



The Role of Generative Adversarial Network in Medical Image Analysis: An In-depth Survey

MANAL ALAMIR and **MANAL ALGHAMDI**, Department of Computer Science, Umm Al-Qura University, Saudi Arabia

96

A generative adversarial network (GAN) is one of the most significant research directions in the field of artificial intelligence, and its superior data generation capability has garnered wide attention. In this article, we discuss the recent advancements in GANs, particularly in the medical field. First, the different medical imaging modalities and the principal theory of GANs were analyzed and summarized, after which, the evaluation metrics and training issues were determined. Third, the extension models of GANs were classified and introduced one-by-one. Fourth, the applications of GAN in medical images including cross-modality, augmentation, detection, classification, and reconstruction were illustrated. Finally, the problems we needed to resolve and future directions were discussed. The objective of this review is to provide a comprehensive overview of the GAN, simplify the GAN's basics, and present the most successful applications in different scenarios.

CCS Concepts: • Computing methodologies → Computer vision; Image segmentation; Reconstruction;
Additional Key Words and Phrases: Generative adversarial network, medical image analysis, computer vision

ACM Reference format:

Manal AlAmir and Manal AlGhamdi. 2022. The Role of Generative Adversarial Network in Medical Image Analysis: An In-depth Survey. *ACM Comput. Surv.* 55, 5, Article 96 (December 2022), 36 pages.

<https://doi.org/10.1145/3527849>

1 INTRODUCTION

Recent years have seen huge progress in **computer-aided diagnosis (CAD)** in medical imaging and diagnostic radiology, thanks to the advancement of deep learning systems [92, 191]. With the constant improvement of medical technology, a diversity of medical imaging means has appeared [197]. Medical imaging is critical in modern clinics to give guidance to the precise detection and treatment of different diseases. It plays an irreplaceable role in treatment planning, disease detection, and clinical monitoring. The most commonly used medical imaging methods include **magnetic resonance imaging (MRI)**, **computer tomography (CT)**, X-rays, and **ultrasound (US)**. These images enable a quantitative and qualitative evaluation of the symptoms at the lesion position. They are utilized in the human body parts such as heart, brain, liver, lung, chest, kidney, and so on. Medical image analysis is used to extract essential information and enhance the level of clinical diagnosis [50]. It usually relies on the radiologist's experience to recognize the image with naked eyes and identify the lesion region [197]. This way is exhausting for the human

Authors' address: M. AlAmir and M. AlGhamdi, Department of Computer Science, Umm Al-Qura University, Saudi Arabia; emails: manalalamir994@gmail.com, maalghamdi@uqu.edu.sa.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0360-0300/2022/12-ART96 \$15.00

<https://doi.org/10.1145/3527849>

eyes, leading to mistakes, and requires a lot of time and effort. To overcome these challenges, researchers have been investigating how to substitute people through computer technology, so as to promote the efficiency and precision of medical image analysis.

Artificial intelligence (AI) has resulted in a revolution of pattern recognition and image analysis and has a tremendous ability to be implemented for more effective and creative medical image computing in a broad range of medical fields [51]. **Deep learning (DL)** is a part of AI that has dominated many vision tasks such as detection, image-to-image translation, semantic segmentation, and so on. DL medical systems have produced much interest over the last few years and have been employed deeply in all fields of medicine, from drug identification to medical decision-making, significantly changing the way medicine is practiced [92, 177]. Usually, the algorithm's architecture is designed to simulate the layers of neurons in the human brain to treat and extract information, which helps to enable machines to learn without being explicitly programmed. The trained models can be used to diagnose diseases and recognize risk factors related to them as the classification of Alzheimer's, breast cancer detection, retinal disease detection, and so on. DL techniques are beneficial for a diversity of medical diagnostic purposes and have even exceeded human experts on some of those.

GANs and their extensions have opened many interesting ways to handle the known and hard medical image analysis problems including medical reconstruction, image resolution enhancement, segmentation, lesion detection, data simulation, or classification [90]. Their capability to generate realistic-looking images provides a chance that the chronic lack of annotated data in the medical field can be solved with the aid of these generative models. GAN is two neural networks that are trained simultaneously. The first network focuses on generating synthesized data to fool the other, and the other discriminates the sample and classifies it as real or fake [152]. The principle of GAN is inspired by the game theory wherein the **generator (G)** (Section 2.1.1) and **discriminator (D)** (Section 2.1.2) will integrate with each other to accomplish the Nash equilibrium in the training procedure. Figure 2 illustrates the overall architecture of GANs. The G input is a random noise vector (z) that usually follows the Gaussian distribution. The G then maps (z) noise to new data with a multidimensional vector to achieve a fake sample. The D's output is the probability that the received sample is fake or real.

There are many publications that employ GANs and their extensions coming out every month for different medical tasks. This motivates us to describe GANs briefly and provides a high-level glance of their usage and role in medical image analysis. This survey will present recent and relevant research on GAN mainly in medical imaging. To achieve this, we searched databases encompassing arXiv, **International Symposium on Biomedical Imaging (ISBI)**, computer-assisted intervention (MICCAI), IEEE, ScienceDirect, and CVPR. To focus on the recent publications, we specified the cut-off time of the search as January 1, 2021. We have adopted the systematic literature review approaches to identify the main and updated literature amongst GAN functionality and applications. Figure 1 represents descriptive statistics of these papers based on task and year.

The remainder of the article is structured as follows: We begin with a summarized introduction followed by an overall review of related surveys and medical image modalities. Section 2 describes GAN in general and its training issues and evaluation metrics. Section 3 explains the GAN's extensions one-by-one in detail. Section 4 illustrates the applications of the GAN in medical fields. Section 5 synopsizes the survey and discusses potential applications and recognizes open challenges followed by the Section 6 conclusion.

1.1 Related Survey

Many reviews have emerged to describe and offer a summary of the field of GANs, highlighting the main concepts, methods, algorithms, and different GANs applications. The related surveys are

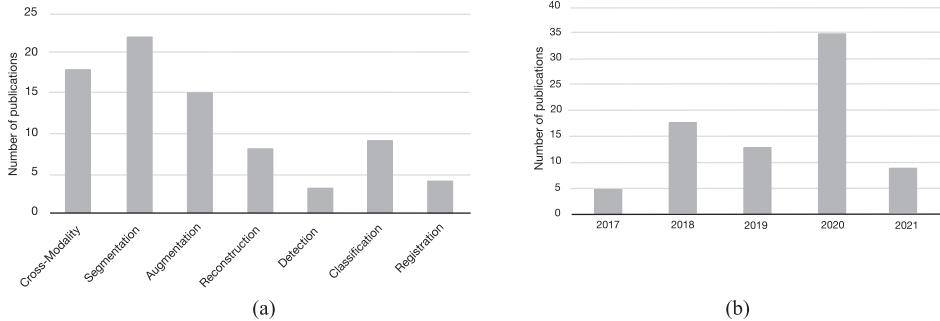


Fig. 1. Categorization of GANs' related papers: (a) Categorize the papers according to the performing task. (b) Categorize the papers according to published year.

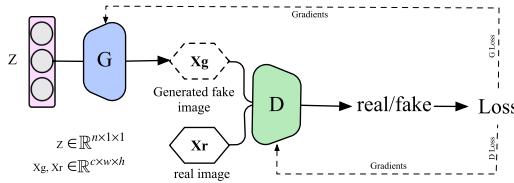


Fig. 2. Schematic view of the general structure of GAN, where it consists of two main part generator G and discriminator D. G generates fake images by using latent vector z . However, D tries to differentiate the generated and real images that work as binary classifiers. Both networks are trained through backpropagation.

summarized in Table 1. For the reviews that have been emerging in non-medical fields, Wang et al. [204] presented a descriptive overview of the GAN taxonomy in the area of computer vision, emphasizing the G and D architecture and the objective function variants. Wiatrak et al. [206] presented a summarized explanation for realizing GAN training stabilization methods and classified the issues in the training of GANs. Wang et al. [198] focused on the GAN's application in the image synthesizing field, i.e., image-to-image translation, image reconstruction, and data augmentation. Pan et al. [152] summarized different GAN derivatives and their application in many fields, such as natural language processing, image translation, super-resolution, and so on. Saxena et al. [173] focused on models that handled GANs problems, i.e., training instability by changing the loss function, the architecture, or the optimization method. Liu et al. [115] focused on applications and algorithms for generating realistic-looking images and videos. Gui et al. [60] provided a broad survey of different GAN models from the aspects of algorithms, theories, and applications, including the theoretical problems. Cao et al. [24] reviewed many GAN models and proved that LSGAN (detailed in Section 3.1) was able to generate high-quality images and also partially emphasizes the instability of the original GAN.

In the medical field, Yi et al. [219] provided a broad review of the recent methods in medical imaging analysis using the GAN with extensive evaluation results and categorization application. Kazeminia et al. [90] presented the use of GAN for medical image analysis includes synthesizing, segmentation, reconstruction, detection, and so on.

In this review, we will focus on the medical imaging analysis and the GAN's successful applications in improving the medical field. We will discuss the usage of the GAN's extension in the different medical tasks such as detection, classification, reconstruction, data augmentation, segmentation, and cross-modality. We will start with the standard architecture of the GAN, evaluation matrices, and training issues.

Table 1. Comparison with Other GAN Surveys That Were Displayed Using Following Points: Covering Training Issues, Loss Function, Evaluation Matrices, the Number of Mentioned GAN Types, and the Focusing Field

Publication	Year	Covering GAN training issue	Loss function	Evaluation Metrics	Number of mentioned GAN Types	Medical image	Remarks
Wanget al. [204]	2020	✓	✓	✓	29	✗	Focusing on the variations of GAN architectures and loss functions.
Wiatrak et al. [206]	2020	✓	✓	✗	N/A	✗	Focusing on GAN training stabilization methods.
Wang et al. [198]	2020	✓	✓	✓	37	✗	Focusing on the applications of GAN in image synthesis.
Pan et al. [152]	2019	✓	✗	✓	5	✗	Recent progress on GAN in different field.
Saxena et al. [173]	2020	✓	✓	✓	67	✗	Focusing on GANs problem and tried to handle.
Liu et al. [115]	2020	✗	✓	✓	N/A	✗	Focusing on the GAN's application in image and video synthesis.
Gui et al. [60]	2020	✓	✓	✓	21	✗	Discussing GANs model from the aspects of theories.
Cao et al. [24]	2018	✗	✓	✓	23	✗	Concentrating on GAN in the computer vision field.
Yi et al. [219]	2019	✓	✓	✓	12	✓	Focusing on medical image analysis using GAN.
Kazemina et al. [90]	2020	✓	✓	✗	7	✓	Concentrating on the usage of GAN's models in different medical applications.
This survey	2021	✓	✓	✓	14	✓	A broad survey about the medical image analysis using GAN, including a brief explanation about its architecture and training problems.

Note that N/A refers to the survey not discussing any types of GAN in detail.

1.2 Imaging Modality

Medical Imaging is the process of generating images of the internal organs in the human body and creating visual representations for detecting and treating medical conditions [165]. Images are generated by employing physical phenomena such as radioactivity, electromagnetic radiation, sound, and nuclear magnetic resonance.

Sound waves and high-frequency magnetic are passed through the human body and bounce off behind the tissue creating visual representations. The commonly utilized imaging modalities are: CT, X-ray radiography, MRI, **positron emission tomography (PET)**, mammograms, **optical coherence tomography (OCT)**, and fundus imaging [233].

The represented information in the medical image reflects the status of a human body [203]. Each type of medical image has its own purposes, strengths, and limitations. Besides, it has various responses to organ tissue and human body structure and only reflects specific information about the human body. The CT image can discriminate tissues with more accuracy and with various densities and can identify the bone structures, e.g., blood vessels and bones, with high resolution, whereas MRI can show smooth tissues, e.g., abdomen, pancreas, liver, instead of bones. However, a PET image presents functional information correlated to metabolism. However, these images are often presented in pseudo color and they usually have a low resolution. Mammogram images are used for screening and diagnosis of the human breast by using low-energy X-ray images [36]. It is utilized for the detection of any abnormality and lesions in breast tissues.

Moreover, a diverse range of imaging modalities has emerged to capture the anatomic structure of the eye, e.g., OCT, and the fundus imaging technique. OCT is a non-invasive technique for cross-sectional volumetric imaging of tissues utilized to envision eye structures in ophthalmology [58, 127]. OCT is utilized for the detection of ocular diseases, e.g., macular edema, age-related macular degeneration, whereas the fundus image is a direct visual capture of the eye and includes the anatomic structures, i.e., **Optic Disc (OD)**, vasculature, macula regions, and lesions as red lesions, including hemorrhages, and microaneurysms [5]. The various medical imaging modalities are shown in Figure 3.

2 GENERATIVE ADVERSARIAL NETWORK

The generative models consist of the following three categories: **Variational Autoencoder (VAE)**, **Generative Adversarial Networks (GANs)**, and AutoRegressive Networks. GANs are a kind of

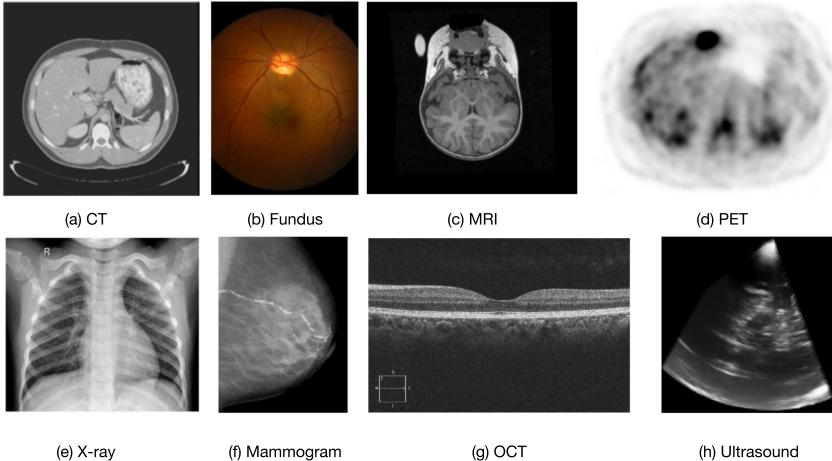


Fig. 3. Medical image modalities: (a) CT image for liver [214]. (b) Colored retinal fundus image [129]. (c) brain MRI [91]. (d) PET image for liver [15]. (e) Chest X-ray for normal case [66]. (f) Mammogram image with BAC (Breast arterial calcification) [6]. (g) Healthy retina OCT image[174]. (h) Ultrasound [134].

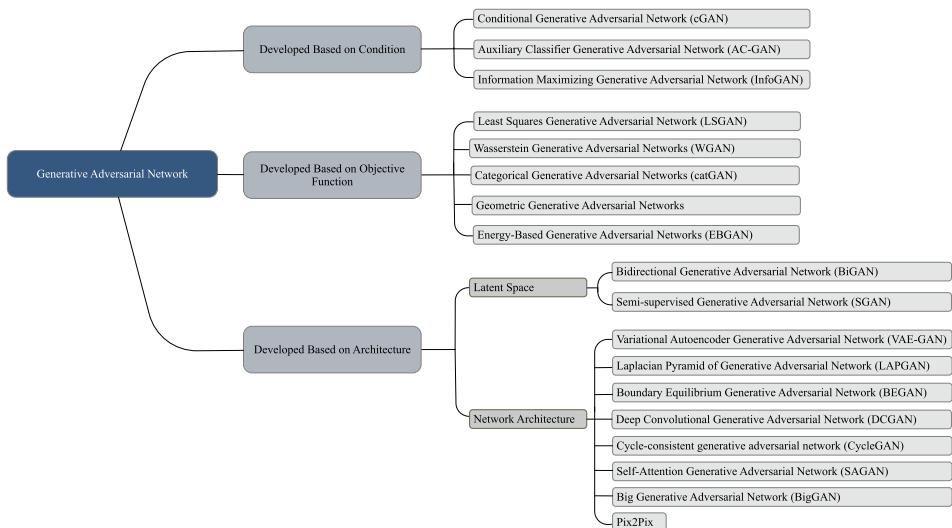


Fig. 4. GAN taxonomy. Note that this article does not describe some of the GANs mentioned here, as these are out of this article's scope.

generative model that have received significant attention due to their capability of modeling complex real-world data. Many different types of GAN architectures have been proposed for different application purposes. Figure 4 shows the splitting of different types of GANs into three categories, i.e., developed based on the condition, developed based on the objective function, and developed based on architecture. In this section, we will introduce the principle and architecture of GANs. The main idea of GANs is inspired by the mini-max two-person zero-sum game, as the two players are the G and the D.

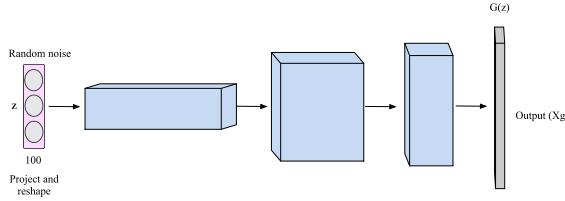


Fig. 5. A schematic view of generator architecture in vanilla GAN, illustrating the input, output, and internal feature representation. G maps the noise vector to fit the real data distribution.

2.1 Vanilla GAN

Goodfellow et al. [57] presented vanilla GAN, which is a generative model used for directly generating samples from specific data distributions without the need for Markov chains. GAN contains two networks: **generator (G)** and **discriminator (D)**, where the generator creates a fake sample and the discriminator tries to distinguish between fake and real samples. During the training, these two networks must be synchronized well, so the gradient information is backpropagated from discriminator to generator. The updating of the G and D parameters is performed iteratively. This means the D's parameters are fixed while training G and vice versa. The loss function for both G and D is presented as shown in Equation (1). Vanilla GAN trained on many datasets including MNIST [103] TFD (Toronto face database), and CIFAR-10.

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

Vanilla GAN has been successfully applied to many applications, e.g., image editing [235], feature learning [48, 170], and image generation [44, 163, 232].

