imaging, where precise local details are essential for accurate decision-making. High-quality medical images are expected to have uniform illumination and well-defined structural details to support effective diagnosis.

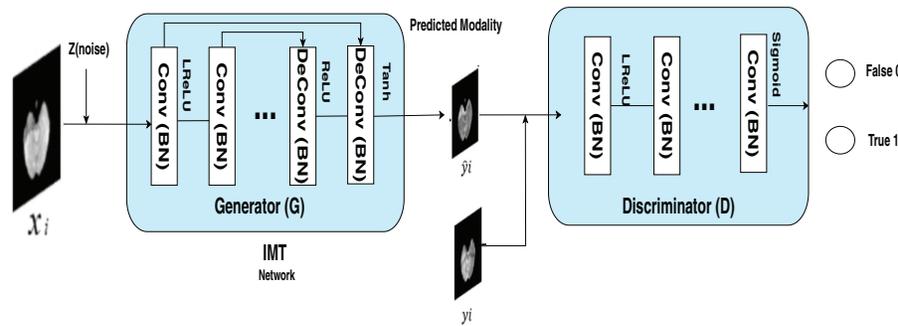


**Figure 10:** The *CycleGAN* architecture enables unpaired image-to-image translation between two domains, A and B. It employs two generators,  $G_{AB}$  and  $G_{BA}$ , to convert images from one domain to the other, resulting in translated outputs  $x_{AB}$  and  $x_{BA}$ . To enforce cycle consistency, the model reconstructs the original inputs as  $x_{ABA}$  and  $x_{BAB}$  from these translations. Reconstruction fidelity is encouraged through cycle consistency losses  $L_{cycle1}$  and  $L_{cycle2}$ , which penalize discrepancies between the original and reconstructed images. Discriminators  $D_A$  and  $D_B$  are tasked with distinguishing real images from synthetic ones, and are trained using adversarial losses  $L_{adv1}$  and  $L_{adv2}$ . Additionally, identity losses  $L_{identity1}$  and  $L_{identity2}$  are incorporated to ensure that if a generator processes an image already from its target domain, the output remains unchanged, thus helping to preserve the image's content [100]

A skin lesion synthesizer based on GAN was proposed in [98], having a coarse to fine generator, multi scale discriminator and a robust objective function for learning. This method synthesized images

from a semantic label map and an instance map [60]. Images of resolution  $1024 \times 512$  were synthesized with proposed generator which incorporated convolutional layers, few residual blocks and deconvolutional layers. The multi scale discriminator consisted of three discriminators, with same input of different resolutions. The training was stabilized using the feature matching loss function, that compared the real and synthetic image features from all the discriminators. The loss function for the proposed network included conditional GAN loss and feature matching loss. Clinically meaningful 2594 skin lesion images were split into two-2346 training images and 248 testing images. Quantitative analysis was performed using AUC and *p*-value. Multiple images of same anatomy with different contrast improves disease diagnosis easier. Multi contrast MR image synthesizer with conditional GAN (*cGAN*) was proposed in [61] for spatially registered images. *cGAN* was compared with cyclic GAN for registered and unregistered *T1*, *T2MRI* images. *T1* weighted MRI image described the gray and white matters, *T2* images described cortical tissue fluid. The method based on *pGAN* was also evaluated in the same. *pGAN* consists of a generator, pre-trained VGG network and a discriminator. The generator learned to synthesize *T2*-weighted image from the *T1*-weighted image. Simultaneously the discriminator differentiated between the real and synthetic image. Pixel by pixel, adversarial and perceptual losses were minimized by the generator and discriminator maximized the adversarial loss. *pGAN* was trained with pixel and perceptual loss. Likewise *T1*-weighted image was synthesized from *T2*-weighted image. *cGAN* implementation used cycle loss instead of pixel wise loss in the *pGAN*. *cGAN* used two generators-*GT1* and *GT2* and two discriminators-*DT1* and *DT2*; one pair of generator and discriminator for *T1* image and other for *T2* image. *GT1* synthesized a *T1* weighted image from the respective *T2* image, meanwhile *DT1* discriminated between real *T1* image and synthetic *T1* image. Three datasets were used-healthy images from MIDAS, IXI and abnormal images from BRATS. Out of the 66 subjects analyzed from MIDAS dataset, 48 were used for training, 5 for validation and 13 for testing. From each subject, 75 axial cross sections of brain tissues without artifacts were selected manually. IXI dataset used 40 subjects with 25 subjects for training, 5 for validation and 10 for training. 41 low grade glioma patients from BRATS dataset were evaluated where 24 for training, 2 for validation and 15 for testing. To avoid biases on dataset, 40 subjects with 4000–5000 images from each dataset were analyzed. Dataset normalization ensured the absence of bias in quantitative evaluation. Armanious et al. in [90] presented *MedGAN* which is a framework designed for end to end medical image to image translation at the image level. It leverages recent progress in Generative Adversarial Networks (GANs) by integrating the adversarial approach with a novel blend of non-adversarial loss functions. A key component of the system is a discriminator network that functions as a trainable feature extractor, enforcing alignment between the generated medical images and their target modalities. To ensure accurate reproduction of textures and detailed structures, style transfer losses are incorporated. Furthermore, the architecture introduces a new generator design called *CasNet*, which progressively refines medical image outputs using a series of encoder-decoder modules to improve image clarity and detail. Also, five experienced radiologists evaluated and verified *MedGAN*'s outputs quality based on subjective assessment.

Yang et al. in [9] introduced a method to perform Image Modality Translation (*IMT*) using a deep learning approach grounded in Conditional Generative Adversarial Networks (*cGANs*). This framework captured low-level pixel information and high-level semantic features such as brain tumors and anatomical structures across different imaging modalities showcasing strong potential as a supportive tool in medical diagnostics. A cross-modality registration technique that integrated deformation fields was introduced, enabling the model to incorporate information from the translated imaging modalities. A Translated Multichannel Segmentation (*TMS*) approach for MR data. In this method, both original and translated modalities were processed together using Fully Convolutional Networks (*FCNs*) to perform segmentation in a multichannel fashion. These two approaches effectively leverage cross-modality information to enhance performance without the need for additional data. Fig. 11 illustrates the above method.



**Figure 11:** The figure outlines the structure of an end-to-end *IMT* network designed for cross-modality image generation. The training dataset is defined as  $S = (x_i, y_i) \mid i = 1, 2, 3, \dots, n$ , where each  $x_i$  represents an image from the source (given) modality, and  $y_i$  is its corresponding image in the target modality. The training process comprises two main components. First, the generator  $G$  receives the input image  $x_i$  along with a random noise vector  $z$  and is trained to generate an output  $\hat{y}_i$  that closely resembles the true image  $y_i$ . Second, the discriminator  $D$  is responsible for distinguishing between the real target images  $y_i$  and the generated ones  $\hat{y}_i$  produced by  $G$ . The discriminator outputs either 1 or 0, where 1 indicates a real image and 0 indicates a synthetic one. During inference, the generator utilizes the learned parameters to produce translated-modality images from new input samples  $G$  [9]

A novel *GAN* architecture named *MMTrans*, which leverages a Swin Transformer backbone for multi-modal medical image translation, is introduced in [137]. The proposed system is composed of three primary components: a generator, a registration module, and a discriminator. The registration module employs a Swin Transformer to estimate a deformable vector field (*DVF*) based on the *SwinIR* framework, which aligns the generated image with the target [138]. For paired datasets, spatial inconsistencies between the input and target images are corrected using this registration network. In scenarios involving unpaired data, the generator  $G(x)$  produces outputs that preserve the anatomical structure of *T1* images and adopt the visual style of *T2*. The registered output  $R(G(x), y)$  further the generated image to match both the morphology and appearance of *T2*. A convolutional neural network-based discriminator evaluates whether the generated images are indistinguishable from actual target modality images. Through extensive testing on both publicly available and clinical datasets—covering both paired and unpaired cases—*MMTrans* demonstrated superior performance compared to existing methods and showed strong potential for clinical adoption. Nonetheless, there are some limitations to this approach. Although processing *2D* slices offers computational efficiency, the spatial context provided by *3D* medical volumes is crucial for many diagnostic tasks. Consequently, future work should focus on extending *MMTrans* to operate effectively on *3D* data to fully leverage its clinical applicability.

Chen et al. in [139] introduced *MI-GAN*, a novel multi-domain medical image translation algorithm that incorporates a key transfer branch. By analyzing the imbalance present in medical imaging datasets, the approach identified critical target domain images and constructed a specialized transfer branch. Utilizing a single generator, the method facilitates multi-domain image translation in the medical context. This structure enhanced both the model's attention mechanism and the quality of the generated images. Additionally, a lung image classification model was presented, leveraging synthetic image data for augmentation. The training dataset combines both synthetic lung *CT* scans and original real-world images to evaluate the effectiveness of the model in diagnosing normal individuals, as well as patients with mild and severe cases of *COVID-19*. The method was seen to outperform the state of art techniques. Ozbey et al. have introduced a method in which adversarial diffusion modelling is used for obtaining improved results in image translation [140].

### 2.5 Reconstruction

Medical image reconstruction is a crucial process for generating high-quality images needed for accurate analysis. However, the quality of these images is often affected by noise and artifacts [141–143]. To address these limitations, there has been a paradigm shift from traditional analytical and iterative reconstruction methods to data-driven machine learning approaches [144,145].

A review of the literature reveals that frameworks like *pix2pix* and *CycleGAN* are widely employed for *MR* image reconstruction [120]. In this method 3D neural network architecture called the multi-level densely connected super-resolution network (*mDCSRN*), which incorporates training guidance from a *GAN* is proposed. The proposed *mDCSRN* is designed for efficient training and inference, while the *GAN* component enhances the realism of the super-resolved images, making them nearly indistinguishable from the original high-resolution counterparts. An illustrative representation of this method is given in Fig. 12. Liao et al. [146] explored sparse-view *CBCT* reconstruction to reduce artifacts, proposing a feature pyramid network for the discriminator and computing a modulated focus map to preserve anatomical structures during reconstruction. Alongside reconstruction from undersampled data, maintaining domain data accuracy is essential. In *MR* reconstruction, undersampled *k*-space data in the frequency domain has also been addressed [147,148]. Various types of loss functions have been applied in image reconstruction to capture local image structures effectively. For example, cycle-consistency and identity loss have been utilized together for denoising cardiac *CT* [149]. Wolterink et al. proposed a method for low-dose *CT* denoising by excluding some domain loss, but this approach resulted in compromised local image structure [106]. Bhadra et al. in [150] introduced an image-adaptive *GAN* based reconstruction approach (*IAGAN*), designed to enhance data fidelity by adjusting the pretrained generative model parameters based on the acquired measurement data. The *IAGAN* framework is applied to reconstruct images from undersampled *MR* data. A cutting-edge generative adversarial model, Progressive Growing of *GANs* (*ProGAN*), was trained using a large dataset of high-quality images from the *NYU fastMRI* repository. The trained generator was then integrated into the *IAGAN* architecture to reconstruct high-resolution images from retrospectively undersampled *k* – *space* data in the validation set. The results demonstrate that this *GAN* driven reconstruction method can recover intricate anatomical details from noisy or incomplete measurements, offering a level of detail that conventional reconstruction techniques—typically dependent on sparsity-based regularization—may struggle to achieve. This highlights the potential of *IAGAN* in improving the diagnostic value of *MR* scans.

