Library Progress International *Print version ISSN 0970 1052*
Vol.44 No.3, July-December 2024: P.8182-8149 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

A Systematic Review On Generative Adversarial Networks (GANs) For Biometrics

[1]Shweta Sinha, [2]Satya Bhushan Verma

[1,2] Department of Computer Science and Engineering,
Shri Ramswaroop Memorial University, Deva Road,
Barabanki, U.P., India

How to cite this article: Shweta Sinha, Satya Bhushan Verma (2024). A Systematic Review On Generative Adversarial Networks (GANs) For Biometrics. *Library Progress International,* 44(3), 8182-8149.

Abstract
Deep learning, a subset of machine learning, aims to imbue machines with human-like perception, learning, and intelligence through advanced technology. It has made significant strides in fields such as speech recognition, computer vision and NLP. GAN (Generative Adversarial Network) is an innovative domain within deep learning that has made significant growth in areas like image processing, art, music, data analysis, drug discovery, and gaming industry applications. Biometric systems, which use distinct physiological features like iris patterns, fingerprints, and facial characteristics for authentication, have increasingly integrated deep learning models. As biometric technology becomes widespread, ensuring information security in sectors like banking, education, and airports is paramount. Deep learning-based generative networks have revolutionized synthetic biometric data generation, produced high-quality artificial data while preserved the statistical characteristics of the original dataset. These synthetic biometric datasets are invaluable for testing and developing biometric systems, especially under high-demand conditions. The purpose of this paper is to present a broad overview of GANs, its loss function, highlighting the popularly used architectures and application domains of the most well-known variations. The optimal biometric application area and the efficacy of various model designs will be discussed.

Keyword: Biometrics, Deep Learning, Generative Adversarial Network, Image Synthesis, Variants of GAN

Paper Layout:
1. Introduction
2. GAN network structure
3. Variants of GAN
4. GAN in biometrics
5. Discussion
6. Conclusion

1. Introduction

Generative models in machine learning, particularly GANs (Generative Adversarial Networks), have gained prominence for their versatility. They find applications in diverse fields such as picture-to-picture interpretation, picture super-resolution, and video edge prediction. Generative models offer solutions to AI challenges like spatial variation and semi-supervised learning by generating fresh samples of unknown data through unsupervised learning. Deep learning models, employing Deep Neural Networks (DNNs), enhance representation and deduction capabilities, excelling at feature extraction. However, they face challenges in applications like image recognition, medical diagnosis, and biometric authentication due to data scarcity [1]. Despite their ability to train features and classifiers simultaneously, deep learning models require substantial training examples. Generative models, including autoregressive models, are built on the foundations of the maximum likelihood principle and parametrized models. While some methods have successfully learned generative models in various fields [2] they may struggle with the complexity of authentic data distribution. To address these limitations, Generative Adversarial Networks (GANs) were introduced, presenting adversarial learning between a discriminator and a generator [1]. GANs outperform other generative models due to their innovative adversarial process. GANs excel in reproducing data distributions and generating diverse information. They have significantly impacted computer vision by synthesizing data of high quality.

Generative Adversarial Networks exhibit significant potential in diverse domains, including music, semi-supervised learning, art, handling missing data, unsupervised learning, and drug discovery. They play a role as both attackers and defenders against privacy risks as data concerns grow [3]. Research on GANs has expanded, focusing on improving model training and applying GANs to various emerging areas. [4][5][6].

Library Progress International| Vol.44 No.3 |July-December 2024

GANs, have revolutionized the creation of synthetic biometric data. This synthetic data is valuable for developing and testing biometric systems under various scenarios, ensuring performance analysis under extreme load conditions. GANs prove to be resource-efficient and capable of generating diverse biometric samples compared to existing methods [7]. This diversity is crucial in overcoming challenges posed by distortion and corruption in data. GANs are instrumental in generating realistic fake biometric data, addressing privacy concerns associated with real biometric data.

A new class of generative model called GAN was created as a way to learn in both supervised and unsupervised environments. The model is referred to be adversarial as a result of training pairs of neural networks in opposition to one another. To provide real samples that are comparable to one another, the discriminator and generator networks compete. Game theory-based GANs were familiarized by Ian Goodfellow in 2014 [1]. A GAN is an ANN that can duplicate a data distribution to produce crisper synthesised data and employs a certain standard deviation to produce fresh data that has never been seen before. GAN is based on the minmax game, in which the two organisations should enhance their comparative objective capabilities. This results in a situation in which the two organisations' combined consequences are zero, and each player gain or loss is perfectly well-adjusted by the gain or loss of the opposing organisation. Numerous issues, like creating pictures from text descriptions, getting high-resolution photos, detecting objects, returning images that fit a specific pattern, etc. have been effectively solved with GAN. The family of generative models, which includes generative adversarial networks, offers a number of benefits over other generative models. This involves leveraging parallel generation rather than serial generation to build samples in Fully Visible Belief Networks. Unlike Boltzmann machines, GANs do not need Markov chains. Furthermore, it has been discovered that GANs provide better samples than other models.

The primary focus of the survey is to contextualize recent progress in the GAN field, examining various recently introduced variants and their approaches to addressing the main challenges in training GANs. The survey provides an overview of the original GAN model, its challenges, modified versions, and applications, with a specific focus on biometrics. Throughout the sections, the paper outlines the various architectures of GAN.

2. GAN network structure

The Discriminator model attempts to discriminate between samples acquired from training samples and data produced from the Generator model, whereas the Generator model makes a sample in the domain, such as an image, where a fixed-length random vector is used as an input. The networks that represent the discriminator and generator are frequently implemented using multi-layer networks with convolutional and/or fully linked layers. The generator and discriminator networks need to be differentiable, although they don't have to be directly invertible. GANs are built on a minimax game,[1], where the Generator generates samples on its own. In order to provide clamour for the generative interaction, a vector is randomly extracted from Gaussian circulation. After preparation, foci in this complicated vector space will correspond to foci in the problem space, framing a condensed representation of the information dispersion. This vector space is also known as a latent space or a vector space containing latent variables (hidden variables), variables that are essential to a domain but cannot be directly seen [4]. If the picture is correctly organised in response to the Discriminator's critique, the Discriminator will be paid; nevertheless, the Generator will be penalised and forced to rebalance its weight. If the data are erroneously categorised, the discriminator is penalised and forced to adjust its weights. As a consequence, the distribution of false photos starts to match that of actual images, and the Discriminator starts to categorise an image as fake or real with a probability of 0.5.

Loss function of GAN

The training objective function ($\min_G \max_D$) is as follows:

$$V(G, D) = E_{x\sim Pdata(x)}[\log D(x)] + E_{zPz(z)}[\log(1 - D(G(z)))] \quad (1)$$

where $p_{z(z)}$ is the probability distribution of G's latent space and $p_{data(x)}$ is the distribution of actual data. When modelling fresh data samples, $D(G(z))$ capacity is to distinguish the actual conveyance $D(x)$ with the combined appropriation $D(G(x))$. $p_{z(z)}$ is generally a uniform or Gaussian noise.

As per the equation (1), the initial presentation of the GANs [1] indicated the presence of a solitary solution. When neither of the players can recover from their defeat, what is known as a Nash Harmony (NE) result. A few investigations have revealed that NE attainment is probably not going to be conceivable in practice [8].
With regard to visual data, imagine that one network is an authority on art, while the another one is an art forger. The G (generator), sometimes referred to as the forger, makes forgeries in the GAN literature with the goal of creating realistic visuals. The master, the discriminator, D, accepts both the real (true) and falsified photographs with the intention of separating them. The generator does not directly access the pictures; instead, it learns via interacting with the discriminator. The real-image samples from the stack as well as the fictitious samples are both available to the discriminator. The discriminator task is to precisely discriminate between the samples produced by the Generator and the real samples using a binary classifier. The Generator then makes an effort to misrepresent the output's authenticity in an effort to raise the chance that the Discriminator will make a mistake. Knowing whether the picture originated from the real or the generator provides a simple ground truth that serves as the discriminator's error signal. The two players are becoming better, and the production quality is improving as well. The discriminator enables the generator to be taught with the same error signal, resulting in higher-quality forgeries.