2.1.1 The Generator(G). The G is a neural network that tries to generate fake samples similar to the real one from random variables $z \sim p_z$ using forward propagation only [24]. G receives a random noise, z , vector and transfers them to pseudo-sample distributions $G(z)$, using an upsampling process. This noise typically utilizes a uniform distribution or Gaussian noise that is a random variable or a random variable in latent space. The G learns to maximize the predicted log-probability for which the discriminator classifies the generated sample as real $\mathbb{E}_{z \sim p_z(z)} [\log(D(G(z)))]$. The loss is calculated from the discriminator's output, where it backpropagates to update the G's parameters. Figure 5 represents the general architecture of G. The training objective of G can be denoted mathematically as:

$$\mathcal{L}_G^{GAN} = \min_G \mathbb{E}_{x_g \sim P_g(x)} [\log(1 - D(x_g))]. \quad (2)$$

2.1.2 The Discriminator(D). The D aims to distinguish between the real and the generated sample and supports a feedback mechanism that updates weight parameters of G. The D's result is the probability that the received sample is a real sample. A higher probability means that the sample is more likely to be from real data. On the contrary, the lower probability and the closer to 0, the more likely the sample is fake. The optimal solution occurs when the probability approximates to $D = \frac{1}{2}$, which means that the discriminator failed to distinguish between the real sample and the fake one. The D learns to maximize the predicted log-probability $\mathbb{E}_{x \sim P_{data}(x)} [\log D(x)]$ while minimizing $\mathbb{E}_{z \sim p_z(z)} [\log(D(G(z)))]$ at the same time. Figure 6 represents the general architecture of D. The training objective of D can be denoted mathematically as:

$$\mathcal{L}_D^{GAN} = \max_D \mathbb{E}_{x_r \sim P_r(x)} [\log D(x_r)] + \mathbb{E}_{x_g \sim P_g(x)} [\log(1 - D(x_g))]. \quad (3)$$

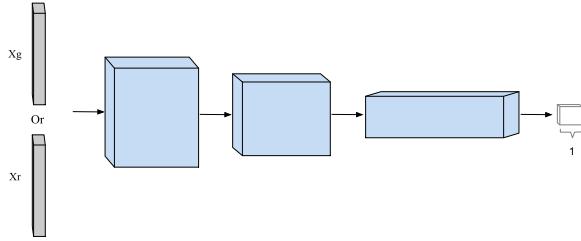


Fig. 6. A schematic view of discriminator architecture in vanilla GAN, illustrating the input, output, and internal feature representation. D tries to classify the received sample to real or fake.

2.2 Problems in GAN Training

Although there is proof of the presence of the unique equilibrium in the GAN game, the training dynamics of GANs have been recognized as being unstable [206], where they need to reach the Nash equilibrium during the training. However, that is not easy. The main causes of this instability are convergence, vanishing gradients, and mode collapse. These GAN issues remain important problems for GAN research [192]. In this section, we will discuss these three reasons for training instability.

2.2.1 Mode Collapse. The persistent GAN's mode collapse training problem occurs when the generator maps different distinct inputs to the same output [206]. This means that the generator creates samples of low diversity that are not considered to be useful for training purposes. Accordingly, G learns to create a sample from just a few modes of the data distribution and ignores the other modes even if it is trained on them [60]. This problem is not straightforward, since the reason for mode collapse is established deeply within the concept of GANs [206]. Addressing mode collapse can be obtained by creating a discriminator with high generalization capabilities. A crucial aspect in addressing mode collapse is a well-trained discriminator able to robustly determine the symptoms of mode collapse and with high generalization capabilities. The first method for handling the problem is based on the objective function, since **deep regret analytic GAN (DRAGAN)** [95] argues that the mode collapse occurs because of the presence of a fake local Nash equilibrium in the nonconvex problem. DRAGAN handles this problem by restraining gradients of the discriminator around the real data space. By adding a gradient penalizing term, the discriminator is biased to have a gradient norm of 1 around the real data space [60]. The second method is based on structure, where PACGAN [113] mitigates the mode collapse by changing the discriminator's input.

2.2.2 Convergence. Despite the fact that the existence of the equilibrium has been proved [57], in practice, reaching this equilibrium is not easy [206]. G and D are actually neural networks. Hence, the optimization process runs in the network's parameters, rather than learning the probability density function directly. Besides, the game is characterized as a non-convex-concave, which makes it particularly difficult for the **gradient descent-ascent (GDA)** algorithm to converge, usually causing diverging, cyclic behavior, or oscillating. SN-GAN [137] suggested a weight normalization technique, called spectral normalization, to delay the convergence of the discriminator. Unrolled GAN [133] attempted to change the optimization process to handle the mode collapse and convergence. There is still a lack of convergence behaviors in training GANs despite many promising practical applications [55].

2.2.3 Vanishing Gradients. When the D is close to optimality, it would provide significant feedback to the generator, thus leading to an improvement in the generator [206]. A highly precise D,

where $D(x) = 1$ and $D(G(z)) = 0$ reduces the loss function to 0, resulting in gradients approaching zero, which gives a little update to the G's parameters. As mentioned in Reference [144], minimization of minimax GAN's objective function leads to vanishing gradient issues, such as where the distributions of source and target are not correctly aligned, the D will approach the optimal level, which leads to the gradient for the GAN's objective function to be almost zero. This results in little feedback to G. It happens when the discriminator can recognize extremely well the difference between real and fake samples before the generator can approximate the distribution of data. Goodfellow et al. [57] suggest utilizing an alternative loss for the generator – $-\log D(G(z))$, which is known as **non-saturating loss (NS)**. Although the vanishing gradient problem has been alleviated, changes in the loss have not completely solved it.

2.3 Evaluation Metrics

While the selection of the suitable model is significant for getting high performance for an application, the selection of a suitable evaluation metric is also significant for getting high accuracy results. This is needed to beat the qualitative measure limitations for designing a good GANs model by improving or utilizing appropriate quantitative metrics. Lately, GANs have different applications, and each one has applied its own evaluation metric. Nevertheless, there are no standard evaluation metrics, which will result in an ambiguity about how researchers can specify evaluation metrics for various tasks. In this section, we will discuss the most commonly used evaluation metrics.

2.3.1 Inception Score (IS). IS is a broadly used evaluation metric proposed by Salimans et al. [170]. A high IS result means that the generative model can produce high-quality and diverse samples where it measures the diversity and quality of the produced images using the external model. IS utilizes pre-trained InceptionNet on ImageNet to capture the properties of produced samples. However, it has many limitations [21]: (i) IS fails to identify mode collapse problem and provides a contrary result. (ii) The IS utilizes InceptionNet, which is pre-trained on ImageNet, and it may prefer the models that generate a good sample instead of a realistic one. (iii) The image resolution impacts the IS result.

2.3.2 Mode Score (MS). MS is based on IS, introduced by Che et al. [30], where it addresses the main disadvantage of IS, which is discarding the prior distribution of a real dataset sample. MS represents the visual quality and variety of the produced samples simultaneously. The MS equation is shown in (4):

$$\exp(\mathbb{E}_x[\text{KL}(p(y|x)||p(y^{train}))] - \text{KL}(p(y)||p(y^{train}))), \quad (4)$$

where $p(y^{train})$ is the empirical distribution of labels calculated from training data.

2.3.3 Fréchet Inception Distance (FID). FID was introduced by Heusel et al. [67] and was utilized to identify the intra-class mode dropping. The produced samples are embedded into the feature space given by a particular layer of the InceptionNet or any **convolutional neural network (CNN)**. FID relies on the supposition that the produced samples follow a multidimensional Gaussian, the covariance and mean are estimated between the produced samples and real data. Following that, the Fréchet distance among these two Gaussians (Wasserstein-2 distance) computes to estimate the quality of produced samples. Nevertheless, the FID and IS cannot handle the overfitting issue well. Bińkowski et al. [20] handled this problem by proposing **Kernel Inception Distance (KID)**. The FID equation is illustrated in Equation (5), where (M_r, Σ_r) and (M_g, Σ_g) are the covariance and mean of the real and generated data, respectively. Low FID refers to the smaller distance

between produced and real data distribution.

$$FID(r, g) = \|M_r - M_g\|_2^2 + T_r \left(\sum_r + \sum_g -2 \left(\sum_r \sum_g \right)^{\frac{1}{2}} \right) \quad (5)$$

3 TYPES OF GENERATIVE ADVERSARIAL NETWORKS

There are many kinds of GAN architectures proposed for different application purposes, such as facial manipulation, image-to-image translation, image super-resolution, and so on. The types of GAN vary in three aspects including (i) differing architecture, (ii) differing objective of the discriminator, and (iii) differing objective of the generator. These extensions follow the overall concept of GANs with improving the performance and image generation. In this survey, we focused on the types that were successfully employed in the medical field and which are utilized as inspiring models for different applications. A schematic view of GAN types is presented in Figure 10 with the activation function in the D and the expected output of each network.

3.1 Least Squares Generative Adversarial Network (LSGAN)

The main idea of LSGAN was to replace the cross-entropy loss used in the discriminator with the least-squares loss function proposed by Mao et al. [128]. The regular GAN considers the discriminator as a classifier with the sigmoid cross-entropy loss function, where cross-entropy loss leads to a vanishing gradient. When the fake sample is classified as a real image, it will cause no error because it is on the correct side from the decision boundary, although, it is still far from the real samples. The least-squares loss function penalizes the data samples that are far away on the correct side of the decision boundary to create samples that are closer to real data. LSGAN is capable of generating a better-quality image than vanilla GAN and is more stable during the training process. However, it suffers from mode collapse, which affects the diversity and quality of images generated [29]. The performance of LSGAN was evaluated on HWDB1.0, LSUN [222], and CIFAR-10. LSGAN is successfully used in References [74, 237] for stable training. The G and D objective functions shown in Equations (6) and (7), respectively.

$$\min_G \mathcal{L}_G = \frac{1}{2} \mathbb{E}_{z \sim P_z} [(D(G(z)) - c)^2], \quad (6)$$

$$\min_D \mathcal{L}_D = \frac{1}{2} \mathbb{E}_{x \sim P_r} [(D(x) - b)^2] + \frac{1}{2} \mathbb{E}_{z \sim P(z)} [(D(G(z)) - a)^2], \quad (7)$$

where a is the label for the produced data, b is the label for the real data, and c is the hyperparameter that G wants D to classify the produced data as the real one by mistake. The overall architecture of LSGAN is similar to the vanilla GAN in Figure 10.

3.2 Deep Convolutional Generative Adversarial Network (DCGAN)

The first piece of work that used a convolutional decoder neural network for a generator was presented by Radford et al. [163]. This approach has certain constraints on the architecture to make it the best model for unsupervised learning. DCGAN employs the spatial upsampling capability of a convolutional decoder operation for G, which allows the generation of higher image resolution using GAN. Some modifications have been applied to the architecture of DCGAN, resulting in stable training and producing a high resolution. At first, DCGAN replaces any pooling layers with fractional-strided convolutions for the generator and strided convolutions for the discriminator. Second, batch-normalization is used in both G and D, which helps to set the produced samples and the real one centering at zero. Third, DCGAN eliminates fully connected hidden layers for a deeper architecture. Finally, the ReLU activation function is utilized in the G for all of the layers except

for the output layer where Tanh is used, and LaekyReLU is used for all of D's layers. DCGAN generates visual quality images quantitatively superior to other GANs proposed later. However, mode collapse is a major DCGAN weakness [29]. DCGAN trained on LSUN [222] and ImageNet. DCGAN has been used for data augmentation as in Reference [53], and also classification as in References [108, 121]. The overall architecture of DCGAN is similar to the vanilla GAN in Figure 10.

3.3 Conditional Generative Adversarial Network (cGAN)

cGAN is a conditional model that allows for adding extra information to control the generated data in the generative model such as class labels or data from modalities proposed by Mirza et al. [135]. cGAN is an extension of GAN where the generator and discriminator receive an extra vector of information (c) as input. This approach shows better representations of multi-modal data generation. The additional information is usually encoded as a one-hot vector inside the discriminator and generator before being concatenated with an encoded z (noise vector) in G and x in D . However, the discriminator can not output the class of the input data. Usually, cGAN demands pairs of input and output images for training, which is not always available in the domain adaptation problems. cGAN conditions the image generation on set of attributes [155, 212], images [76, 201], texts [166], and class labels [76]. cGAN can easily handle with unimodal, i.e., MNIST [103] and multi-modal, i.e., Flickr dataset. The loss function of CGAN as seen in Equation (8). Figure 10(b) shows the architecture of cGAN framework.

$$\min_G \max_D \mathbb{E}_{x \sim P_x} [\log[D(x|c)] + \mathbb{E}_{z \sim P_z} \log(1 - D(G(z|c)))] \quad (8)$$

cGAN is adapted for image-to-image translation [76, 82, 89, 171], synthesizing photo-realistic [227], data augmentation [139, 141], and segmentation [179].

3.4 Information Maximizing Generative Adversarial Network (InfoGAN)

InfoGAN is an information-theoretic extension to the GAN introduced beyond the cGAN [135]. InfoGAN, proposed by Chen et al. [31], maximizes the mutual information between conditional variables and the generative data to learn meaningful and interpretable representations in an unsupervised manner. They decompose the noise vector into two parts: incompressible noise z and latent code c . They present a new classifier named Q that tries to estimate the c value given by $Q(c|x)$. Q and D share whole convolution layers to save computational cost except the fully connected layer. The loss of InfoGAN is a regularization of cGAN's loss as shown in Equation (9), where $I(\cdot)$ is the mutual information. InfoGAN architecture is used in References [71, 223]. Figure 10(c) shows the architecture of the InfoGAN framework.

$$\min_G \max_D V_I(D, G) = V(D, G) - \lambda I(c; G(z, c)), \quad (9)$$

where $V(D, G)$ is the objective of cGAN except that the discriminator does not take c as input. One limitation of InfoGAN is that the category of the latent variable is uncontrollable.

3.5 Self-Attention Generative Adversarial Network (SAGAN)

Traditional convolutional GAN has many difficulties in learning some image classes over other classes. Moreover, it captures only spatial information. Although the state-of-the-art GAN model outperforms at a generated image with a few constraints (i.e., sky or ocean), it fails to capture an image with a geometric pattern. Proposed by Zhang et al. [226], SAGAN employs a self-attention technique into a convolutional GAN that helps to guarantee a large receptive field without harming the computational efficiency and uses global, long-range dependencies for synthesizing images. Self-attention computes a response at a region as a weighted sum of the feature at all other regions where the attention vector (weight) is computed with a less computational cost. They use spectral

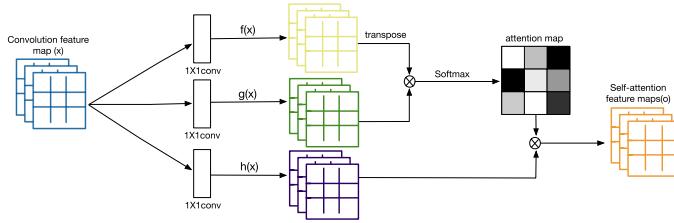


Fig. 7. The self-attention mechanism architecture f , g , and h letters are related to the query, key, and value, respectively, and \otimes is matrix multiplication. The attention map shows the long-range spatial dependencies.

normalization to the weights in both G and D , which helps to stabilize the training. The generator creates images with fine details at each location and the discrimination imposes complex geometric constraints on the image structure. SAGAN outperformed other works in the image generation. It has been used to improve human pose performance [200] and super-resolution [110]. The self-attention mechanism architecture is shown in Figure 7. This model minimizes the hinge version of adversarial Loss as shown in the equations below:

$$\mathcal{L}_D = -\mathbb{E}_{(x, y) \sim P_{data}} [\min(0, -1 + D(x, y))] - \mathbb{E}_{z \sim P_z, y \sim P_{data}} [\min(0, -1 - D(G(z), y))], \quad (10)$$

$$\mathcal{L}_G = -\mathbb{E}_{z \sim P_z, y \sim P_{data}} D(G(z), y). \quad (11)$$

3.6 Semi-supervised Generative Adversarial Network (SGAN)

Semi-supervised learning relies on both unlabeled and labeled samples to train the model. SGAN was proposed by Odena [149]. The SGAN's discriminator is considered to be multi-headed, adopting softmax for classification and sigmoid discriminative. They trained the G and D on the datasets with N classes. The discriminator predicts that the input belonged to the specific $N+1$ class where the additional class is the Fake class. However, this uncomplicated architecture of the multi-headed discriminator contributes to limiting the diversity of the generated images. They trained SGAN on the MNIST [103] dataset in the same way as GAN. SGAN is used in Reference [176]. Figure 10(e) shows the architecture of the SGAN.

3.7 Cycle-consistent Generative Adversarial Network (CycleGAN)

CycleGAN architecture allows transferring the image from one domain to another without requiring a paired image dataset, which was introduced by Zhu et al. [236]. CycleGAN utilizes two generators, G_1 and G_2 and two discriminators, D_1 and D_2 . The generator G_1 learns to map from the A domain to B and the inverse for generator G_2 . The discriminators measure how the generated images are real and match the distribution of the target domain ($G(X_a) \approx B$ or $G(X_b) \approx A$) by adversarial loss in Equation (12), and how the accuracy of the reconstructed image after two generator networks ($G_2(G_1(X_a)) \approx A$ or $G_1(G_2(X_b)) \approx B$) by cycle consistency loss in Equation (13). D_1 attempts to differentiate between images X_r and generated images $G(y)$, and in the same way, D_2 tries to distinguish between y and $G(x)$.

$$\mathcal{L}_{GAN}(G, D_1, A, B) = \mathbb{E}_{x_b \sim P_{data}(b)} [\log D_1(x_b)] + \mathbb{E}_{x_a \sim P_{data}(a)} [\log(1 - D_1(G_1(x_a)))] \quad (12)$$

$$\mathcal{L}_{cyc}(G_1, G_2) = \mathbb{E}_{x_a \sim P_{data}(a)} [\|G_2(G_1(x_a)) - x_a\|_1] + \mathbb{E}_{x_b \sim P_{data}(b)} [\|G_1(G_2(x_b)) - x_b\|_1] \quad (13)$$

The overall objective is illustrated in Equation (14):

$$\mathcal{L}(G_1, G_2, D_1, D_2) = \mathcal{L}_{GAN}(G_1, D_1, A, B) + \mathcal{L}_{GAN}(G_2, D_2, B, A) + \lambda \mathcal{L}_{cyc}(G_1, G_2). \quad (14)$$

CycleGAN is commonly used in medical images for cross-modality [68, 84, 213] and image reconstruction [3, 88, 221]. Figure 10(k) shows the architecture of CycleGAN framework. However, in some cases, the CycleGAN generates blurry images and decreases the contrast at boundaries [22].