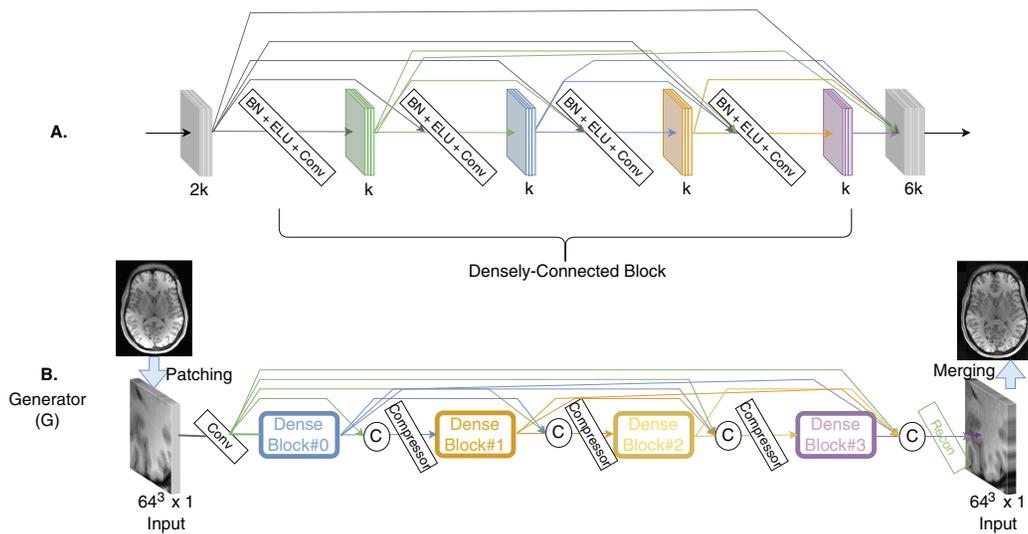
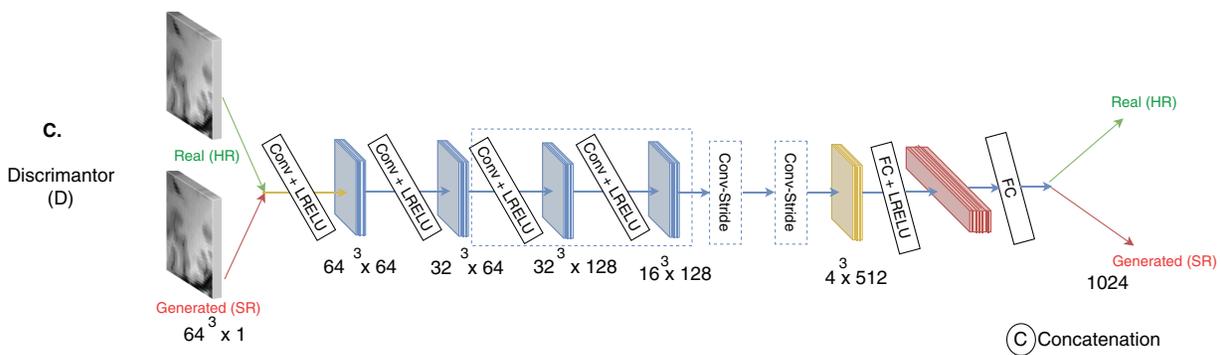


Figure 12: (Continued)



**Figure 12:** The architecture consists of (A) a *DenseBlock* utilizing  $3 \times 3 \times 3$  convolutions, and (B, C) the *mDCSRN-GAN* framework. The generator (G) follows a *b4u4* configuration, meaning it comprises 4 blocks with 4 units each. The initial convolutional layer produces  $2k$  feature maps, where  $k = 16$ . Each compressor within the network reduces the number of feature maps to  $2k$  using a  $1 \times 1 \times 1$  convolution. The final image reconstruction is performed through an additional  $1 \times 1 \times 1$  convolution layer. The discriminator (D) mirrors the architecture of *SRGAN*, with the exception that *BatchNorm* layers are replaced by *LayerNorm*, as recommended in the *WGAN-GP* framework [120]

Additionally, conditional *GANs* (*cGAN*) with skip connections in the generator have been used to synthesize full-dose equivalent *PET* scans from low-dose data was explored by in Rashid et al. in [151]. The primary aim of this study was to assess the effectiveness of a *cGAN*-based approach in enhancing image quality, minimizing noise, and accelerating reconstruction time, in comparison to conventional methods such as Maximum Likelihood Expectation maximization (*MLEM*) and Total Variation (*TV*). The approach involved iterative training of a *U-Net* based generator with a full-image discriminator. Results demonstrated that the proposed *cGAN* framework significantly improved image sharpness, reduced noise, and offered faster reconstruction, outperforming traditional techniques.

Ahn et al. in [152] utilized 10,000 anteroposterior (*AP*) knee radiographs to develop a more cost-effective and balanced medical imaging dataset. Two convolutional neural network models were implemented: Deep Convolutional *GAN* (*DCGAN*) and *StyleGAN2* with Adaptive Discriminator Augmentation (*StyleGAN2-ADA*). To assess the images generated by *StyleGAN2-ADA* compared to actual radiographs, a Visual Turing Test was conducted involving two computer vision specialists, two orthopedic surgeons, and a musculoskeletal radiologist. For evaluation, the Fréchet Inception Distance (*FID*) and Principal Component Analysis (*PCA*) were employed. The synthetic images successfully replicated key pathological features such as osteophyte formation, joint space narrowing, and subchondral sclerosis. Expert classification accuracy when distinguishing real from generated images varied, with scores of 34%, 43%, 44%, 57%, and 50%. The *FID* score between the generated and authentic images was 2.96, significantly lower than that of another medical dataset (*BreCaHAD* = 15.1), indicating higher image fidelity. *PCA* results revealed no statistically significant differences between principal components of the real and generated images ( $p > 0.05$ ). Overall, this research highlights the potential of *GANs* in generating realistic images.

A novel framework named *DualMMP-GAN* was introduced in [153] for generating high fidelity medical images from a given source modality. This method enhances the traditional *CycleGAN* by incorporating dilated residual blocks, a dual-scale patch based discriminator, and a perceptual consistency loss to improve generation quality, particularly in regions containing lesions. *DualMMP-GAN* excels in preserving both contextual information and fine structural details, making it more effective in reconstructing lesion areas. Instead of relying on the standard single-scale discriminator used in *CycleGAN*, *DualMMP-GAN* adopts a dual-scale discriminator that operates on image patches. This design allows the network to learn features

at both fine and coarse levels. The synergy between these two scales improves the model's adaptability to lesions of varying sizes, shapes, and locations. To enhance the network's ability to capture contextual features without significantly increasing computational cost, dilated residual blocks are introduced in place of standard residual blocks. These dilated blocks widen the receptive field, enabling extraction of spatial and structural information from *MR* data. This contributes to more accurate preservation of lesion boundaries, morphology, and overall image continuity, resulting in more detailed and high-resolution outputs. Furthermore, the model leverages perceptual consistency loss, an improvement over the traditional cycle consistency loss employed by *CycleGAN*. By comparing feature representations across different layers, this loss function ensures that the generated images closely resemble the target modality at multiple levels of abstraction, thereby enhancing visual clarity and detail in the synthesized *MR* scans.

Synthetic pterygium images were produced using the default configuration of the *StyleGAN3* architecture in [154], chosen for its strong performance and advanced generative capabilities, which are well-suited for producing photorealistic outputs. The generator employed a 512-dimensional latent and intermediate space, with a mapping network consisting of two layers specifically adjusted for image synthesis. To maximize the model's representational power, the channel base was set to 32, 768, capped at 512 channels. An Exponential Moving Average (*EMA*) value of 0.99 was applied to control parameter updates and promote training stability.

The discriminator also followed the *StyleGAN3* design without freezing any layers. A minibatch standard deviation group size of four was used to encourage variation across samples during training. It mirrored the generator in terms of channel base and maximum. Both models were optimized using the Adam optimizer, configured with beta parameters [0, 0.99] and an epsilon of  $1e-08$ . Learning rates were assigned as 0.0025 for the generator and 0.002 for the discriminator to maintain stable learning. Loss was utilized alongside an *R1* regularization term with a weight of 8.0 to ensure controlled gradient updates within the discriminator.

Despite the architectural strengths of *StyleGAN3*, generating highly realistic medical images presented certain limitations. Initially, a lack of diversity in the training set contributed to elevated Fréchet Inception Distance (*FID*) scores, reflecting a degree of synthetic bias. Additionally, the complexity and variability found in medical imagery required meticulous tuning of model parameters and implementation of sophisticated augmentation techniques to improve visual authenticity. Evaluation through confusion matrices demonstrated that the synthetic images reached a high level of realism, though clinician performance varied—highlighting the subjective nature of visual interpretation and the difficulty in achieving universally indistinguishable synthetic outputs.

These outcomes suggest that *cGAN* based reconstruction holds strong potential to enhance diagnostic precision and streamline clinical imaging workflows. Although *GANs* have been widely applied in various medical imaging tasks, their direct integration into clinical diagnostics and decision-making remains challenging. A significant portion of current research in image reconstruction relies on conventional quantitative metrics to evaluate performance. However, when *GANs* are trained using additional loss functions, optimizing for visual quality becomes difficult, particularly in the absence of a specialized reference metric tailored to assess perceptual image quality. This presents a key obstacle to aligning *GAN* generated outputs with clinical standards [65].

Understanding how diseases evolve over time is essential for early detection and effective treatment planning. This is particularly important for severe conditions like Idiopathic Pulmonary Fibrosis (*IPF*), a chronic and progressive lung disease with a survival rate similar to that of various cancers. *CT* scans are widely recognized as a dependable method for diagnosing *IPF*. Predicting future *CT* images for patients in the early stages of *IPF* can play a significant role in enhancing treatment strategies and improving patient

outcomes. Zhao et al. introduced a novel model in [155] named 4D Vector Quantised Generative Adversarial Networks (4D-VQ-GAN) designed to synthesize realistic CT volumes of IPF patients across different time points. The model training process involved two main stages. Initially, a 3D-VQ-GAN was trained to reconstruct CT volumes. Subsequently, a temporal model based on Neural Ordinary Differential Equations (ODEs) was trained to learn the temporal progression of the quantized latent embeddings produced by the first stage's encoder. Various configurations of the models were tested to generate sequential CT scans and compare to actual data using both quantitative metrics and qualitative assessments. Survival analysis based on imaging biomarkers extracted from the synthetic CT volumes were performed to validate the clinical relevance of our generated scans. The resulting concordance index (*C-index*) was on par with that of biomarkers obtained from real CT scans, indicating the potential of generated CT data for accurate survival prediction and real-world clinical application.

## 2.6 Data Augmentation

Deep learning models typically require large datasets, making it difficult to apply them in situations where limited data is available. One common solution is data augmentation, which involves creating training examples by generating new data. This approach involves basic techniques like random rotations, flipping, cropping, and adding noise. However, such transformations often fall short when applied to complex datasets like medical images. To address this, researchers have developed more strategies for medical imaging. The primary aim is to produce synthetic data that closely mirrors the original distribution. The emergence of Generative Adversarial Networks (GANs) has significantly enhanced data augmentation capabilities [70].

GAN based data augmentation involves training a generator network to produce synthetic images from a latent space, thereby increasing the dataset's diversity and variability beyond simple transformations. This method is especially advantageous in situations with limited labeled data or class imbalance, where creating additional samples of underrepresented classes can greatly enhance classifier performance [156]. For instance, in medical image analysis, GANs have been employed to generate rare pathological cases, facilitating better diagnostic model training without the need for extensive manual data collection [148].

A data augmentation technique was presented by Frid et al. in [157] data augmentation approach that integrates traditional image perturbation techniques with the generation of synthetic liver lesions using Generative Adversarial Networks (GANs) to enhance liver lesion classification accuracy. The main contributions of this work include: (1) the generation of high-quality synthetic focal liver lesions from CT scans using GANs, (2) the development of a convolutional neural network (CNN)-based model for liver lesion classification, and (3) the enrichment of the CNN training dataset with synthetic samples to achieve improved classification performance. The work compares the results obtained with classical augmentation (not involving the usage of GAN). It could be seen that the classification performance improved progressively with the increase in training data, reaching a plateau at approximately 78.6%, beyond which the inclusion of additional augmented samples did not lead to further enhancement in accuracy.

Gan et al. introduced a generative adversarial network (GAN) model, named *HieGAN* in [158], which employs a hierarchical structure to produce high-quality synthetic knee images. This model is intended to enhance data augmentation strategies for deep learning tasks. During the training phase, *HieGAN* incorporates an attention mechanism within both the generator and discriminator networks, specifically before the  $256 \times 256$  image scale, to better extract critical features from knee images. To ensure stable training, a novel approach combining pixelwise and spectral normalization was applied. *HieGAN* was assessed using a large-scale dataset of knee images, with performance measured by *AmScore* and *ModeScore*. Experimental results demonstrated that *HieGAN* surpassed existing state-of-the-art methods. Consequently, *HieGAN*

represents a promising advancement toward developing more reliable deep learning models for knee image segmentation. Future research should explore clinical validation through Visual Turing Tests.

Furthermore, datasets augmented with *GANs* have demonstrated improved performance of deep learning models by reducing overfitting and increasing robustness to noise and variations in data. The adversarial training process pushes the generator to create samples that challenge the discriminator, resulting in realistic synthetic data that enriches the training set [159]. However, challenges persist in ensuring the quality and diversity of *GAN* generated samples, as mode collapse and training instability can hinder their effectiveness.

### 3 Results and Discussion

In medical imaging applications of *GANs*, result analysis is carried out by assessing the generated images through various performance metrics such as *SSIM*, *PSNR*, accuracy, *AUC*, Dice coefficient, *MAE*, *ROC* curve, *IoU*, entropy, and normalization. These metrics are used to evaluate aspects like image quality, classification performance, segmentation accuracy, and overall model efficiency. Depending on the specific objective—whether synthesis, denoising, translation, or segmentation—researchers choose appropriate metrics to ensure a thorough evaluation of the *GAN* model's effectiveness. The consolidated findings of applications of *GANs* in medical images using standardized metrics is as given in Tables 8 and 9. In examining the table outlining *GAN* based applications in medical image processing, it becomes clear that different algorithms perform optimally in specific tasks. For instance, *cGAN* demonstrates strong performance in image denoising, achieving a high Peak to Signal Ratio (*PSNR*) and Structural Similarity Index (*SSIM*). In the domain of image super resolution, Deep *CycleGAN* achieves a notable *PSNR* of 30.72 dB and *SSIM* of 0.924. *TGAN*, which focuses on image reconstruction, records a *PSNR* of 34.69 dB and *SSIM* of 0.953.

