IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 January 2024, accepted 21 February 2024, date of publication 26 February 2024, date of current version 12 March 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3370848*

SURVEY

# Generative Adversarial Networks (GANs) in Medical Imaging: Advancements, Applications, and Challenges

**SHOWROV ISLAM[1], MD. TAREK AZIZ[1], HADIUR RAHMAN NABIL[1], JAMIN RAHMAN JIM[1], M. F. MRIDHA[1], (Senior Member, IEEE), MD. MOHSIN KABIR[2], NOBUYOSHI ASAI[3], AND JUNGPIL SHIN[3], (Senior Member, IEEE)**

[1]Department of Computer Science and Engineering, American International University-Bangladesh, Dhaka 1229, Bangladesh
[2]Faculty of Informatics, Eövös Loránd University, 1117 Budapest, Hungary
[3]School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu 965-8580, Japan

Corresponding author: Jungpil Shin (jpshin@u-aizu.ac.jp)

This work was supported by the Competitive Research Fund of The University of Aizu, Japan.

**ABSTRACT** Generative Adversarial Networks are a class of artificial intelligence algorithms that consist of a generator and a discriminator trained simultaneously through adversarial training. GANs have found crucial applications in various fields, including medical imaging. In healthcare, GANs contribute by generating synthetic medical images, enhancing data quality, and aiding in image segmentation, disease detection, and medical image synthesis. Their importance lies in their ability to generate realistic images, facilitating improved diagnostics, research, and training for medical professionals. Understanding its applications, algorithms, current advancements, and challenges is imperative for further advancement in the medical imaging domain. However, no study explores the recent state-of-the-art development of GANs in medical imaging. To overcome this research gap, in this extensive study, we began by exploring the vast array of applications of GANs in medical imaging, scrutinizing them within recent research. We then dive into the prevalent datasets and pre-processing techniques to enhance comprehension. Subsequently, an in-depth discussion of the GAN algorithms, elucidating their respective strengths and limitations, is provided. After that, we meticulously analyzed the results and experimental details of some recent cutting-edge research to obtain a more comprehensive understanding of the current development of GANs in medical imaging. Lastly, we discussed the diverse challenges encountered and future research directions to mitigate these concerns. This systematic review offers a complete overview of GANs in medical imaging, encompassing their application domains, models, state-of-the-art results analysis, challenges, and research directions, serving as a valuable resource for multidisciplinary studies.

**INDEX TERMS** Generative adversarial networks, medical imaging, medical image synthesis, medical image enhancement, medical image augmentation, medical image segmentation.

## I. INTRODUCTION

Genenerative Adversarial Networks (GANs) stand out as a powerful class of neural networks that use parallel training of two networks for image generation and discrimination. It is known for its ability to produce realistic images and handle domain shifts [1]. The computed distribution of probabilities is what Generative Adversarial Networks use to generate more instances. While there are other deep-learning generative models available, GANs are particularly impressive at producing detailed, real-life visuals. GANs have gained significant popularity in academia and industry, offering unique research opportunities due to their foundation in game theory, which distinguishes them from traditional optimization-based generative modeling techniques [2].

GANs in medical imaging are revolutionizing diagnostic accuracy and image enhancement by generating high-quality medical images from limited datasets. For

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Giannelli.

example, CycleGAN is used for domain adaptation by translating images in different modalities [3], while pix2pix focuses on translation between images by supporting tasks such as resolution enhancement or denoising of medical images [4]. Another notable GAN algorithm in the area of medical imaging is UNIT GAN, which enables cross-modal image fusion in medical imaging by learning a shared latent space between different modalities, whereas ProGAN makes it easier to produce high-resolution medical images, creating detailed and genuine images that are critical for diagnostics and healthcare research [5]. These GAN algorithms have significantly improved the quality and diversity of medical images, enabling better analysis and diagnosis in healthcare settings.

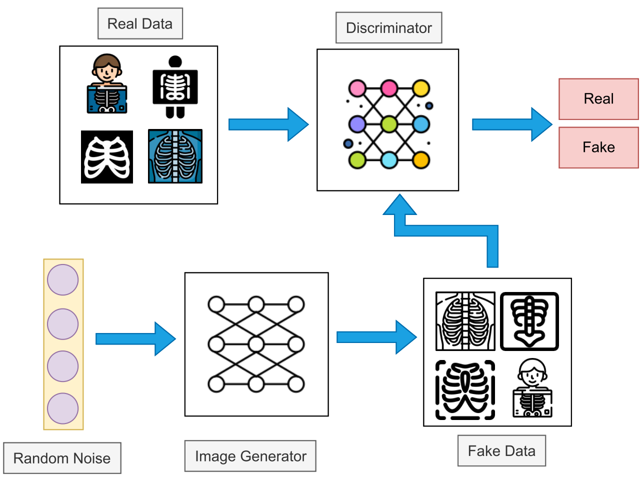


**FIGURE 1. Workflow of GANs in medical imaging.**

GANs in medical imaging involve two neural networks where one is the generator (G) and another is the discriminator (D). The generator produces synthetic medical images from random noise, aiming to replicate real data, while the discriminator evaluates and differentiates between genuine and generated images. The generator refines its output through iterative feedback to create more realistic data, contributing to applications like augmentation, synthesis, classification, detection, reconstruction, segmentation, etc. GANs enhance diagnostic image quality, address data limitations, and support machine learning model development in healthcare. Figure 1 demonstrates the basic workflow of GAN in medical imaging.

As GANs can produce high-quality medical images even in the face of limited datasets, they have revolutionized diagnostic precision and image enhancement in the area of medical imaging [6]. Over time, GANs have evolved significantly, beginning with the basic adversarial process in 2014 but facing challenges in fully covering data distribution [7]. The introduction of DCGAN in 2015 improved image quality [8], followed by WGAN in 2017 addressing mode stability [9]. CycleGAN enabled image-to-image translation without paired training data [3], while PGAN introduced a progressive training approach [5]. Later, SAGAN in 2019 focused on relevant image regions and long-term relationships [3]. Recent advancements include RANDGAN, emphasizing segmentation for anomaly detection and outperforming traditional GANs in medical imaging [10]. DGGAN generates anonymous brain vascular images for medical imaging from MRA patches [11], while ED-GAN combines VAEs and GANs [12]. These improvements indicate the remarkable evolution of GANs, expanding their applications in medical imaging, anomaly detection, and complex data synthesis. In medical imaging, GAN applications include segmentation, reconstruction, detection, classification, augmentation, registration, and image synthesis. To understand GANs in medical imaging, a comprehensive review of recent advancements in GANs in medical imaging is essential, covering application domains, datasets, pre-processing techniques, GAN algorithms, and recent experimental results. Existing surveys lack a comprehensive scope, demonstrating a research gap. This survey focuses specifically on recent research articles, offering an up-to-date overview of GAN advancements in medical imaging. It explores diverse applications, commonly used datasets, pre-processing methods, GAN algorithms, and recent experimental findings. The study identifies challenges and recommends forthcoming research approaches. Table 1 provides a comparison with existing surveys. This article assists investors and researchers in developing GAN-based medical imaging ecosystems and understanding their progress. The paper's key contributions include a thorough analysis of recent developments, providing valuable insights for the field's advancement.

The overall contributions of this paper are:

- An in-depth Exploration into the various application domains of GANs in Medical Imaging. This exploration discusses the researchers' progress in this field, encompassing a broad spectrum of practical areas of Medical Imaging where GANs are applied.
- Provides an overview of commonly used datasets, data pre-processing techniques, and evaluation metrics for optimal model training in GANs.
- Analyzes different GAN models used in Medical Imaging, elucidating their suitability for specific healthcare applications and aiding researchers in informed model selection.
- A comprehensive examination scrutinizes recent advancements and contributions of GANs by researchers in Medical Imaging, offering insights into the state-of-the-art experimental results that influence the landscape of the subject.
- Finally, the paper concludes by emphasizing the challenges and future research directions for GANs in medical imaging, guiding researchers toward potential areas of exploration and innovation, and facilitating their future research endeavors.

The remaining article is structured as follows: First, section II discusses the methodology employed for conducting this review of GANs. Then, section III explores the various applications of GANs in Medical Imaging. Section IV

**TABLE 1. Comparative analysis of recent GANs surveys in medical imaging.**

| Ref. | Applications | Datasets | Preprocessing Methods | Algorithms | Result Analysis | Challenges and Future Work | Contribution |
|---|---|---|---|---|---|---|---|
| [15] | ✓(Focused on Classification and Segmentation) | ✗ | ✗ | ✓ | ✓ | ✗(Only Challenges) | This paper addressed a systematic review of recent GANs architectures for medical image analysis between 2015-2020, emphasizing their potential to address dataset limitations and improve classification and segmentation tasks while drawing attention to the need for rigorous validation of GAN generated images to ensure clinical reliability. |
| [16] | ✓(Focused on Augmentation) | ✗ | ✗ | ✓ | ✓ | ✓ | This paper gives an extensive review and analysis of GAN-based medical image augmentation methodologies from 2018 to 2021, offering insights into benchmark models, loss functions, and evaluation metrics, aiming to guide and inspire future research in the domain. |
| [17] | ✓(Focused on Segmentation) | ✗ | ✗ | ✓ | ✓ | ✗ | This paper issued a comprehensive review of over 120 GAN-based architectures for medical image segmentation, highlighting their advantages, challenges, and pointing to forthcoming research directions for enhancing accuracy and clinical adoption. |
| [18] | ✓ | ✗ | ✗ | ✓ | ✓ | ✓ | This article offers an in-depth review of the advancements and applications of GANs in medical imaging, clarifying its basics, extensions, and role in tasks such as cross-modality, augmentation, and lesion segmentation, while also addressing training challenges and prospects. |
| [19] | ✓(Focused on Augmentation) | ✗ | ✗ | ✓ | ✗ | ✗(Only Future Work ) | This article provides a detailed analysis of GAN-based models for medical image augmentation, discussing popular architectures, imaging modalities, and body organs while underscoring challenges, evaluation metrics, and future directions for implementation in clinical settings. |
| Our Paper | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Provides a comprehensive review of the state-of-the-art advancement of GANs in Medical Imaging by covering its applications, datasets, preprocessing methods, evaluation metrics, GAN models, experimental results analysis, challenges, and future research opportunities. |

discusses some of the most commonly used datasets, preprocessing methods, and evaluation metrics GANs in Medical Imaging experiments. After that, section V offers an analysis of the commonly utilized GAN models in Medical Imaging. Section VI some of the state-of-the-art experimental results are then analyzed, and section VII outlines the research challenges and potential areas for improvement in the field of GANs. Finally, section VIII concludes the paper.

## II. METHODOLOGY

This study employs a systematic literature review (SLR) methodology based on the framework proposed by Keele et al. [17] and Kitchenham et al. [18]. The survey focuses specifically on high-quality academic articles from reputable databases such as ScienceDirect, SpringerLink, ACM Digital Library, and IEEE Xplore. The research methodology adheres to the guidelines outlined in Figure 2 of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), ensuring a systematic and transparent approach to data collection. The critical resources for this survey were gathered following the recommendations of PRISMA, a recognized reporting guideline for systematic reviews. Additionally, the inclusion and exclusion criteria utilized in PRISMA are thoroughly documented in Table 2. This table provides a comprehensive overview of the criteria employed to determine whether a paper is eligible for inclusion in the review or should be excluded from consideration.

We selected 725 articles for the review at the start of this study. Table 3 displays all of the keywords we used to choose articles from various databases and the number

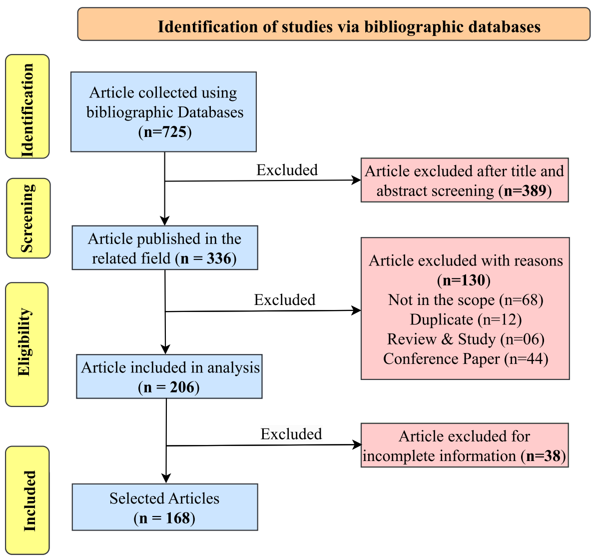


**FIGURE 2. PRISMA flow diagram of the article selection procedure.**

of articles we chose for analysis for each term. These items were chosen throughout four distinct periods. These include December 2022–January 2023, March 2023–April 2023, September 2023–October 2023, and final selection was conducted from 4th December 2023–25th December 2023.

After a comprehensive assessment, 174 articles focused on applications, difficulties, and recent developments were finally chosen for review. The bibliometric analysis of our chosen articles is shown in Figures 3 based on the Scopus index. Our purposeful decision to only consider publications

**TABLE 2.** **The table discusses the including and excluding criteria for selecting articles.**

| | Inclusion Criteria | Exclusion Criteria |
|---|---|---|
| Types of study | Original and review articles. | Thesis, white papers, communication letters, reports, and editorials. |
| Language | Research articles conducted in English. | Duplicate and non-English articles. |
| Publication Year | Articles published in 2020 to 2023 (For applications and results analysis part). | Not related to the theme of the review. |
| Source | Articles published only in academic journals and conferences. | Articles that lack information and review papers. |
| Intervention | Machine Learning and Generative method. | Traditional and Statistical methods. |
| Region | Not restricted to a particular region. | - |
| Settings | Generative Adversarial Networks in Medical Imaging. | Not in medical imaging settings. |

**TABLE 3.** **The table discusses the keywords that we used for article selection in different databases and the paper count for respective keywords.**

| Keyword | Paper count |
|---|---|
| GANs in medical imaging. | 176 |
| GANs in medical image segmentation | 108 |
| GANs in medical image augmentation | 72 |
| GANs in medical image synthesis | 74 |
| GANs in medical image reconstruction | 41 |
| GANS in medical image registration | 36 |
| GANS in medical image detection | 92 |
| GANs in medical image classification | 126 |

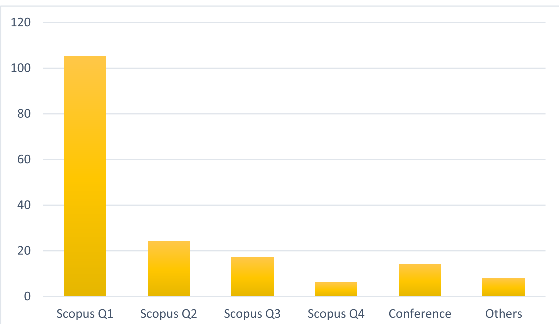


**FIGURE 3.** **Number of articles based on scopus index.**

published recently emphasizes our dedication to offering an advanced and leading-edge evaluation.

## III. APPLICATIONS OF GANS IN MEDICAL IMAGING

Generative Adversarial Networks are making significant strides in medical imaging applications. One critical use is in generating synthetic images that closely mimic real patient data, which is crucial for augmenting limited labeled datasets. GANs also excel at enhancing medical images by creating high-resolution versions from lower-resolution inputs, improving diagnostic precision. Additionally, GANs facilitate cross-modal image transformations, converting CT scans to MRIs and vice versa. Their capacity to produce realistic and diverse medical images contributes to image analysis and segmentation advancements, aiding healthcare professionals in making more accurate clinical decisions [19]. Figure 4 shows some applications of GAN in medical imaging

### A. SEGMENTATION

Image segmentation is used as a method to improve picture reconstruction, allow multi-modal fusion, and provide uncertainty estimation in medical imaging. Medical diagnoses are reliable more accurate and when using segmentation-based techniques [15]. Table 4 highlights some key applications in the segmentation of medical images.

### B. RECONSTRUCTION

Image reconstruction is a process in which an algorithm or model reconstructs an image from incomplete or degraded data. It is typically used in medical imaging, computer vision, and remote sensing. In Generative Adversarial Networks, image reconstruction often refers to producing high-fidelity images from low-resolution or noisy input through the use of Generative Adversarial Networks. GANs consist of a generator that learns to create realistic images from random noise or imperfect data. They are used to increase the quality of images, making them particularly valuable for tasks like super-resolution imaging, denoising, and enhancing the visual quality of medical scans, where clear and detailed images are critical for accurate diagnosis and analysis [29]. Table 5 shows some applications of GANs in image reconstruction.

### C. DETECTION

Generative Adversarial Networks have become a potent tool in medical imaging for tasks related to detection. GANs identify medical conditions such as cancer, anomalies, pneumonia, tumors, etc., by generating synthetic images of healthy tissues or normal anatomy and then discerning anomalies in actual medical images. This approach enhances the accuracy of early disease detection, reducing the risk of missed diagnoses and improving patient outcomes [35]. The utilization of GANs in medical image detection is illustrated in table 6.