3.8 Laplacian Pyramid of Generative Adversarial Network (LAPGAN)

LAPGAN was introduced by Denton et al. [44]. This approach aims to generate high-quality natural images in a coarse-to-fine fashion by combining the cGAN [135] with a Laplacian pyramid. At each level of the pyramid, a separate deep convolutional networks model was trained utilizing cGAN. LAPGAN utilizes a Laplacian pyramid to upsample the image. The first generator is used to generate a very small image that can help to reduce the instability issue for the generator, then upsample the generated image using the Laplacian pyramid. For producing the image difference, the upsampled image is fed to the next generator and the summation is applied for the image difference and upsampled image. The image difference produced by the generator is less complicated than the same size raw image. LAPGAN has been utilized in Reference [14]. Figure 10(f) shows the architecture of the LAPGAN framework. The main drawback of LAPGAN is that the network consists of sub-networks that are trained independently and have no weight-sharing among them. Thus, the capacity of the network is restricted by the depth of each sub-network.

3.9 Auxiliary Classifier Generative Adversarial Network (AC-GAN)

The auxiliary classifier GAN introduced by Odena et al. [150] contains an auxiliary classifier in the architecture same as InfoGAN [31] and cGAN [135]. In AC-GAN, the additional information only refers to the class label, contrary to the previous architecture of InfoGAN [31] and cGAN [135], where the previous two can be other data domains. The discriminator contains a classifier that classifies the sample into different categories, which contribute to stabilizing the training. Although the capability of the AC-GAN for producing high-quality images, it is found in the literature [138] that as the number of labels increases, the model tends to produce less diverse images for most classes. Miyato et al. [138] expected that this was caused by the auxiliary classifier. The model was trained on CIFAR-10 and ImageNet. AC-GAN was proposed for data augmentation [194], image synthesis [150], and so on. Figure 10(i) shows the architecture of AC-GAN framework.

3.10 Pix2pix

Pix2pix was proposed by Isola et al. [76] and is a successful variant of cGAN [135] for image-to-image translation. Pix2pix framework demands a pair of images, the input, and the desired output. The generator follows the U-net [168] architecture with a skip connection. The addition of skip connections is useful for the global coherence of the generated image. The discriminator uses a **patch-based fully convolutional network (PatchGAN)**. Pix2pix applies L1 distance loss between the generated image and the ground truth. PatchGAN penalizes structure at the scale of local image patches, meaning that it basically tries to differentiate if each $N \times N$ patch in the synthesis image is real or fake where the discriminator is run convolutionally across the image. It models images as a Markov random field and supposes the independence between pixels detached by more than a patch diameter. Pix2pix has been utilized for image reconstruction [116], cross-modality [35], and data augmentation [56]. Figure 10(l) shows the architecture of pix2pix framework. Pix2pix requires paired images to train models, which becomes hard for some applications such as object replacement.

3.11 Bidirectional Generative Adversarial Networks (BiGAN)

The learning of inverse mapping not being applied in vanilla GAN means projecting the data back to the latent space. **Bidirectional generative adversarial networks (BiGAN)** was proposed

by Donahue et al. [48] to capture this purpose. BiGAN architecture contains the following parts: **encoder** (G_e), **decoder** (G_d), and discriminator. The encoder maps real samples into latent representations z , while the decoder (G_d) decodes the z latent representations to x . The discriminator aims to distinguish between pairs (X_r, Z_r) and (Z_g, X_r) . The encoder and decoder parts does not communicate directly and must train to invert one another to fool the discriminator. The BiGAN module has been trained on MNIST and the ImageNet. BiGAN is used in a human motion prediction as in Reference [79]. Figure 10(d) shows the architecture of BiGAN framework. Although the BiGAN is able to reconstruct the original image from the latent space, the visual quality of the generated data is generally worse than DCGAN.

3.12 Boundary Equilibrium Generative Adversarial Networks (BEGAN)

BEGAN adopts the auto-encoder architecture for the discriminator, which was first proposed in Energy-based GAN by Zhao et al. [232]. BEGAN was proposed by Berthelot et al. [17]. The objective while training the auto-encoder is to maximize reconstruction loss for both the real and generated images. The optimization of reconstruction losses is equivalent to the Wasserstein distance. To control the balance between generator and discriminator losses, they proposed the utilize of hyperparameter $\gamma = \frac{\mathbb{E}[\mathcal{L}(G(z))]}{\mathbb{E}[\mathcal{L}(x)]}$, $\gamma \in [0,1]$, which helps to balance the effort assigned to both G and D , i.e., control the diversity of synthesized faces. Despite BEGAN's impressive performance, it has limitations in terms of visual quality and diversity due to its lack of reconstruction capability in the discriminator, the inherent structure of the generator, and instability during training [94]. BEGAN has been introduced for more stable training behavior. The overall loss function is summarized in Equation (15):

$$\begin{aligned}\mathcal{L}_D &= \mathcal{L}(x) - k_t \mathcal{L}(G(z_D)), && \text{for updating } \theta_D, \\ \mathcal{L}_G &= \mathcal{L}(G_{zG}), && \text{for updating } \theta_G, \\ K_{t+1} &= K_{t+1} + \lambda_k (\gamma \mathcal{L}(x) - \mathcal{L}(G(zG))), && \text{for each training iteration } t,\end{aligned}\quad (15)$$

where $\mathcal{L}(\cdot)$ refers to the auto-encoder reconstruction loss, $k_t \in [0,1]$ is used to control how much emphasis of $\mathcal{L}(G)$ is penalized for the loss where the k is initialized as 0 and is controlled by λ_k . λ_k can be considered as the learning rate for k . Figure 10(h) shows the architecture of the BEGAN framework.

3.13 Categorical Generative Adversarial Networks (CatGAN)

CatGAN, which was proposed by Springenberg [184], utilizes multiclass classification instead of the binary classification in the vanilla GAN in an unsupervised manner. The D is trained to differentiate between real and fake data in addition to classifying all data samples into a previously chosen number of categories. The D classifies the data into categories while being uncertain of category assignments for samples generated by G . Moreover, it demands that G synthesizes samples belonging to a certain class rather than generating samples belonging to the dataset. CatGAN is able to learn a classifier from unlabeled data, but it still has difficulty clustering input features. The objective function for generator and discriminator is shown in Equation (16). The objective of CatGAN is to maximize $H[p(c|x, D)]$ and $H[p(c|D)]$, and to minimize $H[p(c|G(z), D)]$. CatGAN is used for image classification in Reference [218].

$$\begin{aligned}\mathcal{L}_D &= \max_D H_X[p(c|D)] - \mathbb{E}_{x \sim X}[H[p(c|x, D)]] + \mathbb{E}_{z \sim P(z)}[H[p(c|G(z), D)]], \\ \mathcal{L}_G &= \min_G -H_G[P(c|D)] + \mathbb{E}_{z \sim P(z)}[H[p(c|G(z), D)]],\end{aligned}\quad (16)$$

where the X is the dataset distribution and $H[\cdot]$ represents the empirical entropy. Figure 10(j) shows the architecture of CatGAN framework.

3.14 Variational Autoencoder Generative Adversarial Network (VAE GAN)

To train a generative model with higher-quality samples, the researchers have aimed to bridge the gap between VAEs (**Variational Autoencoders**) and GANs. The combination of VAEs with GANs contributed to enhancing image generation. Larsen et al. [101] proposed VAE-GAN, which combines a variational auto-encoder objective with a GAN and uses the learned features from the GAN's discriminator for enhancing image similarity metrics. VAE-GAN utilizes feature-wise distance for the reconstruction. VAE is trained jointly with the GAN model and the pixel-wise loss function that is employed for training VAE is replaced with the feature-wise loss function learned from the GAN's discrimination. GAN creates more realistic samples than VAE, thanks to an adversarial regime that allows the network to learn more complex distributions [107]. However, GAN can learn only a unidirectional mapping for generating data and does not allow to infer latent codes from the given samples. Moreover, GAN suffers from the mode collapse problem, since many modes of the data distribution are not captured in the generated samples. Figure 10(g) shows the architecture of the VAE-GAN framework.

4 APPLICATION OF THE GAN

The GAN was first used in non-medical images after which researchers in medical fields were inspired by the promising results achieved in non-medical images and extended these models to be useful for medical images. GANs became widely used for real image processing because of their performance in synthesizing realistic photo images. Image processing includes many applications such as style transfer [72], super-resolution [118], image-to-image translation [76, 231, 236, 237], data augmentation [215], and face manipulation [65]. Furthermore, GAN has been used for medical image processing across many tasks such as segmentation, detection, image translation, and so on. Figure 8 provides examples of GAN performance in various applications. The following section will review the applications of the GAN in medical images. Figure 9 shows the distribution of papers among the different applications.

4.1 Cross-modality Synthesis

The image-to-image translation task can be considered as mapping the input tensor to output with a different appearance, yet of the same underlying structure [9]. Medical image translation is more challenging when compared to regular images because of the amount of associated medical information contained in detailed structures in the images, which can be distorted or lost during the translation processes. There are several basic physical principles, providing imaging data of various dimensionalities and of differing contrasts. The diversity offers several diagnostic choices, but is also considered a challenge when translating image information between various modalities or various acquisitions within one modality. Many researchers tried to estimate one modal image from another modality image, such as 2D X-ray-to-3D CT [220], and MRI-to-CT [68, 148, 151]. Different GAN types have been employed to perform image synthesis for inter- or intra-modality, including cGAN, CycleGAN, StarGAN, and InfoGAN. CycleGAN has been broadly utilized for cross-domain medical image synthesis purposes particularly because of its capability to deal with unpaired data.

Zhang et al. [231] proposed an approach that is able to learn to translate and segment 3D images simultaneously. For segmenting MRI and CT images, they trained two auxiliary CNN's, respectively, and defined the loss to constrain the segmentation of the created image. They utilized a shape consistency loss that is acquired from two segmentation networks where using only cycle loss is insufficient in the cross-modality to guarantee the geometric invariance of the synthesis image. The segmentation network slices image modality into semantic labels and applies implicit

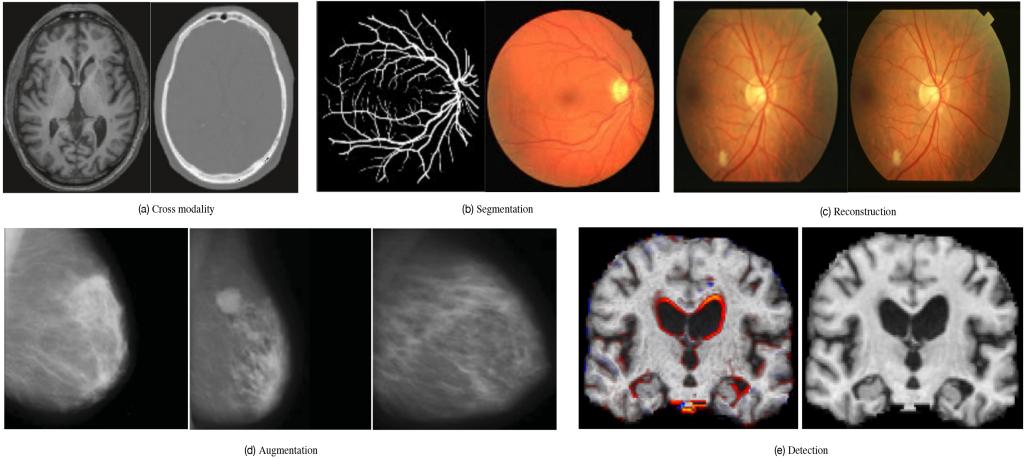


Fig. 8. Example applications employing GANs. Figures are taken from the corresponding papers as the following: (a) Cross-modality [148], (b) Segmentation [61], (c) Reconstruction [19], (d) Augmentation[46], and (e) Detection [13].

shape constraints on anatomy during translation. The generator used CycleGAN [236] with modifications, the discriminator used PatchGAN [76], and the segmentors utilized the U-net [168]. Yang et al. [213] introduced structure-constrained CycleGAN [236] for synthesizing brain MRI to CT utilizing unpaired data. They defined structure-consistency loss based on **MIND (modality independent neighborhood descriptor)** being used between the synthetic and input images. The MIND compares each patch with all its neighbors in a non-local region to extract the distinctive structure of the image. The structure-consistency loss was defined to constrain the extracted feature in the created image to be voxel-wise close to the extracted feature from the input image.

Armaniousa et al. [9] proposed the MedGAN approach for medical image translation, which is an adversarial framework merging cascaded U-net [168] in generator architecture (CasNet). The generating task is executed utilizing the collective capacity of the U blocks in an end-to-end manner. The discriminator is used as a feature extractor that penalizes the discrepancy between the generated image and the desired modality. MedGAN applies to the three tasks: PET image denoising, PET-CT translation, correction of MR motion artifacts. Ben-Cohen et al. [15] introduced an approach based on the combination of VGG-16 and cGAN [135] to produce simulated PET images from CT images in an unsupervised manner. The created PET can be utilized for a false-positive reduction in liver lesion detection solution. They used VGG-16 to produce initial PET-like images from CT input and then moved to cGAN [135] with the real CT image to improve the VGG-16's output. The U-net [168] architecture was utilized as generator. Studies related to cross-modality synthesis are summarized in Table 2.

4.2 Segmentation

Medical image segmentation is an important task that aims to recognize the exact object boundaries as organs and lesion regions (e.g., tumor) [167]. Automating segmentation in the medical image is a challenging task as a result of the high variability in the appearance of organ tissues among different people and the high similarities between the healthy and non-healthy tissues.

The **deep transfer learning (DTL)** approach is based on fine-tuning models trained on different datasets or tasks before substituting the last layer of the trained model with a randomly initialized layer that fits the purpose of the new task and then optimizing all network weights. DTL consists of various layers of pooling and the convolution layer that performs feature extraction

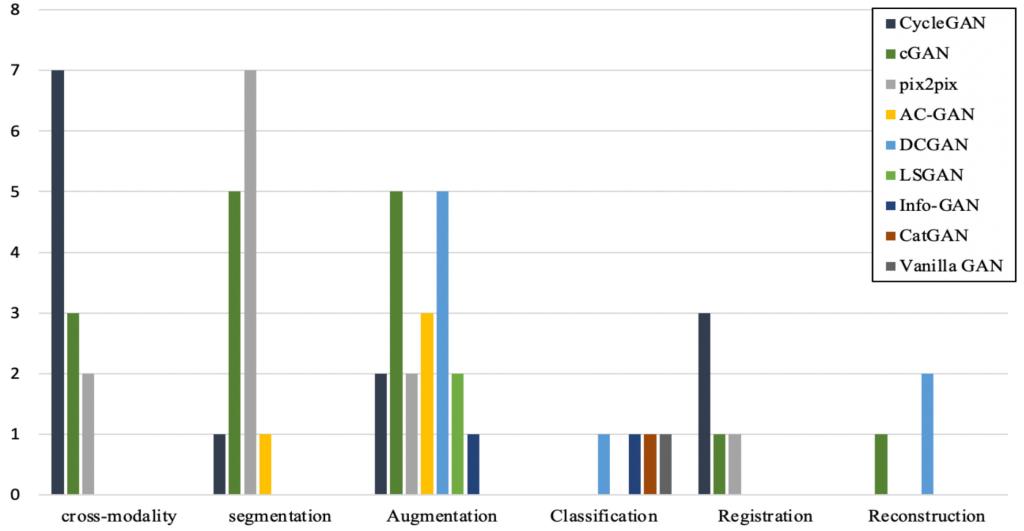


Fig. 9. The chart represents the distribution of papers among the different applications. Note that some uncounted papers adopt a modified GAN structure where the modifications are applied to the original architecture.

Table 2. Summary of Cross-modality in Medical Image Applications for Both Inter- and Intra-modality

Publication	Approach	Remarks	Dataset	Result
Inter-modality				
MRI → CT				
Oublacha and Kadoury [151]	CycleGAN*	[lumbar spine] Unsupervised approach for the construction of CT images using pseudo-3D CycleGAN.	SpineWeb	dice score = 0.83
Koike et al. [96]	cGAN*	[brain] Synthesizing CT images from multi-sequence MRI and evaluating images utilizing voxel-wise mean absolute errors of the CT.	TCIA [37]	MAE = 38.9 ± 10.7
Nie et al. [148]	Cascade GAN	[brain, pelvic] A patch-based GAN, using CNN as a generator and proposed an auto-context model for image refinement.	Disease Neuroimaging Initiative (ADNI)	MAE = 0.92
Hiasa et al. [68]	CycleGAN*	[musculoskeletal] Enhancing the accuracy at the boundaries by incorporating gradient consistency loss in training.	N/A	Dice score = 0.8
Lie et al. [106]	CycleGAN*	[brain, pelvic] Dense block-based networks utilized to build generator of cycle GAN.	N/A	MAE = 0.009
Tie et al. [190]	GAN*	[nasopharyngeal] Using three MRI images with different contrast in a multi-path multi-channel structure.	N/A	Dice score = 0.86
Qi et al. [158]	cGAN*	[head,neck] Utilizing a four-channel input that encompasses T1w, T2w, contrast-enhanced T1w Dixon water, and contrast-enhanced T1w images.	N/A	SSIM = 0.84
CT → MRI				
Dong et al. [49]	CycleGAN*	[pelvic] CycleGAN was utilized to estimate synthetic MRI from CT images to segment MRI later using deep attention U-net.	N/A	Dice score = 0.89
Lei et al. [105]	CycleGAN*	[pelvic] Translating image to improve soft tissue segmentation.	N/A	Dice score = 0.91
MRI ↔ CT				
Zhang et al. [231]	CycleGAN*	[3D heart] Incorporating CycleGAN with shape consistency loss.	N/A	Dice = 0.74
PET → CT				
Armaniousa et al. [9]	cGAN*	[brain] Incorporating a cascaded U-net generator architecture with non-adversarial losses.	N/A	SSIM = 0.91
CT → PET				
Ben-Cohen et al. [15]	cGAN*	[liver] Combination of VGG-16 and cGAN.	N/A	MAE = 0.63
Intra-modality				
T1 ↔ T2 MRI				
Yang et al. [216]	cGAN	[brain] Translating between three modalities T1 (spin-lattice relaxation), T2 (spin-spin relaxation), and T2-Flair (fluid attenuation inversion recovery).	BraTS [132], Iseg2017 [199], MRBrain13 [131]	SSIM = 0.910, 0.902, 0.863

Note that the * symbol refers to some modifications applied on the original architecture or losses. MAE refers to mean absolute error. N/A means not public dataset.