**Table 8:** Consolidated findings of applications of *GANs* in medical images using standardized evaluation metrics

Application	Study/Model	Dataset used	<i>GAN</i> architecture	Results
Denoising	Li et al. [53]	<i>JSRT</i> and <i>LIDC</i> datasets	<i>cGAN</i>	<b>JSRT</b> PSNR: 33.26 dB, SSIM:0.9206 <b>LIDC</b> PSNR: 35.11 dB SSIM 0.9328
Denoising	Yang et al. [94]	Clinical low-dose <i>CT</i> scans	<i>WGAN</i>	PSNR: 22.01 dB SSIM: 0.7745
Denoising	Huang et al. [107]	Simulated and real-world low-dose <i>CT</i> datasets	<i>DU-GAN</i> with dual <i>U-Net</i> discriminators	PSNR: 22.12 dB SSIM: 0.7454
Segmentation	Li et al. [95]	<i>I3A</i> datasets	<i>cGAN</i>	SEG: 87.29%
Segmentation	Yang et al. [94]	Clinical low-dose <i>CT</i> scans	<i>WGAN</i>	PSNR: 22.01 dB SSIM: 0.7745

(Continued)

**Table 8 (continued)**

Application	Study/Model	Dataset used	GAN architecture	Results
Segmentation	Han et al. [57]	Spinal MR datasets	<i>Spine-GAN</i>	Accuracy: 96.2% Dice Coefficient: 87.1% Sensitivity: 89.1% Specificity: 86.0%
Super resolution	Gu et al. [124]	ADNI (589 T1 weighted images)	SRGAN with 3DConv layers	PSNR: 29.59 dB SSIM: 0.61
Super resolution	You et al. [93]	Tibia (25) Abdominal (5936)	Deep <i>CycleGAN</i> (GAN-CIRCLE)	PSNR: 27.74 dB SSIM: 0.895 PSNR: 30.72 dB SSIM: 0.924
Super resolution	Wang et al. [125]	50 simulated pairs, 40 phantom pairs, 72 <i>in vivo</i> pairs	SSCU-Net (Sparse Skip U-Net)	PSNR: 25.64 db SSIM:0.38
Translation	Kong et al. [160]	<i>BraTS</i> 2018	RegGAN	PSNR: 26.0 dB SSIM: 0.86
Translation	Ozbey et al. [140]	MR brain datasets	Adversarial Diffusion Models ( <i>SynDiff</i> )	PSNR: 30.09 dB SSIM: 0.94
Translation	Chen et al. [139]	Chest CT images	Multi domain GAN with key transfer	GAN test: 92.19% GAN train: 85.07% Accuracy: 93.85% Sensitivity: 96.69% Specificity: 89.70% AUC: 96.17%
Translation	Chen et al. [120]	Brain structural MR dataset	<i>mDCSRN + GAN</i>	PSNR: 35.88 dB SSIM: 0.9424

**Table 9:** Consolidated findings of applications of GANs in medical images using standardized evaluation metrics

Application	Study/Model	Dataset used	GAN architecture	Results
Reconstruction	Ran et al. [102]	Clinical data and Simulated data	Residual encoder decoder wasserstein GAN( <i>RED-WGAN</i> )	<b>Clinical result</b> PSNR: 31.15 dB SSIM: 0.8624 <b>Simulated result</b> PSNR: 30.22 dB SSIM: 0.8323
Reconstruction	Seitzer et al. [143]	Cardiac MR (8 fold undersampling)	Two stage CNN with GAN + perceptual loss	PSNR: 31.89 dB MOS: 3.24 ± 0.63 SIS: 0.941

(Continued)

**Table 9 (continued)**

Application	Study/Model	Dataset used	GAN architecture	Results
Reconstruction	Guo et al. [67]	Blister, Demodicosis, Parakeratosis, Molluscum	<i>MEDGAN</i>	MAP = 0.96 Accuracy = 62%
Reconstruction	Du et al. [161]	Abdominal <i>MR</i> image	<i>T-GAN</i>	PSNR: 34.69 dB SSIM: 0.9353
Reconstruction	Zuo et al. [162]	ADNI Dataset	<i>VAE-GAN</i>	ACC: 85.18 SEN: 84.44 SPE: 85.92 AUC: 86.16
Data augmentation	Frid-Adar et al. [157]	<i>CT</i> images of v 182 liver lesions	<i>GAN</i> based Synthetic Medical Augmentation	Accuracy: 85.7%
Data augmentation	Antoniou et al. [156]	Omniglot dataset, <i>EMNIST</i> dataset, <i>VGGFace</i> dataset	Data Augmentation <i>GAN (DAGAN)</i>	<b>Omniglot</b> Accuracy: 82% <b>EMNIST</b> Accuracy: 76% <b>VGG-Face</b> Accuracy: 12%
Data augmentation	Zhang et al. [163]	Thyroid ultrasound image	<i>RFI-GAN</i>	PSNR: 32.78 dB SSIM: 0.9598
Data augmentation	Zhang et al. [164]	Heart disease cleveland	<i>WGAN-GP</i>	AUC: 0.902 SPE: 0.82

Key datasets such as *OASIS*, *SRPBS*, and *ABIDE* are frequently employed in these studies, highlighting their critical role in advancing research within the field. Their extensive use demonstrates their value in benchmarking and evaluating algorithm performance across a range of medical imaging applications. Ultimately, the diverse nature of medical imaging tasks requires a strategic approach to selecting algorithms. While certain *GAN* models excel in specific areas, the variation in results underscores the need for task-specific algorithm selection. This tailored approach is essential for advancing the capabilities and accuracy of medical image analysis.

#### 4 Challenges, Ethics and Future Research Directions of GAN for Medical Images

Although *GANs* have demonstrated significant potential in creating lifelike medical images and enhancing diagnostic tools, they face notable challenges. The following sub-section discusses some of the major challenges while using *GANs* for clinical use under regulatory frameworks [165].

##### 4.1 The Non Convergence Problem

In *GANs*, achieving convergence between the generator and discriminator at a global optimum—known as the Nash equilibrium—is essential. The training process follows a minimax game framework aimed at reaching this equilibrium. Effective training strategies for both networks are crucial for optimal

performance. As the generator becomes more proficient, it generates synthetic images that closely resemble real ones, making it increasingly difficult for the discriminator to tell them apart. When the generator reaches peak performance and produces highly realistic images, the discriminator's accuracy drops to around 50%, indicating it can no longer differentiate between real and fake data. At this stage, the feedback provided to the generator becomes uninformative, hindering further improvement in image quality. This imbalance can ultimately result in non-convergence during GAN training [166]. The issue of non-convergence significantly impacts the quality of synthetic image generation. This problem becomes evident when examining the characteristics of the generated outputs. In many cases, non-convergence causes the generator to fail, resulting in the production of flat, single-color images—such as entirely black or white—particularly in gray scale image synthesis.

The issue of non-convergence is a significant challenge encountered during the training of GANs. To address the persistent issue of non-convergence in biomedical imaging applications of GANs, various technical studies have been examined. One proposed solution involves guiding the training process toward achieving a Nash equilibrium, which can help stabilize GAN training. However, maintaining this equilibrium during training is notably challenging. With this concept as a foundation, the surveyed literature is organized into three principal categories [166]:

- optimization of update algorithms [167]
- adversarial learning [168]
- tuning of hyperparameters [169].

#### 4.1.1 Optimization of Update Algorithms

The evolution of updating algorithms has been examined across different GAN architectures, including the original Vanilla GAN [42], the Wasserstein GAN (WGAN) introduced by [40], and the more recent *uGAN* model designed by [167]. The update algorithms introduced in the original Vanilla GANs are mostly restricted to their initial experimental contexts. While the update mechanism in WGANs [170] data has shown some success in achieving Nash equilibrium for specific applications, its applicability remains limited. While the update mechanisms in Vanilla GAN and WGAN were developed for general image generation tasks, *uGAN* specifically targets applications in biomedical image synthesis. Each of these approaches introduces methods for adjusting the frequency of discriminator updates relative to generator updates during training. These strategies have demonstrated improved training stability and a greater ability to reach equilibrium in GAN learning.

#### 4.1.2 Adversarial Learning

Achieving balance in GAN training is closely linked to adjusting the learning rates of the generator and discriminator. This approach was adopted by [168] to mitigate non-convergence issues in biomedical image synthesis. The underlying concept of stabilizing GAN training through learning rate control was originally proposed by [171], who introduced the Two Time-scale Update Rule (*TTUR*). *TTUR* employs separate learning rates for the generator and discriminator, enabling the model to approach a local Nash equilibrium without relying on multiple update steps. In their study, Abdelhalim et al. incorporated both *TTUR* and a custom discriminator update strategy into the *SPGGAN* framework for synthesizing skin lesion images. Specifically, they updated the discriminator five times for each generator update, promoting greater training stability. This adjustment aimed to slow down discriminator learning just enough to allow the generator to keep pace and improve image quality without mode collapse.

### 4.1.3 Tuning of Hyperparameters

Selecting suitable hyperparameters for controlling the generator and discriminator in GANs remains a significant challenge. To tackle this issue, optimization techniques have been explored to derive adaptive loss functions that effectively guide the generator’s weight updates. Goel et al. [169] introduced an optimized GAN framework designed to generate synthetic chest CT images for COVID-19 cases. Their approach integrates a Conditional GAN (CGAN) with the Whale Optimization Algorithm (WOA), a nature-inspired metaheuristic based on the bubble-net hunting behavior of humpback whales [172]. Within this framework, the behavior of whales in locating prey is modeled to guide the generator’s hyperparameter search. The optimization process is governed by three main rules:

- Encircling strategy: The lead whale locates the prey and simulates encircling it. Analogously, the generator’s candidate solutions (search agents) evaluate a fitness function during each iteration and refine their positions accordingly.
- Distance-based updating: The proximity between the prey (optimal solution) and each search agent is calculated, and agent positions are adjusted based on this measure.
- Exploration through random search: Unlike the first rule which focuses on the best-known position, this rule updates the agents’ positions based on a randomized strategy to encourage exploration of the solution space.

The use of this optimized strategy enhances both the generator’s performance and the discriminator’s ability to distinguish between real and synthetic images. As a result, the model achieves adaptive loss tuning, leading to the generation of higher-quality and more diverse images. In terms of performance, the optimized GAN outperformed the baseline CGAN in classification tasks using the synthesized and original images. Specifically, it achieved an *F1-score* of 98.79% and accuracy of 98.78%, compared to 90.99% *F1-score* and 91.60% accuracy from the baseline CGAN. These results indicate that the optimized GAN effectively balances GAN training through hyperparameter tuning.

Researchers have explored the use of Jensen-Shannon (*JS*) divergence [173] to maintain a training balance. Alternative strategies, such as utilizing *f*-divergence and refined Wasserstein loss functions, have been proposed, but they still require further refinement. Research should aim to enhance the stability of GAN training by refining *JS* divergence and utilizing strategies like stochastic gradient descent and Pareto optimization. Also, new game-theoretic frameworks combined with divergence measures may provide promising solutions to address the non-convergence issue in GANs. A summary of existing solutions to address the non-convergence problem in GANs for medical images is given in Table 10.

**Table 10:** summary of existing solutions to address the non-convergence problem in GANs for medical images

GAN variant	Proposed solution	Summary of solution
SPGGANTTUR [168]	Adversarial training	The Two Time-Scale Update Rule (TTUR) helps the achievement of a local Nash equilibrium by using different learning rates to the generator and discriminator. TTUR ensures stable training dynamics and the creation of synthetic skin lesion images through this strategy.

(Continued)

**Table 10 (continued)**

GAN variant	Proposed solution	Summary of solution
Optimized GAN [169]	Adversarial training	The Whale Optimization Algorithm (WOA) is utilized to stabilize the training of a Conditional GAN (CGAN) to synthesize CT images. It adjusts the hyperparameters by mimicking the unique hunting strategy of humpback whales using which they locate the optimal location of the prey. This is modeled to locate the generator's effective search agents in coordination with a fixed discriminator, hence ensuring balanced and efficient GAN training.
$\mu$ GAN [167]	Adversarial training	The generator and discriminator are updated with the same number of training iterations. This enables the production of high-quality synthetic retinal images.

#### 4.2 Mode Collapse & Hallucinated Features

Mode collapse represents a significant challenge during the training of GANs. This issue often results in outputs that lack the diversity observed in real medical images. When mode collapse occurs, the generator overlooks important features and generates identical patterns, leading to a loss of meaningful variability in the synthetic data. This repetition reduces the utility of the generated images. Training a GAN to completely eliminate mode collapse remains a difficult task. For instance, when generating segmented chest radiographs using ground truth data and corresponding segmentation masks, mode collapse can lead to repetitive or incomplete outputs. The high complexity and dimensionality of 3D brain MR data further contribute to the occurrence of mode collapse during image synthesis, making it a persistent challenge in this domain.

GANs may also introduce feature hallucinations while generating synthetic data [174]. Feature hallucination refers to the creation of artificial elements or the omission of critical features in generated images, which can lead to diagnostic errors [175]. These hallucinations often arise as a result of mode collapse during training. This issue is prominent in image to image translation tasks, where synthetic outputs may include inaccurate or misleading features. Addressing mode collapse effectively can help minimize the occurrence of hallucinated features in GAN generated images. Various strategies have been employed in GANs to tackle the issue of mode collapse. Ensuring training of both the generator and discriminator is necessary. The generator must be able to capture the complete range of feature distributions and anatomical structures present, while the discriminator must provide feedback so that the generator can produce diverse outputs. Exploring advanced optimization methods to improve the stability of the training process by resolving challenges like mode collapse can significantly boost the performance of GANs.

The mode collapse problem can be alleviated by using different methods such as

- regularization
- modified architectures
- adversarial training.