GAN Challenges.
Although Generative Adversarial Networks (GANs) are powerful instruments for creating artificial data, they also present certain difficulties, including nonconvergence because of unstable behavior between the discriminator and generator, delayed learning, and mode collapse. [9] Some of the key challenges are:

Mode Collapse. To produce synthesised data from a latent space, high-quality data and generalization are required. GAN models can reproduce previously undiscovered data, but mode collapse occurs when multiple inputs in the latent space produce the same output, leading to a lack of diversity in generated samples [10]. Mode collapse has been the focus of several efforts, but remains an open problem. Partial mode collapse happens when a vast number of results are the same, but complete mode collapse occurs when the GAN produces the similar output with a variety of inputs. Despite NE identification, mode collapse does not result in GAN convergence [11]. Several recently suggested GAN variations, such as WGAN (Wasserstein GAN), have been demonstrated to lessen mode collapse [12][13].

Gradient Vanishing. Adjusting GAN's preparation is necessary, and D and G should be synchronised to learn rationally together [14][10]. The condition $D(x) = 1$ and another $D(G(z)) = 0$ show that an extremely precise D can discriminate between actual and synthetic data. In this case, the loss function is nearing zero, resulting in nearly negligible gradients and little input to the G. An erroneous D, on the other hand, cannot discriminate between genuine and simulated data, supplying the G with nonsensical information.

Stopping Problem. In theory, classic neural networks must optimise a loss function that monotonically decreases as the cost function increases. This does not occur since GANs must optimise the minimax game [15][16]. It is not feasible to infer the networks' state from their loss function, since neither of any specific pattern is followed by the loss function throughout GAN training. As a result, determining when the models have hit their maximum degree of training optimisation is difficult.

Evaluation Metrices. Due to the uniqueness of GAN, there isn't a singular criterion for assessing the quality of the synthesised data [17]. There is no agreement across research since each GAN application is unique, which is one of the causes. No universal metric is there for evaluating the quality of the artificially created data due to unique nature of GANs and diversity of applications. GANs may be used to mimic any data distribution, several metrics are used to compare the origin and synthesised distributions [18]. Various metrics have been proposed, but there is no consensus, and the choice of metrics depends on the specific application and provide a bigger picture of GAN performance [11].

Instability. GAN planning involves combining two learned models to improve the total loss function. The model design is based on a loss scenario where two organizations compete to find their unique layout in a minimax game. The organization structure and goal-capability can lead to significant changes in one organization, causing further modifications. The timings at which both networks desynchronize their states are sensitive, as major variations in gradients can cause a model to lose its learning. Unstable periods can affect training efficiency [19][4]. Recent GAN designs focus on stabilizing their training, as improving network performance often involves stabilizing training. Most recent developments include more stable training to address these challenges.

3. Variants of GAN

GANs have some drawbacks, including the mode collapse problem, oscillation between G and D during training, and instability as a player gains strength. The likelihood that the generated samples are real approaches zero when the quality of the samples is low because D gains the capacity to distinguish between authentic and fraudulent samples. As a consequence, G stops updating and produces a minor gradient of log(1 - D(G(z))). The selection of hyperparameters, including batch size, momentum, weight decay, and learning rate, is essential for the convergence of GAN training. The need for more diversified models to solve certain problems is one of the main concerns linked with GANs, which have led researchers to propose several variants of the concept. Some of these research advocate changing a GAN's training process, while others advocate changing a G or D model's structure. This section discusses the various GAN versions that have been put out over the years with their need for development. (Ref. Table 1)

Conditional GAN. Synthetic data need not always be random. GANs may require modifications to create desired data sets. CGANs [20] can produce specific image components or marked photographs based on requirements. Mirza et al. used class-conditioning to make both the discriminator and generator networks conditional in a 2D GAN architecture. A conditional GAN is a learning strategy that focuses on using helper data. To generate realistic data, the C-GAN expands the standard GAN by taking auxiliary information (c) and input noise (z) for the generator to produce conditional data (G(z|c)), such as class labels, text, or images. C-GAN discriminator distinguishes generator samples from real data using auxiliary info c and real data x. Conditional variables control generator output, which a standard GAN cannot. Hidden representation 'C' adaptability assessed using framework with info from labels or other data modalities. For single-labeled CGANs, a spliced label is inputted into both the Generator and Discriminator to identify the required photos. C-GAN has limitations - labeled datasets are crucial and objects created vary based on conditions. They prohibit different conditions for producing a comparable object.

DCGAN. DCGAN [21] is a GAN network design mostly made up of completely connected or convolution layers without max pooling. Downsampling and upsampling are accomplished by using transposed convolution and convolutional stride. To convert a 3x64x64 input image into a scalar probability that the input is from the genuine data distribution, the discriminator employs batch norm layers, strided convolution layers, and LeakyReLU activations. Bunch standardisation layers are used in both G and D to reduce noise and improve test results [22]. Redressed Straight Unit (LeakyReLU) improves DCGAN results on complex datasets like human faces and LSUN. DCGAN may collapse on some datasets with long training. The DC-GAN sets a benchmark for other GANs, with consistent training and excellent performance in creating high-quality images.

InfoGAN. InfoGANs [23] are an extension of GANs that learn disentangled features with no supervision, making them effective for recognizing objects and faces. InfoGANs change GANs' purpose by increasing common data between fewer elements and perceptions. GANs have unlimited noise capability. The noise could be tangled with data, while deconstructing a domain matches semantic features. In InfoGANs, the noise vector is split into the portions: z (noise source) and c (latent code focusing on structured properties). This results in G(z) generator network having both c and z. To prevent disregard of latent code c, information-theoretic regularisation and extended information I are used. Mutual Information maximizes latent space and code by using tricky calculations of P(c|x) (posterior). To boost the training, define conditional distribution (Q(c|x)) as an auxiliary distribution. InfoGANs offer output manipulation capability. InfoGANs can utilize various factors, including font tilt, thickness, and illumination direction, to manipulate the MNIST dataset.

PixtoPix. GANs, such as the pix2pix [24] architecture, are more effective in producing high-quality images. The generator uses an encoder-decoder architecture with convolutional layers that are down-sampled and up-sampled to resize images. The generator network maintains the image's dimensions and size using only one convolutional layer block before downsampling. The U-Net architecture has undergone significant advancements since then. The proposed discriminator design penalizes structure at the patch scale using a patch-wise approach. The final output of the patch GAN determines whether each patch of $N \times N$ in an image is authentic or forge. The Patch GAN discriminator requires fewer training parameters, operates faster, and can be used with any image size. Pix2pix GAN projects include images converted from black and white to color, animated drawings to photographs of realistic people, semantic labels to images translated from photos of cityscapes, and architectural labels to images of facades.

Cycle GAN. A DL architecture called CycleGAN [6] makes it possible to translate images to images without the requirement for paired training sets. To learn mapping between two image domains, it makes use of two GANs. The model is trained to extract features from the source domain and create new images from the target domain that have those same features. CycleGAN is a method that trains two GANs' generators independently, using the discriminator of the second GAN to create a source domain image from a target domain image, and the generator of the first GAN to create a target domain image from a source domain image. A cycle-consistent loss is used to produce images identical to genuine images from the target domain. To make sure that the features of the source domain images are maintained during translation to the target domain, an identity mapping component is included. Two generator models and two discriminator models are trained simultaneously in CycleGAN, an expansion of the Pix2Pix architecture. The same model can be used to convert images in the opposite direction (from source to target imagery) using unpaired datasets. The two generator models that make up the model architecture are Generator-A, that produces images for the primary domain and Generator-B, which produces images for the second domain.