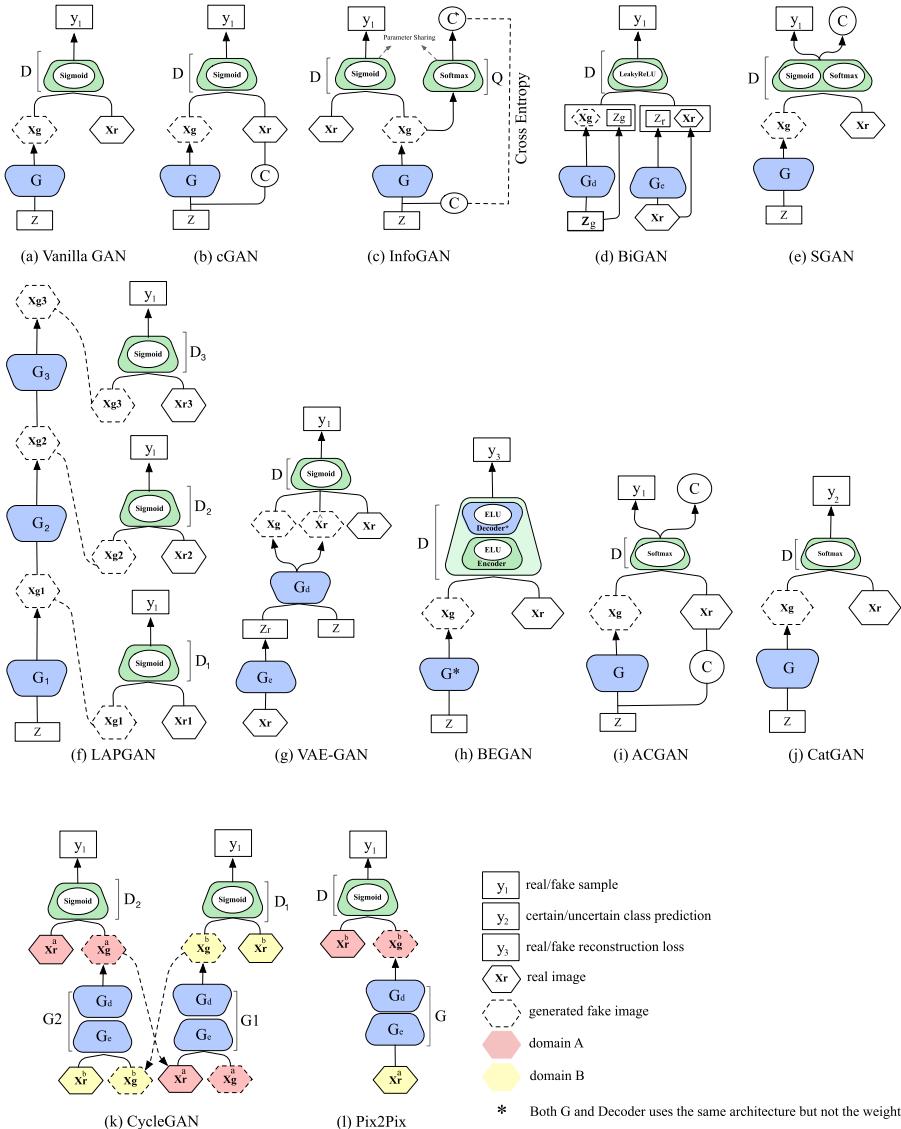


Fig. 10. A schematic view of GANs models architecture. c is conditional vector, where in cGAN and AC-GAN represents a discrete categorical code that encodes class label. However, in the case of Infogan it can also be a continuous code that encodes attributes. Pix2pix demands aligned training data. In turn, this constraint is not required in CycleGAN, yet commonly suffers from performance loss. G_e related to encoder and G_d to decoder. BiGAN follows the architecture used by DCGAN. The activation function is according to the original paper and relevant github page. Inspired by Yi et al. [219].

from images and more complicated features in deeper layers. Majurski et al. [126] proposed an approach based on transfer learning that combines the GAN and U-net [168] to enhance the accuracy of the cell boundary detection. The GAN trained to learn the abstract representation of the data in an unsupervised manner from unannotated data. After that, Majurski transferred the discriminator optimized weights to the encoder part in U-net. The U-net was trained on

manually annotated images. The architecture of the GAN utilized the encoder-decoder structural elements from U-net where the U-net encoder was utilized as the discriminator and decoder as the generator. Li et al. [112] proposed an approach of combining the AC-GAN [150] and pix2pix [76] models to solve the overfitting problem in deep learning networks. The generator architecture consisted of RU-Net, which is a combination of ResNet and U-net [168]. RU-Net contains seven **residual modules (RM)** and seven deconvolutional layers. It replaces the U-net convolutional connection with RMs, which helps to reduce gradient vanishing. This model used three losses: L1 loss, GAN loss, and Softmax loss. Xue et al. [211] proposed using multiscale L1 loss to enforce G and D to learn the hierarchical features that capture spatial pixel-dependencies. Moreover, multiscale L1 loss was used to constrain the segmentation map by taking into account the difference between the ground truth and predicted mask at multiple layers. The generator tried to maximize the multiscale L1, and the discriminator tried to minimize it. The U-net [168] has been used in the generator and the discrimination had a similar structure to the decoder.

Rezaei et al. [167] proposed RNN-GAN for semantic segmentation using a complementary segmentation mask. To alleviate the imbalanced pixel label, they mixed the adversarial with a weighted categorical loss and a complementary label. The network was trained with both the ground truth and a complementary mask. For the generator, they used a modified U-net [168], where it trained on a sequence of medical images to produce the segmentation map and the complementary label at the pixel level. The discriminator used a fully convolutional encoder. Both the generator and discriminator were replaced with bi-directional LSTM units to improve temporal consistency. Singh et al. [178] presented an approach for segmenting breast tumors in ultrasound images based on contextual-information-aware cGAN [135]. The framework consisted of a generator used to extract the tumor-relevant features and a discriminator to differentiate whether the predicted mask is real or just synthesized segmentation of the input image. To reinforce the tumor-relevant features, they utilized an **atrous convolution (AC)** block in the generator, which ensured that the high-level feature of the small size contains features from the breast tumor region from expanding the field-of-view of the filters. In addition, AC helps to alleviate the issue of loss of small tumor-relevant features because of the usage of consecutive downsampling layers. Moreover, they added a channel weighting and a channel attention mechanism in the generator to improve the segmentation accuracy and enhance the tumor-relevant features in an unsupervised manner. Han et al. [64] proposed a semi-supervised approach that applied a modified DeepLap for a generator that was trained on annotated data, and then the unannotated was passed to the pre-trained network to extract multiscale features. The discriminator integrates a dual-attention-fusion (DAF) block to differentiate the poor mask from good ones. DAF aims to obtain representative features from the tumor region and background separately, with two separate spatial attention paths designed to improve the discriminative ability. Moreover, they employed an **atrous spatial pyramid pooling (ASPP)** block to capture the multiscale features. Several other architecture-based improvements for skin lesion segmentation have also been proposed using GAN [172, 210]. Lei et al. [104] used a dual discriminator to differentiate between the ground truth images and the segmented one to increase the segmentation accuracy of the generator. One discriminator combined the real image and the mask as input to focus on spatial information between two images regardless of the boundary. The second discriminator extracted the boundary information utilizing only the masks as input. Adversarial losses were used in both discriminators to capture the different information. The U-net-SCDC module was used as a generator, which is the combination between the skip connection and dense convolution U-net. Moreover, GAN had also been used for lung segmentation, as seen in References [145, 154, 188]. Pawar et al. [154] used cGAN for lung segmentation where the generator consisted of the following three blocks: encoder, multiscale dense feature extraction, and decoder. The encoder was used to encode the CT image to feature using a downsample and

Table 3. Summary of Segmentation in Medical Image Applications

Publication	Remarks	Dataset	Result
Majurski et al. [126]	[cell] GAN-based transfer learning for a U-net segmentation.	Warwick-QU [180]	Dice score = 0.63
Li et al. [112]	[cell] Combining pix2pix and AC-GAN models by attaching a classification branch to the pix2pix's discriminator.	I3A [69], MIVIA [52]	Accuracy = 0.75
Aida et al. [4]	[cell] Employing cGAN to segment cancer stem cells.	N/A	Recall = 0.80
Hu et al. [73]	[cell] Using GAN to identify the cell centroid and produce a likelihood map.	MoNuSeg [99]	Dice score = 0.664
Xue et al. [211]	[brain] Using multiscale L1 loss to enforce G and D to learn the hierarchical features.	BRATS	Dice score = 0.85
Nema et al. [147]	[brain] An unpaired 2D GAN-based training is applied by a RescueNet model to address the limited annotated dataset.	BraTS [132]	Dice score = 0.94
Singh et al. [178]	[breast] Adding channel attention and channel weighting mechanisms in the G network to promote the tumor-relevant features.	Breast ultrasound [75]	Dice score = 0.9376
Han et al. [64]	[breast] Using DeepLap in the segmentation network.	Breast ultrasound [217]	Dice score = 0.79
Lei et al. [104]	[skin] Using dense convolution and skip connection U-net for generator and dual discrimination module.	ISIC	Accuracy = 0.92
Sarker et al. [172]	[skin] Integrating 1-D kernel factorized networks, position, and channel attention modules with a GAN to boost skin lesion feature.	ISBI, ISIC [38]	Dice score = 0.90
Xue et al. [210]	[skin] Combining CNN with residual blocks/skip connections and adaptive logistic activation function as G.	ISIC [38]	Dice score = 0.86
izadi et al. [78]	[skin] Refining the boundary accuracy by adversarial training.	DermoFit [12]	Dice score = 0.898
Munawar et al. [145]	[lung] Utilizing cGAN for enhancing lung segmentation.	JSRT, Montgomery Shenzhen	Dice score = 0.97
Tan et al. [188]	[lung] Incorporating GAN with Earth Mover distance-based loss function.	LIDC-IDRI [10]	IOU = 0.9018
Pawar and Talbar [154]	[lung] Using (MSDDE) module in cGAN generator.	ILD [45]	Dice score = 0.9899
Decourt and Duong [42]	[heart] Penalizing incorrect predictions by using loss function relying on distance transform and pixel-wise cross-entropy.	ACDC [16]	Dice score = 0.89
Zhang et al. [225]	[heart] Utilizing the conditional convolution generative adversarial network, which receives the extracted ROI as input.	ACDC [16]	Dice score = 0.96
Guo et al. [61]	[eye] Employing Dense U-net architecture utilizing Inception module for retinal vessel segmentation.	DRIVE [185]	Dice score = 0.82
Son et al. [181]	[eye] Deep structure leads to less false positives with fine vessels and much better for distinguishing whole images.	DRIVE [185], STARE [70]	Dice score = 0.82, 0.83
Son et al. [182]	[eye] Several models have been explored for the discriminator, including pix-GAN, image-GAN, and patch-GAN.	DRIVE [185], STARE [70]	Dice score = 0.8275, 0.8378

The table categorized according to the image of the organs used in the research. N/A means not public dataset.

the decoder upsampled the feature. The multiscale dense feature extraction block contained four inception blocks linked serially with dense connections to extract multiscale features and reutilize the learned features of the first inception block. They used six **multiscale dense feature extraction (MSDDE)** blocks aimed at identifying lungs with different shapes, sizes, and textures. In addition, GAN was also implemented for segmentation of the heart [42, 225]. Decourt and Duong [42] proposed a method for segmentation of the left ventricle in a semi-supervised manner. They suggested using a new loss function based on pixel-wise cross-entropy and distance transform to improve the accuracy of boundary pixels. This loss was used in a pre-trained network by penalizing the wrong predictions during the training on the unlabeled data. Modified DeepLap-v2 was used for the generator to produce the binary mask. The discriminator's output is a confidence map utilized as a signal to guide the cross-entropy loss.

The diagnosis of some diseases is sometimes founded on retinal blood vessels that contain significant information related to human health. Automatic segmentation of retinal vessels is an important and challenging task. Many automatic segmentation algorithms for retinal vessels have been proposed using GAN [61, 181, 182]. Guo et al. [61] introduced a retinal blood vessel segmentation approach by combining GAN and Dense U-net using the Inception module. The author applied two modifications to the basic U-net [168] architecture and then used it as a generator: (i) replacing the skip connection with Dense blocks; (ii) applying the Inception module instead of the traditional convolution operation. Studies related to medical image segmentation are summarized in Table 3.

4.3 Augmentation

Data Augmentation is a straightforward technique to increase the training set based on producing synthetic images [59]. Deep learning algorithm performance relies on training data availability, which is a barrier to improvement in many fields such as medical image processing. However, some of the medical datasets contain more negative than positive samples [80]. To increase the number

of training samples, data augmentation is one of the most common ways of improving classification performance and balancing the dataset. It includes a classical image transformation such as rotation, zooming, cropping, and so on, as well as advanced methods based on deep learning such as style transfer and GAN. Typically, methods based on deep learning produce more realistic results that are also new to the human eye. GAN synthesizes dataset images and transforms them from one class to another in a realistic way [46, 54, 62, 97, 143, 157, 186, 194].

Qi et al. [157] augmented the data utilizing cycle-consistency GAN to improve the classifier performance. The proposed approach added attention-guided CycleGAN to create tumors in normal images and return normal images from tumor images. They trained the attention module with pixel-wise loss and adversarial loss. Pixel-wise loss motivated the attention module to triangulate the tumor's location accurately. Motamed et al. [143] proposed the **Inception-augmentation GAN (IAGAN)** model. In generator architecture, before concatenating the input image with the noise vector, they encoded it to a lower-dimensional representation utilizing convolution and attention layers to achieve a more accurate representation. Moreover, the inception and residual module are used after each convolution and pooling layer to improve the ability of GAN to capture more details from training while maintaining the spatial information. Waheed et al. [194] presented the CovidGAN model based on AC-GAN for generating synthetic chest X-ray images to improve the CNN performance. Madani et al. [120] found that the adversarial loss helps to reduce over-fitting by feeding unannotated test images to the discriminator in recognizing cardiac abnormalities in a chest X-ray. Liu et al. [114] proposed an approach called the active cell appearance model for computing the statistical distribution of intensity and shape of the cells. Later on, he used this model to guide cGAN to produce more realistic images. Han et al. [63] introduced a 3D multi-conditional GAN with two discriminators for constraining both the context and the lung nodule incorporated with LSGAN to produce realistic and diverse nodules in CT images. Ghorbani et al. [56] proposed the DermGAN framework (an adaptation of the pix2pix architecture), which could transform skin conditions to a generated realistic image with varied locations, sizes, and skin colors. Bailo et al. [11] utilized cGAN to generate blood smear image data from the segmentation mask according to microscopic images. Tekchandani et al. [189] implemented different GANs architecture, including cGAN, DCGAN, WGAN, AC-GAN, and InfoGAN, to augment benign and malignant mediastinal lymph node images. Chaudhari et al. [28] proposed a modified generator for data augmentation to improve cancer classification utilizing gene expression data. The GAN was based on feeding both the original data and the noise to G for producing the data with Gaussian distribution. Frid-Adar et al. [53] showed that utilizing separate DCGAN for each lesion class led to improved performance in lesion classification when compared to utilizing a unified AC-GAN for all classes. Papers related to medical image augmentation are summarized in Table 4.

4.4 Classification

Image classification is considered a challenging task where there was a large amount of intra-class variability, illumination, occlusion, and so on [26]. However, it was one of the most successful tasks applied by deep learning. The hierarchical features extracted from the image can be extracted from the deep neural network. GANs have been utilized for classification tasks as well, either utilizing a part of the discriminator and generator as a feature extractor or using the discriminator directly as a classifier by adding an additional class corresponding to the synthesized images.

Lahiri et al. [100], Madani et al. [120], and Lecouat et al. [102] trained GAN in a semi-supervised manner for learning patch-based retinal vessel classification, cardiovascular abnormality classification, and cardiac disease diagnosis, respectively. Yi et al. [218] proposed CatWGAN, which combined WGAN and CatGAN to learn the feature representation of dermoscopy images in an unsupervised and semi-supervised way. The features were extracted from the discriminator and a

Table 4. Summary of Data Augmentation Applications in Medical Image with a Brief Description

Publication	Remarks	Dataset	Result
Qi et al. [157]	[brain] Considering tumor and normal image as two domains, and incorporating semi-supervised attention mechanism to GAN.	BraTS [132]	Accuracy = 0.95
Motamed et al. [143]	[chest] Incorporating inception and residual architectures helps to improve GAN's capability to capture more details and maintains spatial information after each convolution and pooling layer.	Chest X-ray [93], Covid-chest X-ray [39]	Accuracy = 0.80, 0.69
Waheed et al. [194]	[chest] Using AC-GAN for data augmentation.	IEEE COVID Chest X-ray [39], COVID-19 Radiography [86] and COVID-19 Chest X-ray [85]	Accuracy = 0.95
Hammami et al. [62]	[liver] Combining CycleGAN with You Only Look Once (YOLO).	Visceral Anatomy [83]	mean average distance = 7.95 ± 6.2 mm.
Desai et al. [46] hline Chaitanya et al. [25]	[breast] Using DCGAN for data augmentation. [cardiac, prostate, and pancreas] Using Semi-supervised task-driven approach for performing data augmentation of image-label pair.	DDSM Pancreas, prostate, and cardiac	Accuracy = 0.87 Dice score = 0.84
Abdelhalim et al. [1]	[skin] Combining self-attention modules into progressive GAN (PGAN).	ISIC2018	Accuracy = 0.70
Abhishek and Hamarneh [2]	[skin] Using cGAN for synthesizing skin lesion images from binary masks.	ISIC, ISBI	Accuracy = 0.93
Sun et al. [186]	[liver] Using GAN incorporated with LSGAN to translate the label maps to a 3D MRI image.	BRATS17	Dice score = 0.69
Liu et al. [114]	[optics retinal] Using active cell appearance model (ACAM) to control cGAN.	Adaptive optics ophthalmoscope [169]	Recall = 0.79
Han et al. [63]	[lung] Attempting to generate 3D multi-conditional images using GAN with two discriminators.	LIDC [10]	Average sensitivity = 0.699
Ghorbani et al. [56]	[skin] Implementing pix2pix to create synthetic image.	N/A	FID score = 122.4 ± 3.4
Bailo et al. [11]	[blood] Using cGAN for generating a sample form segmentation mask.	N/A	Average precision = 0.895
Tekchandani et al. [189]	[lung] Synthesize the data using different GAN models as CGAN, DCGAN, AC-GAN, WGAN, and so on. AC-GAN approach with Inception V4 outperforms the other method.	Lymph nodes [156]	Accuracy = 0.94
Chaudhari et al. [28]	[gene expression] Feeding to G both noise and original dataset.	gene expression microarray [146]	Accuracy = 0.93
Frid-Adar et al. [53]	[liver] Using three DCGANs to produce artificial samples for three classes of liver lesions (cysts, metastases, and hemangiomas).	N/A	Accuracy = 0.85

N/A means not public dataset.