##### 4.2.1 Regularization

In deep learning, minimizing the loss function is a primary objective; however, achieving this becomes difficult when the model contains excessively large weight values. Large weights can lead to overfitting, where

the model performs well on training data but generalizes poorly to new, unseen data. To counteract this, regularization techniques are employed to constrain the size of the weights or limit the overall capacity of the model [42].

In the context of *GANs*, both the generator and discriminator are neural networks and are therefore susceptible to similar challenges. Mode collapse often occurs when the discriminator provides inconsistent or vague gradient feedback. To mitigate this issue, weight normalization is applied as a form of regularization. Unlike traditional regularization methods that introduce additional loss terms, Weight Normalization (*WN*) directly modifies the training process by adjusting how weights are updated. During training, gradients are backpropagated based on normalized weight matrices, improving training stability without adding to the loss function [176]. Several normalization strategies have been proposed for *GANs*, including:

1. Spectral Normalization [177]
2. Batch Normalization [38]
3. Self-Normalization [178]

Among these, spectral normalization has proven particularly effective in stabilizing *GAN* training by controlling the Lipschitz constant of the discriminator network.

Xu et al. [179] addressed the problem of mode collapse in *GANs* applied to low-dose *X-ray* image super-resolution. They introduced a model called Spectral Normalization Super-Resolution *GAN* (*SNSRGAN*), which incorporated spectral normalization in the discriminator to constrain the Lipschitz constant to a value of 1. This was accomplished by applying the spectral norm—the largest singular value of a weight matrix, equivalent to its L2 norm—during training. To assess the performance of their model, the authors employed the Inception Score (*IS*) and Multi-Scale Structural Similarity Index (*MS-SSIM*), both of which measure the quality and diversity of the generated images. *SNSRGAN* achieved an *IS* of 6.56 and an *MS-SSIM* of 0.986, outperforming the baseline *SRGAN* and demonstrating enhanced image diversity and resolution quality.

#### 4.2.2 Modified Architectures

In the context of *GANs*, architectural changes involving the generator, discriminator, or both—relative to the original (vanilla) *GAN*—are referred to as modified architectures. A common strategy to mitigate the mode collapse issue is to employ multiple generators rather than a single one, as used in the vanilla *GAN*. This method has been shown to improve the diversity of generated outputs [180]. However, managing and training multiple generators introduces considerable complexity and demands significant computational resources. To overcome this challenge, Wu et al. [181] introduced a novel approach that utilizes multiple data distributions rather than multiple generators for synthesizing human cell images. Their model incorporates a generator based on a Gaussian Mixture Model (*GMM*), allowing it to capture various data distributions within the latent space. This structure, called *MDGAN*, enables the generation of a wide variety of image samples by drawing from a mixture of distributions. It was observed that while increasing the number of distributions can enhance sample diversity, it also substantially increases computational requirements. The synthetic human cell images generated by *MDGAN* were used to augment training data for classification purposes. Although the paper does not provide quantitative evaluation metrics for the generated images, it highlights that incorporating these synthetic samples led to a 4.6% improvement in *CNN* classification precision.

#### 4.2.3 Adversarial Training

Creating segmentation masks and corresponding ground truth images separately using *GANs* can be a resource-intensive and time-consuming process. To streamline this, Neff et al. [182] introduced a modified

version of *DCGAN* designed to simultaneously generate both chest *X-ray* images and their associated segmentation masks within a single generation step. During adversarial training, the generator may begin to produce repetitive image-segmentation pairs with minor variations, leading to a mode collapse issue. To tackle this, the researchers employed a perceptual image hashing technique to filter out duplicate synthetic image-segmentation pairs. This approach involves computing hash values based on distinct visual features of both real and generated images. By comparing values, the similarity between two images can be determined. The effectiveness of the generated images were verified by using them to augment training data for a segmentation task. Specifically, a *U-Net* model was trained using a dataset composed of 30 real images and 120 synthetic images. The evaluation pointed a Hausdorff distance of 7.2885, which was lower than when the model was trained exclusively on real or synthetic data. Despite these improvements, the authors acknowledged the presence of a mild form of mode collapse, indicating limited diversity among the generated samples. A summary of existing solutions to address the mode collapse problem in *GANs* for medical images is shown in [Tables 11](#) and [12](#).

**Table 11:** summary of existing solutions to address the mode collapse problem in *GANs* for medical images

<b>GAN variant</b>	<b>Proposed solution</b>	<b>Summary of solution</b>
<i>SPGGAN</i> [168]	Modified architecture	A self-attention mechanism is presented. This module enables attention maps to highlight lesion characteristics of the dermoscopic images, enabling the generator to produce a more diverse set of synthetic images.
<i>DCGAN</i> [182]	Adversarial training	A perceptual image hash function is employed which eliminates duplicate pairs generated during training. The differences between the real and generated images are compared by computing hash values.
<i>SNSRGAN</i> [179]	Modified architecture	Spectral normalization is applied in <i>SNSRGAN</i> which constrains the spectral norm of the discriminator's weight matrices, thus limiting the Lipschitz constant to 1 which inturn stabizes the training.
<i>AiIN</i> <i>DCGAN</i> [183]	Pre-processing	<i>AiIN</i> is utilized in <i>DCGAN</i> adjusting image contrast and emphasizing key features within chest <i>X-ray</i> images.
<i>MSG</i> <i>SAGAN</i> [184]	Modified architecture	Self attention mechanism is incorporated in generator nd discriminator <i>MSG-GAN</i> to enhance the diversity of synthetic <i>X-ray</i> images. This mechanism utilizes attention feature maps to enable the model to capture long-range dependencies among key features the images.
Modified <i>CGAN</i> [185]	Modified architecture	Minibatch discrimination is incorporated into the discriminator of <i>CGAN</i> . The discriminator aligns its gradients with the training data, fostering more stable learning. This technique penalizes the generator when it defaults to a single mode.
<i>CGAN</i> [186]	Modified architecture	A discriminator incorporating conditional information is introduced. This discriminator utilizes specific details from the <i>CT</i> scans to direct the generator. By using this information, the generator is guided to produce a diverse set of images.

(Continued)

**Table 11 (continued)**

<b>GAN variant</b>	<b>Proposed solution</b>	<b>Summary of solution</b>
<i>DCR</i> <i>AEGAN</i> [187]	Modified architecture	<i>AEGAN</i> employs a deep convolutional refiner to produce diverse <i>MR</i> images. This refiner, built on a <i>ResNet</i> architecture, incorporates skip connections alongside deep convolutional layers. They help in regulating the bypassing of certain training layers, which contributes to refining the shapes of generated images and enhances the overall diversity in image synthesis.
<i>AEGAN</i> [188]	Modified architecture	A Variational Autoencoder ( <i>VAE</i> ) into <i>AEGAN</i> which models the probability distributions of the training data.
<i>SLStyleGAN</i> [189]	Modified architecture	The generator of <i>StyleGAN</i> 's is fine-tuned by adjusting the number of fully connected layers. A configuration with four fully connected layers is identified as the most effective, resulting in improved diversity.

**Table 12:** summary of existing solutions to address the mode collapse problem in *GANs* for medical images

<b>GAN variant</b>	<b>Proposed solution</b>	<b>Summary of solution</b>
<i>ScarGAN</i> [189]	Adversarial training	An experience replay buffer strategy is implemented. A portion of the generated masks is saved in a buffer for future reuse. The discriminator selects half of its training batches from this buffer, allowing it to revisit previously generated scar tissue samples. This helps to discourage the generator from repeatedly producing scar tissue masks with similar shape.
<i>MDGAN</i> [181]	Modified architecture	A generator based on a Gaussian Mixture Model ( <i>GMM</i> ) is employed in <i>MDGAN</i> . This enables the <i>GAN</i> to represent multiple data distributions within the latent space.
<i>CycleGAN</i> [190]	Modified architecture	A modified architecture featuring a $34 \times 34$ patch discriminator is implemented in <i>CycleGAN</i> to address mode collapse in <i>MR</i> to <i>MR</i> image translation. This limited receptive field enables the discriminator to focus on fine grained details, enabling the generator to preserve sharpness and high-frequency features.

### 4.3 Metrics for Quantitative Evaluation

In addressing the training challenges of *GANs*, evaluation metrics play a vital role in assessing the diversity and quality of the generated images. To measure image diversity, particularly in the context of mode collapse, metrics such as Inception Score (*IS*), Multi-Scale Structural Similarity Index (*MS-SSIM*), and Maximum Mean Discrepancy (*MMD*) are commonly used. For issues like non-convergence and instability, Peak Signal-to-Noise Ratio (*PSNR*) and Fréchet Inception Distance (*FID*) are frequently applied. Among these, *IS* and *FID* are standard metrics for evaluating image quality; however, both are based on models pretrained on the ImageNet dataset [191], which does not include biomedical image classes. As a result, these

metrics are not ideally suited for applications in the biomedical imaging field. Additionally, *MS-SSIM* is a perceptual metric focused primarily on luminance and contrast, offering a limited view of image similarity. *PSNR*, while widely adopted for assessing image quality, is primarily effective for grayscale images and lacks robustness in more complex scenarios. In the field of biomedical image analysis, researchers often rely on conventional pixel-wise evaluation metrics to assess the performance of *GANs*. These metrics are typically designed for supervised learning scenarios that depend on the presence of reference images. However, in biomedical imaging, acquiring reference data is challenging due to privacy concerns and the errors associated with manual annotations. Hence, unsupervised learning methods are commonly used. Also, assessing the performance of *GANs* is crucial due to factors such as random initialization, optimization variability, and technical complexities. Comparing generated images to real ones remains a difficult task, highlighting the need for further investigation into reliable evaluation techniques. Consequently, developing evaluation metrics that capture both subjective perceptions and objective measures presents a significant challenge for the field.

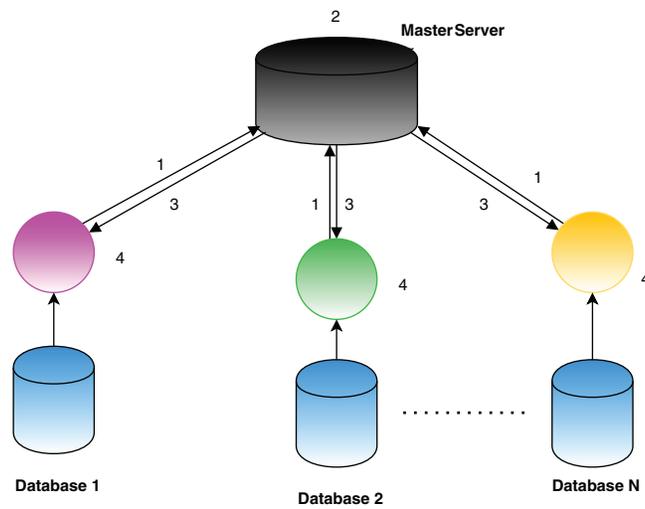
#### 4.4 Privacy Concerns and Ethics

Medical imaging concerns handling highly sensitive patient data, which raises important privacy concerns during both data collection and application. While *GANs* are capable of generating visually convincing medical images, a major issue is the introduction of artificial elements known as hallucinated features. These artifacts may mimic real pathological signs, leading to incorrect diagnostic algorithms. These can compromise clinical reliability. Therefore, it is essential to validate synthetic outputs through both expert evaluation and robust quantitative testing to ensure they reflect true features. To improve the integrity of these images, it is essential to integrate physics-informed simulations and conduct thorough experimental evaluations aimed at understanding the convergence behavior of *GANs* within the medical imaging domain. Using *GAN* generated data in clinical settings requires adherence to strict regulatory standards. Health authorities such as the Food and Drug Administration (*FDA*), USA and European Medicines Agency (*EMA*) demand clear evidence of performance, safety, and reliability to show that the synthetic images are clinically valid, reproducible, and free from bias. Documentation of the development process, ethical oversight, and compliance with data protection laws (e.g., *HIPAA*, *GDPR*) are also crucial. Additionally, synthetic images must be clearly labeled, and their use justified with comprehensive risk assessments to ensure transparency.

Federated learning presents a promising approach to reduce privacy concerns while using *GANs* in medical imaging. However, it encounters several challenges. Achieving effective model convergence under the constraints of restricted communication bandwidth in decentralized environments remains a major difficulty. Also, combining federated learning with differential privacy requires a careful balance between safeguarding sensitive data and maintaining high model performance. The complexity of medical data because of its heterogeneity and the variety of imaging modalities complicates the collaborative model training. Future research should lay priority on overcoming communication inefficiencies, improving the resilience of models, and customizing federated learning techniques to accommodate the characteristics of medical datasets. Furthermore, exploring advanced privacy preserving methods and striving to create a fully decentralized, privacy centric *GAN* capable of generating high-quality medical images without compromising patient confidentiality continues to be a challenging yet vital endeavor in the healthcare sector.

[Fig. 13](#) gives an illustration of the steps in federated learning. In the first phase, each participant independently calculates the model gradients on their local data. To ensure data confidentiality, cryptographic methods such as homomorphic encryption are applied before transmitting the encrypted gradients to the central server. In the second phase, the central (master) server performs secure aggregation of the encrypted

gradients. In the third phase, the aggregated results are sent back to the participants. During the fourth phase, each participant decrypts the aggregated gradients and updates their local model accordingly. This cycle is repeated iteratively until either the loss function reaches convergence or a predefined number of iterations is completed. Throughout this process, the participants' data remains stored locally, maintaining privacy and offering an advantage over centralized approaches such as those based on Hadoop. Federated learning facilitates collaborative model training across multiple databases without the need to centralize the data, enabling scalability with growing datasets while minimizing communication overhead, as only gradients—not raw data—are shared.