SRGAN. SRGAN [25] uses GANs to improve the quality of the images of upscaled photos. This GAN architecture consists of a G and D, which generate high-resolution images and differentiate between the real and produced high-resolution images. SRGAN adds a perceptual loss function to enhance the resulting photos' perceived quality. This function computes the perceptual loss using features from a pre-trained deep neural network to ensure high-level semantic characteristics in the produced images. SRGAN has applications in creative picture improvement, super-resolution of images, and video super-resolution, improving the visual experience. Despite challenges such as generalization, model artifacts, and training complexity, SRGAN offers a complex method for producing realistic, high-quality images from lower-resolution inputs, making it a significant development in image super-resolution.

BigGAN. BigGAN [26] is a Generative Adversarial Network (GAN) designed for large-scale high-fidelity natural image synthesis. It excels in generating diverse, realistic images conditioned on specific classes due to its hierarchical class-conditional structure and a novel truncation trick. BigGAN surpasses previous models in Inception Scores and addresses challenges in large-scale training of GAN, offering potential for improved image synthesis and a deeper understanding of large-scale model training dynamics in artificial intelligence. The model trains larger neural networks with more parameters, resulting in impressively detailed images. The related model, Bi-Directional BigGAN (BigBiGAN), aims to expand the display by increasing the age-free picture age limit. The work scales up with an eightfold increase in batch size for GAN training, demonstrating that larger batches yield increased gradients per iteration, resulting in better outcomes in fewer stages. Despite BigGAN's outstanding performance, the variety of generated images is lower than real images of comparable size, especially for vast datasets like ImageNet. BigBiGAN enhances unconditional picture creation and representation learning capabilities showing greater accuracy scores in freshet inception distance and inception-score for unconditional findings.

PROGAN. Tero et al. introduced the Progressive Growing GAN (PG-GAN), [4] a generative adversarial network with multiple scales. In this approach, both the D and the G start training with images of low-resolution and progressively move to higher resolutions. The PG-GAN has demonstrated superior performance over non-progressive GANs in terms of quality, stability, and variance. It benefits from quickly convergent first layers, the ability to train a small number of layers at once, and a significant reduction in training time. Despite its success, the PG-GAN struggles with the mode collapse problem, where imbalances in training can lead to the generator producing identical samples. The Progressive Growing GAN tackles this by starting training with a low resolution, such as 4 x 4, and gradually increasing complexity and depth by adding more layers. This iterative process results in a network capable of generating high-quality, large-scale images, such as 1024 × 1024 pixels. Throughout the training phase, the resolution progresses from 4 x 4 to 8 x 8, and then to 16 x 16, ensuring a stable learning curve for both networks. This concept allows for the creation of higher-quality images in less time. However, when dealing with limited GPU resources, a bottleneck arises as the network expands, requiring more memory, despite the initial lower resolution's attempt to avoid GPU memory limits.

Style GAN. A major modification to the generator model is introduced by StyleGAN [5], an extension of the GAN architecture. It adds noise (a source of variation), applies style control at each iteration, and uses a mapping network to map latent space points to a middle latent space. With the use of noise and style vectors that can be adjusted at different levels of detail, this model produces faces that are highly lifelike and photorealistic. A continuously expanding training regime is used by StyleGAN, which follows Progressive GAN in its use of an alternate generator architecture taken from style transfer literature. Instead, then using randomly generated latent variables, it produces from a fixed value tensor. It also uses a regularization technique called mixing regularization, which blends two different styles of latent variables together during training. It has been used for advanced applications like artifact removal and creating superior facial pictures.

EBGAN. EBGAN [27] focuses on addressing the assessment metric issue by using an energy value as the estimate metric. An energy-based model always provides a mapping between the scalar values (also known as "energy") and each unique point in the input space. An energy function, upon which the discriminator is based, displays low energy on actual data and high energy on fake data. The energy function is therefore dissimilar from the discriminator probability function of a standard GAN. Furthermore, EBGAN incorporates an additional calamity capacity for the Generator. When both model capabilities are considered, the Discriminator strives to generate low energy values, whilst the Generator attempts to do the opposite.

BEGAN. BEGAN [28] addresses the assessment problem using boundary balance architectures. BEGAN uses the Wasserstein distance to calculate real loss between real and simulated image reconstruction losses, while BE-GAN achieves equilibrium between supremacy and diversity using their 'boundary equilibrium GAN'. BEGAN's training is speedy and stable, balancing quality and diversity by maintaining an equilibrium threshold. It also uses an autoencoder to stabilize the training process and ensure quality sample creation.

WGAN. An extension of the GAN that improves model stability and offers a loss function correlated with the quality of generated images is the Wasserstein Generative Adversarial Network (WGAN) [19]. By reducing the Earth-Mover distance approximation, WGAN produces less mode collapse and more stable training. The probability distribution of the generator (P_g) and the probability distribution derived from actual images (P_r) are the two main probability distributions that are used. To get realistic, high-quality output, the objective is to make sure that these probability distributions are close to one another. Three main techniques are proposed by mathematical statistics in machine learning to determine the distance amid these probability distributions: the JS divergence, the Wasserstein distance, and the Kullback-Leibler divergence. Initially employed in basic GAN networks, Jensen-Shannon divergence has gradient-related problems that might cause training to become unstable. In order to solve these recurrent problems and enhance the stability of WGAN training, the Wasserstein distance is employed.

LSGAN. LSGAN (Least Squared Generative Adversarial Networks) [29] address difficulties that develop after applying WGANs. This method produces higher-quality pictures than previous GANs and has a steadier training process. Even if the outlier samples are accurately identified, the least-square function penalises them more, helping eliminate vanishing gradient issues. LSGANs were presented to address difficulties that develop after applying WGANs, thus enhancing picture quality. LSGAN presented a novel loss function for the Discriminator model, allowing it to run smoothly over unsaturated gradients. Using the sigmoid cross entropy loss function, LSGAN solves the vanishing gradient issue that emerges during the learning stage, during backpropagation, which was present in all basic structure GANs. Using a least-square loss function, LSGANs penalized distant data on the same side of the decision boundary, producing picture quality improvements with noticeably better stability throughout the course of the learning phase.

BIGAN. A method BIGAN [30] is proposed for reprojecting data into the latent space utilizing these acquired feature representations in order to learn the semantics of data distribution and its inverse mapping. The BIGAN discriminator D complements the G in the conventional GAN model by differentiating simultaneously in the data and latent spaces (tuples (x,E(x)) versus (G(z),z)), in addition to differentiating in the data space (x versus G(z)), where the encoder output is represented by E(x) and z represents the generator input. In this case, GANs are used to teach the BidirectionalGAN encoder E how to flip the generator G. Many aspects, including learning, gradient, network topology, technique, and many performance measurements, can be used to compare the outcomes.

DUALGAN. The DualGAN and CycleGAN architectures are remarkably similar [31]. For the purpose of training its models, DualGAN and CycleGAN do not require paired data. The network design and the goal function's evolution both demonstrate this. An error term for reconstruction is defined and used to train every pair of D and G. Similar to how it is done in CycleGAN for the reconstruction error goal, the difference between the factual data sample and the equivalent recovered sample is determined. DualGAN has been modified and used multiple times [32][33].