Table 5. Summary of Data Classification Applications in Medical Image with a Brief Description

Publication	Remarks	Dataset	Result
Pasupa et al. [153]	Using semi-supervised DCGAN for Red Blood Cell morphology classification.	N/A	Accuracy = 0.89
Yi et al. [218]	Incorporating WGAN and CatGAN.	ISIC, ISBI	Accuracy = 0.81
Hu et al. [71]	Combining WGAN and InfoGAN for unsupervised learning of cell-level feature representation.	Bone marrow [87]	Recall = 0.84
Rashid et al. [164]	Augmenting data set using G and classifying it using discriminator into many classes.	ISIC	Accuracy = 0.86
Kuang et al. [98]	Unsupervised multi-discriminator GAN incorporated with an encoder for benign and malignant classification of lung nodules.	LIDC [10]	Accuracy = 0.95

N/A means not public dataset.

classifier was built on top. Hu et al. [71] combined WGAN and InfoGAN for unsupervised learning of cell-level feature representation in histopathology images to perform tasks, i.e., nuclei segmentation, cell-level classification, and cell counting. InfoGAN has been utilized for feature extraction to minimize the mutual information between the synthesized images and a pre-defined subset of latent codes. Rashid et al. [164] adopted GAN's generator for data augmentation of a skin lesion image, and the discriminator was used as a classifier that was trained to identify seven skin lesion categories. Kuang et al. [98] proposed the MDGAN approach for lung nodule malignancy classification in an unsupervised manner. MDGAN consists of multiple discriminator networks. The approach calculates the feature loss and reconstruction loss to score benign and malignant lung nodules where high scores are assigned to malignant nodes and lower scores are assigned to benign nodes. Papers related to classification are summarized in Table 5.

The majority of research is centered around performing classification based on a two-phase approach. The first phase is applying GAN to augment training samples before classifying the data by adopting the classification model as in References [8, 18, 43, 160, 187, 193]. These two phases are trained disjointedly. Qin et al. [160] used GAN for data augmentation and transfer-ResNet50 as a classifier, which constructed rely on deep transfer neural network utilizing transfer learning.

Table 6. Summary of Registration Applications in Medical Image with a Brief Description

Publication	Remarks	Dataset	Result
Zhang et al. [230]	Integrating gradient loss into the network.	BraTS	Dice = 0.69
Qiao et al. [159]	Utilizing a single G and D for multi-contrast of registrations amount different modalities.	IXI [77]	SSIM = 0.75
Mahapatra et al. [122]	Using CycleGAN for multi-modal (retinal) and uni-modal (MR) deformable registration.	Sunnybrook cardiac [162]	Dice score = 0.887
Mahapatra et al. [125]	Utilizing CycleGAN for joint registration and segmentation.	NIH chest X-ray [202]	Dice score = 0.88
Wang et al. [195]	Handling domain-specific deformations using CycleGAN.	XI dataset [77], MA3RS [119]	SSIM = 0.74

4.5 Registration

Image registration (a.k.a. image fusion, warping, or matching) is a process of aligning two or more images to obtain the optimal transformation that better aligns the structures of interest in the source image. It is the essential step for image analysis in which important information is conveyed in more than one image. Most medical image registration studies concentrate on two modalities (e.g., MRI and CT). Zhang et al. [230] proposed a non-rigid registration approach for 3D medical images in an unsupervised learning manner. They utilized gradient loss in deep learning-based registration. Qiao et al. [159] introduced a model for multi-contrast modalities medical image registration using GAN in an unsupervised manner. Mahapatra et al. [122] utilized CycleGAN for multimodal (retinal) and unimodal (MR) deformable registration where the generator directly produced the registered image with the deformation field. Mahapatra et al. [125] proposed using CycleGAN for joint registration and segmentation of lung X-ray images. They found that their approach worked better than the separate methods for lung X-ray images. Table 6 summarizes the papers related to medical image registration.

4.6 Reconstruction

Medical image reconstruction is one of the most basic and significant components of medical imaging, and its main objective is to obtain high-quality medical images for clinical purposes at minimal risk and cost to the patients. Enhancement of lower-quality images can be performed by using the deep learning technique by reducing noise and improving sharpness and contrast along with the resolution. Super-resolution is an approach to produce higher-resolution images from lower-resolution data. The super-resolved images can be utilized for improving the accuracy of the detection of diseases and landmarks where small details are critical to performing accurate disease analysis. GAN showed a promising result in both local image restoration (such as super-resolution and inpainting) and image reconstruction. GAN has been proven to work well in image restoration tasks [19, 33, 81, 109, 117, 123, 124, 142, 183, 196, 209, 234].

Bing et al. [19] presented an approach for enhancing image resolution by embedding improved squeeze and excitation blocks in GAN's generator and discriminator. Moreover, they utilized new fusion loss, which can enhance the constraints on low-level features. The loss function was obtained by combining L1 loss, perceptual loss, mean square error loss, and relativistic adversarial loss. Wang et al. [196] adopted modified GAN architecture to tackle the task of 3D single image super-resolution brain MRI. For a generator, 3D memory-efficient residual in the residual-dense block was used, which contained 3D convolutions and took 3D patches as input. For the discriminator, instead of using PatchGAN, they used new architecture, namely, a pyramid pooling discriminator, to recover details on various size scales simultaneously. Chen et al. [33] proposed a 3D approach, called a **multi-level densely connected super-resolution network (mDCSRN)** with a GAN to produce a high-resolution MRI image from a low-resolution one. The approach combined WGAN and a modified DenseNet model. Similarly, a ResNet model was utilized as a part of a GAN to enhance the resolution of ultrasound images [136]. In Reference [207], a developed model for CT image noise reduction was used with a CNN in the generator. Moran et al.

Table 7. Summary of Medical Image Reconstruction

Publication	Remarks	Dataset	Result
Xu et al. [209]	Combining cGAN with spectral normalization for chest X-ray image super-resolution recovery.	CXR1 [40], CXR2 [41]	SSIM = 0.98, 0.925
Moran et al. [142]	Using GAN and transfer learning to improve resolution.	Dental X-ray [161], Pneumonia chest X-ray [34]	MSE = 14.79
Zhang et al. [229]	Enhancing CT resolution by incorporating multiple dense residual block structure-based GAN.	N/A	SSIM = 0.86
Bing et al. [19]	Embedding enhanced squeeze and excitation blocks in the G and the D.	DRIVE [185], STARE [70]	SSIM = 0.99
Wang et al. [196]	The architecture of G relied on memory-efficient implementation of residual dense connections and fully convolutional pyramid pooling used for D.	Human connectome project (HCP).	SSIM = 0.96
Li et al. [111]	Using dual D for edge enhancement.	FastMRI [224]	SSIM = 0.985
Shaul et al. [175]	GAN has been used for estimating the missing k-space samples.	IXI [77]	Dice score = 0.92
Mahapatra et al. [124]	Combining the triplet loss with the PGANs' cost function.	Sunnybrook cardiac [162]	SSIM = 0.83

Table 8. Summary of Lesion Detection Applications in Medical Image with a Brief Description

Publication	Remarks	Dataset	Result
Chen and Konukoglu [32]	Utilizing an adversarial auto-encoder to identify the sample distribution of healthy brain images.	BraTS [132] and human connectome project (HCP)	Area under the curve = 0.923
Xie et al. [208]	The G contains two parts: attention encoder and generation flow with the residual block.	SLO-1, SLO-2	Accuracy = 0.84, 0.97
Alex et al. [7]	Utilizing the D to classify MRI brain into lesion and non-lesion.	BraTS [132]	Dice score = 0.69
Ding et al. [47]	Using BEGAN for the generation of the spinal canal.	N/A	Accuracy = 0.93

[142] evaluated the resolution enhancement using transfer learning. Li et al. [111] proposed an edge-enhancement framework using dual D and one G with a holistic-image discriminator, multiscale edge discriminator, and a generator. The architecture aimed to maintain the edge details, the holistic image content, and stabilize the training procedure in an altogether better way. Shaul et al. [175] proposed a framework that employed the strengths of the GAN and U-net architectures for accelerating MRI acquisition. This was achieved by MRI subsampling followed by the estimated missing k-space samples using GAN. Table 7 summarizes the papers related to medical image reconstruction.

4.7 Detection

Lesion detection from images demands a tremendous amount of labeled training data [219]. GAN handles this issue in two ways: (i) by improving the dataset with generated images and then applying traditional detection models, as in References [63, 189], or (ii) by modeling distribution from which lesions can be recognized as outliers. The GAN's discriminator can be used to identify lesions by training on images representing normal pathology and learning the probability distribution of these images. Chen and Konukoglu [32] utilized an **adversarial auto-encoder (AAE)** and VAE to identify abnormalities according to the learned data distribution of healthy brain MRI images. The learned latent space can be used to map the lesion image to an image without a lesion before computing the residual of these two images to highlight the lesion. Xie et al. [208] proposed an approach for detecting the fundus disease. The generator was designed in two parts: (i) **attention encoder (AE)** module and (ii) generation module. The AE module encodes the real images to extract the feature of shallow layers, whereas the generation module produces the fake images by handling the input random noise by a set of the residual blocks with upsampling (RU) operations. The discriminator is developed by using a multi-branch ResNet-34 frame to extract the deep feature and for high-level feature extraction, the **deep-wise asymmetric dilated convolution (DADC)** module has been used. The discriminator's last layer has been modified to build a classifier to detect the diseased and normal images. Alex et al. [7] utilized GAN for the detection of a brain lesion on an MRI image, where the G produces the sample by modeling the distribution of normal patches and the D allocates a higher posterior probability of being real when compared to patches from various distributions (non-lesion patches). Table 8 summarizes the papers related to medical image detection.

5 DISCUSSION

Recent years have witnessed tremendous progress in GANs and their extensions, with their successful employment in the different medical fields. The modifications of the GAN architecture led to more stable training and an important impact on generated image diversity and quality [204]. The latest research proves that the capability and the performance of GANs are associated with the network size and batch size, which means that a well-designed structure is significant for great performances of GANs. Nevertheless, changes in the GAN's architecture are not only capable of removing all the inherent training issues for GANs. Reconstruction of the loss function as normalization and regularization can also help to yield better stable training for GANs. The reviewed papers in this survey can be found on the GitHub¹ repository. Around 28% of the studied papers were used in segmentation tasks in medical imaging for detecting abnormalities in the image, which has proved that segmentation plays an essential role in medical image analysis. GAN forces strong constraints on generator's output by applying adversarial training to ensure close to-real generation from the predicted mask, which enhance the output of the segmentation tasks. Additionally, about 23% of the reviewed papers used GAN in cross-modality tasks to transfer one image to another domain. CycleGAN has been broadly utilized for cross-modality medical image synthesis tasks especially due to its capacity to deal with the unpaired domain. The most commonly used modalities are MRI and CT, as reviewed in GAN-related literature, due to a large number of publicly available MRI and CT datasets. Another 19% of these studies fall into augmentation, which is considered the main task in GAN due to its capacity to generate a diversity of samples not distinguishable from the real one. The training of the models depends on the availability of training data to achieve the model's generalizability on unseen testing data and avoiding over-fitting. The augmented dataset was used to enhance the model's accuracy and has been adopted as the first phase in some detection and classification models. Most of these models concentrated on producing little objects that can be easily aligned, i.e., lesions, cells, and nodules [219]. This technique leads the training to become more stable because of the comparatively smaller content variation of these images compared to the whole context image. Moreover, it saves computation time where the training on high-resolution images demands a lot of GPU time. The importance of data augmentation emerged in rare diseases where it is difficult to collect sufficient training data. Thus, GAN has the ability to generate repeated samples through conditional information feeds to the model defined by medical experts. Around 12% of studies pertain to reconstruction, because the ability of GAN to synthesize a realistic-looking sample gives a new solution for enhancing image resolution and edges, which contributes to facilitating the clinical decisions. Furthermore, approximately 12% of review papers fall into the classification. GAN has also been used in classification applications, either utilizing part of G and D as a feature extractor or using D directly as a classifier. The remaining studies are related to detection and registration tasks where GANs have limited use in these tasks.

5.1 Future Challenges

Despite the successful utilities of GAN, there are still challenges that need to be handled to achieve its effective employment in medical imaging analysis.

- **Optimization process:** The most common GAN challenge is the complexity of the optimization process. The Nash equilibrium is the optimal point in such a mini-max game, as proved in Reference [57], where every player has to accomplish optimal cost. However, arriving at this equilibrium is not easy.

¹https://github.com/ManalMohammed1994/GAN_Review.

- **Class leakage:** Another challenge that faces GAN is class leakage, as found by way of experiments [23] where the generated image from one class included a set of properties from another class. This problem is not easy, since it requires defining appropriate metrics for better constraining the generated classes and preventing mixing the properties among them.
- **Traditional metrics:** Most models used traditional metrics as a **mean absolute error (MAE)**, peak **signal-to-noise ratio (PSNR)**, or **structural similarity (SSIM)** for quantitative evaluation in cross-modality and reconstruction tasks. However, these metrics are not compatible with the visual quality of the image. This is because direct optimization of pixel-wise loss generates a blurry result, yet gives higher numbers than utilizing adversarial loss. This problem can be solved by validating the quality of the generated sample using segmentation or classification. Zhang et al. [228] introduced **learned perceptual image path similarity (LPIPS)**, which achieved better agreement with human judgments than previous metrics. It has been applied in References [9] and [205] for evaluating the produced image quality and the normalized image quality.
- **No guarantee of preservation of small abnormality:** Another problem in cross-modality models that uses unpaired data is where the protection of small abnormality regions through the transformation process is not ensured. As mentioned previously, cross-modality can be performed on both paired and unpaired images, although training the model on the unpaired training samples does not assure the protection of tiny abnormality regions during the transformation process where the pre-trained CycleGAN can be subject to bias because of matching the produced data to the distribution of the target domain. The model bias occurs when target domain images in the training data have an over- or under-representation of specific classes. Moreover, this bias also occurs in paired data when the model was trained on the normal images, yet tested on the abnormal images. Modanwala et al. [140] handle this by altering the architecture of the discriminator and enforcing the penalization according to the structure at the scale of smaller patches, which allows the network to focus more on the image features.

5.2 Open Research Problems

Incremental developments of the GANs' architecture results in a powerful network that has a great impact on healthcare and clinical decision-making. However, GAN has limited use in some fields, although there is still an open research application that needs GAN's abilities. In the information security domain, GANs will also have significant uses. Maintaining the patients' critical information from malicious copying or even from the manipulation of medical evidence is demand. Recently, Chang et al. [27] suggested a framework, named AsynDGAN, for data privacy-protecting and effective communication and distribution of GAN. It is used to learn the generative distribution of the real datasets in various health entities without sharing or direct access to patients' data. Another open research area is applying LAPGAN in medical data, where it has been proved to work for synthesizing 96×96 px-sized images of realistically looking scenes. But, it is rarely used in the medical field. Lately, Baur et al. [14] tried to generate a high-resolution image for skin lesions using LAPGAN. Automatic 3D image modality colorization is another topic that has not been studied much and is still subject to manual methods. This problem has been heavily explored for 2D images. Colorization requires the synthesis of colors while maintaining structural content and the semantics of the target image. More recently, Mathur et al. [130] proposed a framework for colorizing multi-modal 3D medical data, utilizing the 2D-style exemplars and the generalization capabilities of a GAN. Elsewhere, most present data synthesis techniques only transfer the image from a source domain to a target domain without strong geometric correlations. Recently, Zhuang and Wang [238] proposed a framework using BiGAN to synthesize different

microscopic images from multiple domains with various geometric features. Finally, we would like to note that, despite there having been many successful GANs applications mentioned in this article, the usage of GANs in the medical domain is still in the beginning and needs to improve more.

6 CONCLUSION

GAN has emerged in the last few years and shows promising results in image processing for different purposes. Nowadays, GAN has become an essential generative model in the medical imaging field and helps to solve different medical images' problems, including augmenting datasets, transferring images from one domain to another, segmenting lesions, and so on. This article synopsizes the studies of GANs and broadly explains their basics, extensions, and their successful application in the medical field. Furthermore, the article provides a glance at different medical imaging modalities and GAN training issues. As presented earlier in the literature review, GAN has shown great results in many tasks, and its architecture has also been improved to alleviate training instability.