**Figure 13:** Steps in federated learning [165]

#### 4.5 Need for Human in the Loop Studies

GAN based image generation models can produce photorealistic images; however, in medical applications, caution is essential as these images play a critical role in disease detection and diagnosis. It is imperative for medical professionals, as the end users, to thoroughly evaluate and validate GAN generated images to ensure their reliability and clinical utility. Trust and acceptance from doctors are crucial for integrating such models into healthcare workflows. While *PSNR* and *SSIM* are standard for evaluating image processing techniques [192], they often fail to reflect perceptual quality or clinical relevance, particularly in the presence of subtle distortions. This underscores the importance of expert-driven evaluation, as automated metrics alone may not capture diagnostic integrity. In medical imaging, radiologists and clinicians are best positioned to perform this qualitative assessment, as their expertise ensures that reconstructed images are not only visually plausible but also diagnostically accurate.

Realistic full-field digital mammograms were generated using a progressive GAN architecture [193], achieving high resolution and realism indistinguishable from real images, even by domain experts. Despite the specialized nature of medical imaging, both experts and non-experts in a reader study showed random success probabilities, emphasizing the critical role of human validation in ensuring clinical applicability. Similarly, the progressive growing GAN (PGGAN) [194] was employed to generate high-resolution chest radiographs (CXR), with two models trained separately on normal and abnormal CXRs, producing synthetic images of  $1000 \times 1000$  pixels. Six radiologists performed binary Turing tests on two validation sets, attaining mean accuracies of 67.42% and 69.92% in the first and second trials, respectively. Majority voting and Cohen's score were used to further assess diagnostic agreement and reliability. Another notable

contribution is *MedGAN* [90], an end-to-end image translation framework tailored for medical applications. By integrating a conditional adversarial setup with multiple non-adversarial losses and a *CasNET* generator, *MedGAN* enhances global consistency and preserves high-frequency details. Its effectiveness was confirmed through expert radiologist evaluations, indicating strong clinical fidelity.

Further advancing *GAN* applications, *GANCS* [195] presents a compressed sensing framework that models the low-dimensional manifold of high-quality MR images using a least-squares *GAN* (*LSGAN*) to capture fine textures, combined with  $\ell_1/\ell_2$  loss to suppress high-frequency noise. Expert radiologist ratings on a large paediatric contrast-enhanced *MR* dataset consistently preferred *GANCS* over traditional wavelet, dictionary learning, and pixel-wise deep learning methods. Diagnostic reliability was supported by normalized Radiologist Opinion Scores (*ROS*) for image quality, artifact presence, and sharpness, aligning with *SSIM* and *SNR* metrics. Additionally, a hybrid deep learning reconstruction strategy was proposed by integrating *CycleGAN* with projection onto convex sets (*POCS*) [196]. The initial *CycleGAN* output was refined using *POCS* and reused in a second training iteration to enhance performance. Radiologists evaluated five reconstruction methods (*U-Net*, *GAN*, *CycleGAN*, *RefineGAN*, and *POCS-CycleGAN*) across brain and knee *MRI* datasets. *POCS-CycleGAN* achieved the highest average scores for both knee (3.3 and 2.7) and brain *MR* images (3.7 and 4.05), consistently demonstrating superior image quality across different modalities and readers.

## 5 Other Generative Models

This section provides an insight into various generative models, including auto-encoders and diffusion models, each tailored for specific tasks in data generation and representation learning. The auto-encoder [197] has encoder-decoder network maps input data to a low-dimensional latent space, enabling the decoder to accurately reconstruct the input. This latent space also facilitates systematic analysis and manipulation of input properties, making the architecture essential for biomedical tasks like image reconstruction, data augmentation, and modality transfer. Diffusion models, a class of deep generative models [198], learn the prior probability distribution of images (e.g., brain *PET* or cardiac *MRI*) from training data and generate new samples by sampling from this distribution. Recently, they have emerged as state-of-the-art in generative modelling, producing higher-fidelity samples compared to auto-encoders and normalizing flows. The comparison between the generative models are shown in Table 13. Diffusion models are extensively applied in medical image processing tasks such as reconstruction, registration, classification, image-to-image translation, segmentation, denoising, and image generation. A detailed explanation regarding this is discussed in the following sub-section.

**Table 13:** Comparison of generative models

Feature	Auto-encoder	GAN	Diffusion model
Training stability	Good	Poor	Good
Image generation quality	Medium	Good	Good
Generation speed	Fast	Fast	Slow
Generation diversity	Good at approximating diversity	Limited due to mode collapse	High, due to variability in input samples
Generation technique	One-shot	One-shot	Iterative

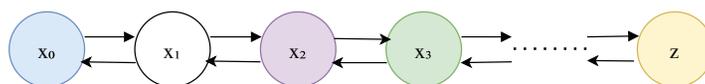
(Continued)

**Table 13 (continued)**

Feature	Auto-encoder	GAN	Diffusion model
<b>Computational complexity</b>	Medium	High, due to training both generator and discriminator networks.	Very high due to iterative noise removal process

**Diffusion Models**

Emerging as promising alternatives to GANs, diffusion models are gaining attention in image generation tasks. These models operate on score-based methodologies, where training involves progressively adding Gaussian noise to images and learning the data distribution that underlies this transformation. Rather than directly reversing the noise, the goal is to accurately capture and reproduce the complex structure of the data distribution. Diffusion models are capable of producing highly realistic images, offering advantages such as stable training dynamics and comprehensive mode coverage. However, one of the primary limitations lies in the prolonged sampling time required during image generation. Although this limitation may be acceptable in fields like medical imaging, where real-time performance is not always essential, reducing this latency remains an active area of research. Current advancements aim to optimize diffusion models for faster inference without compromising image fidelity. This includes the development of new model architectures, improvements in computational efficiency, and enhancements in their suitability for time-sensitive applications. In the context of medical imaging, such innovations could significantly improve both the practicality and impact of diffusion-based approaches. Fig. 14 illustrates the process of diffusion models. Diffusion models operate through two distinct phases: the forward and reverse diffusion processes. In the forward phase, a Markov chain progressively corrupts the input data by adding Gaussian noise over a series of steps—typically around 1000 until the data resembles pure white noise. This phase is fixed and non-trainable. Conversely, the reverse phase is designed to gradually remove the added noise, effectively reconstructing the original data. This denoising process is guided by a neural network that is trained specifically to approximate the reverse of the forward diffusion.



**Figure 14:** Diffusion model [165]

**6 Conclusion**

Generative Adversarial Networks (GANs) have shown significant promise in transforming medical image analysis by enabling high-quality image synthesis, cross-modality translation, data augmentation, and aiding diagnostic tasks. Their capability to learn complex data distributions makes them valuable in scenarios where annotated medical data is limited or imbalanced. This study has explored the diverse applications of GANs in the medical imaging domain, highlighting their role in generating synthetic data to overcome data limitations and improve the performance of diagnostic models. Furthermore, it outlines how GANs contribute to modality translation, enabling the transformation of images between different imaging techniques. By identifying the advancements and the inherent limitations of GANs, this study provides a balanced view on the current landscape and future potential and existing research gaps of GAN in healthcare.

Despite the advancements in GANs, several challenges still hinder their full integration into clinical practice. Key difficulties include unstable training processes, lack of standard evaluation protocols for image generation, and the potential for generated images to introduce clinically misleading artifacts. Although GANs have shown notable advancements in image quality metrics like *SSIM*, *PSNR*, and *FID*, it is crucial to separate these technical achievements from their actual clinical usefulness. Assertions about improved diagnostics or better treatment planning need to be backed by empirical research, including clinical trials, expert assessments, or regulatory benchmarks. Without such validation, enhancements in image realism or reconstruction quality should not be presumed to equate to clinical readiness. Ensuring that synthetic images retain diagnostic relevance and do not compromise patient safety remains a critical concern. Additionally, the opaque nature of GAN models limits interpretability, which is essential for medical decision making and clinician acceptance. Ethical issues, such as data misuse and patient privacy, further complicate the deployment of GAN generated data.

Looking to the future, research should focus on building more stable and explainable GAN architectures tailored specifically to medical applications. Developing robust validation frameworks that incorporate expert clinical feedback will be vital to ensuring safety and effectiveness. Future studies should focus on interpretability, reliability, and adherence to clinical and regulatory standards to confirm the practical applicability of GAN based tools in healthcare. There is also potential for GANs to work synergistically with other AI techniques, including self-supervised learning and multimodal models, to further improve diagnostic support systems.

**Acknowledgement:** The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/540/46.

**Funding Statement:** The research was supported by Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/540/46.

**Author Contributions:** The authors confirm contribution to the paper as follows: Conceptualization, Sameera V. Mohd Sagheer and U. Nimitha; methodology, P. M. Ameer; software, Sameera V. Mohd Sagheer; validation, Sameera V. Mohd Sagheer, U. Nimitha and P. M. Ameer; formal analysis, P. M. Ameer; investigation, Muneer Parayangat; resources, Mohamed Abbas; writing—original draft preparation, Sameera V. Mohd Sagheer; writing—review and editing, U. Nimitha; visualization, P. M. Ameer; supervision, Muneer Parayangat; project administration, Mohamed Abbas; funding acquisition, Mohamed Abbas and Krishna Prakash Arunachalam. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** Data openly available in a public repository.

**Ethics Approval:** Not applicable.

**Conflicts of Interest:** The authors declare no conflicts of interest to report regarding the present study.

**Supplementary Materials:** The supplementary material is available online at <https://www.techscience.com/doi/10.32604/cmcs.2025.067108/sl>.

## References

1. Yu Y, Feng T, Qiu H, Gu Y, Chen Q, Zuo C, et al. Simultaneous photoacoustic and ultrasound imaging: a review. *Ultrasonics*. 2024;139:107277. doi:10.1016/j.ultras.2024.107277.
2. Dangoury S, marghichi El M, Sadik M, Fail A. The design of an efficient low-cost FPGA-based unit for generation ultrasound beamforming. *Pertanika J Sci Technol*. 2023;31(6):3077–92. doi:10.47836/pjst.31.6.24.

3. Ali M, Magee D, Dasgupta U. Signal processing overview of ultrasound systems for medical imaging. In: White paper. Vol. SPRAB12. Dallas, TX, USA: Texas Instruments; 2008.
4. Tao Z, Tagare HD, Beatty JD. Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images. *IEEE Transact Med Imag.* 2006;25(11):1483–91. doi:10.1109/tmi.2006.881376.
5. Weng L, Reid JM, Shankar PM, Soetanto K. Ultrasound speckle analysis based on the K distribution. *J Acoust Soc America.* 1991;89(6):2992–5. doi:10.1121/1.400818.
6. Sagheer SVM, George SN. An approach for despeckling a sequence of ultrasound images based on statistical analysis. *Sens Imag.* 2017;18(1):29. doi:10.1007/s11220-017-0181-8.
7. Henkelman RM. Measurement of signal intensities in the presence of noise in MR images. *Medical Physics.* 1985;12(2):232–3. doi:10.1118/1.595711.
8. Zeng Y, Zhu J, Wang J, Parasuraman P, Busi S, Nauli SM, et al. Functional probes for cardiovascular molecular imaging. *Quantit Ima Med Surg.* 2018;8(8):838. doi:10.21037/qims.2018.09.19.
9. Yang Q, Li N, Zhao Z, Fan X, Chang EIC, Xu Y. MRI cross-modality image-to-image translation. *Scient Rep.* 2020;10(1):3753. doi:10.1038/s41598-020-60520-6.
10. Mredhula L, Dorairangasamy M. An extensive review of significant researches on medical image denoising techniques. *Int J Comput Applicat.* 2013;64(14):1–12. doi:10.5120/10699-1551.
11. Ding Q, Long Y, Zhang X, Fessler JA. Statistical image reconstruction using mixed Poisson-Gaussian noise model for X-ray CT. arXiv:1801.09533. 2018.
12. Borsdorf A, Raupach R, Flohr T, Hornegger J. Wavelet based noise reduction in CT-images using correlation analysis. *IEEE Transact Med Imag.* 2008;27(12):1685–703. doi:10.1109/tmi.2008.923983.
13. Rahiman MF, Rahim RA, Zakaria Z. Design and modelling of ultrasonic tomography for two-component high-acoustic impedance mixture. *Sens Actuat A: Phys.* 2008;147(2):409–14. doi:10.1016/j.sna.2008.05.024.
14. Chen Y, Shi L, Feng Q, Yang J, Shu H, Luo L, et al. Artifact suppressed dictionary learning for low-dose CT image processing. *IEEE Transact Med Imag.* 2014;33(12):2271–92. doi:10.1109/isbi.2014.6868073.
15. Ollinger J, Fessler J. Positron-emission tomography. *IEEE Signal Process Mag.* 1997;14(1):43–55. doi:10.1109/79.560323.
16. Coxson PG. Consequences of using a simplified kinetic model for dynamic PET data. *J Nucl Med.* 1995;38(4):660–7.
17. Abdallah NG, Rashdan M, Khalaf AA. High resolution time-to-digital converter for pet imaging. In: International Conference on Innovative Trends in Communication and Computer Engineering (ITCE); 2020 Feb 8–9; Aswan, Egypt. p. 295–8.
18. Shabbir J, Anwer T. Artificial intelligence and its role in near future. arXiv:1804.01396. 2018.
19. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Transact Patt Analy Mach Intell.* 2013;35(8):1798–828. doi:10.1109/tpami.2013.50.
20. Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning requires rethinking generalization. arXiv:1611.03530. 2016.
21. Schmidhuber J. Deep learning in neural networks: an overview. *Neur Netw.* 2015;61(3):85–117. doi:10.1016/j.neunet.2014.09.003.
22. Talaei Khoei T, Ould Slimane H, Kaabouch N. Deep learning: systematic review, models, challenges, and research directions. *Neu Comput Applicat.* 2023;35(31):23103–24. doi:10.1007/s00521-023-08957-4.
23. Pan Z, Yu W, Yi X, Khan A, Yuan F, Zheng Y. Recent progress on generative adversarial networks (GANs): a survey. *IEEE Access.* 2019;7:36322–33. doi:10.1109/access.2019.2905015.
24. Aljuaid A, Anwar M. Survey of supervised learning for medical image processing. *SN Comput Sci.* 2022;3(4):292. doi:10.1007/s42979-022-01166-1.
25. Rani V, Kumar M, Gupta A, Sachdeva M, Mittal A, Kumar K. Self-supervised learning for medical image analysis: a comprehensive review. *Evolv Syst.* 2024;15(4):1607–33. doi:10.1007/s12530-024-09581-w.
26. Halevy A, Norvig P, Pereira F. The unreasonable effectiveness of data. *IEEE Intell Syst.* 2009;24(2):8–12. doi:10.1109/mis.2009.36.