DGGAN. The Dynamically Grown GAN (DGGAN) [34] method is a revolutionary approach that combines gradient-based training with architecture search approaches to determine the best growing strategy, network unit selection, and generator-discriminator balance. It alternates between training the new design and optimizing the generator and discriminator architecture. The DGGAN approach makes optimization easier by gradually expanding the architecture and design space through architecture searches. It enables the discriminator and generator to grow independently or in tandem dynamically, producing various unusual balances. DGGAN complements progressive growth with architecture search, making it easier to train GANs with complicated structures and high resolutions. The method extends parent architectures to new, child architectures, using weight inheritance during the training phase. This weight inheritance approach makes the optimization challenges of child designs easier and shortens the time needed for training each new candidate.

Self-Attention GAN(SAGAN). The issue of images' local spatial information is addressed by the SAGAN [35] architecture. For example, because the network's receptive field isn't large enough, covering pictures with distinct components associated in diverse portions of the image might be tough. In SAGAN, age of various highlights is calculated by bearing in mind signals from all photographs. Also, SAGAN D may assess the image's feature consistency. Self-awareness layers are used in SAGAN and they may be able to capture the geometric and structural characteristics of multiclass datasets. Each convolution's feature maps are divided into a one-one convolution in terms of query, key, and value to create the layer's output, could teach the network long-term dependencies.

StarGAN. StarGAN [36] is a scalable method for translating images between several domains using a single architecture. StarGAN allows for concurrent training of various datasets with distinct domains within a only network, resulting in high-quality translated images and flexibility to translate an input image to any desired target domain. The method has been proven in tasks like face expression synthesis and facial attribute transfer. StarGAN's multi-task learning setting allows it to produce images with superior visual quality. It also uses multiple datasets with different domain labels to handle all available labels. The goal is to create engaging image translation applications across various fields using a simple mask vector to access datasets with different domain labels.

S.No.	Authors /	Variants	Applications	Objective Function	Architecture	Type of Learning
1.	Mehndi Mirza, Simon Osindero (2014) [20]	Conditional GAN	Image to Image Translation, Convolution Face Generation, Video Generation, Text to	Minimize adversarial loss, ensuring realistic samples, with additional task-specific objectives for conditioning	Multilayer Perceptron Structure	Supervised Learning
2.	Radford et al. (2015) [21]	DCGAN	Image Generation, Image to Image Translation, Super Resolution	Discover the representation hierarchy in both G and D, from object pieces to scenes.	Convolution Network with constraint	Unsupervised Learning
3.	Xi Chen et al. (2016) [23]	InfoGAN	Disentangled representations, facial expressions,	maximum mutual information to create a disentangled representation.	Multilayered Perceptron	Unsupervised Learning
4.	Phillip Isola et al. (2017) [24]	PixtoPix	Maps from satellite images, colorizes grayscale photos,	Adversarial loss for realism, loss for pixel-wise similarity, guiding paired image translation	U-Net generators, PatchGAN discriminator	Supervised Learning
5.	Zhu et al. (2017) [6]	Cycle GAN	Unpaired image translation, style transfer,	blend of adversarial loss, cycle-consistency loss, and potentially identity loss.	Convolutional	Unsupervised Learning
6.	Christian Ledig et al. (2017) [25]	SRGAN	Enhancing image resolution in medical imaging, surveillance, artistic restoration.	Minimize adversarial, perceptual, content losses for realistic super-resolution images.	Deep residual network with adversarial and perceptual loss.	Unsupervised learning.

7.	Andrew Brock, et al. (2019) [26]	BigGAN	High-fidelity image synthesis, diverse class-conditional generation.	Minimize adversarial loss, generate high-fidelity images, conditioned on specific classes, utilizing large-scale	Scaled GAN architecture, hierarchical class embeddings.	Unsupervised learning
8.	Tero Karras et al., (2018) [4]	ProGAN	High-resolution image synthesis, progressive	Adversarial Loss: Similar to vanilla GANs	multi-resolution layers.	Unsupervised Learning
9.	Tero Karras et al. (2019) [5]	Style GAN	Photorealistic image synthesis, deepfake generation, artistic style transfer	Minimize perceptual loss, match feature statistics, enforce disentangled latent space	PG-GAN with style-based generator, mapping network, and stochastic	Unsupervised learning
10.	Junbo Zhao et al. (2017) [27]	EBGAN	Anomaly detection, image reconstruction, unsupervised learning, feature learning	Minimize autoencoder reconstruction error, energy function, and adversarial loss.	autoencoder architecture, reconstructive loss, and discriminator emphasizing energy function	Unsupervised learning
11.	David Berthelot (2017) [28]	BEGAN	Facial expression synthesis, image-to-image translation	Minimize generator and discriminator losses with equilibrium.	encoder-decoder architecture, emphasizing balance in G and D	Unsupervised Learning
12.	Martin Arjovsky (2017) [19]	WGAN	higher-fidelity image synthesis	Wasserstein distance as the training objective, leading to better convergence properties	Critic instead of discriminator, weight clipping or gradient penalty to enforce	Unsupervised Learning
13.	Mao et al. (2017) [29]	LSGAN	Realistic image synthesis, stable training for GANs, improved quality in medical image	Least Squares Loss replaces the traditional adversarial loss, aiming to address mode collapse and instability issues.	Improved convergence and stable adversarial training with the Least Squares Loss-based discriminator	Unsupervised Learning

14.	Donahue et al. (2016) [30]	BIGAN	Cross-modal learning, image-to-image translation, joint generation and recognition	Includes Encoder that maps data to latent representation	Convolutional	Unsupervised Learning
15.	Li et al. (2017) [31]	DUAL GAN	Medical image synthesis for diverse datasets, weather condition translation for autonomous vehicles,	Minimize domain gap through adversarial loss, cycle consistency, and identity preservation for effective cross-domain image translation.	Two generators and discriminators for bidirectional image translation, minimizing domain gap	Unsupervised Learning
16.	Liu et al. (2021) [34]	DGGAN	Enhancing image realism, medical image synthesis, style transfer, anomaly	optimizes generator and discriminator adversarial losses, high-quality image synthesis with dynamically grown capacity and diverse representations.	a generator with dynamically increasing capacity	Unsupervised adversarial and dynamic learning.
17.	Zhang et al. (2019) [35]	SAGAN	synthesizing realistic facial features, aiding in facial recognition, identity	optimize a hybrid objective function combining adversarial loss and self-attention regularization for improved image synthesis quality	a generator with self-attention modules, enhancing image synthesis by capturing	Unsupervised adversarial and attention learning.
18.	Choi et al. (2018) [36]	STARGAN	multimodal image translation, where it can be used for diverse tasks.	blend of the adversarial loss, domain classification loss, and reconstruction loss.	a complex architecture including a discriminator, a generator and a domain	Unsupervised Learning

Table 1. Variants Of GAN

Depending on the focus of their improvements, GAN variants have their own features. Labels are used by the latent space in cGAN as supplementary information to improve discrimination and data production. While ACGAN (Auxiliary Classifier Generative Adversarial Network) [37] learns representation with labels by means of an auxiliary classifier, InfoGAN seeks to optimize the mutual info between labels and generating data. While LSGAN employs the a-b coding strategy in the least squares approach to solve gradient vanish in GAN, WGAN computes loss using Wasserstein distance to solve mode collapse. BEGAN employs an autoencoder-based architecture for fair adversarial training, while DCGAN uses a deep CNN-based architecture for high-quality pictures and videos. SAGAN employs a self-attention mechanism for global long-range dependency and ProGAN progressively increases the depth of D and G in drill to produce images of high-resolution.