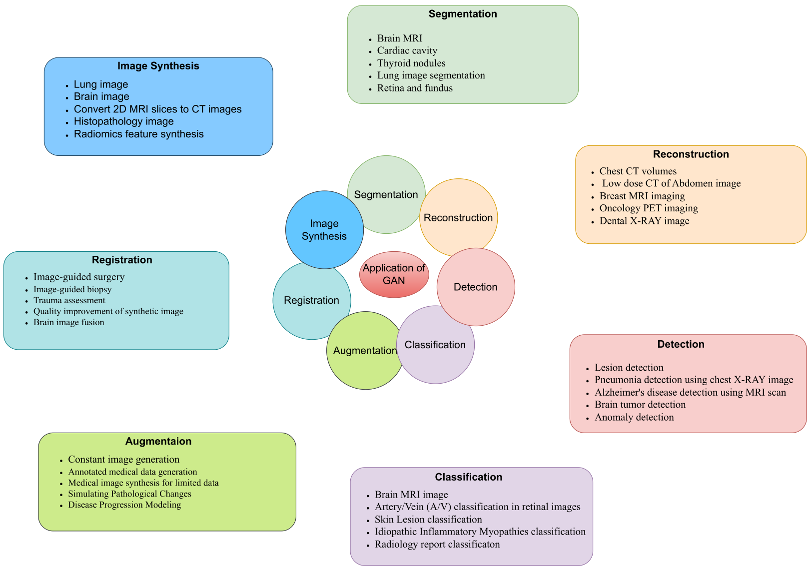


**FIGURE 4. The figure shows various applications, including Segmentation, Reconstruction, Detection, Classification, Augmentation, Registration, and Image Synthesis of GAN in medical imaging.**

**TABLE 4. Application of GANs in medical image segmentation.**

| Ref. | Description |
|---|---|
| Zhang *et al.* [24] | Proposed an image segmentation is used to aid medical diagnosis, pointing out the challenges in generalizing deep learning methods due to the differences in the distribution of regions. To improve medical image segmentation performance, it is proposed to mitigate this with bidirectional GANs, particularly CycleGAN. |
| Güven *et al.* [25] | Introduced SSimDCL, a novel GAN tailored for brain MRI segmentation to improve accuracy, which is critical for diagnosing brain diseases such as cancer, and validates its performance through four detailed studies . |
| Al Khalil *et al.* [26] | Utilized Deep learning-based image segmentation and GAN-fused CMR images to enhance the robustness of cardiac cavity segmentation models in MRI by addressing performance degradation across different image domains. |
| Al Khalil *et al.* [27] | Also introduced an optimized framework for cardiac structures in CMR image segmentation, incorporating key components such as heart detection, image fusion using GAN for realism, and late fusion segmentation with intensity transformations to enhance segmentation accuracy and consistency of clinical metrics. |
| Kunapinun *et al.* [28] | Proposed a hybrid approach for thyroid nodules segmentation in ultrasound images that combines semantic segmentation with GANs to improve performance while introducing closed-loop supervision of output gain loss in the discriminator to reduce GAN training instability, using StableSeg GAN with the generator DeeplabV3+ and discriminator as ResNet18, applying PID control for GAN training stability. |
| Rezaei *et al.* [29] | Introduced the lung segmentation approach, which uses snake optimization to isolate the lung region from the background accurately. Subsequently, the segmented lung images are confronted with three fuzzy data-based clustering methods (KFCM, FCM, FRFCM and SAFCM) using evaluation metrics such as intersection over union (IoU) and Hausdorff distance to evaluate the quality and accuracy of the segmentation results. |
| Narayanan *et al.* [30] | Used image segmentation to identify brain tumors in medical images, and the proposed approach uses generative networks to generate synthetic data pairs to improve training and segmentation accuracy. |
| Zhu *et al.* [31] | Proposed an improved segmentation performance of the DualMMP-GAN framework that contributes to better clinical diagnosis and medical image analysis in brain tumor detection and treatment planning. |
| Munawar *et al.* [32] | Introduced a Generative Adversarial Network (GAN) for lung segmentation in chest X-ray (CXR) images, where the generator is trained to produce realistic segmented masks emphasizing the importance of reliable CXR segmentation for computer-aided diagnostic systems. |

## D. CLASSIFICATION

Generative Adversarial Networks indirectly aid image classification by diversifying training data through data augmentation, aligning domain distributions for robustness, improving image quality for enhanced feature extraction, detecting anomalies, and providing unlabeled samples for semi-supervised learning. This boosts classifier accuracy and adaptability across various applications [42]. The table 7 shows some applications of GAN in image classification.

## E. AUGMENTATION

Image augmentation is commonly employed in data and image processing, particularly in medical image analysis. It involves applying various transformations like rotation, scaling, and noise addition to expand a dataset's size and diversity. In medical imaging, where datasets can be limited and diverse clinical cases are crucial, augmentation becomes essential. By artificially generating new training examples or modifying existing ones, augmentation improves the

**TABLE 5.** Application of GANs in medical image reconstruction.

| Ref. | Description |
|---|---|
| Du *et al.* [34] | Proposed T-GANs, utilizing deep learning to enhance low-resolution medical images. Particularly in low-field MRI, which focuses on preserving texture details. |
| Wang *et al.* [35] | Introduced TRCT-GAN, a GAN network designed to recreate new chest CT volumes using biplane X-ray pictures. It utilizes a Transformer network module and a dynamic attention module to boost the feature representation and contextual association . |
| Mishra *et al.* [36] | An approach for synthesizing visual stimuli from EEG data, focusing on images of objects, digits, and characters, was introduced. The proposed architecture combines an attention and auxiliary classifier-based GAN. |
| Jiang *et al.* [37] | presented a Generative Adversarial Network (PLA-GAN) based on Proximal Linear ADMM framework for Low-dose Computed Tomography (LDCT) reconstruction . |
| Ramanathan *et al.* [38] | Introduced Low-Dose CT image reconstruction using an autoencoder network with vector quantization is a novel method . |

**TABLE 6.** Application of GANs in medical image detection.

| Ref. | Description |
|---|---|
| Guan *et al.* [40] | Proposed lesion detection case studies in medical imaging is efficiently achieved through the practical application of the Texture-Constrained Multichannel Progressive Generative Adversarial Network (TMP-GAN). By utilizing joint training across multiple channels, TMP-GAN successfully overcomes the typical drawbacks of current-generation methods. |
| Srivastav *et al.* [41] | Highlighted the extensive integration of deep learning in medical diagnostics, specifically in pneumonia detection through the utilization of chest X-ray images. This involves the application of transfer learning, employing the VGG-16 model, and incorporating synthetic image augmentation through GANs. |
| Zhang *et al.* [42] | The challenges in detecting lesions in computed tomography (CT) images, such as image quality degradation, noise interference, complex lesion shapes, and indistinct object-background differentiation. The authors suggest a symmetrical GAN detection network grounded on a one-stage network for detection. They employ the Deep Lesion dataset to overcome data limitations commonly encountered in medical datasets. |
| Vashisht *et al.* [43] | Highlighted Alzheimer's disease detection with MRI scans, focusing on classifying it into four categories: healthy, very mild demented, mild demented, and moderately demented. The study utilizes deep learning algorithms, with CNN as the base model, augmented with GAN-generated data. |
| Reddy *et al.* [44] | Application of GANs to augment small and fragmented medical image datasets, particularly for brain tumor detection, aiming to improve the CNN model accuracy in diagnosing brain tumors was introduced. |
| Liu *et al.* [45] | Introduced Skip-Attention GAN (SAGAN), an anomaly detection network that enhances the accuracy of latent image representations by incorporating attention modules to capture local information and employs depth-wise separable convolutions to reduce model parameters . |

**TABLE 7.** Application of GAN in medical image classification.

| Ref. | Description |
|---|---|
| Jeong *et al.* [47] | Presented the utilization of GANs in medical image analysis, highlighting their relevance and potential in addressing limited dataset sizes and class imbalances. |
| Tan *et al.* [48] | Demonstrated the effectiveness of incorporating GANs with an attentional mechanism in the augmentation of ultrasound data for the classification of idiopathic inflammatory myopathies (IIMs) . |
| Alrashedy *et al.* [49] | Introduced BrainGAN, a framework utilizing Generative Adversarial Networks to produce and classify brain MRI images. |
| Mao *et al.* [50] | Introduced an innovative semi-supervised classification algorithm for medical imaging, named Pseudo-Labeling Generative Adversarial Networks (PLGAN). Enhancing the training set. The algorithm incorporates MixMatch, contrastive learning, self-attention mechanisms, and cyclic consistency loss to improve classification performance . |
| Chen *et al.* [51] | Introduced TW-GAN, a novel approach for automatic artery/vein (A/V) retinal image classification. TW-GAN integrates vessel width and topology connectivity details into a deep-learning framework. |
| Teodoro *et al.* [52] | Presented EfficientAttentionNet, a CNN architecture designed for early identification of skin lesions, distinguishing between melanoma and non-melanoma. The model incorporates pre-processing, GAN-generated synthetic images, and a mask-based attention mechanism, demonstrating promising potential as a reference for future research. |

robustness and generalization of ML models. This helps train models to be more effective in disease detection, classification, and analysis of medical images, where diverse and comprehensive datasets are essential for accurate results and reliable model performance [14].The table 8 highlights some applications of GAN in image augmentation.

### F. REGISTRATION

Image registration in medical imaging involves aligning multiple images of the same subject or structure, whether from different time points or modalities, to ensure spatial correspondence. It facilitates multi-modal fusion in healthcare, combining data from MRI, CT, and other sources for comprehensive diagnoses. Temporal alignment aids in tracking disease progression and treatment efficacy. Image registration is paramount in image-guided and radiation therapy, enabling accurate targeting while minimizing collateral damage. GANs are applied in image registration by generating synthetic images to assist in the alignment process, improving registration accuracy in challenging cases [52]. The table 9 presents some applications of GAN in image registration.

### G. IMAGE SYNTHESIS

Medical image synthesis is a technique that utilizes Generative Adversarial Networks (GANs) to generate synthetic medical images with characteristics similar to real patient data. Generative Adversarial Networks (GANs) consist of a generator and a discriminator network that operate in an adversarial manner, generating realistic images, while the role of the discriminator is to differentiate between real and synthetic instances. In healthcare, medical image synthesis is invaluable for generating diverse and privacy-compliant datasets, enabling the training of robust machine learning models. It aids in addressing data scarcity issues as well as ensuring patient data confidentiality. Additionally, it can be used for data augmentation, anomaly detection, and simulating various medical conditions, enhancing diagnostic and research capabilities in medical imaging [58]. The table 10 shows some applications of GAN in image synthesis.

## IV. DATASETS, EVALUATION METRICS AND PREPROCESSING METHODS

This section provides insights into popular medical image datasets, prevalent preprocessing methods, and key evaluation metrics for assessing GAN performance. The discussion aims to understand better the dataset characteristics, preprocessing techniques, and evaluation metrics.

### A. DATASETS

The GAN framework thrives on diverse and robust datasets to generate realistic content. The MC dataset, CORN-2, and BraTs2019 exemplify the variety essential for training GANs effectively. With its rich multimedia content, the MC dataset offers a broad spectrum of visual data. BraTs2019 centred on brain tumour images, contributes to medical imaging advancements. These datasets empower GANs to understand and replicate complex patterns across domains, fostering innovation in fields ranging from multimedia generation to medical diagnostics. Their diversity ensures that the GAN model generalizes well and produces authentic outputs across various domains. Here are some common datasets used in GAN for medical imaging:

### B. EVALUATION METRICS

Selecting a suitable model is essential for optimal application performance, but choosing an appropriate evaluation metric is equally vital to ensuring the best result. GANs have diverse applications; each field has adopted its own evaluation criteria. However, the absence of standardized metrics has led to uncertainty among researchers in defining the assessment criteria for various tasks. In the table 12, we will delve into the evaluation metrics frequently used for assessment purposes.

### C. PRE-PROCESSING METHODS

Pre-processing is crucial in Generative Adversarial Networks to enhance training efficiency and generate image quality. Methods like image crop, random flipping, normalization, size adjustment, contrast enhancement, etc., are discussed in Table 13.

## V. EXPLORING POPULAR GAN VARIANTS

This comprehensive algorithm section discusses several subsections within the Generative Adversarial Networks in the medical field. It analyzes various applications, including image segmentation, registration, anomaly detection, regression with time-series data, data augmentation, image synthesis, enhancement, super-resolution, image-to-image translation, multimodal GANs, disease progression modeling, and additional information. These applications correlate with a specific GAN model, such as DCGAN, SRGAN, Pix2Pix, RGAN, SAGAN, and CGAN.

### A. DATA AUGMENTATION AND IMAGE SYNTHESIS

Generative Adversarial Networks are an effective framework for synthesizing images and augmenting data. GANs can produce realistic synthetic data or images by competing between two neural networks: discriminators and generators. GANs are helpful for augmenting data because they produce new, diverse samples that closely imitate the original data distribution, thus helping to expand the limits of data sets. GANs are excellent at producing high-quality, realistic images in image synthesis, allowing new possibilities for image editing and other creative applications.

#### 1) DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS (DCGAN)

The pioneering work introducing a convolutional decoder neural network for a generator was showcased by Radford et al. [8]. This particular approach entails specific architectural constraints designed to optimize its performance within the realm of unsupervised learning. DCGAN, or Deep Convolutional GAN, capitalizes on the spatial upsampling capacity inherent in the convolutional decoder operation of the generator (G). This distinctive capability facilitates the generation of images with higher resolutions, thereby leveraging the power of GANs. Refinements have been applied to the original DCGAN architecture, leading to both improved stability in training and the production of high-quality images. The main modifications are the use of strided convolutions in the discriminator (D) and fractional-strided convolutions in place of the pooling layers in the generator (G). Secondly, the incorporation of batch normalization in both the G (generator) and the D (discriminator) aids in aligning the distribution of generated and actual samples

**TABLE 8.** Application of GAN in Medical Image Augmentation.

| Ref. | Description |
|---|---|
| Tan *et al.* [48] | Introduced a method using generative adversarial networks for improving the classification of idiopathic inflammatory myopathies (IIMs) with notable improvement in classification accuracy data augmentation with GAN-based . |
| Zhang *et al.* [54] | Presented a novel approach called Reference-guided Fuzzy Integral GAN (RFI-GAN) to address the challenge of generating anatomically consistent ultrasound images. RFI-GAN uses reference sets to enforce logical constraints on generated images implicitly and employs Fuzzy Integral Modules for nonlinear fusion of texture and structure features . |
| Zhang *et al.* [55] | Addressed WGAN-GP algorithm to produce radionics data in the medical imaging field, addressing the challenge of obtaining annotated medical data samples for data augmentation in the medical domain . |
| Zhang *et al.* [56] | Proposed Minimal Generative Adversarial Network (MinimalGAN) for synthesizing medical images with less training data . |
| Al *et al.* [57] | Presented a model for segmentation of cardiac magnetic resonance (CMR) images. The proposed pipeline involves the detection of the heart region., image synthesis for augmentation, and segmentation of late-fusion. |
| Waheed *et al.* [58] | Introduces CovidGAN, an Auxiliary Classifier Generative Adversarial Network (ACGAN), to generate synthetic chest X-ray (CXR) images for COVID-19 detection combining these synthetic images with CNN offering a promising method to enhance the speed and robustness in the detection of COVID-19 in radiology systems. |

**TABLE 9.** Application of GAN in medical image registration.

| Ref. | Description |
|---|---|
| Liu *et al.* [60] | Presented an unsupervised image registration framework that employs adversarial techniques for multiple modalities., addressing challenges in multi-modal similarity metrics and contrasting differences between modalities. The suggested approach utilizes image-to-image translation to facilitate registration and incorporates a geometry-consistent training scheme and a multi-scale registration network with partial sharing. |
| Suwanraksa *et al.* [61] | Proposed a GAN with a registration network (RegNet) to improve the quality of synthetic CT (sCT) generated from Cone-Beam Computed Tomography (CBCT). The incorporation of RegNet led to reduced errors, improved image metrics, and sCT images maintaining anatomical accuracy compared to GANs without RegNet . |
| Zhou *et al.* [39] | Provided a thorough overview of Generative Adversarial Network models within the context of medical image fusion. |
| Mi *et al.* [62] | Introduced a medical image-fusion model called KDE-GAN, which utilizes a Generative Adversarial Network incorporating knowledge distillation and an explainable AI module, featuring dual discriminators.. |
| Fan *et al.* [63] | presented the development of the U-Patch GAN model for self-supervised fusion of multimodal brain images, focusing on enhancing fusion quality. |
| Huang *et al.* [64] | Introduced an innovative approach to medical image registration using MGMDcGAN, successfully integrating structural and functional information from images of different resolutions (e.g., MRI-PET, MRI-SPECT, CT-SPECT). Additionally, the proposed approach ensures the maximal retention of information from source images during the fusion process, |