27. Raza K, Singh NK. A tour of unsupervised deep learning for medical image analysis. *Current Med Imag Rev.* 2021;17(9):1059–77. doi:10.2174/1573405617666210127154257.
28. Ganeshkumar M, Sowmya V, Gopalakrishnan E, Soman K. Unsupervised deep learning-based disease diagnosis using medical images. In: *Cognitive and soft computing techniques for the analysis of healthcare data.* Amsterdam, The Netherlands: Elsevier; 2022. p. 203–20. doi:10.1016/b978-0-323-85751-2.00011-6.
29. Croitoru FA, Hondru V, Ionescu RT, Shah M. Diffusion models in vision: a survey. *IEEE Transact Pattern Anal Mach Intell.* 2023;45(9):10850–69. doi:10.1109/tpami.2023.3261988.
30. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. *Communicat ACM.* 2020;63(11):139–44. doi:10.1145/3422622.
31. De Souza VLT, Marques BAD, Batagelo HC, Gois JP. A review on generative adversarial networks for image generation. *Comput Graph.* 2023;114:13–25.
32. Gui J, Sun Z, Wen Y, Tao D, Ye J. A review on generative adversarial networks: algorithms, theory, and applications. *IEEE Transact Know Data Eng.* 2023;35(4):3313–32. doi:10.1109/tkde.2021.3130191.
33. Krichen M. Generative adversarial networks. In: *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT); 2023 Jul 6–8; Delhi, India.* p. 1–7.
34. Saxena D, Cao J. Generative adversarial networks (GANs) challenges, solutions, and future directions. *ACM Comput Surv (CSUR).* 2021;54(3):1–42. doi:10.1145/3446374.
35. Aggarwal A, Mittal M, Battineni G. Generative adversarial network: an overview of theory and applications. *Int J Inform Manag Data Insights.* 2021;1(1):100004. doi:10.1016/j.jjime.2020.100004.
36. Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA. Generative adversarial networks: an overview. *IEEE Signal Process Magaz.* 2018;35(1):53–65. doi:10.1109/msp.2017.2765202.
37. Bisht A, Rawat K, Bhati JP, Vats S, Sharma V, Singh S. Generating images using vanilla generative adversarial networks. In: *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS). IEEE; 2024 Nov 13–15; Tashkent, Uzbekistan: IEEE.* p. 463–8.
38. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434. 2015.
39. Kim HJ, Lee D. Image denoising with conditional generative adversarial networks (CGAN) in low dose chest images. *Nuclear Instrum Meth Phys Res Sec A Accel Spectrom Detect Assoc Equip.* 2020;954(8):161914. doi:10.1016/j.nima.2019.02.041.
40. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC. Improved training of wasserstein gans. In: *Advances in neural information processing systems.* Cambridge, MA, USA: The MIT Press; 2017.
41. Chu C, Zhmoginov A, Sandler M. CycleGAN, a master of steganography. arXiv:1712.02950. 2017.
42. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial. arXiv:1406.2661. 2014.
43. Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of gans for improved quality, stability, and variation. arXiv:1710.10196. 2017.
44. Motamed S, Rogalla P, Khalvati F. RANDGAN: randomized generative adversarial network for detection of COVID-19 in chest X-ray. *Scient Rep.* 2021;11(1):8602. doi:10.1038/s41598-021-87994-2.
45. Kossen T, Subramaniam P, Madai VI, Hennemuth A, Hildebrand K, Hilbert A, et al. Synthesizing anonymized and labeled TOF-MRA patches for brain vessel segmentation using generative adversarial networks. *Comput Biol Med.* 2021;131:104254. doi:10.1016/j.combiomed.2021.104254.
46. Ahmad B, Sun J, You Q, Palade V, Mao Z. Brain tumor classification using a combination of variational autoencoders and generative adversarial networks. *Biomedicines.* 2022;10(2):223. doi:10.3390/biomedicines10020223.
47. Krithika alias Anbu Devi M, Suganthi K. Review of medical image synthesis using GAN techniques. In: *ITM Web of Conferences.* Vol. 37. Les Ulis, France: EDP Sciences; 2021. 01005 p.
48. Zhou T, Li Q, Lu H, Cheng Q, Zhang X. GAN review: models and medical image fusion applications. *Inform Fus.* 2023;91(11):134–48. doi:10.1016/j.inffus.2022.10.017.
49. Ma Y, Liu J, Liu Y, Fu H, Hu Y, Cheng J, et al. Structure and illumination constrained GAN for medical image enhancement. *IEEE Transact Med Imag.* 2021;40(12):3955–67. doi:10.1109/tmi.2021.3101937.

50. Tightiz L, Jung MH, Song I, Lee K. Trustworthy TAVR navigator system, I: a generative adversarial network-driven medical twin approach for Post-TAVR pacemaker implantation prediction. *Expert Syst Applicat.* 2025;275(1):126973. doi:10.1016/j.eswa.2025.126973.
51. Kazemina S, Baur C, Kuijper A, van Ginneken B, Navab N, Albarqouni S, et al. GANs for medical image analysis. *Artif Intelli Med.* 2020;109(12):101938. doi:10.1016/j.artmed.2020.101938.
52. Fu B, Zhang X, Wang L, Ren Y, Thanh DN. A blind medical image denoising method with noise generation network. *J X-Ray Sci Technol.* 2022;30(3):531–47. doi:10.3233/xst-211098.
53. Li Y, Zhang K, Shi W, Miao Y, Jiang Z. A novel medical image denoising method based on conditional generative adversarial network. *Computat Mathem Meth Med.* 2021;2021(2):9974017. doi:10.1155/2021/9974017.
54. Geng M, Meng X, Yu J, Zhu L, Jin L, Jiang Z, et al. Content-noise complementary learning for medical image denoising. *IEEE Transact Med Imag.* 2021;41(2):407–19. doi:10.1109/tmi.2021.3113365.
55. Gondara L. Medical image denoising using convolutional denoising autoencoders. In: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW); 2016 Dec 12–15; Barcelona, Spain. p. 241–6.
56. Xue Y, Xu T, Zhang H, Long LR, Huang X. SegAN: adversarial network with multi-scale L1 loss for medical image segmentation. *Neuroinformatics.* 2018;16(3–4):383–92. doi:10.1007/s12021-018-9377-x.
57. Han Z, Wei B, Mercado A, Leung S, Li S. Spine-GAN: semantic segmentation of multiple spinal structures. *Med Image Anal.* 2018;50(1):23–35. doi:10.1016/j.media.2018.08.005.
58. Trinh DH, Luong M, Dibos F, Rocchisani JM, Pham CD, Nguyen TQ. Novel example-based method for super-resolution and denoising of medical images. *IEEE Transact Image Process.* 2014;23(4):1882–95. doi:10.1109/tip.2014.2308422.
59. Sánchez I, Vilaplana V. Brain MRI super-resolution using 3D generative adversarial networks. arXiv:1812.11440. 2018.
60. Wang TC, Liu MY, Zhu JY, Tao A, Kautz J, Catanzaro B. High-resolution image synthesis and semantic manipulation with conditional gans. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition; 2018 Jun 18–23; Salt Lake City, UT, USA. p. 8798–807.
61. Dar SU, Yurt M, Karacan L, Erdem A, Erdem E, Cukur T. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. *IEEE Transact Med Imag.* 2019;38(10):2375–88. doi:10.1109/tmi.2019.2901750.
62. Platscher M, Zopes J, Federau C. Image translation for medical image generation: ischemic stroke lesion segmentation. *Biomed Signal Process Cont.* 2022;72(4):103283. doi:10.1016/j.bspc.2021.103283.
63. Upadhyay U, Chen Y, Hepp T, Gatidis S, Akata Z. Uncertainty-guided progressive GANs for medical image translation. In: Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference. Cham, Switserland: Springer; 2021. p. 614–24.
64. Han C, Hayashi H, Rundo L, Araki R, Shimoda W, Muramatsu S, et al. GAN-based synthetic brain MR image generation. In: IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 2018 Apr 4–7; Washington, DC, USA; 2018. p. 734–8.
65. Singh NK, Raza K. Medical image generation using generative adversarial networks: a review. In: Health informatics: a computational perspective in healthcare. Singapore: Springer; 2021. p. 77–96.
66. Fan C, Lin H, Qiu Y. U-Patch GAN: a medical image fusion method based on GAN. *J Digit Imag.* 2023;36(1):339–55. doi:10.1007/s10278-022-00696-7.
67. Guo K, Chen J, Qiu T, Guo S, Luo T, Chen T, et al. MedGAN: an adaptive GAN approach for medical image generation. *Comput Biol Med.* 2023;163(7):107119. doi:10.1016/j.compbiomed.2023.107119.
68. Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. *J Med Imag Radiat Oncol.* 2021;65(5):545–63. doi:10.1111/1754-9485.13261.
69. Garcea F, Serra A, Lamberti F, Morra L. Data augmentation for medical imaging: a systematic literature review. *Comput Biol Med.* 2023;152(1):106391. doi:10.1016/j.compbiomed.2022.106391.
70. Kebaili A, Lapuyade-Lahorgue J, Ruan S. Deep learning approaches for data augmentation in medical imaging: a review. *J Imag.* 2023;9(4):81. doi:10.3390/jimaging9040081.

71. Sun Y, Yuan P, Sun Y. MM-GAN: 3D MRI data augmentation for medical image segmentation via generative adversarial networks. In: 2020 IEEE International Conference on Knowledge Graph (ICKG); 2020 Aug 9–11; Nanjing, China. p. 227–34.
72. Goceri E. Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev.* 2023;56(11):12561–605. doi:10.1007/s10462-023-10453-z.
73. Tschuchnig ME, Gadermayr M. Anomaly detection in medical imaging—a mini review. In: *International Data Science Conference*. Cham, Switzerland: Springer; 2021. p. 33–8.
74. Xia X, Pan X, Li N, He X, Ma L, Zhang X, et al. GAN-based anomaly detection: a review. *Neurocomputing.* 2022;493:497–535. doi:10.1016/j.neucom.2021.12.093.
75. Esmaeili M, Toosi A, Roshanpoor A, Changizi V, Ghazisaeedi M, Rahmim A, et al. Generative adversarial networks for anomaly detection in biomedical imaging: a study on seven medical image datasets. *IEEE Access.* 2023;11:17906–21. doi:10.1109/access.2023.3244741.
76. Vyas B, Rajendran RM. Generative adversarial networks for anomaly detection in medical images. *Int J Multidisci Innova Res Method.* 2023;2(4):52–8.
77. Han C, Rundo L, Murao K, Noguchi T, Shimahara Y, Milacski ZÁ, et al. MADGAN: unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction. *BMC Bioinformatics.* 2021;22(S2):31. doi:10.1186/s12859-020-03936-1.
78. Zhang H, Guo W, Zhang S, Lu H, Zhao X. Unsupervised deep anomaly detection for medical images using an improved adversarial autoencoder. *J Digital Imag.* 2022;35(2):153–61. doi:10.1007/s10278-021-00558-8.
79. Guan H, Liu M. Domain adaptation for medical image analysis: a survey. *IEEE Transact Biomed Eng.* 2022;69(3):1173–85. doi:10.1109/tbme.2021.3117407.
80. Xie X, Chen J, Li Y, Shen L, Ma K, Zheng Y. MI 2 GAN: generative adversarial network for medical image domain adaptation using mutual information constraint. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham, Switzerland: Springer; 2020. p. 516–25.
81. Sun Y, Yang G, Ding D, Cheng G, Xu J, Li X. A GAN-based domain adaptation method for glaucoma diagnosis. In: 2020 International Joint Conference on Neural Networks (IJCNN); 2020 Jul 19–24; Glasgow, UK. p. 1–8.
82. Sagheer SVM, George SN. A review on medical image denoising algorithms. *Biomed Signal Process Cont.* 2020;61(11):102036. doi:10.1016/j.bspc.2020.102036.
83. Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*; 2017 Jul 21–26; Honolulu, HI, USA. p. 1125–34.
84. Kitchenham B, Brereton OP, Budgen D, Turner M, Bailey J, Linkman S. Systematic literature reviews in software engineering—a systematic literature review. *Informati Software Technol.* 2009;51(1):7–15. doi:10.1016/j.infsof.2008.09.009.
85. Ali M, Ali M, Hussain M, Koundal D. Generative adversarial networks (GANs) for medical image processing: recent advancements. *Arch Comput Methods Eng.* 2025;32(2):1185–98. doi:10.1007/s11831-024-10174-8.
86. Tu H, Wang Z, Zhao Y. Unpaired image-to-image translation with diffusion adversarial network. *Mathematics.* 2024;3178. doi:10.21203/rs.3.rs-4502713/v1.
87. Wang G, Shi H, Chen Y, Wu B. Unsupervised image-to-image translation via long-short cycle-consistent adversarial networks. *Appl Intell.* 2023;53(14):17243–59. doi:10.1007/s10489-022-04389-0.
88. Sun L, Chen J, Xu Y, Gong M, Yu K, Batmanghelich K. Hierarchical amortized GAN for 3D high resolution medical image synthesis. *IEEE J Biomed Health Inform.* 2022;26(8):3966–75. doi:10.1109/jbhi.2022.3172976.
89. Xun S, Li D, Zhu H, Chen M, Wang J, Li J, et al. Generative adversarial networks in medical image segmentation: a review. *Comput Biol Med.* 2022;140:105063. doi:10.1016/j.compbiomed.2021.105063.
90. Armanious K, Jiang C, Fischer M, Küstner T, Hepp T, Nikolaou K, et al. MedGAN: medical image translation using GANs. *Comput Med Imag Grap.* 2020;79:101684. doi:10.1016/j.compmedimag.2019.101684.
91. Zhang Q, Sun J, Mok GS. Low dose SPECT image denoising using a generative adversarial network. *arXiv:1907.11944.* 2019.