4. GAN in biometrics

Generative Adversarial Networks (GANs) are revolutionizing biometric security by improving human trait authentication and recognition. GANs are broadly used in biometrics for various tasks, including image completion, quality enhancement, style transfer, random realistic biometric sample creation, and identity-aware image reconstruction. Biometric characteristics like iris patterns, fingerprints, and facial features are crucial in contemporary security systems. Combining GANs with biometrics promises more private, effective, and safe solutions. GANs address issues like heterogeneous datasets, privacy concerns, and reliable recognition models. They make it easier to transfer domain pieces from an existing dataset and produce realistic patterns, which are difficult for model-based creation. Biometric traits use GANs to generate synthetic datasets, which are important resources for biometric recognition model training and testing. GANs solve issues with scarce real-world datasets, privacy concerns, and the requirement for representative and diverse samples. They create artificial samples that maintain key components of biometric traits without disclosing private information about actual people, protecting privacy and enhancing performance. GANs can handle small amounts of data and create diversity by producing variances in biometric features, essential for training reliable biometric identification models across diverse populations. They can also help with cross-modal biometrics by creating synthetic samples in one modality using information from another modality.

Related work

Various recent GAN variants and their modified versions have been used for the field of biometric traits, few of the related works are given in the section.

Minaee et al, for the purpose of creating an artificial handprint, fingerprint, and iris, respectively presented IrisGAN [38], FingerGAN [39] and PalmGAN [40] The foundation of all these models is DCGAN. FingerGAN and PalmGAN employ a total variation loss in addition to the DCGAN loss. It takes a while to train the models, and the output results are heavily noisy. G-GANISR was projected by Shamsolmoali et al. [41] to provide high-resolution data by utilizing the concept of super-resolution. Two generators were employed in place of a network of single generators. This strengthens the model and facilitates improved training. The network is skilled using a loss function that is derived from the least squares approach. This approach produces biometric data that are blurry and have a bad structure. Woung et al. [42] improved images to help identify subjects even when they are far from the camera's front. Cascaded SR-GAN was employed in order to improve. The SALR-VIPeR, SALR-PRID, and CAVIAR databases were utilized to train the GAN. Hu et al. [43] used GiGGAN to create a gait model that they suggested from every angle. This makes it possible to alter elements that could impact the identification process, such as clothing, lighting, etc. To train their GAN, they employed the OU-ISIR, MultiView Large Demographic Dataset.

Model-based reconstruction from minutiae with added styletransfer (e.g. Cycle GAN) and conditional GANs can be used to achieve fingerprint synthesis with a pre-defined identity [44]. Using a minute template, a conditional GAN (pix2pix) reconstructs a fingerprint. StyleGAN2 and a convolutional minutiae-to-vector encoder together enable attributes-aware, identity-preserving fingerprint reconstruction from minutiae. Due to their ability to learn faces from a single still image, refined GAN models allow for unconstrained face detection. A deep convolutional GAN is used in the Iris-GAN method to generate an iris from random data. For iris synthesis, conditional GAN (pix2pix) is utilized with the goal of enhancing data to increase iris recognition accuracy.

Face Recognition Using GAN Variants. GANs have been extensively used for face recognition. One common application is generating synthetic face images to augment datasets for improved training and to address challenges related to imbalanced data. Studies employ GANs for pose-invariant face recognition. By generating images with different poses, GANs enhance the strength of face recognition systems to variations in head orientation. GANs have enabled facial attribute manipulation, allowing researchers to modify attributes like gender, age and facial expressions in facial images for various applications, including age progression and emotion recognition. (Ref. Table 2)

Zhu et al. [24] introduced CycleGAN, a framework used for image-2-image translation tasks without matching data. It's significant in applications where mapping between different facial attributes is needed. In facial recognition, it aids in translating facial attributes across images without the need for paired datasets, enabling applications like expression synthesis. Antipov, G., et al. [45] specifically explores the use of GANs for face aging, a crucial component in various applications including forensics and missing person identification. Zhang, Z., et al. [46] discusses the use of Conditional Adversarial Autoencoders, a variant of GANs, for age progression and regression in facial images. The authors demonstrate the ability to synthesize facial images of different ages aiding age-invariant face recognition.

Huang, R., [47] addresses the challenge of synthesizing frontal views of faces, which is vital for facial recognition systems. It focuses on both local and global features of the face crucial for various facial recognition applications. Zhang, et al. [48] introduced StackGAN++ that utilizes a stacked GAN architecture for producing images with high resolution. In face recognition, it contributes to generating detailed and realistic facial images, enhancing the quality of training datasets. Choi et al. [36] proposed StarGAN that enables the translation of facial attributes across different domains, offering a unified framework for diverse facial recognition applications. It allows the transformation of facial features like age, gender, and expression seamlessly. He, R., et al. [49] proposed Wasserstein CNN, while primarily focusing on face recognition, this paper uses GANs to learn invariant features, making it more robust to various lighting conditions, including near-infrared (NIR) and visible light. Karras, T., et al. [4] introduces PGGAN (Progressive Growing GANs) that offers high-resolution image synthesis. In facial recognition, this technique enables the production of detailed and realistic facial images, contributing to the improvement of recognition models. Wu et al. [50] proposes a PP-GAN, Privacy-Protective-GAN architecture to address privacy issues during de-identification. It includes validation and regulator modules, maintains structural similarity, and offers a useful framework for customization.

GANs have revolutionized the creation of realistic face images but despite their potential, effective detection faces challenges such as exposing the wide variance in these images, resisting adversarial attacks, and ensuring interpretable decision-making for non-experts. Human performance in GAN-face identification is limited, achieving only 50% to 60% accuracy compared to AI technologies [51]. Various sophisticated GAN models, including BigGAN, PG-GAN, StyleGAN, and StyleGAN, have been developed to create realistic face images with diverse characteristics. However, GANs lack encoders or inference functions, restricting them to generated images. GAN inversion techniques have emerged to bridge the gap between actual and artificial facial images, enhancing the quality of produced images.
ProGANs can benefit from StyleGANs' results to the extent that some datasets, like human face images, are indistinguishable from genuine ones. StyleGAN allows for style mixing and high-quality image creation through combining existing images and separating image features for greater control. Allows for image feature modification across different levels. StyleGAN texture sticking was a major problem. Generated images had a distinct texture, which was noticeable in interpolated versions - like the hair on a face remained consistent despite movement.

Alias-Free GAN [52] introduced an architecture to solve texture sticking, creating a smooth interpolation of generated images. This results in a set of realistic and continuous images. Upgraded Pseudonym Free GAN and StyleGAN blending enable human face movements such as changing position, orientation, and smile.

Generalized Facial Recognition GAN models [53] exhibit resolution flexibility, but their efficacy hinges on training photos. Handling higher-dimensional images poses challenges due to exponential neural development. GANs' ability to synthesize facial images with novel poses has led to the development of frontalization techniques like DR-GAN and LB-GAN [54]. For instance, DR-GAN creates synthetic faces from input photos in any orientation, while LB-GAN adjusts the yaw angle of an input face image to a target angle using learned poses. However, IP-GAN faces challenges in learning identification and pose disentanglement when trained on unconstrained datasets.