**TABLE 10.** Application of GAN in medical image image synthesis.

| Ref. | Description |
|---|---|
| Jin *et al.* [66] | Highlighted the vulnerability of federated GANs (FedGANs) to backdoor attacks in the federated learning (FL) setting, highlighting the potential contamination of synthetic data and the need for defenses against such attacks. |
| Wang *et al.* [67] | Introduced FedMed-GAN serves as a benchmark for the translation of domains in federated settings on unsupervised brain image synthesis. FedMed-GAN effectively addresses mode collapse and performs well across unpaired and paired data ratios. |
| Zhang *et al.* [56] | Proposed Minimal Generative Adversarial Network (MinimalGAN) for synthesizing medical images with style transfer and target variation techniques. |
| Sun *et al.* [68] | Presented a new approach based on CycleGAN to address spatial inconsistencies in the synthesis of 3D medical images, particularly in converting to computed tomography (CT) images from 2D slices of magnetic resonance (MR) images. |
| Mendes *et al.* [69] | Explored using Generative Adversarial Networks (GAN) to synthesize artificial lung images from real computed tomography scans and semantic annotations. |

around a zero-centered point. Thirdly, DCGAN opts for a deeper architectural approach by eliminating connected hidden layers. Finally, ReLU activation is adopted in the generator for all layers excluding the output layer, where Tanh activation is employed. For the discriminator, LeakyReLU is utilized across all layers. Notably, DCGAN's quantitative generation of visually high-quality images surpasses many subsequently proposed GAN variants. However, it is worth mentioning that DCGAN is susceptible to mode collapse, which is a notable weakness [122]. DCGAN has been successfully trained on datasets such as LSUN [123] and ImageNet, and its applications include data augmentation [124] and classification tasks [125], [126]. The objective function V (D; G) of GAN is as

**TABLE 11. The table discusses the commonly used datasets in GANs.**

| Ref. | Dataset | Feature |
|---|---|---|
| [70] | Harvard Medical College public dataset | A collection of 275 pairs of brain slice images obtained from the Harvard Medical School database constitutes the training dataset, including six different modes (SPECT-TI, CT, SPECT-TC, MRT1, PET-FDG, PET, and MRT2). Each image within the dataset is standardized to a resolution of 256x256 pixels. Augmentation techniques were applied to expand the dataset, specifically involving rotation and random cropping operations, resulting in a total of 4,650 pairs. |
| [71] | CORN-2 | CORN-2 dataset includes 688 confocal microscopic images of the cornea measuring 384x384, of which 340 pairs were identified as high-quality images, while 288 pairs were categorized as low-quality images for training. Additionally, 60 low-quality images were set aside specifically for testing. Attributes like contrast, speckle noise, and diverse illumination distinguished the low-quality images. Two specialists performed the grading of these low-quality images. |
| [72] | The MC dataset | The MC dataset consisted of 138 chest radiographs (80 normal, 58 with TB) from the Montgomery County TB screening program in Maryland, USA, submitted as 12-bit grayscale PNG images at two resolutions (4020x4892 and 4892x4020 pixels) along with pixel-by-pixel segmentation of the lungs under the guidance of a radiologist. |
| [73] | CVC-ClinicDB | The CVC-ClinicDB dataset comprises 612 colorectal colonoscopy images in TIFF format, displaying abnormalities. These images are extracted from 29 videos, each sized at 384x288 pixels. Additionally, binary truth maps corresponding to the abnormalities are included. |
| [74] | IDRiD | The IDRiD dataset consists of fundus retinal images from diabetic patients in Nanded, India, taken with a Kowa VX-10$\alpha$ camera (50° FOV, resolution 4288x2848). It includes 516 jpg images with 5 DR classes and 3 DME classes. It provides annotations at the pixel level, classifications at the image level, and coordinates for the OD and fovea from medical experts., with 81 images having pixel-level OD annotations. |
| [75] | ADNI | Data from ADNI with 270 subjects (NC and EMCI) were used in this study, using 2.0 $mm^3$ DWI voxels and five-fold cross-validation for model estimation. |
| [76] | BraTs2019 | The BraTs2019 dataset comprises 335 cases, each containing four MRI sequences (T1WI, T2WI, FLAIR, and CE-T1WI), and includes manual annotations for tumor core, edema, enhanced tumor, and whole tumor masks. External testing involved ten clinical cases. All images were aligned based on the CE-T1WI modality and uniformly resampled to a size of 240x240x16. Radiologists provided the tumor mask labels. The study received approval from institutional review boards, and explicit informed consent was obtained from patients or their legal guardians. |
| [77] | ISIC 2018 | The training set includes 2594 pairs of images, and the validation set consists of 100 pairs. For MinimalGAN training, all images were resized to 240x320 and later padded to achieve a final size of 384x384. |
| [78] | COVIDx | The COVIDx dataset is composed of RGB images with pixel values ranging from 0 to 255 and varying dimensions. In the training of generative models, the images underwent a grayscale transformation and were resized to 128x128 pixels. Furthermore, pixel intensities were normalized to fall within the range of [-1, 1]. |
| [79] | DRIVE | The DRIVE dataset consists of 40 images obtained from a diabetic retinopathy screening initiative. Each image has a 565 $\times$ 584 pixel resolution, and 8 bits are allocated per color channel. Alongside the retinal images, the dataset provides segmented blood vessel images, serving as ground truth annotations. These annotations were crafted by human experts during the segmentation process. |
| [80] | STARE | The STARE database comprises 20 fundus images, each accompanied by corresponding segmented blood vessel images serving as ground truth. Each image in the dataset has a resolution of 605 $\times$ 700 pixels and is represented in standard RGB format with 24 bits per pixel. |
| [81] | HRF | The HRF dataset includes fundus images representing healthy retinas, glaucomatous retinas, and retinas affected by diabetic retinopathy (DR). This dataset is divided into three sets, each containing 15 colored retinal images, resulting in a total of 45 images. Additionally, the dataset includes corresponding manually segmented images and mask images. The images within the HRF dataset have dimensions of 3,504 $\times$ 2,336 pixels. |
| [82] | CHASE-DB1 | The CHASE-DB1 database consists of 28 retinal fundus images collected from a diverse group of school children. Each image in the dataset has a resolution of 960 $\times$ 999 pixels. |

**TABLE 12. Metric explanation: TP: true positive (where the model correctly predicts the positive class.), TN: true negative (where the model correctly predicts the negative class.), FP: false positive (where the model incorrectly predicts the positive class.), and FN: false negative (where the model incorrectly predicts the negative class).**

| Metric Name | Formula | Studies |
|---|---|---|
| Inception Score (IS) | $\exp\left(\mathbb{E}_x\left[\mathrm{KL}\left(p(y\|x) \parallel p(y)\right)\right]\right)$ | [83] [84] [85] [86] [87] |
| Accuracy (Acc) | $\frac{TP+TN}{TP+TN+FP+FN}$ | [88] [89] [90] [91] |
| Dice Similarity Coefficient/Dice Score(DSC) | $\frac{2TP}{2TP+FP+FN}$ | [92] [93] [94] [95] [96] |
| F-Score | $\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$ | [97] [98] [99] [100] |
| F1-Score | $\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}\times\text{Recall}}$ | [101] [102] [103] [104] [105] |
| Intersection Over Union (IOU)/Jaccard Index (JI) | $\frac{TP}{TP+FP+FN}$ | [106] [107] [108] [109] |
| Precision (Pre) | $\frac{TP}{TP+FP}$ | [110] [111] [112] [113] |
| Recall (Rec)/Sensitivity (Sen) | $\frac{TP}{TP+FN}$ | [110] [111] [112] [113] [114] [115] |
| Specificity (Spe) | $\frac{TN}{FP+TN}$ | [116] [117] [118] [119] [120] [121] |

follows:

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

### 2) VANILLA GENERATIVE ADVERSARIAL NETWORKS (VANILLAGAN)

The Generative Adversarial Network framework, first proposed by Goodfellow et al. in 2014 [7], comprises a generator

**TABLE 13. The table discusses the commonly used pre-processing methods in GANs.**

| Pre-Processing Method | Description | Studies |
|---|---|---|
| Image Crop | Image cropping is a pre-processing technique that involves removing a specific part of an image, typically concentrating on a region of interest or resizing it for further analysis or display. | [122] [123] [124] |
| Contrast Enhancement | Contrast enhancement is a fundamental image processing technique that enhances the visibility of details by adjusting brightness and color levels, making them more distinguishable [125]. | [122] [126] |
| Normalization | Normalization is a data pre-processing method that scales values in a dataset to a standard range, typically between 0 and 1, to facilitate accurate and consistent analysis, particularly in machine learning [127]. | [128] [54] [129] [57] |
| Co-registration | Co-registration is a method used in image processing and remote sensing to align or merge numerous images of the same area that are taken at various times or using distinct sensors to enable comparative analysis and data integration [130]. | [128] [131] |
| Size Adjustment | Size Adjustment involves changing the dimensions or resolution of an image to fit a specific requirement, such as resizing it for display or analysis. | [132] [129] |
| Random Flipping | Random flipping are highly used in machine learning and computer vision, which involve horizontally or vertically mirroring images at random to extend the variety of training data and enhance the robustness of the model. | [132] [124] |
| Texture transform | Texture transform refers to altering an image's perceived texture through various image processing techniques, often used for artistic or analytical purposes to achieve a specific visual or data analysis outcome. | [34] |
| Intensity inhomogeneity correction | Intensity inhomogeneity correction is a computational method used to remove or reduce variations in brightness across an image, typically caused by uneven lighting conditions or sensor artifacts, to enhance the accuracy of subsequent image analysis or computer vision tasks. | [133] |
| Non-rigid registration | Non-rigid registration is a technique that aligns and deforms images to match each other, accounting for complex, non-linear distortions or deformations, allowing precise comparisons or merging of images with varying structures or shapes [134]. | [133] |

and a discriminator network that engages in a competitive training process. The generator yields data from random noise, while the discriminator distinguishes between genuine data and generated samples. Iteratively, the discriminator seeks to correctly classify both data types, while the generator aims to generate samples that can challenge the discriminator's discrimination ability. This dynamic continues until Nash equilibrium is achieved, where the generator generates data resembling the training distribution, and the discriminator struggles to differentiate. Ultimately, GAN training can be seen as a two-player game aiming to minimize the value function V(D, G), reflecting the balance between generator and discriminator performance.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x\sim P_{\text{Data}}(x)}\left[\log(D(x))\right] + \mathbb{E}_{z\sim P_z(z)}\left[1 - \log(D(G(z)))\right] \quad (2)$$

The generator produces synthetic images using random noise, while the discriminator distinguishes between genuine and generated samples. Adversarial training drives both the discriminator's discriminative capacity and the generator's image-generation skills to their optimal levels. This iterative optimization leads to the following representative generative model:

$$G^* = \arg\min_G \max_D V(D, G) \quad (3)$$

GAN operates on the principles of unsupervised learning and finds application in both unsupervised and semi-supervised learning scenarios. It stands out as a generative model that relies solely on backpropagation, avoiding the complexities of Markov chains. The updates to the generator's parameters are facilitated through backpropagation from the discriminator. GAN surpasses other advanced models by generating more realistic and distinct samples. Nevertheless, GAN has shortcomings, such as training fluctuation leading to mode collapse and a limited variety of generated outputs. Numerous optimized approaches have been suggested to handle these issues based on the actual GAN framework. Furthermore, GAN still has limitations, rendering it unsuitable for processing discrete data forms like text.

### 3) AUXILIARY CLASSIFIER GENERATIVE ADVERSARIAL NETWORKS (ACGAN)

The auxiliary classifier Generative Adversarial Network, as introduced by Odena et al. [127], incorporates an auxiliary classifier into its structure, much like the design in InfoGAN [128] and cGAN [129]. In the AC-GAN framework, the added information is limited to the class label, which contrasts with the earlier structures of InfoGAN [128] and cGAN [129], where these two architectures could pertain to other data domains. AC-GAN discriminator encompasses a classifier responsible for categorizing samples into distinct classes. This inclusion of a classifier contributes significantly to enhancing training stability. Despite AC-GAN's demonstrated ability to generate high-quality images, scholarly works [130] have pointed out that the model generates often-used images for most classes as the quantity of labels increases. Miyato et al. [130] suggested that this phenomenon could be attributed to the presence of the auxiliary classifier. The model's training involved datasets such as ImageNet and CIFAR-10. AC-GAN was initially suggested for purposes like data augmentation [51] and image synthesis [127].

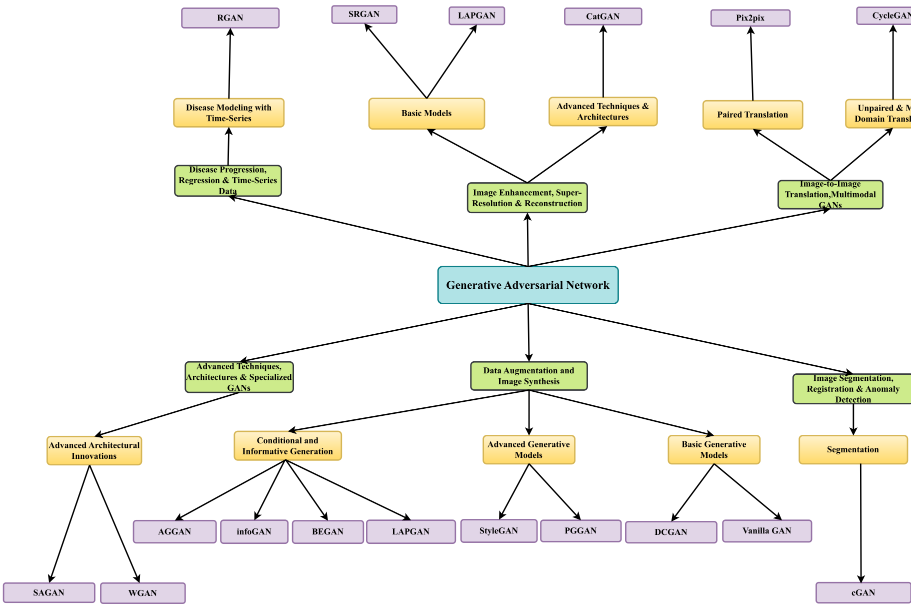


**FIGURE 5.** Taxonomy of generative adversarial networks.

Within the context of ACGAN, the generator G employs inputs of noise Z along with the associated categorical label C sampled from the distribution PC, resulting in the generation of a sample denoted as Xfake = G(c, z). On the other hand, discriminator D provides the possibility of discrimination between actual and fabricated images, considering both the authenticity and the category labels. When it comes to training the discriminator D, the primary aim is to maximize the combined loss $Ls + Lc$. Conversely, during the training of the generator G, the objective is to maximize the difference between losses, denoted as $Ls - Lc$. Notably, in the framework of ACGAN, the representation of the noise Z is acquired autonomously and remains uninfluenced by the associated class label.

### 4) INFORMATION MAXIMIZING GENERATIVE ADVERSARIAL NETWORK (INFOGAN)

Although trained GANs can create new images, Create photos with specific attributes. Chen et al. [3] proposed InfoGAN. A GAN that can learn disentangled representations without supervision. InfoGAN, which is based on information theory, can overcome the explanation difficulties of the hidden variables of GAN. The "info" variable describes the common information between the generated distribution G(z, c) and the implicit encoding c. The information that is shared between the underlying encoding c and the generated distribution G(z, c) is denoted by the variable "info". I(c; G(z, c) should be maximized to close the correlation between x and c, regular-term constraints complement the essential GAN objective function. At the same time, an auxiliary distribution Q(c|x) was defined to approximate the P(c|x) approximation so that the lower bound on the P(c|x) variation could be determined.

$$\min_{G,Q} \max_{D} L_1(\mathrm{D,G,Q}) = V(D, G) - \lambda L_I(G, Q) \tag{4}$$

### 5) PROGRESSIVE GROWING GENERATIVE ADVERSARIAL NETWORK (PGGAN)

The Progressive Growing of GANs (PGGAN) is a prominent GAN variant comprising generator and discriminator networks. PGGAN is specified to produce high-resolution CXR because of its exceptional capacity to capture global structures. This progressive approach enables the simultaneous learning of general image characteristics and fine details as layers grow. Notably, lower-resolution layer outputs significantly influence the high-resolution output through a fade-in mechanism. PGGAN's unique capability lies in its gradual transformation of low-resolution photos into high-resolution ones, factoring in the knowledge acquired in earlier layers.