92. Mahapatra D, Bozorgtabar B, Garnavi R. Image super-resolution using progressive generative adversarial networks for medical image analysis. *Comput Med Imag Grap.* 2019;71(6):30–9. doi:10.1016/j.compmedimag.2018.10.005.
93. You C, Li G, Zhang Y, Zhang X, Shan H, Li M, et al. CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). *IEEE Transact Med Imag.* 2020;39(1):188–203. doi:10.1109/tmi.2019.2922960.
94. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Transact Med Imag.* 2018;37(6):1348–57. doi:10.1109/tmi.2018.2827462.
95. Li Y, Shen L. cC-GAN: a robust transfer-learning framework for HEp-2 specimen image segmentation. *IEEE Access.* 2018;6:14048–58. doi:10.1109/access.2018.2808938.
96. Hu B, Tang Y, Eric I, Chang C, Fan Y, Lai M, et al. Unsupervised learning for cell-level visual representation in histopathology images with generative adversarial networks. *IEEE J Biomed Health Inform.* 2018;23(3):1316–28. doi:10.1109/jbhi.2018.2852639.
97. Nie D, Trullo R, Lian J, Wang L, Petitjean C, Ruan S, et al. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transact Biomed Eng.* 2018;65(12):2720–30.
98. Bissoto A, Perez F, Valle E, Avila S. Skin lesion synthesis with generative adversarial networks. In: *OR 2.0 context-aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures, and skin image analysis.* Cham, Switzerland: Springer; 2018. p. 294–302. doi:10.1007/978-3-030-01201-4\_32.
99. Udrea A, Mitra GD. Generative adversarial neural networks for pigmented and non-pigmented skin lesions detection in clinical images. In: *2017 21st International Conference on Control Systems and Computer Science (CSCS); 2017 May 29–31; Bucharest, Romania.* p. 364–8.
100. Wang Y, Yang N, Li J. GAN-based architecture for low-dose computed tomography imaging denoising. *arXiv:2411.09512.* 2024.
101. Yu X, Luan S, Lei S, Huang J, Liu Z, Xue X, et al. Deep learning for fast denoising filtering in ultrasound localization microscopy. *Phy Med Biol.* 2023;68(20):205002. doi:10.1088/1361-6560/acf98f.
102. Ran M, Hu J, Chen Y, Chen H, Sun H, Zhou J, et al. Denoising of 3D magnetic resonance images using a residual encoder-decoder Wasserstein generative adversarial network. *Med Image Anal.* 2019;55(2):165–80. doi:10.1016/j.media.2019.05.001.
103. Kaur A, Dong G. A complete review on image denoising techniques for medical images. *Neural Process Lett.* 2023;55(6):7807–50. doi:10.1007/s11063-023-11286-1.
104. Yi X, Babyn P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *J Digital Imag.* 2018;31(5):655–69. doi:10.1007/s10278-018-0056-0.
105. Shan H, Zhang Y, Yang Q, Kruger U, Kalra MK, Sun L, et al. 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *IEEE Transact Med Imag.* 2018;37(6):1522–34. doi:10.1109/tmi.2018.2878429.
106. Wolterink JM, Leiner T, Viergever MA, Išgum I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Transact Med Imag.* 2017;36(12):2536–45. doi:10.1109/tmi.2017.2708987.
107. Huang Z, Zhang J, Zhang Y, Shan H. DU-GAN: generative adversarial networks with dual-domain U-Net-based discriminators for low-dose CT denoising. *IEEE Transact Instrument Measure.* 2022;71:1–12. doi:10.1109/tim.2021.3128703.
108. Zhu ML, Zhao LL, Xiao L. Image denoising based on GAN with optimization algorithm. *Electronics.* 2022;11(15):2445. doi:10.3390/electronics11152445.
109. Wang R, Lei T, Cui R, Zhang B, Meng H, Nandi AK. Medical image segmentation using deep learning: a survey. *IET Image Process.* 2022;16(5):1243–67. doi:10.1049/ipr2.12419.
110. Cai L, Fang H, Xu N, Ren B. Counterfactual causal-effect intervention for interpretable medical visual question answering. *IEEE Trans Med Imaging.* 2024;43(12):4430–41. doi:10.1109/TMI.2024.3425533.
111. Song W, Wang X, Guo Y, Li S, Xia B, Hao A. Centerformer: a novel cluster center enhanced transformer for unconstrained dental plaque segmentation. *IEEE Trans Multimedia.* 2024;26:10965–78. doi:10.1109/TMM.2024.3428349.

112. Dong X, Lei Y, Wang T, Thomas M, Tang L, Curran WJ, et al. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Med Phys*. 2019;46(5):2157–68. doi:10.1002/mp.13458.
113. Guo B, Yang Y, Yu Z, Zhu Y. Multi-modality medical image segmentation via adversarial learning with CV energy functional. *Expert Syst Appl*. 2025;283(8):127467. doi:10.1016/j.eswa.2025.127467.
114. Wang G, Ma Q, Li Y, Mao K, Xu L, Zhao Y. A skin lesion segmentation network with edge and body fusion. *Appl Soft Comput*. 2025;170(1):112683. doi:10.1016/j.asoc.2024.112683.
115. Gunasagar P, Durgaprasad V, Abhijith K, Sai RM, Ameer P. Enhancing facial expression synthesis through GAN with multi-scale dilated feature extraction and edge-enhanced facial features. In: *TENCON 2023-2023 IEEE Region 10 Conference (TENCON)*; 2023 Oct 31–Nov 3; Chiang Mai, Thailand. p. 47–52.
116. Nimitha U, Ameer P. MRI super-resolution using similarity distance and multi-scale receptive field based feature fusion GAN and pre-trained slice interpolation network. *Magn Reson Imag*. 2024;110(5):195–209. doi:10.1016/j.mri.2024.04.021.
117. Wang B, Chen W, Qian J, Feng S, Chen Q, Zuo C. Single-shot super-resolved fringe projection profilometry (SSSR-FPP): 100,000 frames-per-second 3D imaging with deep learning. *Light Sci Appl*. 2025;14(1):70. doi:10.1038/s41377-024-01721-w.
118. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*; 2017 Jul 21–26; Honolulu, HI, USA. p. 4681–90.
119. Zhao T, Lin Y, Xu Y, Chen W, Wang Z. Learning-based quality assessment for image super-resolution. *IEEE Transact Multim*. 2022;24:3570–81. doi:10.1109/tmm.2021.3102401.
120. Chen Y, Shi F, Christodoulou AG, Xie Y, Zhou Z, Li D. Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham, Switzerland: Springer; 2018. p. 91–9.
121. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng*. 2009;2:147–71. doi:10.1109/rbme.2009.2034865.
122. Vu CT, Phan TD, Chandler DM. A spectral and spatial measure of local perceived sharpness in natural images. *IEEE Transact Image Process*. 2012;21(3):934–45. doi:10.1109/tip.2011.2169974.
123. Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. *New Engl J Med*. 2007;357(22):2277–84. doi:10.1056/nejmra072149.
124. Gu Y, Zeng Z, Chen H, Wei J, Zhang Y, Chen B, et al. MedSRGAN: medical images super-resolution using generative adversarial networks. *Multim Tools Applcat*. 2020;79(29–30):21815–40. doi:10.1007/s11042-020-08980-w.
125. Wang R, Fang Z, Gu J, Guo Y, Zhou S, Wang Y, et al. High-resolution image reconstruction for portable ultrasound imaging devices. *EURASIP J Adv Signal Process*. 2019;2019(1):1–12. doi:10.1186/s13634-019-0649-x.
126. Luan S, Yu X, Lei S, Ma C, Wang X, Xue X, et al. Deep learning for fast super-resolution ultrasound microvessel imaging. *Phy Med Biol*. 2023;68(24):245023. doi:10.1088/1361-6560/ad0a5a.
127. Nimitha U, Ameer P. Multi image super resolution of MRI images using generative adversarial network. *J Ambient Intell Human Comput*. 2024;15(4):2241–53. doi:10.1007/s12652-024-04751-9.
128. Jia Y, Chen G, Chi H. Retinal fundus image super-resolution based on generative adversarial network guided with vascular structure prior. *Scient Rep*. 2024;14(1):22786. doi:10.1038/s41598-024-74186-x.
129. Wang X, Xie L, Dong C, Shan Y. Real-esrgan: training real-world blind super-resolution with pure synthetic data. In: *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*; 2021 Oct 11–17; Montreal, BC, Canada. p. 1905–14.
130. Zhao Y, Zheng Y, Liu Y, Zhao Y, Luo L, Yang S, et al. Automatic 2-D/3-D vessel enhancement in multiple modality images using a weighted symmetry filter. *IEEE Transact Med Imag*. 2017;37(2):438–50. doi:10.1109/tmi.2017.2756073.
131. Jobson DJ, Rahman Z, Woodell GA. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transact Image Process*. 1997;6(7):965–76. doi:10.1109/83.597272.

132. Chen YS, Wang YC, Kao MH, Chuang YY. Deep photo enhancer: unpaired learning for image enhancement from photographs with gans. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition; 2018 Jun 18–23; Salt Lake City, UT, USA. p. 6306–14.
133. Alotaibi A. Deep generative adversarial networks for image-to-image translation: a review. *Symmetry*. 2020;12(10):1705. doi:10.3390/sym12101705.
134. Lore KG, Akintayo A, Sarkar S. LLNet: a deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit*. 2017;61(6):650–62. doi:10.1016/j.patcog.2016.06.008.
135. Gatys LA. A neural algorithm of artistic style. arXiv:1508.06576. 2015.
136. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the 2017 IEEE International Conference on Computer Vision; 2017 Oct 22–29; Venice, Italy. p. 2223–32.
137. Yan S, Wang C, Chen W, Lyu J. Swin transformer-based GAN for multi-modal medical image translation. *Front Oncol*. 2022;12:942511. doi:10.3389/fonc.2022.942511.
138. Liang J, Cao J, Sun G, Zhang K, Van Gool L, Timofte R. Swinir: image restoration using swin transformer. In: Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision; 2021 Oct 11–17; Montreal, BC, Canada. p. 1833–44.
139. Chen Y, Lin Y, Xu X, Ding J, Li C, Zeng Y, et al. Multi-domain medical image translation generation for lung image classification based on generative adversarial networks. *Comput Meth Prog Biomed*. 2023;229(20):107200. doi:10.1016/j.cmpb.2022.107200.
140. Özbey M, Dalmaz O, Dar SU, Bedel HA, Öztürk Ş, Güngör A, et al. Unsupervised medical image translation with adversarial diffusion models. *IEEE Transact Med Imag*. 2023;42(12):3524–39. doi:10.1109/tmi.2023.3290149.
141. Kim KH, Do WJ, Park SH. Improving resolution of MR images with an adversarial network incorporating images with different contrast. *Med Phy*. 2018;45(7):3120–31. doi:10.1002/mp.12945.
142. Shitrit O, Riklin Raviv T. Accelerated magnetic resonance imaging by adversarial neural network. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. Cham, Switzerland: Springer; 2017. p. 30–8. doi:10.1007/978-3-319-67558-9\_4.
143. Seitzer M, Yang G, Schlemper J, Oktay O, Würfl T, Christlein V, et al. Adversarial and perceptual refinement for compressed sensing MRI reconstruction. In: Medical image computing and computer assisted intervention-MICCAI 2018. Cham, Switzerland: Springer; 2018. p. 232–40. doi:10.1007/978-3-030-00928-1\_27.
144. Armanious K, Mecky Y, Gatidis S, Yang B. Adversarial inpainting of medical image modalities. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2019 May 12–17; Brighton, UK. p. 3267–71.
145. Wang Y, Yu B, Wang L, Zu C, Lalush DS, Lin W, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *Neuroimage*. 2018;174(7):550–62. doi:10.1016/j.neuroimage.2018.03.045.
146. Liao H, Huo Z, Sehnert WJ, Zhou SK, Luo J. Adversarial sparse-view CBCT artifact reduction. In: Medical image computing and computer assisted intervention-MICCAI 2018I. Cham, Switzerland: Springer; 2018. p. 154–62. doi:10.1007/978-3-030-00928-1\_18.
147. Quan TM, Nguyen-Duc T, Jeong WK. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Transact Med Imag*. 2018;37(6):1488–97. doi:10.1109/tmi.2018.2820120.
148. Mardani M, Gong E, Cheng JY, Vasanawala SS, Zaharchuk G, Xing L, et al. Deep generative adversarial neural networks for compressive sensing MRI. *IEEE Transact Med Imag*. 2018;38(1):167–79. doi:10.1109/tmi.2018.2858752.
149. Kang E, Koo HJ, Yang DH, Seo JB, Ye JC. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. *Med Phy*. 2019;46(2):550–62. doi:10.1002/mp.13284.
150. Bhadra S, Zhou W, Anastasio MA. Medical image reconstruction with image-adaptive priors learned by use of generative adversarial networks. In: Medical imaging 2020: physics of medical imaging. Vol. 11312. Bellingham, WA, USA: SPIE; 2020. p. 206–13.
151. Rashid ZM, Alsawaff ZH, Yousif AS, Al-Nima RRO. Optimization of PET image reconstruction for enhanced image quality in various tasks using a conventional PET scanner. *J Elect Comput Eng*. 2025;2025(1):8108611. doi:10.1155/jece/8108611.