Fingerprint Recognition Using GAN Variants. GANs contribute to fingerprint recognition by enhancing the quality of fingerprint images. This includes removing noise, enhancing ridge patterns, and improving minutiae extraction. They are also utilized in fingerprint database augmentation, allowing for the generation of synthetic fingerprint images to enrich training datasets. (Ref. Table 2)

DeepMasterPrint [55], in order to build a master fingerprint, suggested a GAN model built using the evolutionary strategy of covariance matrix adaption and Wasserstein distance; nevertheless, this model was unable to produce clear images. Fahim and Jung [56] introduced a lightweight GAN Network for large-scale fingerprint generation (LGN-LSFG). It employs spectral normalization and conditional loss doping to provide robust model training. The model performs poorly for iris and palmprints, but it produces good results for fingerprint generation. Huang et al. [57], utilized GAN to enhance fingerprint images from criminal scenes. They employed PatchGAN for identification in addition to traditional GAN architecture for recognition. They employed the NISTSD27 latent fingerprint dataset to assess their method. Minaee and Abdolrashidi [39] by using the same DC-GAN architecture, also worked on producing fingerprint images. To create the photos, they employed the PolyU and FVC 2006 Fingerprint Databases. Takahashi et al. [58] employed CycleGAN to normalize images in order to make them suitable for recognition. Joshi et al. [59] investigated latent image enhancement with GAN. They trained their architectures using the IIITD-MOLF and IIITD-MSLFD Datasets. Joshi et al. [60] introduces a fingerprint deblurring model called FDeblur-GAN, using a multi-stage framework of stack GAN and conditional GANs. It includes a ridge extractor model and a database of blurry fingerprints. The model achieved a 95.18% accuracy rate in matching deblurred and ground truth fingerprints. Improvements were achieved through various parameters. Pankaj Bamoriya et al, [61] demonstrated that Generative networks based on deep learning have transformed the creation of synthetic biometric data, opening up new possibilities for training and developing biometric systems. The proposed DSB-GAN uses a convolutional autoencoder and GAN model to generate realistic biometrics for fingerprint, iris, and palmprint, ensuring data availability and diversity compared to state-of-the-art methods. Moon et al. [62] used Pix2Pix model to increase the accuracy of reconstruction. Nonetheless, the reconstruction's matching precision of the fingerprints were noticeably worse than the original photos of fingerprints.

Iris Recognition. GANs are used to generate synthetic iris images, assisting in iris recognition systems' training and testing, and providing a solution for data scarcity in this biometric modality. They have improved iris spoofing detection by generating realistic synthetic images of spoof attacks, facilitating the development of more robust anti-spoofing algorithms. (Ref. Table 2)

Zou et al. [63], in their research on false iris recognition, created fake iris images using 4DCycle-GAN. Minaee and Abdolrashidi [38], developed realistic iris pictures with DCGAN. They generated images using the IIT Delhi Iris Database and the CASIA Iris Dataset. Kashihara [64] prior to identifying the iris, enhanced the pictures using Super-resolution GAN (Biometric4-SRGAN). Kakani et al., [65] presented a technique for creating an ID-preserving synthetic iris database using segmentation, identification, and generative adversarial learning. It validates the accuracy of generated images using AUC and compares them with previous studies. Bhuiyan and Czajka, [66] presented a Conditional StyleGAN-based model for iris synthesis trained on over 350 post-mortem iris samples. Following ISO/IEC 29794-6 guidelines, the model produces multiple within-class and between-class images, providing expert forensic human examiners with unprecedented deformations for various post-mortem intervals, potentially improving dataset accuracy and model weights.

Palmprint Recognition. Palmprint recognition, which uses the unique patterns seen in the skin of the palm, has developed into a non-intrusive and effective biometric technique. The incorporation of GANs adds a new level to palmprint recognition, even if standard recognition systems have seen great success. This is because it makes it possible to create synthetic palmprints for better training and testing. With its specific benefits in terms of stability and uniqueness, palmprint recognition has become a dependable biometric modality. GANs have become more well-known due to their capacity to produce realistic and varied synthetic data with the introduction of deep-learning techniques. This paper investigates the use of GANs in palmprint recognition, going over the difficulties, achievements, and possible directions for using GANs to improve the precision and resilience of palmprint recognition systems. (Ref. Table 2)

Minaee et al. [40] by using DCGANs attempted to produce realistic palm images. The PolyU Plamprint Database was utilized to train the GAN. Wang et al.'s method [67] trained a generative adversarial network using a huge number of palmprint images, but it was time-taken and required a large quantity of images. This paper uses a single palmprint image, which can be easily obtained, as the required image in the attacks, reducing the training time and improving the visual quality.

Table 2: GAN variants used in Biometric Traits

Biometric Trait	Author	Year	Model
Face	Zhu et al. [24]	2017	Cycle GAN
Face	Antipov et al. [45]	2017	Conditional GAN
Face	Zhang et al. [46]	2017	Conditional Adversarial Autoencoder
Face	Huang et al. [47]	2017	Global and Local Perception GAN
Face	Zhang et al. [48]	2017	StackGAN++
Face	Choi et al. [36]	2018	StarGAN

Face	He et al. [49]	2018	WGAN
Face	Karras et al. [4]	2018	PGGAN
Face	Wu et al. [50]	2019	Privacy Protective GAN
Fingerprint	P. Bontrager, et al. [55]	2018	WGAN
Fingerprint	Fahim and Jung [56]	2020	Lightweight GAN
Fingerprint	Huang et al. [57]	2020	PatchGAN
Fingerprint	Minaee and Abdolrashidi [39]	2018	FingerGAN
Fingerprint	Takahashi et al. [58]	2019	Cycle GAN
Fingerprint	Joshi et al. [59]	2019	GAN
Fingerprint	Joshi et al. [60]	2021	CGAN+Stack GAN
Fingerprint	Bamoriya P. et al. [61]	2022	DSBGAN
Fingerprint	Moon et al. [62]	2021	Pix2pix
Iris	Zou et al. [63]	2018	4D Cycle GAN
Iris	Minnae et al. [38]	2018	IrisGAN
Iris	Kashihara [64]	2020	SRGAN
Iris	Kakani et al. [65]	2023	GAN
Iris	Bhuiyan and Czajka [66]	2023	Conditional Style Gan
Palmprint	Minaee et al. [40]	2020	DCGAN
Palmprint	Wang et al. [67]	2020	DCGAN

5. Discussion

Research efforts in accurate detection of GAN-generated fake images have yielded promising results, with categorization into spatial domain and frequency domain methods. In the spatial domain, steganalysis based on photo response non-uniformity (PRNU) patterns, saturation cues, neural-network-based detectors, and spectral artifact patterns have been proposed. Conditional GANs, such as TV-GAN [68] and AP-GAN [69], utilize additional data for image creation, such as infrared images, thermal images, semantic labels, and bi-directional translation between thermal and visible images. Overall, spectral anomaly mitigation for most GANs remains challenging despite these efforts and proposed methods.

Although biometric identification systems are safe and reliable, they can be attacked in a number of ways, according to Dargan and Kumar [70] including direct or indirect assaults directed at template matching modules and feature extraction techniques. Certain biometric identification methods, including face, iris, fingerprint, palmprint, ECG, and voice recognition, are vulnerable to certain types of attacks. Artificial intelligence (AI) produces deepfakes, which are easily fooled by human sight. Countermeasures have been proposed by researchers to guard against gathered or fraudulent data entering biometric recognition systems.

Multimodal systems improve user authentication accuracy, security, and dependability by fusing several biometric features. Multimodal biometric systems are attracting researchers for attack prevention and security, with vein pattern recognition being a recent trend. Choosing the right combination of biometric modalities, determining how many traits to use, utilizing a fusion framework and effective recognition algorithms, taking into account biometric trait capture, device accuracy and reliability, designing a real-time application-specific system, creating application-specific data sets, and creating better data acquisition tools based on the combination of biometric traits used are some of the trade-offs that must be taken into account when implementing multimodal biometric authentication systems. User acceptability is also essential for a successful implementation.