Implementation of PGGAN is accessible through an official website built with TensorFlow in Python, specifically using TensorFlow-gpu 1.6.0 and Python 3.4.0. PGGAN's training process is rooted in game theory, where two players, the generator (G) and discriminator or classifier (C), compete. The discriminator network is specifically trained to distinguish between genuine and artificially created samples, while the generator network learns how to transform random noise into the input space. Theoretically, the minimax objective establishes the loss function.

$$\min_{G} \max_{C} \mathop{\mathbb{E}}_{x \sim \mathbb{P}_r} [\log C(x)] + \mathop{\mathbb{E}}_{\widetilde{x} \sim \mathbb{P}_g} [\log(1 - C(\widetilde{x}))] \tag{5}$$

In this context, authors deal with three significant distributions: $P_r$, which characterizes the data allocation of real images; $P_g$, representing the model allocation implicitly shaped by $x = G(z)$, where $z \sim p(z)$ with $P(z)$ as a Gaussian distribution; and $P_{\hat{x}}$, defined through uniform sampling along linear paths connecting point pairs sampled from $P_r$ and $P_g$. The Progressive Growing of GANs (PGGAN) method adopts the enhanced Wasserstein GAN loss, commonly denoted as the WGAN-GP loss, which surpasses the original Wasserstein GAN (WGAN) due to incorporating a gradient penalty. In this study, PGGAN is selected in conjunction with the enhanced Wasserstein GAN (WGAN-GP) loss for its remarkable training stability, a crucial factor enabling the synthesis of high-resolution images.

$$L = \underset{\widetilde{x}\sim\mathbb{P}_g}{\mathbb{E}}[C(\widetilde{x})] - \underset{x\sim\mathbb{P}_r}{\mathbb{E}}[C(x)] + \lambda \underset{\widehat{x}\sim\mathbb{P}_{\widehat{x}}}{\mathbb{E}}\left[(\|\nabla_{\widehat{x}}C(\widehat{x})\|_2 - 1)^2\right] \tag{6}$$

#### 6) STYLE GENERATIVE ADVERSARIAL NETWORK (STYLEGAN)

StyleGAN builds upon the progressive GAN architecture, gradually expanding image size from low to high resolution. Unlike earlier versions, it employs bi-linear sampling rather than the closest neighbor up/down selection for both the generator and discriminator, integrating a 2nd-order binomial filter to enhance quality. It creates an intermediate latent vector by incorporating a mapping network into the original, influencing diverse visual attributes. This 8-layer MLP-based mapping function converts a 512-dimensional latent vector (z) into another 512-dimensional vector (w), further transformed through a learned affine operation (A). This modified w is then fed into the synthesis network alongside the AdaIN module for adaptive instance normalization. The discriminator's architecture aligns with the baseline while the generator incorporates these changes, enabling StyleGAN to synthesize images with more nuanced control over features, marking a significant advancement in GAN technology [131], [132], [133], [134]. The AdaIN operation takes an input of $y = (y_s, y_b)$, which is yielded by applying $A$ to $w$. Here is how its equation is defined:

$$\text{AdaIN}(x_i, y) = y_{s,i}\frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,i} \tag{7}$$

A plausible explanation for the perceptual path length involves the consideration of an infinitely refined sum; however, in practical implementation, this is approximated by utilizing a small subdivision parameter denoted as epsilon ($\epsilon$). Consequently, the mean perceptual path length within the latent space Z, encompassing all conceivable endpoints, is computed as an average over this approximation:

$$lZ = \mathbb{E}\left[\frac{1}{\epsilon^2}d(G(\text{slerp}(z_1, z_2; t)), G(\text{slerp}(z_1, z_2; t+\epsilon)))\right] \tag{8}$$

#### 7) BOUNDARY EQUILIBRIUM GENERATIVE ADVERSARIAL NETWORK (BEGAN)

BEGAN incorporates the discriminator auto-encoder architecture, a concept first introduced in energy-based GAN by Zhao et al. [135]. Berthelot et al. [136] introduced the BEGAN framework; improving the reconstruction loss for genuine and generated images is the goal of the autoencoder's training. Minimizing the Wasserstein gap between distributions is in accordance with this objective. To adjust the tradeoff between the generator and discriminator losses, they introduce a hyperparameter $\gamma = \frac{\mathbb{E}[L(G(z))]}{\mathbb{E}[L(x)]}, \gamma \in [0, 1]$. This parameter balances the effort distribution between G (generator) and D (discriminator), effectively controlling the diversity of the generated face images. Considering its remarkable performance and efficiency, BEGAN's fundamental generator framework, training instability, and discriminator's lack of reconstruction capabilities limit the visual quality and difference of the framework. To improve the stability of training, BEGAN was introduced with modifications.

### B. IMAGE ENHANCEMENT, SUPER-RESOLUTION & RECONSTRUCTION

GANs are excellent at improving, super-resolving, and reconstructing images in medical imaging. Using a dual-network architecture, in which one generator is used to improve image quality and another to assess authenticity, can provide more accurate and detailed medical image reconstruction, better resolution, and improved diagnostic accuracy.

#### 1) SUPER-RESOLUTION GENERATIVE ADVERSARIAL NETWORK (SRGAN)

The Super-Resolution GAN (SRGAN) improves the realism of generated images by introducing perceptual loss in addition to adversarial loss [137]. While employing mean squared error (MSE) as the loss function can yield a high peak signal-to-noise ratio (PSNR) during GAN training, it often results in a loss of fine, high-frequency details in the recovered images. The perceptual loss [138] is computed by contrasting feature differences obtained from a convolutional neural network (CNN) applied to both generated and genuine images. Numerous investigations have utilized SRGAN to restore image details based on low-quality inputs and rebuild high-quality images. For instance, Gu et al. employed a GAN-based super-resolution framework for synthesizing high-quality images from computed tomography (CT) and magnetic resonance (MR) low-quality images. Their approach preserved texture details and realistic patterns in the generated images [139]. The equation for perceptual loss:

$$\mathcal{L}_{\text{perceptual}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\phi(I_{\text{real}}^{(i)}) - \phi(I_{\text{gen}}^{(i)})\right\|_2^2 \tag{9}$$

where: - $\mathcal{L}_{\text{perceptual}}$ is the perceptual loss. - $\lambda_{\text{perceptual}}$ is a weight parameter. - $WHC$ represents the dimensions of the images. - $\phi(\hat{y})$ and $\phi(y)$ are feature maps removed from the generated and target images $\hat{y}$ and $y$, respectively.

#### 2) LEAST SQUARES GENERATIVE ADVERSARIAL NETWORK (LSGAN)

The fundamental concept behind LSGAN involves replacing the use of the cross-entropy loss from the discriminator with the least-squares loss function introduced by Mao et al. [140]. Traditional GANs consider the discriminator as a classifier using the sigmoid loss from cross-entropy.

This approach leads to problems with gradient vanishing. If a false sample is classified as a true image, it does not carry an error since it is on the right side of the decision boundary, although it is clearly different from the true samples. In contrast, the least-squares loss function penalizes data points located far from the right of the decision boundary. This penalty encourages the generation of samples that more closely resemble the actual data. LSGAN is distinguished by generating images of higher quality compared to standard GANs, and it features increased robustness in training. However, it encounters mode collapse, which negatively influences the variety and effectiveness of the produced images [122]. To evaluate the performance of LSGAN, it was estimated on datasets such as HWDB1.0, LSUN [123], and CIFAR-10. In [141] and [142], LSGAN has been used effectively to achieve reliable learning. The objective functions of the generator (G) and discriminator (D) are represented in equations 10 and 11, respectively.

$$\min_G L_G = \frac{1}{2}\mathbb{E}_{z\sim P_z}\left[(D(G(z)) - c)^2\right] \tag{10}$$

$$\min_D L_D = \frac{1}{2}\mathbb{E}_{x\sim P_r}\left[(D(x)-b)^2\right]+\frac{1}{2}\mathbb{E}_{z\sim P_z}\left[(D(G(z))-a)^2\right] \tag{11}$$

In this context, "a" represents the label assigned to the generated data, "b" corresponds to the label given to the actual data, and "c" denotes the hyperparameter. G, the generator, seeks D, the discriminator, to misclassify the generated data as accurate.

#### 3) CATEGORY-BASED GENERATIVE ADVERSARIAL NETWORK (CATGAN)

CatGAN, introduced by Springenberg [143], uses a different approach compared to traditional GAN. Instead of binary classification, CatGAN uses multiclass classification in an unsupervised method. In this scheme, the discriminator (D) is trained to distinguish between true and false data and categorize all the sample data into pre-selected categories. The process involves D categorizing the data into classes even when it is unsure of the distribution of categories for the samples generated by the generator (G). Notably, CatGAN requires G to create samples associated with specific classes rather than simply reproducing samples from a dataset. Despite its ability to learn from unlabeled data, CatGAN faces challenges in effectively clustering input features. The CatGAN objectives encompass maximizing H[p(c|x,D)] and H[p(c|D)] while minimizing H[p(c|G(z),D)]. Remarkably, CatGAN finds application in image classification, as can be seen from its application in Reference [144]. For the generator and discriminator, the respective objective functions are expressed as follows in Equations 12 and 13.

$$\begin{aligned} LD = \max_D H_X[p(c|D)] - \mathbb{E}_{x\sim X}[H[p(c|x, D)]] \\ + \mathbb{E}_{z\sim P(z)}[H[p(c|G(z), D)]] \end{aligned} \tag{12}$$

$$LG = \min_G - H_G[P(c|D)] + \mathbb{E}_{z\sim P(z)}[H[P(c|G(z), D)]] \tag{13}$$

In these equations, X represents the distribution of the data set, and $H[\cdot]$ denotes empirical entropy.

### C. IMAGE-TO-IMAGE TRANSLATION,MULTIMODAL GANS

To convert images from one domain to another, such as a CT scan to an MRI, Generative Adversarial Networks in Image-to-Image Translation use two neural networks, the discriminator and generator. For example, multimodal GANs in medical imaging focus on producing a variety of outputs, such as multiple views or modalities, from a single input. This helps with diagnosis by offering a variety of perspectives or features that may be lacking for more in-depth medical analysis and explanation.

#### 1) PIXEL-TO-PIXEL GENERATIVE ADVERSARIAL NETWORK (PIX2PIX GAN)

Pix2pix falls under the conditional Generative Adversarial Networks category, where the generation of a result image is contingent upon an input (source) image. The network comprises two primary components: the generator and the discriminator. The role of the generator is to convert the input image into the desired output image. Simultaneously, the discriminator assesses the resemblance of the input image to an unspecified image, which could either be a target image from the dataset or an output image produced by the generator. Its objective is to discern whether the generator generated the image. During training, the generator is continually revised to minimize the loss that the discriminator predicts for the generated images [145]. To moderate the risk of overfitting, the training process involves a strategy where the generator is not explicitly exposed to the training dataset. Instead, dropout layers are employed during both dataset training and prediction phases to introduce randomness; the generator is then directed by the loss functions throughout the training progression [135]. The generator in Pix2Pix learns a mapping that takes the source image x and a random noise image z as inputs and produces the corresponding target image y, which can be expressed as x, z → y. Meanwhile, the discriminator's role is to differentiate between genuine target images y given the source image x and fake ones developed by the generator.

#### 2) CYCLE GENERATIVE ADVERSARIAL NETWORK (CYCLEGAN)

CycleGAN, proposed by Zhu et al. [146], is intended to handle the problem of changing pictures between two independent domains. This method uses a pair of complementary Generative Adversarial Networks (GANs) that form a cyclic network topology. CycleGAN's primary goal is to create a mapping between domain X and domain Y pictures. There are two generators, G and F, and two discriminators, Dy and Dx, in this structure. The generator G takes an image from domain X and attempts to transform it into one from domain Y. At the same time, the discriminator Dy determines if G generates a given image or is an authentic image from the domain. Similarly, the generator F is tasked with converting an image from domain Y to an image from domain X, while the discriminator Dx determines if a picture is generated by generator F or belongs to domain X. CycleGAN utilizes cyclic consistency loss to enhance overall performance further, thus enhancing the model's robustness. Figure 6 shows the schematic diagram of CycleGAN.

### D. DISEASE PROGRESSION, REGRESSION & TIME-SERIES DATA

In disease progression and time-series medical imaging, Generative Adversarial Networks provide realistic images demonstrating how diseases progress or regress over time. GANs help simulate disease histories by utilizing deep learning, which enables the analysis and forecast of the advancement or regression of diseases, improving diagnostic and predictive capabilities in healthcare.

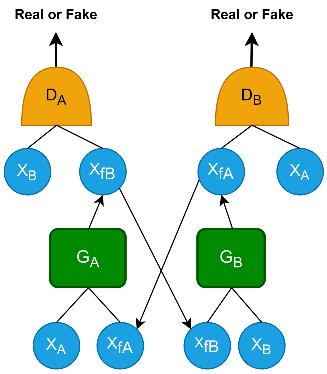


**FIGURE 6.** Schematic diagram of CycleGAN.

#### 1) RELATIVISTIC GENERATIVE ADVERSARIAL NETWORK (RGAN)

Relativistic GAN (RGAN) introduces a fresh training approach to enhance GAN performance. To address the conventional GAN's non-convergence issue, RGAN's discriminator is devised to distinguish that the likelihood of actual data is more genuine compared to randomly generated artificial data [147]. After RGAN, various other training techniques have emerged. To mitigate method failure within GAN, Song et al. suggested incorporating some inaccurate samples as genuine samples during training, thereby curbing the generator's gradient value to prevent gradient explosion [148]. In tackling the training instability inherent in GAN, a point cloud upsampling network (PU-GAN) was introduced, with the discriminator structured as a classifier for positive-unlabeled classification as opposed to the traditional positive and negative classification [149]. In accordance with the work by Goodfellow and colleagues [7], they proceed to establish a discriminator network, denoted as $D\theta D$, which is systematically refined in an alternating fashion alongside the generator network $G\theta G$. This joint refinement process is undertaken to address the adversarial minimax dilemma:

$$\min_{\theta_G}\max_{\theta_D} \mathbb{E}_{I^{HR}\sim p_{\text{train}}(I^{HR})}\left[\log D_{\theta_D}(I^{HR})\right] + \mathbb{E}_{I^{LR}\sim p_G(I^{LR})}\left[\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR})))\right] \quad (14)$$

The fundamental concept underpinning this formulation is to enable the training of a generative model, denoted as G, with the specific objective of deceiving a differentiable discriminator, labeled as D. This discriminator is trained to determine between super-resolved images produced by the generator and authentic images. By adopting this strategy, this generator can acquire the ability to generate solutions that closely resemble genuine images, thereby posing a challenge for discriminator D to classify them accurately.

### E. ADVANCED TECHNIQUES, ARCHITECTURES & SPECIALIZED GANS

Advanced approaches for Generative Adversarial Networks use cutting-edge architectures and versions developed especially for medical imaging. These GANs utilize complicated designs and optimization techniques to produce images with great accuracy and clinical significance, which aids medical research, diagnosis, and treatment planning.

#### 1) SELF-ATTENTION GENERATIVE ADVERSARIAL NETWORK (SAGAN)

Conventional convolutional GAN faces challenges in learning certain classes of images compared to others and captures mostly spatial details. Although the latter GAN model demonstrates superiority in generating images with specific constraints, such as those involving ocean or sky scenes, it is unable to reproduce images with complex geometric patterns. SAGAN, introduced by Zhang et al. [150], introduces a self-monitoring mechanism into a convolutional GAN. This embedding provides an extensive receptive field without compromising computational efficiency by allowing the model to account for global long-range dependencies for image fusion. The self-suggestion mechanism calculates the response within a region by weighting the features from all other regions. Attention weights are computed with minimal computational cost. Spectral normalization of both generator (G) and discriminator (D) weights is applied to improve the robustness of training. The generator produces images with complex details at different locations, while the discriminator imposes complicated geometric conditions on the image structure. Self-Attention GAN outperforms alternative image generation methods by demonstrating remarkable improvements in tasks such as improving human pose representation [151] and achieving superresolution [152]. Figure 7 illustrates the architecture of the self-monitoring mechanism. The figure involves the letters g, f, and h being associated with the key, query, and value sequentially. The symbol ⊗ denotes matrix multiplication. The attention map illustrates spatial dependencies that extend over long distances.

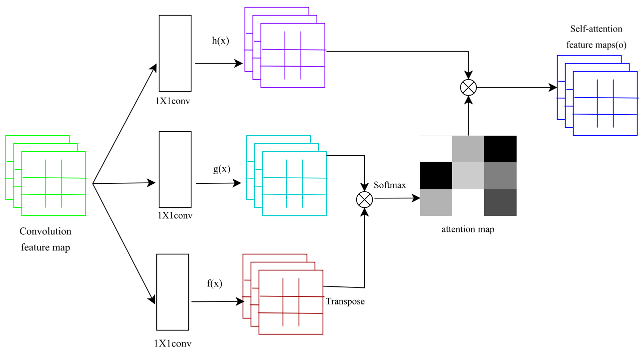


**FIGURE 7.** SAGAN architecture.

The model uses an articulated version of the adversarial optimization loss.

### 2) WASSERSTEIN GENERATIVE ADVERSARIAL NETWORK (WGAN)

The problem of unstable GAN [8] the WGAN effectively solves learning [153] as described in the above text [154]. WGAN solves the problem of training instability faced by GANs [154]. The problems observed in the original GAN can be attributed to two main aspects, as highlighted in the source: (i) the use of impractical difference measurements, such as the Kullback Leibler (KL) and Jensen Shannon (JS) differences, after equivalent optimization, and (ii) the challenge of achieving overlap between the original randomly initialized generator distribution and the true distribution. WGAN addresses these issues by introducing the Wasserstein distance. Unlike the KL and JS distributions, the Wasserstein distance retains its ability to measure differences even when the two distributions do not have overlapping support. Furthermore, the Wasserstein distance provides better smoothing performance than the KL and JS differences. This favorable property solves the gradient vanishing problem by allowing a better balance between the generator and the discriminator in the learning process. The mathematical formulation of the Wasserstein distance allows it to be optimized using a neural network denoted as D. Subsequently, G is fine-tuned with respect to this approximately optimal D to minimize the Wasserstein distance, successfully reducing the discrepancy between the generated and actual distributions. The loss functions for both the generator (G) and the discriminator (D) in the Wasserstein GAN (WGAN) are defined separately as follows:

$$L_G = -\mathbb{E}_{x \in P_g}[D_w(x)] \tag{15}$$

$$L_D = \mathbb{E}_{x \in P_g}[D_w(x)] - \mathbb{E}_{x \in P_r}[D_w(x)] \tag{16}$$

where w denotes parameters that have a finite value.