REFERENCES

- [1] Ibrahim Saad Aly Abdelhalim, Mamdouh Farouk Mohamed, and Yousef Bassyouni Mahdy. 2021. Data augmentation for skin lesion using self-attention based progressive generative adversarial network. *Exp. Syst. Applic.* 165 (2021), 113922.
- [2] Kumar Abhishek and Ghassan Hamarneh. 2019. Mask2Lesion: Mask-constrained adversarial skin lesion image synthesis. In *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, Cham, 71–80.
- [3] David Abramian and Anders Eklund. 2019. Refacing: Reconstructing anonymized facial features using GANs. In *IEEE 16th International Symposium on Biomedical Imaging (ISBI'19)*. 1104–1108.
- [4] Saori Aida, Junpei Okugawa, Serena Fujisaka, Tomonari Kasai, Hiroyuki Kameda, and Tomoyasu Sugiyama. 2020. Deep learning of cancer stem cell morphology using conditional generative adversarial networks. *Biomolecules* 10, 6 (2020), 931.
- [5] Mohamed Akil, Yaroub Elloumi, and Rostom Kachouri. 2021. Detection of retinal abnormalities in fundus image using CNN deep learning networks. In *State of the Art in Neural Networks and their Applications*. Elsevier, 19–61.
- [6] Manal Al Ghamdi, Mohamed Abdel-Mottaleb, and Fernando Collado-Mesa. 2020. DU-Net: Convolutional network for the detection of arterial calcifications in mammograms. *IEEE Trans. Med. Imag.* 39, 10 (2020), 3240–3249.
- [7] Varghese Alex, Mohammed Saawan K. P., Sai Saketh Chennamsetty, and Ganapathy Krishnamurthi. 2017. Generative adversarial networks for brain lesion detection. In *Medical Imaging 2017: Image Processing*. Vol. 10133. International Society for Optics and Photonics, Medical Imaging.
- [8] Ibrahim Saad Ali, Mamdouh Farouk Mohamed, and Yousef Bassyouni Mahdy. 2019. Data augmentation for skin lesion using self-attention based progressive generative adversarial network. *arXiv preprint arXiv:1910.11960*.
- [9] Karim Armanious, Chenming Jiang, Marc Fischer, Thomas Küstner, Tobias Hepp, Konstantin Nikolaou, Sergios Gatidis, and Bin Yang. 2020. MedGAN: Medical image translation using GANs. *Comput. Med. Imag. Graph.* 79 (2020), 101684.
- [10] Samuel G. Armato III, Geoffrey McLennan, Luc Bidaut, Michael F. McNitt-Gray, Charles R. Meyer, Anthony P. Reeves, Binsheng Zhao, Denise R. Aberle, Claudia I. Henschke, Eric A. Hoffman et al. 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys.* 38, 2 (2011), 915–931.
- [11] Oleksandr Bailo, DongShik Ham, and Young Min Shin. 2019. Red blood cell image generation for data augmentation using conditional generative adversarial networks. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*.
- [12] Lucia Ballerini, Robert B. Fisher, Ben Aldridge, and Jonathan Rees. 2013. A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions. In *Color Medical Image Analysis*. Springer, Cham, 63–86.
- [13] Christian F. Baumgartner, Lisa M. Koch, Kerem Can Tezcan, Jia Xi Ang, and Ender Konukoglu. 2018. Visual feature attribution using Wasserstein GANs. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. DOI : <https://doi.org/10.1109/CVPR.2018.00867>
- [14] Christoph Baur, Shadi Albarqouni, and Nassir Navab. 2018. Melanogans: High resolution skin lesion synthesis with GANs. *arXiv preprint arXiv:1804.04338*.
- [15] Avi Ben Cohen, Eyal Klang, Stephen P. Raskin, Shelly Soffer, Simona Ben Haim, Eli Konen, Michal Marianne Amitai, and Hayit Greenspan. 2019. Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. *Eng. Applic. Artif. Intell.* 78 (2019), 186–194.

- [16] Olivier Bernard, Alain Lalande, Clement Zotti, Frederick Cervenansky, Xin Yang, Pheng-Ann Heng, Irem Cetin, Karim Lekadir, Oscar Camara, Miguel Angel Gonzalez Ballester, et al. 2018. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Trans. Med. Imag.* 37, 11 (2018), 2514–2525.
- [17] David Berthelot, Thomas Schumm, and Luke Metz. 2017.Began: Boundary equilibrium generative adversarial networks. *arXiv preprint arXiv:1703.10717*.
- [18] Vedant Bhagat and Swapnil Bhaumik. 2019. Data augmentation using generative adversarial networks for pneumonia classification in chest Xrays. In *5th International Conference on Image Information Processing (ICIIP)*. IEEE, 574–579.
- [19] Xinyang Bing, Wenwu Zhang, Liying Zheng, and Yanbo Zhang. 2019. Medical image super resolution using improved generative adversarial networks. *IEEE Access* 7 (2019), 145030–145038.
- [20] Mikołaj Bińkowski, Dougal J. Sutherland, Michael Arbel, and Arthur Gretton. 2018. Demystifying MMDGANs. *arXiv preprint arXiv:1801.01401* 1050 (2018).
- [21] Ali Borji. 2019. Pros and cons of GAN evaluation measures. *Comput. Vis. Image Underst.* 179 (2019), 41–65.
- [22] Elliott Brion, Jean Léger, Ana M. Barragán-Montero, Nicolas Meert, John A. Lee, and Benoit Macq. 2021. Domain adversarial networks and intensity-based data augmentation for male pelvic organ segmentation in cone beam CT. *Comput. Biol. Med.* 131 (2021), 104269.
- [23] Andrew Brock, Jeff Donahue, and Karen Simonyan. 2018. Large scale GAN training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*.
- [24] Yang-Jie Cao, Li-Li Jia, Yong-Xia Chen, Nan Lin, Cong Yang, Bo Zhang, Zhi Liu, Xue-Xiang Li, and Hong-Hua Dai. 2018. Recent advances of generative adversarial networks in computer vision. *IEEE Access* 7 (2018), 14985–15006.
- [25] Krishna Chaitanya, Neerav Karani, Christian F. Baumgartner, Ertunc Erdil, Anton Becker, Olivio Donati, and Ender Konukoglu. 2020. Semi-supervised task-driven data augmentation for medical image segmentation. *Med. Image Anal.* 68 (2020), 101934.
- [26] Tsung-Han Chan, Kui Jia, Shenghua Gao, Jiwen Lu, Zinan Zeng, and Yi Ma. 2015. PCANet: A simple deep learning baseline for image classification? *IEEE Trans. Image Process.* 24, 12 (2015), 5017–5032.
- [27] Qi Chang, Hui Qu, Yikai Zhang, Mert Sabuncu, Chao Chen, Tong Zhang, and Dimitris N. Metaxas. 2020. Synthetic learning: Learn from distributed asynchronous discriminator GAN without sharing medical image data. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 13856–13866.
- [28] Poonam Chaudhari, Himanshu Agrawal, and Ketan Kotecha. 2019. Data augmentation using MG-GAN for improved cancer classification on gene expression data. *Soft Comput.* 24 (2019), 1–11.
- [29] Fengdi Che, Xiru Zhu, Tianzi Yang, and Tzu-yu Yang. 2019. 3SGAN: 3D shape embedded generative adversarial networks. In *IEEE/CVF International Conference on Computer Vision Workshops*. IEEE Computer Society.
- [30] Tong Che, Yanran Li, Athul Paul Jacob, Yoshua Bengio, and Wenjie Li. 2016. Mode regularized generative adversarial networks. *arXiv preprint arXiv:1612.02136*.
- [31] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. 2016. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets. In *Conference on Advances in Neural Information Processing Systems*.
- [32] Xiaoran Chen and Ender Konukoglu. 2018. Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders. *arXiv preprint arXiv:1806.04972*.
- [33] Yuhua Chen, Feng Shi, Anthony G. Christodoulou, Yibin Xie, Zhengwei Zhou, and Debiao Li. 2018. Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. In *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, Cham, 91–99.
- [34] Partha Chakraborty. Kaggle. 2019. Retrieved from <https://www.kaggle.com/parthachakraborty/pneumonia-chest-x-ray>.
- [35] Hongyoon Choi and Dong Soo Lee. 2018. Generation of structural MR images from amyloid PET: Application to MR-less quantification. *J. Nuclear Med.* 59, 7 (2018), 1111–1117.
- [36] Chiranjit Lal Chowdhary and D. P. Acharya. 2020. Segmentation and feature extraction in medical imaging: A systematic review. *Procedia Comput. Sci.* 167 (2020), 26–36.
- [37] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, et al. 2013. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. *J. Dig. Imag.* 26, 6 (2013), 1045–1057.
- [38] Noel C. F. Codella, David Gutman, M. Emre Celebi, Brian Helba, Michael A. Marchetti, Stephen W. Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin Mishra, Harald Kittler, et al. 2018. Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the international Skin Imaging Collaboration (ISIC). In *IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*. 168–172.
- [39] Joseph Paul Cohen, Paul Morrison, Lan Dao, Karsten Roth, Tim Q. Duong, and Marzyeh Ghassemi. 2020. COVID-19 image data collection: Prospective predictions are the future. *arXiv* 2006.11988. Retrieved from <https://github.com/ieee8023/covid-chestxray-dataset>.

- [40] CXR1. 2020. Retrieved from <https://data.mendeley.com/datasets/rsccb9sj/3>.
- [41] CXR2. 2020. Retrieved from <https://nihcc.app.box.com/v/ChestXray-NIHCC>.
- [42] Colin Decourt and Luc Duong. 2020. Semi-supervised generative adversarial networks for the segmentation of the left ventricle in pediatric MRI. *Comput. Biol. Med.* 123 (2020), 103884.
- [43] S. Deepak and P. M. Ameer. 2020. MSG-GAN based synthesis of brain MRI with meningioma for data augmentation. In *IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*.
- [44] Emily L. Denton, Soumith Chintala, Arthur Szlam, and Robert Fergus. 2015. deep generative image models using a Laplacian pyramid of adversarial networks. *CoRR* abs/1506.05751 (2015), 1486–1494.
- [45] Adrien Depeursinge, Alejandro Vargas, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, and Henning Müller. 2012. Building a reference multimedia database for interstitial lung diseases. *Comput. Med. Imag. Graph.* 36, 3 (2012), 227–238.
- [46] S. D. Desai, S. Giraddi, N. Verma, P. Gupta, and S. Ramya. 2020. Breast cancer detection using GAN for limited labeled dataset. In *12th International Conference on Computational Intelligence and Communication Networks (CICN)*. DOI :<https://doi.org/10.1109/CICN49253.2020.9242551>
- [47] Li Ding, Yu Sun, Shibo Li, Ying Hu, and Wei Tian. 2020. Research on spinal canal generation method based on vertebral foramina inpainting of spinal CT Images by using BEGAN. *J. Imag. Sci. Technol.* 64, 3 (2020), 30505–1.
- [48] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. 2016. Adversarial feature learning. *arXiv preprint arXiv:1605.09782*.
- [49] Xue Dong, Yang Lei, Sibo Tian, Tonghe Wang, Pretesh Patel, Walter J. Curran, Ashesh B. Jani, Tian Liu, and Xiaofeng Yang. 2019. Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network. *Radiother. Oncol.* 141 (2019), 192–199.
- [50] Getao Du, Xu Cao, Jimin Liang, Xueli Chen, and Yonghua Zhan. 2020. Medical image segmentation based on U-net: A review. *J. Imag. Sci. Technol.* 64, 2 (2020), 20508–1.
- [51] Qi Duan, Guotai Wang, Rui Wang, Chao Fu, Xinjun Li, Maoliang Gong, Xinglong Liu, Qing Xia, Xiaodi Huang, Zhiqiang Hu, et al. 2020. SenseCare: A research platform for medical image informatics and interactive 3D visualization. *arXiv preprint arXiv:2004.07031*.
- [52] Pasquale Foggia, Gennaro Percannella, Paolo Soda, and Mario Vento. 2013. Benchmarking HEp-2 cells classification methods. *IEEE Trans. Med. Imag.* 32, 10 (2013), 1878–1889.
- [53] Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. 2018. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 321 (2018), 321–331.
- [54] Maayan Frid-Adar, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. 2018. Synthetic data augmentation using GAN for improved liver lesion classification. In *IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*, 289–293.
- [55] Hao Ge, Yin Xia, Xu Chen, Randall Berry, and Ying Wu. 2018. Fictitious GAN: Training GANs with historical models. In *European Conference on Computer Vision (ECCV)*.
- [56] Amirata Ghorbani, Vivek Natarajan, David Coz, and Yuan Liu. 2020. DermGAN: Synthetic generation of clinical skin images with pathology. In *Machine Learning for Health Workshop*. PMLR, 155–170.
- [57] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Conference on Advances in Neural Information Processing Systems*. The MIT Press, Cambridge, MA, 2672–2680.
- [58] Neha Gour and Pratee Khanna. 2019. Speckle denoising in optical coherence tomography images using residual deep convolutional neural network. *Multim. Tools Applic.* 79 (2019), 1–17.
- [59] Shuyue Guan and Murray Loew. 2019. Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks. *J. Med. Imag.* 6, 3 (2019), 031411.
- [60] Jie Gui, Zhenan Sun, Yonggang Wen, Dacheng Tao, and Jieping Ye. 2020. A review on generative adversarial networks: Algorithms, theory, and applications. *arXiv preprint arXiv:2001.06937*.
- [61] Xiaoyu Guo, Cheng Chen, Yuanyuan Lu, Ke Meng, Hongyu Chen, Kangneng Zhou, Zhiliang Wang, and Ruoxiu Xiao. 2020. Retinal vessel segmentation combined with generative adversarial networks and dense Unet. *IEEE Access* 8 (2020), 194551–194560.
- [62] Maryam Hammami, Denis Friboulet, and Razmig Kechichian. 2020. Cycle GAN-based data augmentation for multi-organ detection in CT images via YOLO. In *IEEE International Conference on Image Processing (ICIP)*. 390–393.
- [63] Changhee Han, Yoshiro Kitamura, Akira Kudo, Akimichi Ichinose, Leonardo Rundo, Yujiro Furukawa, Kazuki Umemoto, Yuanzhong Li, and Hideki Nakayama. 2019. Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection. In *International Conference on 3D Vision (3DV)*. 729–737.

- [64] Luyi Han, Yunzhi Huang, Haoran Dou, Shuai Wang, Sahar Ahamad, Honghao Luo, Qi Liu, Jingfan Fan, and Jiang Zhang. 2020. Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network. *Comput. Meth. Prog. Biomed.* 189 (2020), 105275.
- [65] Zhenliang He, Meina Kan, Jichao Zhang, and Shiguang Shan. 2020. PA-GAN: Progressive attention generative adversarial network for facial attribute editing. *arXiv preprint arXiv:2007.05892*.
- [66] Ezz El-Din Hemdan, Marwa A. Shouman, and Mohamed Esmail Karar. 2020. COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images. *arXiv preprint arXiv:2003.11055*.
- [67] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, Günter Klambauer, and Sepp Hochreiter. 2017. GANs trained by a two time-scale update rule converge to a Nash equilibrium. *CoRR abs/1706.08500* (2017).
- [68] Yuta Hiasa, Yoshito Otake, Masaki Takao, Takumi Matsuoka, Kazuma Takashima, Aaron Carass, Jerry L. Prince, Nobuhiko Sugano, and Yoshinobu Sato. 2018. Cross-modality image synthesis from unpaired data using CycleGAN. In *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, 31–41.
- [69] Peter Hobson, Brian C. Lovell, Gennaro Percannella, Mario Vento, and Arnold Wiliem. 2015. Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset. *Arti. Intell. Med.* 65, 3 (2015), 239–250.
- [70] A. D. Hoover, Valentina Kouznetsova, and Michael Goldbaum. 2000. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imag.* 19, 3 (2000), 203–210.
- [71] Bo Hu, Ye Tang, I. Eric, Chao Chang, Yubo Fan, Maode Lai, and Yan Xu. 2018. Unsupervised learning for cell-level visual representation in histopathology images with generative adversarial networks. *IEEE J. Biomed. Health Inform.* 23, 3 (2018), 1316–1328.
- [72] Chan Hu, Youdong Ding, and Yuhang Li. 2020. Image style transfer based on generative adversarial network. In *IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. 2098–2102.
- [73] Wei Hu, Huanhuan Sheng, Jing Wu, Yining Li, Tianyi Liu, Yonghao Wang, and Yuan Wen. 2020. Generative adversarial training for weakly supervised nuclei instance segmentation. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 3649–3654.
- [74] Xun Huang, Ming-Yu Liu, Serge Belongie, and Jan Kautz. 2018. Multimodal unsupervised image-to-image translation. In *European Conference on Computer Vision (ECCV)*. Springer, Cham, 172–189.
- [75] Paulo Sergio Rodrigues. 2017. Mendeley Data. Breast Ultrasound Image. Retrieved from <https://data.mendeley.com/datasets/wmy84gzngw/1>.
- [76] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. 2017. Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 1125–1134.
- [77] IXI. 2020. Retrieved from <https://brain-development.org/ixi-dataset/>.
- [78] Saeed Izadi, Zahra Mirikhraj, Jeremy Kawahara, and Ghassan Hamarneh. 2018. Generative adversarial networks to segment skin lesions. In *IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*. 881–884.
- [79] Deepak Kumar Jain, Masoumeh Zareapoor, Rachna Jain, Abhishek Kathuria, and Shivam Bachhetty. 2020. GAN-Poser: An improvised bidirectional GAN model for human motion prediction. *Neural Comput. Applic.* 32 (2020), 1–13.
- [80] Lukas Jendele, Ondrej Skopek, Anton S. Becker, and Ender Konukoglu. 2019. Adversarial augmentation for enhancing classification of mammography images. *arXiv preprint arXiv:1902.07762*.
- [81] Xuhao Jiang, Yifei Xu, Pingping Wei, and Zhuming Zhou. 2020. CT image super resolution based on improved SR-GAN. In *5th International Conference on Computer and Communication Systems (ICCCS)*. 363–367.
- [82] Yifan Jiang, Han Chen, Murray Loew, and Hanseok Ko. 2020. COVID-19 CT image synthesis with a conditional generative adversarial network. *arXiv preprint arXiv:2007.14638*.
- [83] Oscar Jimenez-del Toro, Henning Müller, Markus Krenn, Katharina Gruenberg, Abdel Aziz Taha, Marianne Winterstein, Ivan Eggel, Antonio Foncubierta-Rodríguez, Orcun Goksel, András Jakab, et al. 2016. Cloud-based evaluation of anatomical structure segmentation and landmark detection algorithms: VISCERAL anatomy benchmarks. *IEEE Trans. Med. Imag.* 35, 11 (2016), 2459–2475.
- [84] Cheng-Bin Jin, Hakil Kim, Mingjie Liu, Wonmo Jung, Seongsu Joo, Eunsik Park, Young Saem Ahn, In Ho Han, Jae Il Lee, and Xuenan Cui. 2019. Deep CT to MR synthesis using paired and unpaired data. *Sensors* 19, 10 (2019), 2361.
- [85] Rahman T. Kaggle. 2020. COVID-19 Chest X-Ray Dataset Initiative. Retrieved from <https://github.com/agchung/Figure1-COVID-chestxray-dataset>.
- [86] Rahman T. Kaggle. 2020. Covid-19 radiography database. Retrieved from <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>.
- [87] Philipp Kainz, Martin Urschler, Samuel Schulter, Paul Wohlhart, and Vincent Lepetit. 2015. You should use regression to detect cells. In *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, Cham, 276–283.
- [88] Eunhee Kang, Hyun Jung Koo, Dong Hyun Yang, Joon Bum Seo, and Jong Chul Ye. 2019. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. *Med. Phys.* 46, 2 (2019), 550–562.