152. Ahn G, Choi BS, Ko S, Jo C, Han HS, Lee MC, et al. High-resolution knee plain radiography image synthesis using style generative adversarial network adaptive discriminator augmentation. *J Orthop Res*<sup>®</sup>. 2023;41(1):84–93. doi:10.1002/jor.25325.
153. Zhu L, He Q, Huang Y, Zhang Z, Zeng J, Lu L, et al. DualMMP-GAN: dual-scale multi-modality perceptual generative adversarial network for medical image segmentation. *Comput Biol Med*. 2022;144(10):105387. doi:10.1016/j.compbiomed.2022.105387.
154. Che Azemin MZ, Mohd Tamrin MI, Hilmi MR, Mohd Kamal K. Assessing the efficacy of StyleGAN 3 in generating realistic medical images with limited data availability. In: *Proceedings of the 2024 13th International Conference on Software and Computer Applications*; 2024 Feb 1–3; Bali Island, Indonesia. p. 192–7.
155. Zhao A, Xu M, Shahin AH, Wuyts W, Jones MG, Jacob J, et al. 4D VQ-GAN: synthesising medical scans at any time point for personalised disease progression modelling of idiopathic pulmonary fibrosis. arXiv:2502.05713. 2025.
156. Antoniou A, Storkey A, Edwards H. Data augmentation generative adversarial networks. arXiv:1711.04340. 2017.
157. Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H. Synthetic data augmentation using GAN for improved liver lesion classification. In: *IEEE 15th International Symposium On Biomedical Imaging (ISBI 2018)*; 2018 Apr 4–7; Washington, DC, USA; 2018. p. 289–93.
158. Gan HS, Ramlee MH, Al-Rimy BAS, Lee YS, Akkaraekthalin P. Hierarchical knee image synthesis framework for generative adversarial network: data from the osteoarthritis initiative. *IEEE Access*. 2022;10:55051–61. doi:10.1109/access.2022.3175506.
159. Bowles C, Chen L, Guerrero R, Bentley P, Gunn R, Hammers A, et al. Gan augmentation: augmenting training data using generative adversarial networks. arXiv:1810.10863. 2018.
160. Kong L, Lian C, Huang D, Hu Y, Zhou Q. Breaking the dilemma of medical image-to-image translation. *Adv Neural Inform Process Syst*. 2021;34:1964–78.
161. Du W, Tian S. Transformer and GAN-based super-resolution reconstruction network for medical images. *Tsinghua Sci Technol*. 2024;29(1):197–206. doi:10.26599/tst.2022.9010071.
162. Zuo Q, Tian H, Li R, Guo J, Hu J, Tang L, et al. Hemisphere-separated cross-connectome aggregating learning via VAE-GAN for brain structural connectivity synthesis. *IEEE Access*. 2023;11:48493–505. doi:10.1109/access.2023.3276989.
163. Zhang R, Lu W, Gao J, Tian Y, Wei X, Wang C, et al. RFI-GAN: a reference-guided fuzzy integral network for ultrasound image augmentation. *Inform Sci*. 2023;623(5):709–28. doi:10.1016/j.ins.2022.12.026.
164. Zhang Y, Wang Z, Zhang Z, Liu J, Feng Y, Wee L, et al. GAN-based one dimensional medical data augmentation. *Soft Comput*. 2023;27(15):10481–91. doi:10.1007/s00500-023-08345-z.
165. Showrov I, Aziz MT, Nabil HR, Jim JR, Kabir MM, Mridha M, et al. Generative adversarial networks (GANs) in medical imaging: advancements, applications and challenges. *IEEE Access*. 2024;12:35728–53. doi:10.1109/access.2024.3370848.
166. Saad MM, O'Reilly R, Rehmani MH. A survey on training challenges in generative adversarial networks for biomedical image analysis. *Artif Intell Rev*. 2024;57(2):19. doi:10.1007/s10462-023-10624-y.
167. Biswas S, Rohdin J, Drahanský M. Synthetic retinal images from unconditional GANs. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; 2019 Jul 23–27; Berlin, Germany. p. 2736–9.
168. Abdelhalim ISA, Mohamed MF, Mahdy YB. Data augmentation for skin lesion using self-attention based progressive generative adversarial network. *Expert Syst Applicat*. 2021;165(3):113922. doi:10.1016/j.eswa.2020.113922.
169. Goel T, Murugan R, Mirjalili S, Chakrabartty DK. Automatic screening of covid-19 using an optimized generative adversarial network. *Cognit Computat*. 2024;16(4):1666–81. doi:10.1007/s12559-020-09785-7.
170. Wang Z, She Q, Ward TE. Generative adversarial networks in computer vision: a survey and taxonomy. *ACM Comput Surv (CSUR)*. 2021;54(2):1–38. doi:10.1145/3439723.
171. 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*. Cambridge, MA, USA: The MIT Press; 2017. doi:10.18034/ajase.v8i1.9.

172. Mirjalili S, Lewis A. The whale optimization algorithm. *Adv Eng Soft.* 2016;95(12):51–67. doi:10.1016/j.advengsoft.2016.01.008.
173. Dukler Y, Li W, Lin A, Montúfar G. Wasserstein of Wasserstein loss for learning generative models. In: *The 36th International Conference on Machine Learning*; 2019 Jun 10–15; Long Beach, CA, USA. p. 1716–25.
174. Laino ME, Cancian P, Politi LS, Della Porta MG, Saba L, Savevski V. Generative adversarial networks in brain imaging: a narrative review. *J Imaging.* 2022;8(4):83. doi:10.3390/jimaging8040083.
175. Wolterink JM, Mukhopadhyay A, Leiner T, Vogl TJ, Bucher AM, Išgum I. Generative adversarial networks: a primer for radiologists. *Radiographics.* 2021;41(3):840–57. doi:10.1148/rg.2021200151.
176. Lee M, Seok J. Regularization methods for generative adversarial networks: an overview of recent studies. *arXiv:2005.09165.* 2020.
177. Miyato T, Kataoka T, Koyama M, Yoshida Y. Spectral normalization for generative adversarial networks. *arXiv:1802.05957.* 2018.
178. Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-normalizing neural networks. In: *Advances in neural information processing systems.* Cambridge, MA, USA: The MIT Press; 2017.
179. Xu L, Zeng X, Huang Z, Li W, Zhang H. Low-dose chest X-ray image super-resolution using generative adversarial nets with spectral normalization. *Biomed Signal Process Cont.* 2020;55(6):101600. doi:10.1016/j.bspc.2019.101600.
180. Hoang Q, Nguyen TD, Le T, Phung D. MGAN: training generative adversarial nets with multiple generators. In: *The 6th International Conference on Learning Representations*; 2018 Apr 30–May 3; Vancouver, BC, Canada. p. 1–24.
181. Wu Y, Yue Y, Tan X, Wang W, Lu T. End-to-end chromosome Karyotyping with data augmentation using GAN. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*; 2018 Oct 7–10; Athens, Greece. p. 2456–60.
182. Neff T, Payer C, Stern D, Urschler M. Generative adversarial network based synthesis for supervised medical image segmentation. In: *Proceedings of the OAGM and ARW Joint Workshop*; 2017 May 10–12; Wien, Austria. p. 140–5.
183. Saad MM, Rehmani MH, O'Reilly R. Addressing the intra-class mode collapse problem using adaptive input image normalization in gan-based x-ray images. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*; 2022 Jul 11–15; Glasgow, UK. p. 2049–52.
184. Saad MM, Rehmani MH, O'Reilly R. A self-attention guided multi-scale gradient GAN for diversified x-ray image synthesis. In: *Irish Conference on Artificial Intelligence and Cognitive Science*; 2022 Dec 8–9; Munster, Ireland. p. 18–31.
185. Xue Y, Zhou Q, Ye J, Long LR, Antani S, Cornwell C, et al. Synthetic augmentation and feature-based filtering for improved cervical histopathology image classification. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*; 2019 Oct 13–17; Shenzhen, China: Springer. p. 387–96.
186. Kudo A, Kitamura Y, Li Y, Iizuka S, Simo-Serra E. Virtual thin slice: 3D conditional GAN-based super-resolution for CT slice interval. In: *International Workshop on Machine Learning for Medical Image Reconstruction*; 2019 Oct 17; Shenzhen, China. p. 91–100. doi:10.1007/978-3-030-33843-5\_9.
187. Segato A, Corbetta V, Di Marzo M, Pozzi L, De Momi E. Data augmentation of 3D brain environment using deep convolutional refined auto-encoding alpha GAN. *IEEE Transact Med Robot Bionics.* 2020;3(1):269–72. doi:10.1109/tmrb.2020.3045230.
188. Kwon G, Han C, Kim DS. Generation of 3D brain MRI using auto-encoding generative adversarial networks. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*; 2019 Oct 13–17; Shenzhen, China. p. 118–26.
189. Qin Z, Liu Z, Zhu P, Xue Y. A GAN-based image synthesis method for skin lesion classification. *Comput Meth Prog Biomed.* 2020;195(2):105568. doi:10.1016/j.cmpb.2020.105568.
190. Modanwal G, Vellal A, Mazurowski MA. Normalization of breast MRIs using cycle-consistent generative adversarial networks. *Comput Meth Prog Biomed.* 2021;208(4):106225. doi:10.1016/j.cmpb.2021.106225.
191. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: a large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*; 2009 Jun 20–25; Miami, FL, USA. p. 248–55.
192. Sara U, Akter M, Uddin MS. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. *J Comput Communicat.* 2019;7(3):8–18. doi:10.4236/jcc.2019.73002.

193. Korkinof D, Harvey H, Heindl A, Karpati E, Williams G, Rijken T, et al. Perceived realism of high-resolution generative adversarial network-derived synthetic mammograms. *Radi Artif Intell.* 2021;3(2):e190181. doi:10.1148/ryai.2020190181.
194. Wang Z, Lim G, Ng WY, Tan TE, Lim J, Lim SH, et al. Synthetic artificial intelligence using generative adversarial network for retinal imaging in detection of age-related macular degeneration. *Front Med.* 2023;10:1184892. doi:10.3389/fmed.2023.1184892.
195. Mardani M, Gong E, Cheng JY, Vasanawala S, Zaharchuk G, Alley M, et al. Deep generative adversarial networks for compressed sensing automates MRI. *arXiv:1706.00051.* 2017.
196. Li Y, Yang H, Xie D, Dreizin D, Zhou F, Wang Z. POCS-augmented CycleGAN for MR image reconstruction. *Appl Sci.* 2021;12(1):114. doi:10.3390/app12010114.
197. Ehrhardt J, Wilms M. Autoencoders and variational autoencoders in medical image analysis. In: Burgos N, Svoboda D, editors. *Biomedical image synthesis and simulation: methods and applications.* Cambridge, MA, USA: Academic Press; 2022. p. 129–62. doi:10.1016/b978-0-12-824349-7.00015-3.
198. Kazerouni A, Aghdam EK, Heidari M, Azad R, Fayyaz M, Hacihaliloglu I, et al. Diffusion models in medical imaging: a comprehensive survey. *Med Image Anal.* 2023;88(1):102846. doi:10.1016/j.media.2023.102846.