## F. IMAGE SEGMENTATION, REGISTRATION & ANOMALY DETECTION

Generative Adversarial Networks are excellent at segmenting, registering, and identifying anomalies in medical imaging. GANs utilize their dual-network structure to precisely identify structures, align images for specific comparison, and highlight variations for improved anomaly detection in medical images to enable more accurate and reliable diagnoses.

### 1) CONDITIONAL GENERATIVE ADVERSARIAL NETWORK (CGAN)

The cGAN constitutes a conditional model designed to introduce supplementary information into the generative model, thereby enabling the controlled synthesis of generated data. This supplemental data could encompass class designations or data originating from different modalities, as Mirza et al. [129] suggested. An extension of the GAN framework, the cGAN mandates both the generator and the discriminator to receive an additional information vector (c) as part of their input. This innovation has demonstrated enhanced capabilities in generating multi-modal data representations. The supplementary information is typically encoded as a one-hot vector within both the discriminator and the generator. Subsequently, it is combined with an encoded noise vector (z) within the generator and with the actual data (x) within the discriminator. However, it's noteworthy that the discriminator itself cannot produce the class label for the input data. It is common for cGANs to necessitate pairs of input and output images during training, a requirement that might not always be met in domain adaptation scenarios. The conditioning aspect of cGAN encompasses a variety of factors, including attributes [155], [156], images [4], [157], texts [158], and class labels [4]. With a versatility that spans both unimodal datasets like MNIST [159] and multi-modal datasets like the Flickr dataset, cGANs exhibit adaptability across various data types. The loss function associated with the cGAN framework can be observed in Equation.

$$\min_G \max_D \mathbb{E}_{x \sim P_x}\left[\log D(x|c) + \mathbb{E}_{z \sim p_z} \log(1 - D(G(z|c)))\right] \tag{17}$$

Several studies have shown that the cGAN model is helpful in image-to-image transformation issues [4], [160], [161], [162]. Its capabilities extend to generating images with a photo-realistic quality [163], facilitating data augmentation [164], [165], and supporting segmentation tasks [166]. Figure 8 demonstrates the architecture of CGAN.

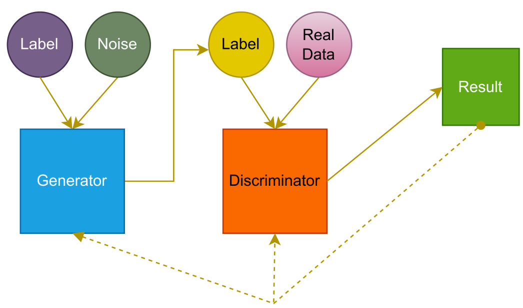


**FIGURE 8.** Schematic diagram of CGAN.

## VI. RESULT ANALYSIS

Result analysis in GANs for medical imaging involves evaluating generated images using metrics like SSIM, PSNR, accuracy, AUC, Dice coefficient, MAE, ROC curve, IoU, entropy, and normalization. These metrics assess image quality, classification accuracy, segmentation precision, and overall model performance. Researchers select metrics based on specific tasks, ensuring a comprehensive evaluation of GAN effectiveness in enhancing medical image synthesis, registration, classification, and segmentation. The table 14 analyses some results from recent papers on GAN in medical imaging:

In the context of our analysis of the discussed table of GAN-based applications in medical image processing, it becomes evident that specific algorithms excel in distinct medical image processing tasks. MMFGAN has shown prowess in image registration, boasting a high Structural Similarity Index (SSIM). Acc is the short form of Accuracy and Amplitude Gain (AG). Transitioning to brain image classification, ED-GAN stands out with an impressive accuracy rate of 96.25%, making it a valuable tool for categorizing diverse brain conditions. However, MedGAN, which specializes in image reconstruction, reports a relatively lower accuracy of 0.62. In the realm of chest CT image classification, MI-GAN emerges as a robust performer, showcasing high accuracy, sensitivity, specificity, and Area Under the Curve (AUC). Shifting focus to abdominal MRI image reconstruction, T-GAN exhibits commendable results, boasting elevated Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Throughout these investigations, datasets like OASIS, SRPBS, and ABIDE are prominently utilized, emphasizing their significance in shaping the landscape of medical image processing. The widespread adoption of these datasets underlines their role as valuable resources, enabling researchers to explore algorithmic capabilities across diverse medical imaging studies. In conclusion, the multifaceted nature of medical imaging demands a nuanced approach to algorithm selection, considering the specific intricacies of each task. While certain algorithms excel in particular applications, the variability in performance underscores the importance of tailored choices in algorithmic deployment, contributing significantly to the evolving landscape of medical image processing.

## VII. CHALLENGES AND FUTURE RESEARCH DIRECTION

Among machine learning models, GANs have emerged as one of the most powerful classes. Their application in medical imaging holds great scope for tasks like image synthesis, segmentation, augmentation, classification, registration, reconstruction, denoising, and detection. However, applying GANs to medical imaging poses unique challenges, necessitating future research to enhance productivity. This section will address key challenges and explore potential research areas for advancing GANs in medical imaging.

### A. OPTIMIZATION PROCESS

The optimization process of GANs is complex. The optimal point in the optimization scenario is the Nash equilibrium, and reaching the Nash equilibrium is challenging due to the nature of the process, which is considered to be a mini-max game nature. There are several conceptual and numerical reasons why standard GAN training is unstable. Mode collapse or mode-hopping are two convergence problems that might arise from this. Investigating more robust optimization techniques and strategies to stabilize the training process, including addressing convergence issues such as mode collapse, Can improve the overall efficiency of GANs. Unrolled GANs aim to mitigate the risk of the generator being excessively optimized for a specific discriminator, which minimizes mode collapse and enhances overall stability in the model. Adding noise to discriminator inputs and penalizing discriminator weights are potential strategies to enhance GAN convergence. Exploring advanced algorithms, regularization methods, and novel loss functions could accelerate convergence and enhance training. This, in turn, could lead to the development of more robust models capable of achieving the Nash equilibrium.

### B. PRIVACY CONCERNS AND TRUST-ABILITY OF GENERATED DATA

Privacy concerns, particularly related to the sensitive nature of patient information, arise in medical imaging collection and application. In the domain of healthcare data privacy, the framework known as Asynchronized Discriminator GAN (AsynDGAN) addresses the critical need to protect patients' sensitive information while enabling effective communication and distribution of GAN-generated medical data. AsynDGAN achieves this by learning generative distributions from real medical datasets across various health entities without the necessity of sharing or directly accessing patients' confidential information. Establishing trust in generated medical images is a significant challenge. Incorporating physics-based simulations and rigorous experimentation to understand GAN convergence in the medical imaging context can enhance the trustworthiness of generated data. Here, Figure 9 represents the architecture of the optimization process of AsynDGAN.

The source mask is denoted by the green ellipse, and the actual image of the target is represented by the brown ellipse. In our iterative update process, the forward pass is illustrated by solid arrows, while the gradient direction during the backward pass is depicted by dotted arrows. Solid blocks signify active updates, whereas dotted blocks imply they remain unchanged during that particular update.

### C. LIMITED INTEGRATION WITH MODELS FROM OTHER DOMAINS AND MULTIMODAL MEDICAL IMAGE AUGMENTATION

Although GANs have been integrated with models from diverse domains, including optimization algorithms, signal graphs, and natural language, but still their integration with

**TABLE 14.** he table analyzes the experimental results of some state-of-the-art research articles published on GANs in Medical Imaging.

| Reference | Doamin | Dataset | Pre-processing Methods | Model | Results |
|---|---|---|---|---|---|
| Guo *et al.* [122] | Registration | Randomly selected a pair of registered fused images and fused them with nine algorithms | Image crop, Contrast enhancement | MMFGAN | SSIM: 0.9594, VIF: 0.7321, FMI: 0.7902, AG: 8.3516, CE: 0.3601 |
| Ahmad *et al.* [123] | Classification | Images of Glioma, Meningioma, Pituitary | Image Crop, Image Enhancement | ED-GAN | Accuracy: 96.25%. |
| Guo *et al.* [190] | Reconstruction | Blister, Demodicosis, Parakeratosis, Molluscum | N/A | MedGAN | MAP: 0.96, Accuracy: 62% |
| Cackowski *et al.* [128] | Registration | OASIS, SRPBS, ABIDE | Normalization, Co-registration | VAE-GAN | SSIM: 0.951±0.013 |
| Zhang *et al.* [54] | Augmentation | Thyroid ultrasound image | Normalization | RFI-GAN | PSNR: 32.78, SSIM: 0.9598 |
| Yu *et al.* [132] | Synthesis | Corneal Confocal microscope (CCM) ( CORN-2) | Size Adjustment, Random Flipping | FS-GAN | Entropy: 6.785 ± 0.138, AvG: 7.332 ± 0.024, Brisque: 0.484 ± 0.005, NIQE: 28.107 ± 6.073, PIQE: 1.774 ± 0.235 |
| Chen *et al.* [129] | Classification | Chest CT image | Size Adjustment, Normalization | MI-GAN | Accuracy: 93.85 %, Sensitivity: 96.69%, Specificity: 89.70%, AUC: 96.17% |
| Liu *et al.* [191] | Registration | BraTs2019, Clinical | N/A | BTMF-GAN | **BraTs2019:** AG: 10.317, SF: 27.655, SSIM: 0.543 PSNR: 17.885 **Clinical:** AG: 8.199, SF: 22.023, SSIM: 0.762, PSNR: 20.985 |
| Du *et al.* [34] | Reconstruction | Abdominal MRI image | Texture transform, Normalization | T-GAN | PSNR: 34.69, SSIM: 0.9353 |
| Zhao *et al.* [133] | Synthesis | Multi-modal pelvic MR-CT | Intensity inhomogeneity correction, Non-rigid registration | C-GAN | MAE: 48.28, SSIM: 0.8882, PSNR: 27.37 |
| Zhang *et al.* [55] | Augmentation | Heart disease Cleveland | Normalization | WGAN-GP | AUC: 0.902, SPE: 0.82 |
| Al *et al.* [57] | Segmentation | ES images | Normalization | C-GAN | Dice: 0.959, HD: 9.37 |
| Zuo *et al.* [192] | Synthesis | ADNI dataset | Image resizing | VAE-GAN | ACC: 85.18, SEN: 84.44, SPE: 85.92, AUC: 86.16 |
| Alauthman *et al.* [124] | Classification | Medical dataset | Image crop, Random Flipping | GAN | Accuracy: 91.13% |
| Kadri *et al.* [193] | Detection | Multi-channel EEG epilepsy | Normalization | GAN | MAE: 61.722, MDAE: 37.322, EV: 87%, GAN-P: 0,272 |
| Li *et al.* [194] | Segmentation | I3A dataset | Contrast Enhancement, Normalization | cC-GAN | SEG: 87.29% |
| Park *et al.* [195] | Segmentation | DRIVE, STARE, HRF, CHASE-DB1 | Automatic color equalization | M-GAN | **DRIVE:** Acc: 97.06%, **STARE:** Acc: 98.7%, **HRF:** Acc: 97.36%, **CHASE-DB1:** Acc: 97.36% |

superior models in the medical image augmentation field remains limited. Absorbing and combining models from other domains, such as attention mechanisms, Transformers, and graph convolution, could offer promising avenues for

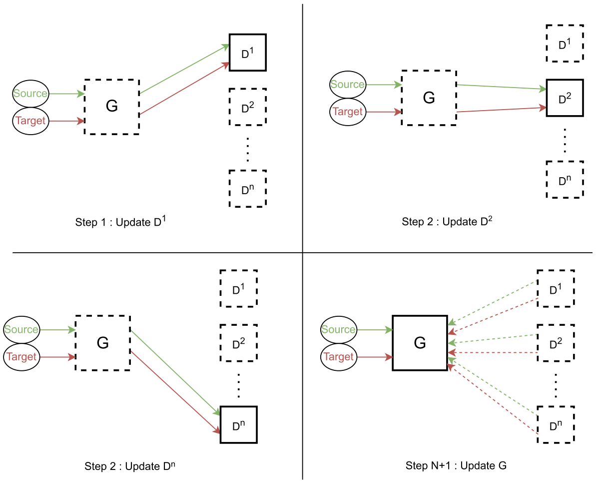


**FIGURE 9.** Optimization process of AsynDGAN.

enhancing GANs in medical image augmentation. Transformers have shown promising results in the field of medical image analysis, specifically in tasks like semantic segmentation for the detection and classification of skin lesions. These models bring distinctive features and capabilities that have the potential to enhance the overall performance of GANs in this particular context. Cross-disciplinary collaboration is expected to facilitate the exchange of ideas and advancements. The growing availability of multimodal images from advanced medical imaging equipment poses challenges to developing effective multimodal medical image augmentation methods. These challenges may include issues related to data fusion, model complexity, and ensuring the preservation of diagnostic information across different imaging modalities. Future research should focus on developing GAN-based techniques for handling multimodal medical images. This includes addressing the complexities of information fusion across different imaging modalities to enhance diagnostic capabilities.

### D. TRADITIONAL METRICS FOR QUANTITATIVE EVALUATION

Traditional pixel-wise metrics like MAE, MSE, PSNR, and SSIM may not effectively capture the accurate visual quality of medical images. Advanced evaluation metrics, such as learned perceptual image path similarity (LPIPS), have been created and embraced to provide more meaningful and human-aligned assessments of the quality of generated medical images. However, it is not always apparent to prioritize the entire medical image during the diagnostic process. Medical diagnosis often requires attention to specific, small regions of interest. Thus, creating evaluation metrics that align with both subjective and objective assessments, while considering the semantic context, introduces a new challenge for the field.

### E. GEOMETRIC CORRELATIONS AND MICROSCOPIC IMAGING

Existing data synthesis techniques often lack consideration for strong geometric correlations, limiting their applicability. Future advancements will likely involve refining architectures to handle specific geometric characteristics unique to different medical imaging tasks and modalities to overcome these limitations. This effort addresses the challenges of validation and real-time application requirements. The Bidirectional Generative Adversarial Network (Bi-GAN) architecture presents a promising solution for synthesizing diverse microscopic images across multiple domains. This approach takes into consideration various geometric features. This strategy might considerably increase the scope of GAN applications in the medical imaging domain by effectively addressing specific geometric complexities implicit in diverse medical imaging tasks.

### F. LACK OF INTERPRETABILITY AND HIGH TRAINING COST

The challenges of interpretability in GAN model training and the associated high costs present significant obstacles in the realm of deep learning, especially in the domain of medical image augmentation. To address this challenge, a promising direction involves integrating classical methods with deep learning techniques. For instance, developing augmentation methods based on techniques like Temporal-Spectral Fusion Generative Adversarial Network (tsf-GAN) and Blending Generative Adversarial Network (blend-GAN) could enhance interpretability. Another issue is that the training procedure for GAN models demands robust GPU hardware, especially when training a substantial quantity of image data. Unfortunately, these systems can be expensive. Furthermore, the continuous advancement of computer hardware, with increased computational power and efficiency, is expected to reduce the training time of deep networks significantly.

### G. INTERDISCIPLINARY COLLABORATION AND SEMI-AUTOMATIC CREATION OF MEDICAL IMAGE REPORTS

The application of GANs specifically designed for text-generation tasks has been investigated for the semi-automatic generation of medical image reports. GANs with attention mechanisms can prioritize critical disease-related information during report generation, ultimately boosting diagnostic report creation and reducing the workload on physicians. As we look toward the future, it is clear that collaboration between medical professionals and AI researchers is essential. With their clinical expertise, physicians can offer specific guidance in feature selection, data tagging, and problem-solving. This interdisciplinary collaboration is critical to fully unlocking AI's potential in medical imaging. It can potentially reshape the roles of physicians, offering specific guidance in feature selection, data tagging, and problem-solving, thereby further advancing the field.

### H. AUTOMATIC 3D IMAGE MODALITY COLORIZATION AND 4D MEDICAL IMAGE AUGMENTATION

Automatic colorization of 3D medical image modalities represents an unexplored area in the field of medical imaging. By leveraging the generalization capabilities of GANs and 2D-style exemplars to colorize multi-modal 3D medical data, potential opportunities for improvement become apparent. This innovative approach aims to enhance the utility of such data. However, further research and refinement are required in this area. While most research projects in medical imaging concentrate on 2D and 3D pictures, utilizing the power of GANs for 4D-based images remains challenging. Augmented models capable of 4D time series prediction represent a promising future direction. In the future, techniques that take advantage of GANs success in 4D imaging might improve 4D time series prediction.