- [89] Levent Karacan, Zeynep Akata, Aykut Erdem, and Erkut Erdem. 2016. Learning to generate images of outdoor scenes from attributes and semantic layouts. *arXiv preprint arXiv:1612.00215*.
- [90] Salome Kazeminia, Christoph Baur, Arjan Kuijper, Bram van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. 2020. GANs for medical image analysis. *Artif. Intell. Med.* 109 (2020), 101938.
- [91] Qiao Ke, Jiangshe Zhang, Wei Wei, Robertas Damaševicius, and Marcin Woźniak. 2019. Adaptive independent subspace analysis of brain magnetic resonance imaging data. *IEEE Access* 7 (2019), 12252–12261.
- [92] Justin Ker, Lipo Wang, Jai Rao, and Tchoyoson Lim. 2017. Deep learning applications in medical image analysis. *IEEE Access* 6 (2017), 9375–9389.
- [93] Daniel Kermany, Kang Zhang, Michael Goldbaum, et al. 2018. Labeled optical coherence tomography (OCT) and chest X-ray images for classification. *Mendeley Data* 2, 2 (2018).
- [94] Hanseok Ko et al. 2020. Unsupervised generation and synthesis of facial images via an auto-encoder-based deep generative adversarial network. *Appl. Sci.* 10, 6 (2020), 1995.
- [95] Naveen Kodali, Jacob Abernethy, James Hays, and Zsolt Kira. 2017. On convergence and stability of GANs. *arXiv preprint arXiv:1705.07215*.
- [96] Yuhei Koike, Yuichi Akino, Iori Sumida, Hiroya Shiomi, Hirokazu Mizuno, Masashi Yagi, Fumiaki Isohashi, Yuji Seo, Osamu Suzuki, and Kazuhiko Ogawa. 2020. Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy. *J. Radiat. Res.* 61, 1 (2020), 92–103.
- [97] Dimitrios Korkinof, Tobias Rijken, Michael O'Neill, Joseph Yearsley, Hugh Harvey, and Ben Glocker. 2018. High-resolution mammogram synthesis using progressive generative adversarial networks. *arXiv preprint arXiv:1807.03401*.
- [98] Yan Kuang, Tian Lan, Xueqiao Peng, Gati Elvis Selasi, Qiao Liu, and Junyi Zhang. 2020. Unsupervised multi-discriminator generative adversarial network for lung nodule malignancy classification. *IEEE Access* 8 (2020), 77725–77734.
- [99] Neeraj Kumar, Ruchika Verma, Deepak Anand, Yanning Zhou, Omer Fahri Onder, Efstratios Tsougenis, Hao Chen, Pheng-Ann Heng, Jiahui Li, Zhiqiang Hu, et al. 2019. A multi-organ nucleus segmentation challenge. *IEEE Trans. Med. Imag.* 39, 5 (2019), 1380–1391.
- [100] Avisek Lahiri, Kumar Ayush, Prabir Kumar Biswas, and Pabitra Mitra. 2017. Generative adversarial learning for reducing manual annotation in semantic segmentation on large scale microscopy images: Automated vessel segmentation in retinal fundus image as test case. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 42–48.
- [101] Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, and Ole Winther. 2016. Autoencoding beyond pixels using a learned similarity metric. In *International Conference on Machine Learning*. PMLR, New York, NY, 1558–1566.
- [102] Bruno Lecouat, Ken Chang, Chuan-Sheng Foo, Balagopal Unnikrishnan, James M. Brown, Houssam Zenati, Andrew Beers, Vijay Chandrasekhar, Jayashree Kalpathy-Cramer, and Pavitra Krishnaswamy. 2018. Semi-supervised deep learning for abnormality classification in retinal images. *arXiv preprint arXiv:1812.07832*.
- [103] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 11 (1998), 2278–2324.
- [104] Baiying Lei, Zaimin Xia, Feng Jiang, Xudong Jiang, Zongyuan Ge, Yanwu Xu, Jing Qin, Siping Chen, Tianfu Wang, and Shuqiang Wang. 2020. Skin lesion segmentation via generative adversarial networks with dual discriminators. *Med. Image Anal.* 64 (2020), 101716.
- [105] Yang Lei, Xue Dong, Zhen Tian, Yingzi Liu, Sibo Tian, Tonghe Wang, Xiaojun Jiang, Pretesh Patel, Ashesh B. Jani, Hui Mao, et al. 2020. CT prostate segmentation based on synthetic MRI-aided deep attention fully convolution network. *Med. Phys.* 47, 2 (2020), 530–540.
- [106] Yang Lei, Joseph Harms, Tonghe Wang, Yingzi Liu, Hui-Kuo Shu, Ashesh B. Jani, Walter J. Curran, Hui Mao, Tian Liu, and Xiaofeng Yang. 2019. MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. *Med. Phys.* 46, 8 (2019), 3565–3581.
- [107] Alexander Hanbo Li, Yaqing Wang, Changyou Chen, and Jing Gao. 2020. Decomposed adversarial learned inference. *CoRR* abs/2004.10267 (2020).
- [108] Meiyu Li, Hailiang Tang, Michael D. Chan, Xiaobo Zhou, and Xiaohua Qian. 2020. DC-AL GAN: Pseudoprogression and true tumor progression of glioblastoma multiform image classification based on DCGAN and AlexNet. *Med. Phys.* 47, 3 (2020), 1139–1150.
- [109] Wanyue Li, Yi He, Wen Kong, Feng Gao, Jing Wang, and Guohua Shi. 2019. Enhancement of retinal image from line-scanning ophthalmoscope using generative adversarial networks. *IEEE Access* 7 (2019), 99830–99841.
- [110] Yunling Li, Hui Huang, Lijuan Zhang, Guangyi Wang, Honglai Zhang, and Wu Zhou. 2020. Super-resolution and self-attention with generative adversarial network for improving malignancy characterization of hepatocellular carcinoma. In *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 1556–1560.

- [111] Y. Li, J. Li, F. Ma, S. Du, and Y. Liu. 2020. High quality and fast compressed sensing MRI reconstruction via edge-enhanced dual discriminator generative adversarial network. *Magn. Reson. Imag.* 77 (2020), 124–136.
- [112] Yuexiang Li and Linlin Shen. 2018. cC-GAN: A robust transfer-learning framework for HEP-2 specimen image segmentation. *IEEE Access* 6 (2018), 14048–14058.
- [113] Zinan Lin, Ashish Khetan, Giulia Fanti, and Sewoong Oh. 2018. PacGAN: The power of two samples in generative adversarial networks. In *32nd International Conference on Neural Information Processing Systems*. 1505–1514.
- [114] Jianfei Liu, Christine Shen, Tao Liu, Nancy Aguilera, and Johnny Tam. 2019. Active appearance model induced generative adversarial network for controlled data augmentation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*. Springer, Cham, 201–208.
- [115] Ming-Yu Liu, Xun Huang, Jiahui Yu, Ting-Chun Wang, and Arun Mallya. 2020. Generative adversarial networks for image and video synthesis: Algorithms and applications. *arXiv preprint arXiv:2008.02793*.
- [116] Zhengchun Liu, Tekin Bicer, Rajkumar Kettimuthu, Doga Gürsoy, Francesco De Carlo, and Ian T. Foster. 2019. TomoGAN: Low-dose X-ray tomography with generative adversarial networks. *CoRR* abs/1902.07582.
- [117] Jiabo Ma, Jingya Yu, Sibo Liu, Li Chen, Xu Li, Jie Feng, Zhixing Chen, Shaoqun Zeng, Xiuli Liu, and Shenghua Cheng. 2020. PathSRGAN: Multi-supervised super-resolution for cytopathological images using generative adversarial network. *IEEE Trans. Med. Imag.* 39 (2020), 9.
- [118] Ruijun Ma, Bob Zhang, and Haifeng Hu. 2020. Gaussian pyramid of conditional generative adversarial network for real-world noisy image denoising. *Neural Process. Lett.* 51 (2020), 1–16.
- [119] MA3RS. 2020. Retrieved from <https://www.isrctn.com/ISRCTN76413758>.
- [120] Ali Madani, Mehdi Moradi, Alexandros Karargyris, and Tanveer Syeda-Mahmood. 2018. Chest X-ray generation and data augmentation for cardiovascular abnormality classification. In *Medical Imaging 2018: Image Processing*, Vol. 10574. International Society for Optics and Photonics.
- [121] Ali Madani, Mehdi Moradi, Alexandros Karargyris, and Tanveer Syeda-Mahmood. 2018. Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation. In *IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*. 1038–1042.
- [122] Dwarikanath Mahapatra, Bhavna Antony, Suman Sedai, and Rahil Garnavi. 2018. Deformable medical image registration using generative adversarial networks. In *IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*. 1449–1453.
- [123] Dwarikanath Mahapatra and Behzad Bozorgtabar. 2017. Retinal vasculature segmentation using local saliency maps and generative adversarial networks for image super resolution. *arXiv preprint arXiv:1710.04783*.
- [124] D. Mahapatra and Behzad Bozorgtabar. 2019. Progressive generative adversarial networks for medical image super resolution. *ArXiv* abs/1902.02144 (2019).
- [125] Dwarikanath Mahapatra, Zongyuan Ge, Suman Sedai, and Rajib Chakravorty. 2018. Joint registration and segmentation of x-ray images using generative adversarial networks. In *International Workshop on Machine Learning in Medical Imaging*. Springer, Cham, 73–80.
- [126] Michael Majurski, Petru Manescu, Sarala Padi, Nicholas Schaub, Nathan Hotaling, Carl Simon Jr, and Peter Bajcsy. 2019. Cell image segmentation using generative adversarial networks, transfer learning, and augmentations. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*.
- [127] Peter M. Maloca, Aaron Y. Lee, Emanuel R. de Carvalho, Mali Okada, Katrin Fasler, Irene Leung, Beat Hörmann, Pascal Kaiser, Susanne Suter, Pascal W. Hasler, et al. 2019. Validation of automated artificial intelligence segmentation of optical coherence tomography images. *PloS One* 14, 8 (2019), e0220063.
- [128] Xudong Mao, Qing Li, Haoran Xie, Raymond Y. K. Lau, Zhen Wang, and Stephen Paul Smolley. 2017. Least squares generative adversarial networks. In *IEEE International Conference on Computer Vision*. IEEE, 2794–2802.
- [129] Muhammad Mateen, Junhao Wen, Sun Song, Zhoupeng Huang, et al. 2019. Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry* 11, 1 (2019), 1.
- [130] Aradhya Neeraj Mathur, Apoorv Khattar, and Ojaswa Sharma. 2021. 2D to 3D medical image colorization. In *IEEE/CVF Winter Conference on Applications of Computer Vision*. 2847–2856.
- [131] Adriënne M. Mendrik, Koen L. Vincken, Hugo J. Kuijf, Marcel Breeuwer, Willem H. Bouvy, Jeroen De Bresser, Amir Alansary, Marleen De Bruijne, Aaron Carass, Ayman El-Baz, et al. 2015. MRBrainS challenge: Online evaluation framework for brain image segmentation in 3T MRI scans. *Computat. Intell. Neurosci.* 2015 (2015). <https://www.hindawi.com/journals/cin/2015/813696/>.
- [132] Bjoern H. Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. 2014. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imag.* 34, 10 (2014), 1993–2024.
- [133] Luke Metz, Ben Poole, David Pfau, and Jascha Sohl-Dickstein. 2016. Unrolled generative adversarial networks. *arXiv preprint arXiv:1611.02163*.

- [134] Fausto Milletari, Seyed-Ahmad Ahmadi, Christine Kroll, Annika Plate, Verena Rozanski, Juliana Maiostre, Johannes Levin, Olaf Dietrich, Birgit Ertl-Wagner, Kai Bötzl, et al. 2017. Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Comput. Vis. Image Underst.* 164 (2017), 92–102.
- [135] Mehdi Mirza and Simon Osindero. 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.
- [136] Deepak Mishra, Santanu Chaudhury, Mukul Sarkar, and Arvinder Singh Soin. 2018. Ultrasound image enhancement using structure oriented adversarial network. *IEEE Sig. Process. Lett.* 25, 9 (2018), 1349–1353.
- [137] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. 2018. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*.
- [138] Takeru Miyato and Masanori Koyama. 2018. cGANs with projection discriminator. *arXiv preprint arXiv:1802.05637*.
- [139] Aryan Mobiny, Pietro Antonio Cicalese, Samira Zare, Pengyu Yuan, Mohammadsajad Abavisani, Carol C. Wu, Jitesh Ahuja, Patricia M. de Groot, and Hien Van Nguyen. 2020. Radiologist-level COVID-19 detection using CT scans with detail-oriented capsule networks. *arXiv preprint arXiv:2004.07407*.
- [140] Gourav Modanwal, Adithya Vellal, Mateusz Buda, and Maciej A. Mazurowski. 2020. MRI image harmonization using cycle-consistent generative adversarial network. In *Medical Imaging 2020: Computer-Aided Diagnosis*, Vol. 11314. International Society for Optics and Photonics.
- [141] Tony C. W. Mok and Albert C. S. Chung. 2018. Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks. In *International MICCAI Brainlesion Workshop*. Springer, Cham, 70–80.
- [142] Maira B. H. Moran, Marcelo D. B. Faria, Gilson A. Giraldi, Luciana F. Bastos, and Aura Conci. 2020. Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs. *Comput. Biol. Med.* 129 (2020), 104139.
- [143] Saman Motamed and Farzad Khalvati. 2020. Inception augmentation generative adversarial network. *arXiv preprint arXiv:2006.03622*.
- [144] Tanya Motwani and Manojkumar Parmar. 2020. A novel framework for selection of GANs for an application. *arXiv preprint arXiv:2002.08641*.
- [145] Faizan Munawar, Shoib Azmat, Talha Iqbal, Christer Grönlund, and Hazrat Ali. 2020. Segmentation of lungs in chest X-ray image using generative adversarial networks. *IEEE Access* 8 (2020), 153535–153545.
- [146] NCBI. 2020. Retrieved from <https://www.ncbi.nlm.nih.gov/>.
- [147] Shubhangi Nema, Akshay Dudhane, Subrahmanyam Murala, and Srivatsava Naidu. 2020. RescueNet: An unpaired GAN for brain tumor segmentation. *Biomed. Sig. Process. Contr.* 55 (2020), 101641.
- [148] Dong Nie, Roger Trullo, Jun Lian, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen. 2017. Medical image synthesis with context-aware generative adversarial networks. In *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, Cham, 417–425.
- [149] Augustus Odena. 2016. Semi-supervised learning with generative adversarial networks. *arXiv preprint arXiv:1606.01583*.
- [150] Augustus Odena, Christopher Olah, and Jonathon Shlens. 2017. Conditional image synthesis with auxiliary classifier GANs. In *International Conference on Machine Learning*. PMLR, 2642–2651.
- [151] Reda Oulbacha and Samuel Kadoury. 2020. MRI to CT synthesis of the lumbar spine from a pseudo-3D cycle GAN. In *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 1784–1787.
- [152] Zhaoqing Pan, Weijie Yu, Xiaokai Yi, Asifullah Khan, Feng Yuan, and Yuhui Zheng. 2019. Recent progress on generative adversarial networks (GANs): A survey. *IEEE Access* 7 (2019), 36322–36333.
- [153] Kitsuchart Pasupa, Suchat Tungjitnob, and Supawit Vathanavarao. 2020. Semi-supervised learning with deep convolutional generative adversarial networks for canine red blood cells morphology classification. *Multim. Tools Applic.* 79, 45 (2020), 34209–34226.
- [154] Swati P. Pawar and Sanjay N. Talbar. 2020. LungSeg-Net: Lung field segmentation using generative adversarial network. *Biomed. Sig. Process. Contr.* 64 (2020), 102296.
- [155] Guim Perarnau, Joost van de Weijer, Bogdan Raducanu, and Jose M. Álvarez. 2016. Invertible conditional GANs for image editing. *arXiv preprint arXiv:1611.06355*.
- [156] Tuan D. Pham. 2020. T. D. Pham, Mediastinal Lymph Node. Retrieved from <https://sites.google.com/site/professorphanpham/codes/>.
- [157] Chang Qi, Junyang Chen, Guizhi Xu, Zhenghua Xu, Thomas Lukasiewicz, and Yang Liu. 2020. SAG-GAN: Semi-supervised attention-guided GANs for data augmentation on medical images. *arXiv preprint arXiv:2011.07534*.
- [158] Mengke Qi, Yongbao Li, Aiqian Wu, Qiyuan Jia, Bin Li, Wenzhao Sun, Zhenhui Dai, Xingyu Lu, Linghong Zhou, Xiaowu Deng, et al. 2020. Multi-sequence MR image-based synthetic CT generation using a generative adversarial network for head and neck MRI-only radiotherapy. *Med. Phys.* 47, 4 (2020), 1880–1894.
- [159] Jinhao Qiao, Qirong Lai, Ying Li, Ting Lan, Chunyan Yu, and Xiu Wang. 2020. A GAN based multi-contrast modalities medical image registration approach. In *IEEE International Conference on Image Processing (ICIP)*. 3000–3004.