For the purpose of training facial recognition systems, realistic image generator StyleGAN can be used to generate high-quality facial images. Its capacity to manipulate latent space and separate high-level features can improve a variety of facial aspects, supporting the augmentation of biometric datasets. Nevertheless, it encounters difficulties in resolving mode collapse problems. CycleGAN may be used to generate age-progressed photos or rejuvenated facial biometric data, as well as activities related to face aging or deaging. By converting facial traits across age groups, it helps create more comprehensive biometric datasets, albeit it might need to be carefully adjusted for certain biometric features. BigGAN, a large-scale high-fidelity picture synthesis tool, can improve the training of biometric identification systems. Though it may have trouble producing a broad range of biometric qualities in some cases, its capacity to manage larger neural networks with more parameters makes it appropriate for producing precise biometric traits. A multidomain model called StarGAN is able to simultaneously provide multimodal biometric data such as fingerprints, iris patterns, and facial features. It works well in situations where multiple biometric features must be produced at the same time, even though managing different modalities might be challenging. By building complexity from lesser resolutions, PG-GAN's progressive training method may produce high-resolution biometric pictures. By offering a steady learning curve and shortening training times, it can overcome difficulties in producing high-fidelity biometric data; nevertheless, scaling up may present difficulties due to GPU resource limitations.

6. Conclusion

The paper provides understanding of recent developments in the field of GAN. It highlights their challenges and advantages compared to traditional algorithms. GANs, operating on the minimax game concept, have significantly advanced image processing and extended their influence to field of biometrics. The report emphasizes the importance of understanding the strengths and weaknesses of GANs, addressing challenges like gradient disappearance and mode collapse. To optimize GANs for various applications, the paper suggests exploring new objective functions and enhancing traditional network structures. The research given in this paper finds a need of a Universal GAN that works with all biometric datasets and features. Facilitating the creation of multi-modal systems, it must provide a single framework for combining diverse biometric features. Cross-modal integration must be facilitated, increasing the comprehensiveness of biometric recognition systems. In addition to optimizing computational resources, protecting privacy, enhancing adversarial robustness, and streamlining research and development, it can manage a wide range of features and generalize across demographics. Furthermore, by providing protection against potential manipulations, it may enhance the security of synthetic biometric data. While discussing the recent GAN applications and different architectures' performance, the paper suggests the need of the best suitable model for most of the biometric traits as the future work.

1. References

[1] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., WardeFarley, D., Ozair, S., Courville, A., and Bengio, Y. "Generative ad-versarial networks", arXiv preprint arXiv:1406.2661 (2014) p. 2672–2680

[2] Hong, Y., Hwang,U., Yoo,J. and Yoon,S., "How Generative Adversarial Networks and Their Variants Work: An Overview", arXiv:1711.05914v10 [cs.LG] (2019)

[3] Cheng, J., Yang, Y., Tang, X., Xiong, N., Zhang, Y., and Lei, F. "Generative adversarial networks: A literature review". KSII Transactions on Internet & Information Systems 14, 12 (2020)

[4] Karras, T., Aila, T., Laine, S., and Lehtinen, J., "Progressive Growing of GANs for Improved Quality, Stability, and Variation". In ICLR. (2018)

[5] Karras, T., Laine, S., and Aila, T. "A style-based generator architecture for generative adversarial networks", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)

[6] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. "A. Unpaired Image-to Image translation using cycle-consistent adversarial networks." In IEEE International Conference on Computer Vision (ICCV) (2017) p. 2672–2680

[7] Ugot, O., Banjo,C., Misra, S., "Biometric Fingerprint Generation Using Generative Adversarial Networks", DOI:10.1007/978-3-030-72236-4_3 In book: Artificial Intelligence for Cyber Security: Methods, Issues and Possible Horizons or Opportunities (2021) p. 51-83

[8] Farnia, F., and Ozdaglar, A. "Gans may have no nash equilibria", (2020)

[9] Iglesias, G., Talavera, E., D´ıaz-Alvarez, A., "A survey on GANs for computer vision: Recent research, analysis and taxonomy", arXiv:2203.11242v2 [cs.LG] (2023)

[10] Zhang, Z., Luo, C., and Yu, J. "Towards the gradient vanishing, divergence mismatching and mode collapse of generative adversarial nets". In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (New York, NY, USA), CIKM '19, Association for Computing Machinery (2019) p. 2377–2380

[11] Goodfellow, I. Nips tutorial: "Generative adversarial networks" (2017)

[12] Li, W., Fan, L., Wang, Z., Ma, C., and Cui, X., "Tackling mode collapse in multi-generator gans with orthogonal vectors". Pattern Recognition 110, 107646. (2021)

[13] Bang, D., and Shim, H., "MGGAN: Solving mode collapse using manifold guided training" (2018)

[14] Su, J. "Gan-qp: A novel gan framework without gradient vanishing and lipschitz constraint" (2018)

[15] Barnett, S. A. "Convergence problems with generative adversarial networks (gans)" (2018)

[16] Liu, S., Bousquet, O., and Chaudhuri, K. "Approximation and convergence properties of generative adversarial learning", (2017)
[17] Xu, J., Ren, X., Lin, J., and Sun, X. "Diversity-promoting GAN: A crossentropy based generative adversarial network for diversified text generation". In Proceedings of Conference on Empirical Methods in Natural Language (2018)
[18] Borji, A., "Pros and cons of GAN evaluation measures"., Computer Vision and Image Understanding 179 (2019) p. 41–65
[19] Arjovsky, M., Chintala, S., and Bottou,L., "Wasserstein generative adversarial networks," in International conference on machine learning,(2017) p. 214–223
[20] Mirza, M., and Osindero, S., "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784 (2014)
[21] Radford A., Metz L., Chintala S., "Unsupervised representation learning with deep convolutional generative adversarial networks", arXiv preprint arXiv :1511.06434 (2016)
[22] Li, Y., Xiao, N., and Ouyang, W., "Improved boundary equilibrium generative adversarial networks". IEEE Access 6, (2018) p. 11342–11348
[23] Chen, Xi, Duan, Y., Houthooft,R., Schulman,J., Sutskever,I., Abbeel,P., "InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets", arXiv:1606.03657v1 [cs.LG] (2016)
[24] Isola,P., Zhu, J.Y.,Zhou, T. , and Efros, A.A., "Image-to-image translation with conditional adversarial networks," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2017) p. 1125–1134
[25] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W., "Photo-realistic single image super-resolution using a generative adversarial network". In: IEEE Conference on Computer Vision and Pattern Recognition (2017) p. 4681-4690
[26] Brock, A., Donahue, J., and Simonyan, K., "Large scale GAN training for high fidelity natural image synthesis". In International Conference on Learning Representations (2019)
[27] M. M. Y. L. Junbo Zhao, "Energy-based Generative Adversarial Network," in ICLR (2017)
[28] Berthelot,D., Schumm, T., and Met.L., "BEGAN: Boundary Equilibrium Generative Adversarial Networks," in arXiv:1703.10717v4 [cs.LG] (2017)
[29] Mao,X., Li, Q., Xie,H., Lau, R. Y. , Wang, Z., and Smolley, S.P., "Least squares generative adversarial networks," in Proceedings of the IEEE international conference on computer vision (2017) p. 2794–2802
[30] Donahue J., Krahenbuhl P., Darrell T., "Adversarial feature learning". arXiv preprint arXiv :1605.09782 (2017)
[31] Yi, Z., Zhang, H., Tan, P., and Gong, M. "Dualgan: Unsupervised dual learning for image-to image translation". In Proceedings of the IEEE international conference on computer vision (2017) p. 2849–2857
[32] Ye, J., Ji, Y., Wang, X., Gao, X., and Song, M. "Data-free knowledge amalgamation via group stack dual-gan". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2020) p. 12516–12525
[33] Liang, W., Ding, D., and Wei, G. "An improved dualgan for near-infrared image colorization". Infrared Physics & Technology 116, 103764 (2021)
[34] Liu, L., Zhang, Y., Deng, J., and Soatto, S., "Dynamically grown generative adversarial networks". In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35 (2021) p. 8680- 8687
[35] Zhang, H., Goodfellow, I., Metaxas, D., and Odena, A. "Self-attention generative adversarial networks". In International conference on machine learning, PMLR, (2019) p. 7354–7363
[36] Choi, Y., Choi,M., Kim, M. Ha, Jung-Woo., Kim, S., Choo,J., "StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation", arXiv:1711.09020v3 cs.CV (2018)
[37] Odena, A., Olah, C., and Shlens, J., "Conditional image synthesis with auxiliary classifier GANs,'' in Proc. 34th Int. Conf. Mach. Learn., vol. 70, [Online] http://proceedings. mlr.press/v70/odena17a.html, (2017) p. 2642–2651
[38] Minaee,S., Abdolrashidi, A., "Iris-gan: Learning to generate realistic iris images using convolutional gan", arXiv preprint arXiv:1812.04822 (2018)
[39] Minaee,S., Abdolrashidi,A., "Finger-GAN: Generating realistic fingerprint images using connectivity imposed GAN", arXiv e-prints. arXiv: 1812.10482 (2018)
[40] Minaee, S., Minaei,M., Abdolrashidi,A., "Palm-GAN: Generating realistic palmprint images using total-variation regularized GAN", arXiv preprint arXiv:2003. 10834 (2020)
[41] Shamsolmoali,P., Zareapoor,M.,Wang,R., Jain, D.K., Yang,J., "G-GANISR: Gradual generative adversarial network for image super resolution", Neurocomputing 366, (2019) p. 140–153
[42] Wang Z., Ye M., Yang F., Bai X., Satoh S.I.. "Cascaded SRGAN for scale-adaptive low resolution person re-identification". Proceedings of the 27th International Joint Conference on Artificial Intelligence; Stockholm, Sweden: AAAI Press; (2018) p. 3891–7
[43] Hu B., Gao Y., Guan Y., Long Y., Lane N., Plötz T. "Robust Cross-View Gait Identification with Evidence:

A Discriminant Gait GAN (DiGGAN) Approach" on 10000 People. ArXiv. 2018;abs/1811.10493. (2018)
[44] Wyzykowski, A.B.V., Segundo, M. P., and Lemes, R. de Paula, ''Multiresolution synthetic fingerprint generation'', IET Biometrics, vol. 11, no. 4 (2022) p. 314–332
[45] Antipov, G.; Baccouche, M.; Dugelay, J.L., "Face aging with conditional generative adversarial networks". IEEE International Conference on Image Processing (ICIP), Beijing, China, IEEE: Piscataway, NJ, USA, (2017) p. 2089–2093
[46] Zhang, Z.; Song, Y.; Qi, H. "Age progression/regression by conditional adversarial autoencoder." In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA;(2017) p. 5810–5818
[47] Huang R., Zhang S., Li T., "Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis", IEEE International Conference on Computer Vision (ICCV) (2017)
[48] Zhang H., Xu T., Li H., Zhang S., Wang X., Huang X., Metaxas D N. "StackGAN++: realistic image synthesis with stacked generative adversarial networks". IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(8) (2019) p. 1947–1962
[49] He, R., Wu, X., Sun, Z., Tan, T., "Wasserstein CNN: Learning Invariant Features for NIR-VIS Face Recognition." IEEE Trans. Pattern Anal. Mach. Intell., 41, (2018) p. 1761–1773
[50] Wu Y., Yang F., Xu Y., Ling H., "Privacy-Protective-GAN for Privacy Preserving Face De Identification". Journal of Computer Science and Technology.;34(1): (2019) p. 47-60
[51] Wang X., Guo, H., Hu. S., Ming-Ching Chang and Siwei Lyu, "GAN-Generated Faces Detection: A Survey and New Perspectives", ECAI (2023)
[52] Karras, T., Aittala, M., Laine, S., Hark¨onen, E., Hellsten, J., Lehti-¨ nen, J., and Aila, T. Alias-free generative adversarial networks. arXiv preprint arXiv:2106.12423 (2021)
[53] Desentz, D., "Partial facial re-imaging using generative adversarial networks", B.S.C.E., Wright State University (2021)
[54] Zeno, B., Kalinovskiy, I., Matveev, Y., "Pfa-Gan: Pose Face Augmentation Based On Generative Adversarial Network", Informatica (2021) p. 1-16
[55] Bontrager, P., Roy,A., Togelius, J., Memon, N., Ross, A., "DeepMasterPrints: Generating masterprints for dictionary attacks via latent variable evolution", in IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), IEEE, (2018) p. 1–9
[56] M.A.-N.I. Fahim, H.Y. Jung, "A lightweight GAN network for large scale fingerprint generation", IEEE Access 8 (2020) p. 92918–92928
[57] Huang X., Qian P., Liu M., editors. "Latent Fingerprint Image Enhancement Based on Progressive Generative Adversarial Network". IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 1419 (2020)
[58] Takahashi A., Koda Y., Ito K., Aoki T., editors. "Domain Transformation of Fingerprint Image Using CycleGAN"., IEEE 8th Global Conference on Consumer Electronics (GCCE); (2019) p. 5-18
[59] Joshi I., Anand A., Vatsa M., Singh R., Roy S.D., Kalra P., editors. "Latent Fingerprint Enhancement Using Generative Adversarial Networks. 2019 IEEE Winter Conference on Applications of Computer Vision (WACV) (2019) p. 7-11
[60] Joshi, A.S., Dabouei, A., Dawson, J., Nasrabadi, N. M., "FDeblur-GAN: Fingerprint Deblurring using Generative Adversarial Network", arXiv:2106.11354v1 [cs.CV] (2021)
[61] Bamoriya, P., Siddhad, G., Kaur, H., Khanna, P., Ojha., A, "DSB-GAN: Generation of deep learning based synthetic biometric data", Displays, Volume 74, 102267 (2022)
[62] Moon J.H., J.-H. Park, and G.-Y. Kim, "Restore fingerprints using pix2pix," in Advances in Computer Science and Ubiquitous Computing.,Springer, 1, 3 (2021) p. 489–494
[63] Zou H., Zhang H., Li X., Liu J., He Z., editors. "Generation Textured Contact Lenses Iris Images Based on 4DCycle-GAN", 24th International Conference on Pattern Recognition (ICPR) (2018) p. 20-24
[64] Kashihara K., editor "Iris Recognition for Biometrics Based on CNN with Super-resolution GAN", IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS); (2020) p. 27-29
[65] Kakani,V., Cheng-Bin Jin·Hakil Kim, "Segmentation-based ID preserving iris synthesis using generative adversarial networks" (2023)
[66] Bhuiyan,R.A., Czajka,A., "Forensic Iris Image Synthesis", arXiv:2312.04125v1 [cs.CV] (2023)
[67] Duan Y., Lu J., Zhou J., editors. "UniformFace: Learning Deep Equidistributed Representation for Face Recognition".IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019) p. 15-20
[68] Zhang, T., Wiliem, A., Yang, S., and Lovell, B., "TV-GAN: Generative Adversarial Network based Thermal to Visible Face Recognition,", International conference on biometrics (ICB), IEEE (2018) p. 174-181
[69] Di, X., Zhang, H., and Patel, V. M., "Polarimetric Thermal to Visible Face Verification via Attribute Preserved Synthesis", IEEE 9th International Conference on Biometrics Theory,Applications and Systems (BTAS), IEEE (2018) p. 1-10

[70] Dargan, S., Kumar, M., "A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities". Expert Systems with Applications, 143: 113114 (2020)