### *I. CLASS LEAKAGE*

Class leakage occurs when generated images include properties from another class, impacting the fidelity of the data. Developing suitable metrics and constraints in GANs is crucial to prevent class leakage and ensure that generated images faithfully represent the desired class characteristics. Addressing class leakage presents a promising direction for research. This can involve the development of novel loss functions or regularization methods that enforce class-wise separation during training. Moreover, exploring techniques from domain adaptation and transfer learning may offer valuable insights into mitigating the issue of mixing properties among classes.

### *J. ADDRESSING THE SKEWED DISTRIBUTION OF MEDICAL DATA AND DATASET QUALITY WITH AVAILABILITY*

Medical data often exhibits a skewed distribution, with a focus on common diseases. To address this imbalance, researchers can combine GANs and attention mechanisms to intentionally generate pathological cases of rare diseases, aiding in research and the diagnosis of less usual conditions. The quality and availability of medical datasets suitable for GANs pose challenges, including sparsity, age, and heterogeneity. Using GANs-based models in biomedical imaging has been severely limited by the lack of a large medical dataset. Efforts should be directed towards creating standardized and diverse medical datasets for GANs. This involves considering specific challenges posed by medical imaging tasks.

### *K. LACK OF PRESERVATION OF SMALL ABNORMALITIES*

Cross-modality models employing unpaired data lack assurance in preserving small abnormality regions throughout the transformation process. This issue arises due to biases in pre-trained models like Cycle-GAN when matching the produced data to the target domain's distribution. Biases can emerge when the training data shows an over or under-representation of particular classes. This holds even in the case of paired data. This bias can be mitigated through strategies such as modifying the discriminator architecture and implementing penalization based on smaller patches.

### *L. REINFORCEMENT LEARNING*

Reinforcement Learning (RL) applied to GANs faces several challenges that hamper its widespread adoption. One of the primary challenges lies in the complex design of RL. Slight variations in the design of states, rewards, and actions can lead to visible different outcomes, and finding optimal hyperparameters requires extensive experimentation. RL models low stability and reproducibility pose additional concerns, as the same workflows may result in different outcomes, particularly when the input data source is altered. Despite these challenges, there are promising future directions for reinforcement learning in GANs applied to medical imaging. One avenue is the exploration of Hierarchical Reinforcement Learning (HRL). HRL aims to enhance agent efficiency by breaking down complex tasks into hierarchical subtasks. Subclasses of HRL, such as spatiotemporal abstraction, intrinsic motivation, and deep successor RL, show potential in improving agent-based pipelines, particularly when dealing with high-dimensional 3D or even 4D image data. Additionally, multitasking and Transfer Reinforcement Learning can enhance RL agents adaptability. Training time can be significantly reduced by enabling agents to perform similar but different tasks through transfer learning. Active Reinforcement Learning presents another excellent direction for the field. Incorporating active learning strategies involving user interaction, particularly with physicians, can enhance the agent's understanding of user intentions and optimize performance with minimal annotated data. However, implementing active RL in practice may pose challenges due to the necessary involvement of human users. Figure 10 displays the general framework of Reinforcement Learning.

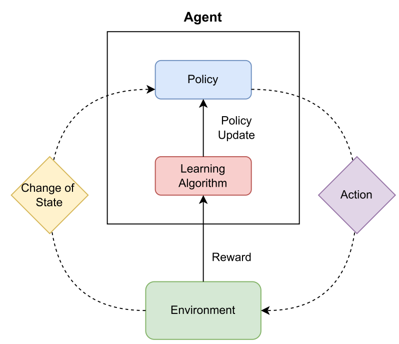


**FIGURE 10. Framework of reinforcement learning.**

Reinforcement Learning operates within a framework where an agent interacts with an environment, making sequential decisions to maximize cumulative rewards. The agent observes the environment's current state, selects actions based on its policy, executes these actions, receives feedback in the form of rewards, and transitions to a new state. The goal is for the agent to learn an optimal policy that guides decision-making for maximizing long-term rewards. This involves refining the policy through iterative learning processes.

### M. FEDERATED LEARNING

Federated learning holds opportunities for mitigating privacy risks in Generative Adversarial Networks for medical imaging. However, it faces challenges. The need for effective convergence within the constraints of limited communication bandwidth in decentralized setups is a significant hurdle. Integrating federated learning with differential privacy requires careful exploration to balance privacy preservation and model performance. The complexity of medical datasets, characterized by heterogeneity and diverse imaging modalities, poses additional challenges for collaborative learning. Future research directions include addressing communication bottlenecks, enhancing model robustness, and adapting federated learning to the intricacies of medical data. Delving into incorporating sophisticated privacy-preserving techniques and pursuing the development of a completely decentralized private GAN that can produce top-notch medical images without compromising data confidentiality continues to be interesting yet complex in the healthcare realm. Figure 11 showcases the workflow of Federated Learning.

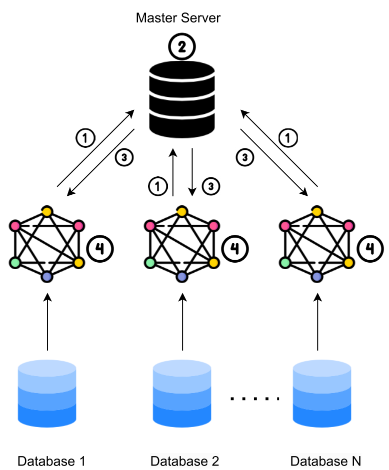


**FIGURE 11. Federated learning working procedure.**

In the initial step, each participant computes the model gradient locally, safeguarding the information with cryptographic techniques such as homomorphic encryption, and transmits the results to the master server. Step 2 entails the master server executing secure aggregation and distributing the aggregated results back to participants in step 3. During step 4, participants decrypt the received gradients and adjust their model parameters. This iterative process continues until the loss function converges or the maximum iterations are met. Notably, participant data remains decentralized, ensuring privacy compared to centralized models on Hadoop. Federated learning allows any number of databases to collaboratively train the model without transferring data to a central location, efficiently handling increasing data volumes without communication bandwidth concerns, as only local gradients are transmitted. ‘

### N. DIFFUSION MODELS

Alternative models to GANs, known as Diffusion models, are currently under development. These score-based techniques involve training by gradually introducing noise, following a Gaussian distribution, to images and learning the underlying data distribution. Instead of explicitly reversing the introduced noise, the primary objective is to model the complex data distribution. Diffusion models have shown the capability to generate realistic images with stable training and good mode coverage. However, a notable challenge is the extended sampling time associated with the diffusion process. While this may be less critical in non-real-time applications like medical imaging, researchers actively address this limitation by optimizing models for faster sampling. Future work in this field includes ongoing efforts to enhance diffusion models by reducing sampling time without compromising image quality. This involves exploring innovative model variants, refining real-time applicability, and advancing the overall effectiveness and applicability of diffusion models, especially in the context of medical imaging. Figure 12 showcases the process of diffusion models.

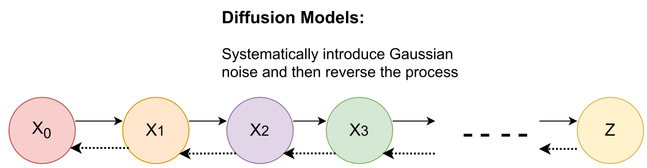


**FIGURE 12. Diffusion models working procedure.**

Diffusion models encompass forward and reverse diffusion processes. The forward diffusion involves a Markov chain that automatically adds noise to the input data until white noise is achieved. This is not a learnable process and usually encompasses 1000 steps. On the other hand, the reverse diffusion process is intended to systematically undo the forward process by eliminating noise step by step to reconstruct the original data. This reversal is facilitated by employing a trainable neural network.

## VIII. CONCLUSION

A specific kind of neural network configuration known as a Generative Adversarial Network trains two networks side by side, generating images and verifying and evaluating them. Using cutting-edge image analysis strategies, GANs improve images, generate essential training data, and advance disease detection and diagnosis in the field of medical imaging. GANs in medical imaging significantly improve the quality of images, enhance diagnostic accuracy, and increase small datasets, advancing patient care and research effectiveness. However, there is a shortage of studies on GANs in medical imaging. Thus, focusing on current advancements, this study offers a systematic review of GAN applications in medical imaging across several domains. It looks at prevalent datasets and preprocessing techniques, evaluates popular GAN algorithms, and does an in-depth evaluation of recent research. It also describes the difficulties that GANs encounter in medical imaging and suggests opportunities for further research. As a result, this study is an essential resource for collaborative research, offering insightful information to assist the development of GANs in medical imaging and helping researchers and investors who want to stay current and foster innovation in this field.

## ACKNOWLEDGMENT

The authors would like to thank the support provided by the Competitive Research Fund of The University of Aizu, Japan, which significantly contributed to the successful completion of this research project, and also would like to thank the Advanced Machine Intelligence Research Laboratory—AMIR Laboratory for their guidance and support.

## REFERENCES

[1] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101552.
[2] I. Goodfellow, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.
[3] C. Chu, A. Zhmoginov, and M. Sandler, "CycleGAN, a master of steganography," 2017, *arXiv:1712.02950*.
[4] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proc. CVPR*, pp. 1125–1134, Jul. 2017.
[5] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, *arXiv:1710.10196*.
[6] M. Alamir and M. Alghamdi, "The role of generative adversarial network in medical image analysis: An in-depth survey," *ACM Comput. Surv.*, vol. 55, no. 5, pp. 1–36, May 2023.
[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.
[8] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.
[9] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–12.
[10] S. Motamed, P. Rogalla, and F. Khalvati, "RANDGAN: Randomized generative adversarial network for detection of COVID-19 in chest X-ray," *Sci. Rep.*, vol. 11, no. 1, p. 8602, Apr. 2021.
[11] T. Kossen, P. Subramaniam, V. I. Madai, A. Hennemuth, K. Hildebrand, A. Hilbert, J. Sobesky, M. Livne, I. Galinovic, A. A. Khalil, J. B. Fiebach, and D. Frey, "Synthesizing anonymized and labeled TOF-MRA patches for brain vessel segmentation using generative adversarial networks," *Comput. Biol. Med.*, vol. 131, Apr. 2021, Art. no. 104254.
[12] B. Ahmad, J. Sun, Q. You, V. Palade, and Z. Mao, "Brain tumor classification using a combination of variational autoencoders and generative adversarial networks," *Biomedicines*, vol. 10, no. 2, p. 223, Jan. 2022.
[13] J. J. Jeong, A. Tariq, T. Adejumo, H. Trivedi, J. W. Gichoya, and I. Banerjee, "Systematic review of generative adversarial networks (GANs) for medical image classification and segmentation," *J. Digit. Imag.*, vol. 35, no. 2, pp. 137–152, Apr. 2022.
[14] Y. Chen, X.-H. Yang, Z. Wei, A. A. Heidari, N. Zheng, Z. Li, H. Chen, H. Hu, Q. Zhou, and Q. Guan, "Generative adversarial networks in medical image augmentation: A review," *Comput. Biol. Med.*, vol. 144, May 2022, Art. no. 105382.
[15] S. Xun, D. Li, H. Zhu, M. Chen, J. Wang, J. Li, M. Chen, B. Wu, H. Zhang, and X. Chai, "Generative adversarial networks in medical image segmentation: A review," *Comput. Biol. Med.*, vol. 140, May 2022, Art. no. 105063.
[16] A. Makhlouf, M. Maayah, N. Abughanam, and C. Catal, "The use of generative adversarial networks in medical image augmentation," *Neural Comput. Appl.*, vol. 35, no. 34, pp. 24055–24068, Dec. 2023.
[17] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," Dept. Comput. Sci., Univ. Durham, Durham, U.K., Tech. Rep. EBSE-2007-01, 2007.
[18] B. Kitchenham, "Procedures for performing systematic reviews," Keele Univ., Keele, U.K., Tech. Rep. TR/SE-0401, 2004, pp. 1–26, vol. 33.
[19] A. You, J. K. Kim, I. H. Ryu, and T. K. Yoo, "Application of generative adversarial networks (GAN) for ophthalmology image domains: A survey," *Eye Vis.*, vol. 9, no. 1, pp. 1–19, Dec. 2022.
[20] Z. Zhang, Y. Li, and B. Shin, "C2-GAN: Content-consistent generative adversarial networks for unsupervised domain adaptation in medical image segmentation," *Med. Phys.*, vol. 49, no. 10, pp. 6491–6504, Oct. 2022.
[21] S. A. Güven and M. F. Talu, "Brain MRI high resolution image creation and segmentation with the new GAN method," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104246.
[22] Y. Al Khalil, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "On the usability of synthetic data for improving the robustness of deep learning-based segmentation of cardiac magnetic resonance images," *Med. Image Anal.*, vol. 84, Feb. 2023, Art. no. 102688.
[23] Y. Al Khalil, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "Reducing segmentation failures in cardiac MRI via late feature fusion and GAN-based augmentation," *Comput. Biol. Med.*, vol. 161, Jul. 2023, Art. no. 106973.
[24] A. Kunapinun, M. N. Dailey, D. Songsaeng, M. Parnichkun, C. Keatmanee, and M. Ekpanyapong, "Improving GAN learning dynamics for thyroid nodule segmentation," *Ultrasound Med. Biol.*, vol. 49, no. 2, pp. 416–430, Feb. 2023.
[25] S. R. Rezaei and A. Ahmadi, "A GAN-based method for 3D lung tumor reconstruction boosted by a knowledge transfer approach," *Multimedia Tools Appl.*, vol. 82, no. 28, pp. 44359–44385, Nov. 2023.
[26] S. J. Narayanan, A. S. Anil, C. Ashtikar, S. Chunduri, and S. Saman, "Automated brain tumor segmentation using GAN augmentation and optimized U-Net," in *Frontiers ICT Healthcare*. Cham, Switzerland: Springer, 2023, pp. 635–646.
[27] L. Zhu, Q. He, Y. Huang, Z. Zhang, J. Zeng, L. Lu, W. Kong, and F. Zhou, "DualMMP-GAN: Dual-scale multi-modality perceptual generative adversarial network for medical image segmentation," *Comput. Biol. Med.*, vol. 144, May 2022, Art. no. 105387.
[28] F. Munawar, S. Azmat, T. Iqbal, C. Grönlund, and H. Ali, "Segmentation of lungs in chest X-ray image using generative adversarial networks," *IEEE Access*, vol. 8, pp. 153535–153545, 2020.
[29] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1310–1321, Jun. 2018.
[30] W. Du and S. Tian, "Transformer and GAN-based super-resolution reconstruction network for medical images," *Tsinghua Sci. Technol.*, vol. 29, no. 1, pp. 197–206, Feb. 2024.
[31] Y. Wang, Z.-L. Sun, Z. Zeng, and K.-M. Lam, "TRCT-GAN: CT reconstruction from biplane X-rays using transformer and generative adversarial networks," *Digit. Signal Process.*, vol. 140, Aug. 2023, Art. no. 104123.
[32] R. Mishra, K. Sharma, R. R. Jha, and A. Bhavsar, "NeuroGAN: Image reconstruction from EEG signals via an attention-based GAN," *Neural Comput. Appl.*, vol. 35, pp. 9181–9192, Dec. 2022.
[33] J. Jiang, Y. Feng, H. Xu, and J. Zheng, "Low-dose CT reconstruction via optimization-inspired GAN," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.
[34] S. Ramanathan and M. Ramasundaram, "Low-dose CT image reconstruction using vector quantized convolutional autoencoder with perceptual loss," *Sādhanā*, vol. 48, no. 2, p. 43, Mar. 2023.
[35] T. Zhou, Q. Li, H. Lu, Q. Cheng, and X. Zhang, "GAN review: Models and medical image fusion applications," *Inf. Fusion*, vol. 91, pp. 134–148, Mar. 2023.
[36] Q. Guan, Y. Chen, Z. Wei, A. A. Heidari, H. Hu, X.-H. Yang, J. Zheng, Q. Zhou, H. Chen, and F. Chen, "Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN," *Comput. Biol. Med.*, vol. 145, Jun. 2022, Art. no. 105444.

[37] D. Srivastav, A. Bajpai, and P. Srivastava, "Improved classification for pneumonia detection using transfer learning with GAN based synthetic image augmentation," in *Proc. 11th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2021, pp. 433–437.

[38] Y. Zhang, S. He, S. Wa, Z. Zong, J. Lin, D. Fan, J. Fu, and C. Lv, "Symmetry GAN detection network: An automatic one-stage high-accuracy detection network for various types of lesions on CT images," *Symmetry*, vol. 14, no. 2, p. 234, Jan. 2022.

[39] S. Vashisht, B. Sharma, and S. Lamba, "Alzheimer detection using CNN and GAN augmentation," in *Proc. World Conf. Commun. Comput. (WCONF)*, Jul. 2023, pp. 1–5.

[40] M. V. K. Reddy, P. K. Murjani, S. Rajkumar, T. Chen, and V. A. Chandrasekar, "Optimized CNN model with deep convolutional GAN for brain tumor detection," in *Congress on Intelligent Systems*. Cham, Switzerland: Springer, 2022, pp. 409–425.