- [160] Zhiwei Qin, Zhao Liu, Ping Zhu, and Yongbo Xue. 2020. A GAN-based image synthesis method for skin lesion classification. *Comput. Meth. Prog. Biomed.* 195 (2020), 105568.
- [161] Abdolvahab Ehsani Rad, Mohd Shafry Mohd Rahim, Amjad Rehman, and Tanzila Saba. 2016. Digital dental X-ray database for caries screening. *3D Research* 7, 2 (2016), 1–5.
- [162] Perry Radau, Yingli Lu, Kim Connelly, G. Paul, A. J. W. G. Dick, and Graham Wright. 2009. Evaluation framework for algorithms segmenting short axis cardiac MRI. *MIDAS J.-Cardiac MR Left Ventric. Segment. Chall.* 49 (2009).
- [163] Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- [164] H. Rashid, M. A. Tanveer, and H. Aqeel Khan. 2019. Skin lesion classification using GAN based data augmentation. In *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 916–919. DOI: <https://doi.org/10.1109/EMBC.2019.8857905>
- [165] Anisha Rebint and Mohan Kumar S. 2019. Importance of manual image annotation tools and free datasets for medical research. *J. Adv. Res. Dynam. Contr. Syst.* 10 (01 2019), 1880–1885.
- [166] Scott Reed, Zeynep Akata, Xincheng Yan, Lajanugen Logeswaran, Bernt Schiele, and Honglak Lee. 2016. Generative adversarial text to image synthesis. *arXiv preprint arXiv:1605.05396* 48 (2016).
- [167] Mina Rezaei, Haojin Yang, and Christoph Meinel. 2020. Recurrent generative adversarial network for learning imbalanced medical image semantic segmentation. *Multimed. Tools Applic.* 79, 21 (2020), 15329–15348.
- [168] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, Cham, 234–241.
- [169] Austin Roorda and Jacque L. Duncan. 2015. Adaptive optics ophthalmoscopy. *Ann. Rev. Vis. Sci.* 1 (2015), 19–50.
- [170] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. 2016. Improved techniques for training GANs. *arXiv preprint arXiv:1606.03498*.
- [171] Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. 2017. Scribbler: Controlling deep image synthesis with sketch and color. In *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 5400–5409.
- [172] Md Sarker, Mostafa Kamal, Hatem A. Rashwan, Mohamed Abdel Nasser, Vivek Kumar Singh, Syeda Furruka Banu, Farhan Akram, Forhad Uh Chowdhury, Kabir Ahmed Choudhury, Sylvie Chambon, et al. 2019. MobileGAN: Skin lesion segmentation using a lightweight generative adversarial network. *arXiv preprint arXiv:1907.00856*.
- [173] Divya Saxena and Jiannong Cao. 2020. Generative adversarial networks (GANs): Challenges, solutions, and future directions. *arXiv preprint arXiv:2005.00065*.
- [174] Philipp Seeböck, José Ignacio Orlando, Thomas Schlegl, Sebastian M. Waldstein, Hrvoje Bogunović, Sophie Klimscha, Georg Langs, and Ursula Schmidt-Erfurth. 2019. Exploiting epistemic uncertainty of anatomy segmentation for anomaly detection in retinal OCT. *IEEE Trans. Med. Imag.* 39, 1 (2019), 87–98.
- [175] Roy Shaul, Itamar David, Ohad Shitrit, and Tammy Riklin Raviv. 2020. Subsampled brain MRI reconstruction by generative adversarial neural networks. *Med. Image Anal.* 65 (2020), 101747.
- [176] Md Fahim Sikder. 2020. Bangla handwritten digit recognition and generation. In *International Joint Conference on Computational Intelligence*. Springer, 547–556.
- [177] Amitojdeep Singh, Sourya Sengupta, and Vasudevan Lakshminarayanan. 2020. Explainable deep learning models in medical image analysis. *J. Imag.* 6, 6 (2020), 52.
- [178] Vivek Kumar Singh, Mohamed Abdel Nasser, Farhan Akram, Hatem A. Rashwan, Md Mostafa Kamal Sarker, Nidhi Pandey, Santiago Romani, and Domènec Puig. 2020. Breast tumor segmentation in ultrasound images using contextual-information-aware deep adversarial learning framework. *Exp. Syst. Applic.* 162 (2020), 113870.
- [179] Vivek Kumar Singh, Hatem A. Rashwan, Farhan Akram, Nidhi Pandey, Md Mostafa Kamal Sarker, Adel Saleh, Saddam Abdulwahab, Najlaa Maaroof, Jordina Torrents-Barrena, Santiago Romani, et al. 2018. Retinal optic disc segmentation using conditional generative adversarial network. In *CCIA*, Vol. 308. 373–380.
- [180] Korsuk Sirinukunwattana, Josien P. W. Pluim, Hao Chen, Xiaojuan Qi, Pheng-Ann Heng, Yun Bo Guo, Li Yang Wang, Bogdan J. Matuszewski, Elia Bruni, Urko Sanchez, et al. 2017. Gland segmentation in colon histology images: The glaS challenge contest. *Med. Image Anal.* 35 (2017), 489–502.
- [181] Jaemin Son, Sang Jun Park, and Kyu-Hwan Jung. 2017. Retinal vessel segmentation in fundoscopic images with generative adversarial networks. *arXiv preprint arXiv:1706.09318*.
- [182] Jaemin Son, Sang Jun Park, and Kyu-Hwan Jung. 2019. Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks. *J. Dig. Imag.* 32, 3 (2019), 499–512.
- [183] Tzu-An Song, Samadrita Roy Chowdhury, Fan Yang, and Joyita Dutta. 2019. Self supervised super-resolution PET using a generative adversarial network. In *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 1–3.
- [184] Jost Tobias Springenberg. 2016. Unsupervised and semi-supervised learning with categorical generative adversarial networks. *arXiv preprint arXiv:1511.06390*.

- [185] Joes Staal, Michael D. Abràmoff, Meindert Niemeijer, Max A. Viergever, and Bram Van Ginneken. 2004. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imag.* 23, 4 (2004), 501–509.
- [186] Yi Sun, Peisen Yuan, and Yuming Sun. 2020. MM-GAN: 3D MRI data augmentation for medical image segmentation via generative adversarial networks. In *IEEE International Conference on Knowledge Graph (ICKG)*. 227–234.
- [187] B. Swiderski, L. Gielata, P. Olszewski, S. Osowski, and M. Kołodziej. 2021. Deep neural system for supporting tumor recognition of mammograms using modified GAN. *Exp. Syst. Applic.* 164 (2021), 113968.
- [188] Jiaxing Tan, Longlong Jing, Yumei Huo, Yingli Tian, and Oguz Akin. 2019. LGAN: Lung segmentation in CT scans using generative adversarial network. *arXiv preprint arXiv:1901.03473*.
- [189] Hitesh Tekchandani, Shrish Verma, and Narendra Londhe. 2020. Performance improvement of mediastinal lymph node severity detection using GAN and Inception network. *Comput. Meth. Prog. Biomed.* 194 (2020), 105478.
- [190] Xin Tie, Sai-Kit Lam, Yong Zhang, Kar-Ho Lee, Kwok-Hung Au, and Jing Cai. 2020. Pseudo-CT generation from multi-parametric MRI using a novel multi-channel multi-path conditional generative adversarial network for nasopharyngeal carcinoma patients. *Med. Phys.* 47, 4 (2020), 1750–1762.
- [191] Daniel S. W. Ting, Yong Liu, Philippe Burlina, Xinxing Xu, Neil M. Bressler, and Tien Y. Wong. 2018. AI for medical imaging goes deep. *Nat. Med.* 24, 5 (2018), 539–540.
- [192] Ngoc-Trung Tran, Tuan-Anh Bui, and Ngai-Man Cheung. 2018. Dist-GAN: An improved GAN using distance constraints. In *European Conference on Computer Vision (ECCV)*. Springer International Publishing, Cham, 370–385.
- [193] Rohit Verma, Raj Mehrotra, Chinmay Rane, Ritu Tiwari, and Arun Kumar Agariya. 2020. Synthetic image augmentation with generative adversarial network for enhanced performance in protein classification. *Biomed. Eng. Lett.* 10, 3 (2020), 443–452.
- [194] Abdul Waheed, Muskan Goyal, Deepak Gupta, Ashish Khanna, Fadi Al-Turjman, and Plácido Rogerio Pinheiro. 2020. CovidGAN: Data augmentation using auxiliary classifier GAN for improved covid-19 detection. *IEEE Access* 8 (2020), 91916–91923.
- [195] Chengjia Wang, Guang Yang, Giorgos Papanastasiou, Sotirios A. Tsaftaris, David E. Newby, Calum Gray, Gillian Macnaught, and Tom J. MacGillivray. 2020. DiCyc: GAN-based deformation invariant cross-domain information fusion for medical image synthesis. *Inf. Fusion* 67 (2020), 147–160.
- [196] Jiancong Wang, Yuhua Chen, Yifan Wu, Jianbo Shi, and James Gee. 2020. Enhanced generative adversarial network for 3D brain MRI super-resolution. In *IEEE Winter Conference on Applications of Computer Vision*. IEEE Computer Society, 3627–3636.
- [197] Jian Wang, Hengde Zhu, Shui-Hua Wang, and Yu-Dong Zhang. 2020. A review of deep learning on medical image analysis. *Mob. Netw. Applic.* 26 (2020), 1–30.
- [198] Lei Wang, Wei Chen, Wenjia Yang, Fangming Bi, and Fei Richard Yu. 2020. A state-of-the-art review on image synthesis with generative adversarial networks. *IEEE Access* 8 (2020), 63514–63537.
- [199] Li Wang, Yaozong Gao, Feng Shi, Gang Li, John H. Gilmore, Weili Lin, and Dinggang Shen. 2015. LINKS: Learning-based multi-source IntegratioN frameworK for segmentation of infant brain images. *NeuroImage* 108 (2015), 160–172.
- [200] Xiangyang Wang, Zhongzheng Cao, Rui Wang, Zhi Liu, and Xiaoqiang Zhu. 2019. Improving human pose estimation with self-attention generative adversarial networks. *IEEE Access* 7 (2019), 119668–119680.
- [201] Xiaolong Wang and Abhinav Gupta. 2016. Generative image modeling using style and structure adversarial networks. In *European Conference on Computer Vision*. Springer, Cham, 318–335.
- [202] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadie Bagheri, and Ronald M. Summers. 2017. Chestx-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *IEEE Conference on Computer Vision and Pattern Recognition*. 2097–2106.
- [203] Zhaobin Wang, Zijing Cui, and Ying Zhu. 2020. Multi-modal medical image fusion by Laplacian pyramid and adaptive sparse representation. *Comput. Biol. Med.* 123 (2020), 103823.
- [204] Zhengwei Wang, Qi She, and Tomas E. Ward. 2019. Generative adversarial networks in computer vision: A survey and taxonomy. *arXiv preprint arXiv:1906.01529*.
- [205] Leihao Wei, Yannan Lin, and William Hsu. 2020. Using a generative adversarial network for ct normalization and its impact on radiomic features. In *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 844–848.
- [206] Maciej Wiatrak and Stefano V. Albrecht. 2019. Stabilizing generative adversarial network training: A survey. *arXiv preprint arXiv:1910.00927*.
- [207] Jelmer M. Wolterink, Tim Leiner, Max A. Viergever, and Ivana Išgum. 2017. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans. Med. Imag.* 36, 12 (2017), 2536–2545.
- [208] Hai Xie, Haijun Lei, Xianlu Zeng, Yejun He, Guozhen Chen, Ahmed Elazab, Guanghui Yue, Jiantao Wang, Guoming Zhang, and Baiying Lei. 2020. AMD-GAN: Attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images. *Neural Netw.* 132 (2020), 477–490.

- [209] Liming Xu, Xianhua Zeng, Zhiwei Huang, Weisheng Li, and He Zhang. 2020. Low-dose chest X-ray image super-resolution using generative adversarial nets with spectral normalization. *Biomed. Sig. Process. Contr.* 55 (2020), 101600.
- [210] Yuan Xue, Tao Xu, and Xiaolei Huang. 2018. Adversarial learning with multi-scale loss for skin lesion segmentation. In *IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*. 859–863.
- [211] Yuan Xue, Tao Xu, Han Zhang, L. Rodney Long, and Xiaolei Huang. 2018. segAN: Adversarial network with multi-scale l 1 loss for medical image segmentation. *Neuroinformatics* 16, 3-4 (2018), 383–392.
- [212] Xinchen Yan, Jimei Yang, Kihyuk Sohn, and Honglak Lee. 2016. Attribute2Image: Conditional image generation from visual attributes. In *European Conference on Computer Vision*. Springer, Cham, 776–791.
- [213] Heran Yang, Jian Sun, Aaron Carass, Can Zhao, Junghoon Lee, Zongben Xu, and Jerry Prince. 2018. Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, Cham, 174–182.
- [214] Junlin Yang, Nicha C. Dvornek, Fan Zhang, Juntang Zhuang, Julius Chapiro, MingDe Lin, and James S. Duncan. 2019. Domain-agnostic learning with anatomy-consistent embedding for cross-modality liver segmentation. In *IEEE International Conference on Computer Vision Workshops*.
- [215] J. Yang, Z. Zhao, H. Zhang, and Y. Shi. 2019. Data augmentation for X-ray prohibited item images using generative adversarial networks. *IEEE Access* 7 (2019), 28894–28902. DOI: <https://doi.org/10.1109/ACCESS.2019.2902121>
- [216] Qianye Yang, Nannan Li, Zixu Zhao, Xingyu Fan, I. Eric, Chao Chang, and Yan Xu. 2020. MRI cross-modality image-to-image translation. *Sci. Rep.* 10, 1 (2020), 1–18.
- [217] Moi Hoon Yap, Gerard Pons, Joan Martí, Sergi Ganau, Melcior Sentís, Reyer Zwiggelaar, Adrian K. Davison, and Robert Martí. 2017. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE J. Biomed. Health Inform.* 22, 4 (2017), 1218–1226.
- [218] Xin Yi, Ekta Walia, and Paul Babyn. 2018. Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by Wasserstein distance for dermoscopy image classification. *arXiv preprint arXiv:1804.03700*.
- [219] Xin Yi, Ekta Walia, and Paul Babyn. 2019. Generative adversarial network in medical imaging: A review. *Med. Image Anal.* 58 (2019), 101552.
- [220] Xingde Ying, Heng Guo, Kai Ma, Jian Wu, Zhengxin Weng, and Yefeng Zheng. 2019. X2CT-GAN: Reconstructing CT from biplanar X-rays with generative adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition*. 10619–10628.
- [221] Chenyu You, Guang Li, Yi Zhang, Xiaoliu Zhang, Hongming Shan, Mengzhou Li, Shenghong Ju, Zhen Zhao, Zhuiyang Zhang, Wenxiang Cong, et al. 2019. CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-circle). *IEEE Trans. Med. Imag.* 39, 1 (2019), 188–203.
- [222] Fisher Yu, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, and Jianxiong Xiao. 2015. LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*.
- [223] Farhad G. Zanjani, Svitlana Zinger, Babak E. Bejnordi, Jeroen Awm van der Laak, et al. 2018. Histopathology stain-color normalization using deep generative models. *Conf. Med. Imag. Deep Learn.* 11 (2018), 1–11.
- [224] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J. Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, et al. 2018. fastMRI: An open dataset and benchmarks for accelerated MRI. *arXiv preprint arXiv:1811.08839*.
- [225] Haoran Zhang, Xuehao Cao, Lisheng Xu, and Lin Qi. 2019. Conditional convolution generative adversarial network for bi-ventricle segmentation in cardiac MR images. In *3rd International Symposium on Image Computing and Digital Medicine*. Association for Computing Machinery, New York, NY, 118–122.
- [226] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. 2019. Self-Attention Generative Adversarial Networks. *arXiv:stat.ML/1805.08318*.
- [227] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, and Dimitris N. Metaxas. 2017. StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In *IEEE International Conference on Computer Vision*. 5907–5915.
- [228] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE Conference on Computer Vision and Pattern Recognition*. 586–595.
- [229] Xiong Zhang, Congli Feng, Anhong Wang, Linlin Yang, and Yawen Hao. 2020. CT super-resolution using multiple dense residual block based GAN. *Sig., Image Vid. Process.* 10.1007 (2020), 1–9.
- [230] Xiaoyue Zhang, Weijian Jian, Yu Chen, and Shihting Yang. 2020. Deform-GAN: An unsupervised learning model for deformable registration. *arXiv preprint arXiv:2002.11430*.
- [231] Zizhao Zhang, Lin Yang, and Yefeng Zheng. 2018. Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network. In *IEEE Conference on Computer Vision and Pattern Recognition*. 9242–9251.

- [232] Junbo Zhao, Michael Mathieu, and Yann LeCun. 2016. Energy-based generative adversarial network. *arXiv preprint arXiv:1609.03126*.
- [233] S. Kevin Zhou, Hayit Greenspan, Christos Davatzikos, James S. Duncan, Bram van Ginneken, Anant Madabhushi, Jerry L. Prince, Daniel Rueckert, and Ronald M. Summers. 2020. A review of deep learning in medical imaging: Image traits, technology trends, case studies with progress highlights, and future promises. *arXiv preprint arXiv:2008.09104*.
- [234] Jin Zhu, Guang Yang, and Pietro Lio. 2019. How can we make GAN perform better in single medical image super-resolution? A lesion focused multi-scale approach. In *IEEE 16th International Symposium on Biomedical Imaging (ISBI'19)*. 1669–1673.
- [235] Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A. Efros. 2016. Generative visual manipulation on the natural image manifold. In *European Conference on Computer Vision*. Springer, Cham, 597–613.
- [236] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *IEEE International Conference on Computer Vision*. IEEE, 2223–2232.
- [237] Jun-Yan Zhu, Richard Zhang, Deepak Pathak, Trevor Darrell, Alexei A. Efros, Oliver Wang, and Eli Shechtman. 2017. Toward multimodal image-to-image translation. In *Conference on Advances in Neural Information Processing Systems*. 465–476.
- [238] Jun Zhuang and Dali Wang. 2020. Geometrically matched multi-source microscopic image synthesis using bidirectional adversarial networks. *arXiv preprint arXiv:2010.13308*.

Received 13 June 2021; revised 5 January 2022; accepted 20 March 2022