[41] G. Liu, S. Lan, T. Zhang, W. Huang, and W. Wang, "SAGAN: Skip-attention GAN for anomaly detection," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2021, pp. 2468–2472.

[42] K. Suganthi, "Review of medical image synthesis using GAN techniques," in *Proc. ITM Web Conf.*, vol. 37, 2021, Art. no. 01005.

[43] H. Tan, X. Lang, B. He, Y. Lu, and Y. Zhang, "GAN-based medical image augmentation for improving CNN performance in myositis ultrasound image classification," in *Proc. 6th Int. Conf. Electron. Technol. (ICET)*, May 2023, pp. 1329–1333.

[44] H. H. N. Alrashedy, A. F. Almansour, D. M. Ibrahim, and M. A. A. Hammoudeh, "BrainGAN: Brain MRI image generation and classification framework using GAN architectures and CNN models," *Sensors*, vol. 22, no. 11, p. 4297, Jun. 2022.

[45] J. Mao, X. Yin, G. Zhang, B. Chen, Y. Chang, W. Chen, J. Yu, and Y. Wang, "Pseudo-labeling generative adversarial networks for medical image classification," *Comput. Biol. Med.*, vol. 147, Aug. 2022, Art. no. 105729.

[46] W. Chen, S. Yu, K. Ma, W. Ji, C. Bian, C. Chu, L. Shen, and Y. Zheng, "TW-GAN: Topology and width aware GAN for retinal artery/vein classification," *Med. Image Anal.*, vol. 77, Apr. 2022, Art. no. 102340.

[47] A. A. M. Teodoro, D. H. Silva, R. L. Rosa, M. Saadi, L. Wuttisittikulkij, R. A. Mumtaz, and D. Z. Rodríguez, "A skin cancer classification approach using GAN and RoI-based attention mechanism," *J. Signal Process. Syst.*, vol. 95, nos. 2–3, pp. 211–224, Mar. 2023.

[48] R. Zhang, W. Lu, J. Gao, Y. Tian, X. Wei, C. Wang, X. Li, and M. Yu, "RFI-GAN: A reference-guided fuzzy integral network for ultrasound image augmentation," *Inf. Sci.*, vol. 623, pp. 709–728, Apr. 2023.

[49] Y. Zhang, Z. Wang, Z. Zhang, J. Liu, Y. Feng, L. Wee, A. Dekker, Q. Chen, and A. Traverso, "GAN-based one dimensional medical data augmentation," *Soft Comput.*, vol. 27, no. 15, pp. 10481–10491, Aug. 2023.

[50] Y. Zhang, Q. Wang, and B. Hu, "MinimalGAN: Diverse medical image synthesis for data augmentation using minimal training data," *Appl. Intell.*, vol. 53, no. 4, pp. 3899–3916, Feb. 2023.

[51] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "COVIDGAN: Data augmentation using auxiliary classifier GAN for improved COVID-19 detection," *IEEE Access*, vol. 8, pp. 91916–91923, 2020.

[52] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: A review," *Phys. Med. Biol.*, vol. 65, no. 20, 2020, Art. no. 20TR01.

[53] Y. Liu, W. Wang, Y. Li, H. Lai, S. Huang, and X. Yang, "Geometry-consistent adversarial registration model for unsupervised multi-modal medical image registration," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 7, pp. 3455–3466, Jul. 2023.

[54] C. Suwanraksa, J. Bridhikitti, T. Liamsuwan, and S. Chaichulee, "CBCT-to-CT translation using registration-based generative adversarial networks in patients with head and neck cancer," *Cancers*, vol. 15, no. 7, p. 2017, Mar. 2023.

[55] J. Mi, L. Wang, Y. Liu, and J. Zhang, "KDE-GAN: A multimodal medical image-fusion model based on knowledge distillation and explainable AI modules," *Comput. Biol. Med.*, vol. 151, Dec. 2022, Art. no. 106273.

[56] C. Fan, H. Lin, and Y. Qiu, "U-patch GAN: A medical image fusion method based on GAN," *J. Digit. Imag.*, vol. 36, no. 1, pp. 339–355, Aug. 2022.

[57] J. Huang, Z. Le, Y. Ma, F. Fan, H. Zhang, and L. Yang, "MGMDcGAN: Medical image fusion using multi-generator multi-discriminator conditional generative adversarial network," *IEEE Access*, vol. 8, pp. 55145–55157, 2020.

[58] Y. Skandarani, P.-M. Jodoin, and A. Lalande, "GANs for medical image synthesis: An empirical study," *J. Imag.*, vol. 9, no. 3, p. 69, Mar. 2023.

[59] R. Jin and X. Li, "Backdoor attack and defense in federated generative adversarial network-based medical image synthesis," *Med. Image Anal.*, vol. 90, Dec. 2023, Art. no. 102965.

[60] J. Wang, G. Xie, Y. Huang, J. Lyu, F. Zheng, Y. Zheng, and Y. Jin, "FedMed-GAN: Federated domain translation on unsupervised cross-modality brain image synthesis," *Neurocomputing*, vol. 546, Aug. 2023, Art. no. 126282.

[61] B. Sun, S. Jia, X. Jiang, and F. Jia, "Double U-Net CycleGAN for 3D MR to CT image synthesis," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 18, no. 1, pp. 149–156, Aug. 2022.

[62] J. Mendes, T. Pereira, F. Silva, J. Frade, J. Morgado, C. Freitas, E. Negrão, B. F. de Lima, M. C. da Silva, A. J. Madureira, I. Ramos, J. L. Costa, V. Hespanhol, A. Cunha, and H. P. Oliveira, "Lung CT image synthesis using GANs," *Expert Syst. Appl.*, vol. 215, Apr. 2023, Art. no. 119350.

[63] *The Whole Brain Atlas—Med.harvard.edu*. Accessed: Oct. 22, 2023. [Online]. Available: http://www.med.harvard.edu/AANLIB/home.html

[64] *CORN: CORneal Nerve Database—Imed.Nimte.Ac.Cn—Imed.Nimte.Ac.Cn*. Accessed: Jan. 4, 2024. [Online]. Available: https://imed.nimte.ac.cn/CORN.html

[65] S. Jaeger, S. Candemir, S. Antani, Y. X. Wang, P. X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quant. Imag. Med. Surg.*, vol. 4, p. 475, Dec. 2014.

[66] *Polyp—Grand Challenge—Polyp.grand-challenge.org*. Accessed: Oct. 22, 2023. [Online]. Available: https://polyp.grand-challenge.org/CVCClinicDB/

[67] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, Jul. 2018.

[68] S. G. Mueller, M. W. Weiner, L. J. Thal, R. C. Petersen, C. R. Jack, W. Jagust, J. Q. Trojanowski, A. W. Toga, and L. Beckett, "Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's disease neuroimaging initiative (ADNI)," *Alzheimer's Dementia*, vol. 1, no. 1, pp. 55–66, Jul. 2005.

[69] *Multimodal Brain Tumor Segmentation Challenge 2019 | CBICA | Perelman School of Medicine at the University of Pennsylvania—Med.upenn.edu*. Accessed: Oct. 22, 2023. [Online]. Available: https://www.med.upenn.edu/cbica/brats-2019/

[70] *ISIC Challenge—Challenge2018.ISIC-Archive.com*. Accessed: Oct. 22, 2023. [Online]. Available: https://challenge2018.isic-archive.com/

[71] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, no. 1, p. 19549, Nov. 2020.

[72] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.

[73] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.

[74] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database," *IET Image Process.*, vol. 7, no. 4, pp. 373–383, Jun. 2013.

[75] C. G. Owen, A. R. Rudnicka, R. Mullen, S. A. Barman, D. Monekosso, P. H. Whincup, J. Ng, and C. Paterson, "Measuring retinal vessel tortuosity in 10-year-old children: Validation of the computer-assisted image analysis of the retina (CAIAR) program," *Investigative Opthalmol.Vis. Sci.*, vol. 50, no. 5, p. 2004, May 2009.

[76] A. Obukhov and M. Krasnyanskiy, "Quality assessment method for GAN based on modified metrics inception score and fréchet inception distance," in *Proc. Comput. Methods Syst. Softw.*, vol. 1. Cham, Switzerland: Springer, 2020, pp. 102–114.

[77] K. Shmelkov, C. Schmid, and K. Alahari, "How good is my GAN?" in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 213–229.

[78] A. Borji, "Pros and cons of GAN evaluation measures: New developments," *Comput. Vis. Image Understand.*, vol. 215, Jan. 2022, Art. no. 103329.

[79] M. J. Chong and D. Forsyth, "Effectively unbiased FID and inception score and where to find them," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Sep. 2020, pp. 6070–6079.

[80] S. Ravuri and O. Vinyals, "Classification accuracy score for conditional generative models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–14.

[81] M. Hossin, M. N. Sulaiman, A. Mustapha, N. Mustapha, and R. W. Rahmat, "A hybrid evaluation metric for optimizing classifier," in *Proc. 3rd Conf. Data Mining Optim. (DMO)*, Jun. 2011, pp. 165–170.

[82] F. J. Moreno-Barea, J. M. Jerez, and L. Franco, "Improving classification accuracy using data augmentation on small data sets," *Expert Syst. Appl.*, vol. 161, Dec. 2020, Art. no. 113696.

[83] Z. Qin, Z. Liu, P. Zhu, and Y. Xue, "A GAN-based image synthesis method for skin lesion classification," *Comput. Methods Programs Biomed.*, vol. 195, Oct. 2020, Art. no. 105568.

[84] L. Gupta, B. M. Klinkhammer, P. Boor, D. Merhof, and M. Gadermayr, "GAN-based image enrichment in digital pathology boosts segmentation accuracy," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Shenzhen, China. Cham, Switzerland: Springer, Oct. 2019, pp. 631–639.

[85] C. Bowles, L. Chen, R. Guerrero, P. Bentley, R. Gunn, A. Hammers, D. A. Dickie, M. V. Hernández, J. Wardlaw, and D. Rueckert, "GAN augmentation: Augmenting training data using generative adversarial networks," 2018, *arXiv:1810.10863*.

[86] S. Nema, A. Dudhane, S. Murala, and S. Naidu, "RescueNet: An unpaired GAN for brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 55, Jan. 2020, Art. no. 101641.

[87] X. Dong, Y. Lei, T. Wang, M. Thomas, L. Tang, W. J. Curran, T. Liu, and X. Yang, "Automatic multiorgan segmentation in thorax CT images using U-Net-GAN," *Med. Phys.*, vol. 46, no. 5, pp. 2157–2168, May 2019.

[88] T. R. V. Bisneto, A. O. de Carvalho Filho, and D. M. V. Magalhães, "Generative adversarial network and texture features applied to automatic glaucoma detection," *Appl. Soft Comput.*, vol. 90, May 2020, Art. no. 106165.

[89] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "F-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," *Med. Image Anal.*, vol. 54, pp. 30–44, May 2019.

[90] B. Hu, Y. Tang, E. I. Chang, Y. Fan, M. Lai, and Y. Xu, "Unsupervised learning for cell-level visual representation in histopathology images with generative adversarial networks," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 3, pp. 1316–1328, May 2019.

[91] X. He, Y. Hua, T. Song, Z. Zhang, Z. Xue, R. Ma, N. Robertson, and H. Guan, "Unsupervised video summarization with attentive conditional generative adversarial networks," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 2296–2304.

[92] H. Mansourifar, L. Chen, and W. Shi, "Virtual big data for GAN based data augmentation," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 1478–1487.

[93] J. Lee and K. Park, "GAN-based imbalanced data intrusion detection system," *Pers. Ubiquitous Comput.*, vol. 25, no. 1, pp. 121–128, Feb. 2021.

[94] Y. Yang, C. Hou, Y. Lang, G. Yue, and Y. He, "One-class classification using generative adversarial networks," *IEEE Access*, vol. 7, pp. 37970–37979, 2019.

[95] F. Calimeri, A. Marzullo, C. Stamile, and G. Terracina, "Biomedical data augmentation using generative adversarial neural networks," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2017, pp. 626–634.

[96] T. Iqbal and H. Ali, "Generative adversarial network for medical images (MI-GAN)," *J. Med. Syst.*, vol. 42, no. 11, pp. 1–11, Nov. 2018.

[97] G. Zhang, Y. Pan, and L. Zhang, "Semi-supervised learning with GAN for automatic defect detection from images," *Autom. Construct.*, vol. 128, Aug. 2021, Art. no. 103764.

[98] E. Agustsson, M. Tschannen, F. Mentzer, R. Timofte, and L. Van Gool, "Generative adversarial networks for extreme learned image compression," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 221–231.

[99] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *Proc. Int. Symp. Vis. Comput.* Cham, Switzerland: Springer, 2016, pp. 234–244.

[100] M. M. R. Siddiquee et al., "Learning fixed points in generative adversarial networks: From image-to-image translation to disease detection and localization," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 191–200.

[101] T. Kynkäänniemi, T. Karras, S. Laine, J. Lehtinen, and T. Aila, "Improved precision and recall metric for assessing generative models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–15.

[102] A. Sharma, P. K. Singh, and R. Chandra, "SMOTified-GAN for class imbalanced pattern classification problems," *IEEE Access*, vol. 10, pp. 30655–30665, 2022.

[103] M. K. Baowaly, C.-C. Lin, C.-L. Liu, and K.-T. Chen, "Synthesizing electronic health records using improved generative adversarial networks," *J. Amer. Med. Inform. Assoc.*, vol. 26, no. 3, pp. 228–241, Mar. 2019.

[104] A. W. Setiawan, "Image segmentation metrics in skin lesion: Accuracy, sensitivity, specificity, dice coefficient, Jaccard index, and Matthews correlation coefficient," in *Proc. Int. Conf. Comput. Eng., Netw., Intell. Multimedia (CENIM)*, Nov. 2020, pp. 97–102.

[105] T. Goel, R. Murugan, S. Mirjalili, and D. K. Chakrabartty, "Automatic screening of COVID-19 using an optimized generative adversarial network," *Cognit. Comput.*, vol. 13, pp. 1–16, Jan. 2021.

[106] C. Zheng, X. Xie, K. Zhou, B. Chen, J. Chen, H. Ye, W. Li, T. Qiao, S. Gao, J. Yang, and J. Liu, "Assessment of generative adversarial networks model for synthetic optical coherence tomography images of retinal disorders," *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 29, May 2020.

[107] A. M. Shaker, M. Tantawi, H. A. Shedeed, and M. F. Tolba, "Generalization of convolutional neural networks for ECG classification using generative adversarial networks," *IEEE Access*, vol. 8, pp. 35592–35605, 2020.

[108] T. Fujioka, K. Kubota, M. Mori, Y. Kikuchi, L. Katsuta, M. Kimura, E. Yamaga, M. Adachi, G. Oda, T. Nakagawa, Y. Kitazume, and U. Tateishi, "Efficient anomaly detection with generative adversarial network for breast ultrasound imaging," *Diagnostics*, vol. 10, no. 7, p. 456, Jul. 2020.

[109] V. Alex, K. P. M. Safwan, S. S. Chennamsetty, and G. Krishnamurthi, "Generative adversarial networks for brain lesion detection," *Proc. SPIE*, vol. 10133, Feb. 2017, pp. 113–121.

[110] K. Guo, X. Hu, and X. Li, "MMFGAN: A novel multimodal brain medical image fusion based on the improvement of generative adversarial network," *Multimedia Tools Appl.*, vol. 81, no. 4, pp. 5889–5927, Feb. 2022.

[111] M. Alauthman, A. Al-Qerem, B. Sowan, A. Alsarhan, M. Eshtay, A. Aldweesh, and N. Aslam, "Enhancing small medical dataset classification performance using GAN," *Informatics*, vol. 10, no. 1, p. 28, Mar. 2023.

[112] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 889–896, May 2000.

[113] B. Xu, D. Zhou, and W. Li, "Image enhancement algorithm based on GAN neural network," *IEEE Access*, vol. 10, pp. 36766–36777, 2022.

[114] S. G. K. Patro and K. K. Sahu, "Normalization: A preprocessing stage," 2015, *arXiv:1503.06462*.

[115] S. Cackowski, E. L. Barbier, M. Dojat, and T. Christen, "ImUnity: A generalizable VAE-GAN solution for multicenter MR image harmonization," *Med. Image Anal.*, vol. 88, Aug. 2023, Art. no. 102799.

[116] Y. Chen, Y. Lin, X. Xu, J. Ding, C. Li, Y. Zeng, W. Xie, and J. Huang, "Multi-domain medical image translation generation for lung image classification based on generative adversarial networks," *Comput. Methods Programs Biomed.*, vol. 229, Feb. 2023, Art. no. 107200.

[117] G. W. Goerres, E. Kamel, T.-N.-H. Heidelberg, M. R. Schwitter, C. Burger, and G. K. von Schulthess, "PET-CT image co-registration in the thorax: Influence of respiration," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 29, no. 3, pp. 351–360, Mar. 2002.

[118] A. Singh, J. Kwiecinski, S. Cadet, A. Killekar, E. Tzolos, M. C. Williams, M. R. Dweck, D. E. Newby, D. Dey, and P. J. Slomka, "Automated nonlinear registration of coronary PET to CT angiography using pseudo-CT generated from PET with generative adversarial networks," *J. Nucl. Cardiology*, vol. 30, no. 2, pp. 604–615, Apr. 2023.

[119] Y.-F. Yu, G. Zhong, Y. Zhou, and L. Chen, "FS-GAN: Fuzzy self-guided structure retention generative adversarial network for medical image enhancement," *Inf. Sci.*, vol. 642, Sep. 2023, Art. no. 119114.

[120] B. Zhao, T. Cheng, X. Zhang, J. Wang, H. Zhu, R. Zhao, D. Li, Z. Zhang, and G. Yu, "CT synthesis from MR in the pelvic area using residual transformer conditional GAN," *Computerized Med. Imag. Graph.*, vol. 103, Jan. 2023, Art. no. 102150.
[121] W. R. Crum, T. Hartkens, and D. L. G. Hill, "Non-rigid image registration: Theory and practice," *Brit. J. Radiol.*, vol. 77, no. 2, pp. S140–S153, Dec. 2004.
[122] F. Che, X. Zhu, T. Yang, and T.-Y. Yu, "3SGAN: 3D shape embedded generative adversarial networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3305–3314.
[123] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao, "LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop," 2015, *arXiv:1506.03365*.
[124] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, Dec. 2018.
[125] M. Li, H. Tang, M. D. Chan, X. Zhou, and X. Qian, "DC-AL GAN: Pseudoprogression and true tumor progression of glioblastoma multiform image classification based on DCGAN and AlexNet," *Med. Phys.*, vol. 47, no. 3, pp. 1139–1150, Mar. 2020.
[126] A. Madani, M. Moradi, A. Karargyris, and T. Syeda-Mahmood, "Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 1038–1042.
[127] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2642–2651.
[128] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–13.
[129] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.
[130] T. Miyato and M. Koyama, "CGANs with projection discriminator," 2018, *arXiv:1802.05637*.
[131] Z. Zhuang, W. Shen, A. Damian, and T. D. Tran, "Incorporating knowledge distillation to the training of medical image synthesis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent (MICCAI)*, 2019, pp. 152–160.
[132] D. Nie, Y. Gao, L. Wang, and D. Shen, "ASDNet: Attention-based semi-supervised deep networks for medical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent (MICCAI)*, 2019, pp. 370–378.
[133] Y. Wang, B. Yu, L. Wang, C. Zu, D. S. Lalush, W. Lin, X. Wu, J. Zhou, D. Shen, and L. Zhou, "3D conditional generative adversarial networks for high-quality PET image estimation at low dose," *NeuroImage*, vol. 174, pp. 550–562, Jul. 2018.
[134] Y. Gao, X. Cao, K. Ma, and D. Shen, "Multi-modal residual network for semi-supervised multi-organ segmentation," *Med. Image Anal.*, vol. 60, Jun. 2020, Art. no. 101612.
[135] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," 2016, *arXiv:1609.03126*.
[136] D. Berthelot, T. Schumm, and L. Metz, "BEGAN: Boundary equilibrium generative adversarial networks," 2017, *arXiv:1703.10717*.
[137] C. Ledig, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.
[138] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.
[139] J. Gu, Z. Yang, H. Zhang, and H. Zhang, "Super-resolution synthesis of computed tomography (CT) and magnetic resonance (MR) images using generative adversarial network (GAN)," *Med. Sci. Monitor, Int. Med. J. Experim. Clin. Res.*, vol. 25, p. 6643, Jan. 2019.
[140] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2813–2821.
[141] X. Huang, "Multimodal unsupervised image-to-image translation," in *Proc. ECCV*, 2018, pp. 172–189.
[142] J.-Y. Zhu, R. Zhang, D. Pathak, T. Darrell, A. A. Efros, O. Wang, and E. Shechtman, "Toward multimodal image-to-image translation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.
[143] J. T. Springenberg, "Unsupervised and semi-supervised learning with categorical generative adversarial networks," 2015, *arXiv:1511.06390*.
[144] X. Yi, E. Walia, and P. Babyn, "Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by Wasserstein distance for dermoscopy image classification," 2018, *arXiv:1804.03700*.
[145] L. Lan, L. You, Z. Zhang, Z. Fan, W. Zhao, N. Zeng, Y. Chen, and X. Zhou, "Generative adversarial networks and its applications in biomedical informatics," *Frontiers Public Health*, vol. 8, p. 164, May 2020.
[146] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. ICCV*, Oct. 2017, pp. 2223–2232.
[147] A. Jolicoeur-Martineau, "The relativistic discriminator: A key element missing from standard GAN," 2018, *arXiv:1807.00734*.
[148] J. Song and S. Ermon, "Improved techniques for training GANs," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2018, pp. 2234–2245.
[149] P. Achlioptas, O. Diamanti, I. Mitliagkas, and L. Guibas, "Learning representations and generative models for 3D point clouds," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2018, pp. 40–49.
[150] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 7354–7363.
[151] X. Wang, Z. Cao, R. Wang, Z. Liu, and X. Zhu, "Improving human pose estimation with self-attention generative adversarial networks," *IEEE Access*, vol. 7, pp. 119668–119680, 2019.
[152] Y. Li, H. Huang, L. Zhang, G. Wang, H. Zhang, and W. Zhou, "Super-resolution and self-attention with generative adversarial network for improving malignancy characterization of hepatocellular carcinoma," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1556–1560.
[153] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 214–223.
[154] Z. Yang, Y. Chen, Z. Le, F. Fan, and E. Pan, "Multi-source medical image fusion based on wasserstein generative adversarial networks," *IEEE Access*, vol. 7, pp. 175947–175958, 2019.
[155] G. Perarnau, J. van de Weijer, B. Raducanu, and J. M. Álvarez, "Invertible conditional GANs for image editing," 2016, *arXiv:1611.06355*.
[156] X. Yan, J. Yang, K. Sohn, and H. Lee, "Attribute2Image: Conditional image generation from visual attributes," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 776–791.
[157] X. Wang and A. Gupta, "Generative image modeling using style and structure adversarial networks," in *Proc. Eur. Conf. Comput. Vis.*, vol. 9908, 2016, pp. 318–335
[158] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1060–1069.
[159] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Jan. 1998.
[160] Y. Jiang, H. Chen, M. Loew, and H. Ko, "COVID-19 CT image synthesis with a conditional generative adversarial network," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 441–452, Feb. 2021.
[161] L. Karacan, Z. Akata, A. Erdem, and E. Erdem, "Learning to generate images of outdoor scenes from attributes and semantic layouts," 2016, *arXiv:1612.00215*.
[162] P. Sangkloy, J. Lu, C. Fang, F. Yu, and J. Hays, "Scribbler: Controlling deep image synthesis with sketch and color," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5400–5409.
[163] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. Metaxas, "StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5908–5916.
[164] A. Mobiny, P. A. Cicalese, S. Zare, P. Yuan, M. Abavisani, C. C. Wu, J. Ahuja, P. M. de Groot, and H. Van Nguyen, "Radiologist-level COVID-19 detection using CT scans with detail-oriented capsule networks," 2020, *arXiv:2004.07407*.
[165] T. C. Mok and A. C. Chung, "Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks," in *Proc. Int. MICCAI Brainlesion Workshop*, Granada, Spain. Cham, Switzerland: Springer, Sep. 2018, pp. 70–80.

[166] V. K. Singh, H. A. Rashwan, F. Akram, N. Pandey, M. M. K. Sarker, A. Saleh, S. Abdulwahab, N. Maaroof, J. Torrents-Barrena, and S. Romani, *Retinal Optic Disc Segmentation Using Conditional Generative Adversarial Network*, vol. 308. Amsterdam, The Netherlands: IOS Press, 2018.
[167] K. Guo, J. Chen, T. Qiu, S. Guo, T. Luo, T. Chen, and S. Ren, "MedGAN: An adaptive GAN approach for medical image generation," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107119.
[168] X. Liu, H. Chen, C. Yao, R. Xiang, K. Zhou, P. Du, W. Liu, J. Liu, and Z. Yu, "BTMF-GAN: A multi-modal MRI fusion generative adversarial network for brain tumors," *Comput. Biol. Med.*, vol. 157, May 2023, Art. no. 106769.
[169] Q. Zuo, H. Tian, R. Li, J. Guo, J. Hu, L. Tang, Y. Di, and H. Kong, "Hemisphere-separated cross-connectome aggregating learning via VAE-GAN for brain structural connectivity synthesis," *IEEE Access*, vol. 11, pp. 48493–48505, 2023.
[170] F. Kadri, A. Dairi, F. Harrou, and Y. Sun, "Towards accurate prediction of patient length of stay at emergency department: A GAN-driven deep learning framework," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 9, pp. 11481–11495, Sep. 2023.
[171] Y. Li and L. Shen, "CC-GAN: A robust transfer-learning framework for HEp-2 specimen image segmentation," *IEEE Access*, vol. 6, pp. 14048–14058, 2018.
[172] K.-B. Park, S. H. Choi, and J. Y. Lee, "M-GAN: Retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks," *IEEE Access*, vol. 8, pp. 146308–146322, 2020.

**SHOWROV ISLAM** is currently pursuing the Bachelor of Science degree in computer science and engineering from American International University Bangladesh (AIUB), in 2023. Throughout his academic journey, he consistently demonstrated outstanding performance and he was honored with the prestigious Deans List recognition for six consecutive times. With an exceptional academic record, he is on track to achieve the esteemed Magna Cum Laude distinction. Currently, he is making significant contributions as a Junior Software Engineer with Techdojo Ltd., a dynamic research and development company committed to innovating and advancing technologies. His role involves actively participating in creating and improving new technologies, showcasing his prowess in the software engineering domain. In addition to his professional role, he is also a Research Intern with the Advanced Machine Intelligence Research Laboratory—AMIR Laboratory. His research interests include various domains, including natural language processing (NLP), deep learning, and medical imaging. This diverse range of interests showcases his commitment to exploring and contributing to the forefront of technological advancements.

**MD. TAREK AZIZ** is currently pursuing the B.Sc. degree in computer science and engineering (CSE) with a major in information systems with American International University-Bangladesh (AIUB). He is familiar with various programming languages with expertise in C, C++, and Python. In addition, he also has a good foundation in data structure and algorithms. He loves problem-solving and critical thinking. His research interests include computational biology, bioinformatics, data science, computer vision, natural language processing, and neural networks. He is proficient in machine learning and deep learning to derive valuable insights from diverse datasets. He is also working on attention mechanisms, transfer learning, and ensemble learning to address challenges in DNA, RNA, protein, gene, and drug-related problems.

**HADIUR RAHMAN NABIL** is currently pursuing the B.Sc. degree in computer science and engineering (CSE) with American International University-Bangladesh. He actively participates in competitive programming and he has competed in various programming competitions, notably the ACM International Collegiate Programming Contest Preliminary (ACM-ICPC) and Intra-University Programming Contests. Proficient in multiple programming languages, such as C++, Python, R, Java, C#, and PHP, he also possesses experience in technologies, such as Tensorflow, ASP.Net, MATLAB, and MySQL. He also serves as a Research Intern with the Advanced Machine Intelligence Research Laboratory. His research interests include artificial intelligence, data science, computer vision, deep learning, and medical imaging.

**JAMIN RAHMAN JIM** received the B.Sc. degree in computer science and engineering (CSE) with a major in information systems from American International University-Bangladesh (AIUB), in 2023. For his academic excellence, he was mentioned in the prestigious Deans list and expected summa cum laude. He finished his internship as a Research Assistant (RA) with Deepchain Labs, a blockchain-based software development company. Currently, he is a Research Assistant and a System Engineer with the Advanced Machine Intelligence Research Laboratory—AMIR Laboratory. He is also working as a Research Assistant at the American International University-Bangladesh. With an extensive research background, he has authored over five articles in high-impact journals, such as IEEE Access, *Computers & Security and Natural Language Processing Journal*. Moreover, some of his research work has been published as a chapter in a few esteemed books related to machine learning and AI. His research interests include deep learning, data science, medical imaging, federated learning, robotics, and natural language processing.

**M. F. MRIDHA** (Senior Member, IEEE) received the Ph.D. degree in AI/ML from Jahangirnagar University, in 2017. He is currently an Associate Professor with the Department of Computer Science, American International University-Bangladesh (AIUB). He is also the Founder and the Director of the Advanced Machine Intelligence Research (AMIR) Laboratory. He was an Associate Professor and the Chairperson of the Department of Computer Science and Engineering, Bangladesh University of Business and Technology (BUBT), from 2019 to 2022, and a CSE Department Faculty Member with the University of Asia Pacific and a Graduate Head, from 2012 to 2019. His research experience, within both academia and industry, results in over 160 journals and conference publications. His research work contributed to the reputed journal of *Scientific Reports* (Nature), *Knowledge-Based Systems*, *Artificial Intelligence Review*, *Engineering Applications of Artificial Intelligence*, *Sensors*, *Cancers Biology*, *Applied Sciences*, and IEEE Access. His research interests include artificial intelligence (AI), machine learning, deep learning, and natural language processing (NLP). For more than ten years, he has been with the master's and undergraduate students as a Supervisor of their thesis work. He has served as a program committee member in several international conferences/workshops. He served as an Editorial Board Member of several journals, including *PLOS One* journal. He has served as a Reviewer of reputed journals, such as IEEE Transactions on Neural Networks, IEEE Access, *Artificial Intelligence Review*, *Knowledge Base System*, *Expert System*, *Bioinformatics*, Springer Nature, and MDPI. He is a Professional Member of ACM.

**MD. MOHSIN KABIR** received the Bachelor of Science degree in CSE from Bangladesh University of Business and Technology (BUBT), Bangladesh, in 2021. He is currently pursuing the joint master's degree in intelligent field robotics systems (IFRoS) with the University of Girona, Spain, and Eötvös Loránd University, Hungary, funded by Erasmus Mundus Scholarship. In addition to his studies, he holds a position as a Lecturer with the Department of Computer Science and Engineering, BUBT (Study-Leave), where he was a Research Assistant and a Researcher with the Advanced Machine Learning Laboratory. Also, he has had the privilege of collaborating with several prominent research labs around the globe, including the Computer Vision and Pattern Recognition Laboratory, University of Asia Pacific, Bangladesh, and the Database System Laboratory, The University of Aizu, Japan. With an extensive research background, he has authored over ten articles in high-impact journals, such as IEEE Access, *Sensors*, *Computer Systems Science and Engineering*, *Biology*, and *Mathematics*. In addition, he has contributed to the scientific community by publishing over ten conference papers and actively participating in well-established conferences, including IEEE HONET, ICCIT, ICIEV, icIVPR, DASA, BIM, and ICSCT. Moreover, some of his research work has been published as a chapter in a few esteemed books related to machine learning and AI. His research interests include artificial intelligence, machine learning, deep learning, computer vision, the IoT, and robotics.

**NOBUYOSHI ASAI** received the B.Eng. degree in information science, and the M.Eng. and Ph.D. degrees in engineering from the University of Tsukuba, Japan, in 1991, 1993, and 1996, respectively. He was with Wave Front Company Ltd., Japan, in 1996, and he was an Associate Professor and a Senior Associate Professor with the School of Computer Science and Engineering, The University of Aizu, Japan, in 2000 and 2010. He was also a Visiting Researcher with the National Institute of Environmental Study, Japan, in 2002 and 2003. His research interests include applied mathematics, numerical analysis, quantum algorithms, high performance computing, synesthesia, and signal processing on EEG or motion. He is a member of JSIAM and IPSJ. He served an Editor for the *Journal of IPSJ*. He serves as the Editor-in-Chief. He also served as the general chair for an IEEE international conference.

**JUNGPIL SHIN** (Senior Member, IEEE) received the B.Sc. degree in computer science and statistics and the M.Sc. degree in computer science from Pusan National University, South Korea, in 1990 and 1994, respectively, and the Ph.D. degree in computer science and communication engineering from Kyushu University, Japan, in 1999, under a scholarship from the Japanese government (MEXT). He was an Associate Professor, a Senior Associate Professor, and a Full Professor with the School of Computer Science and Engineering, The University of Aizu, Japan, in 1999, 2004, and 2019, respectively. His research interests include pattern recognition, image processing, computer vision, machine learning, human–computer interaction, nontouch interfaces, human gesture recognition, automatic control, Parkinson's disease diagnosis, ADHD diagnosis, user authentication, machine intelligence, bioinformatics, and handwriting analysis, recognition, and synthesis. He is a member of ACM, IEICE, IPSJ, KISS, and KIPS. He served as the program chair and a program committee member for numerous international conferences. He serves as an Editor for IEEE journals, Springer, Sage, Taylor & Francis, *Sensors and Electronics* (MDPI), and Tech Science. He serves as an Editorial Board Member for *Scientific Reports*. He serves as a reviewer for several major IEEE and SCI journals. He has coauthored more than 350 published papers for widely cited journals and conferences.