Generative Adversarial Networks : A Review

Avinash H. Hedaoo
Dept of Computer Science
Prerna College of Commerce, Nagpur, (MS) – India

Abstract- *GANs have gained much attention from the research community in recent years in terms of generating quality images and data augmentation. Generative adversarial networks (GANs) present a way to learn deep representations without extensively annotated training data. . The goal of GANs is to estimate the potential distribution of real data samples and generate new samples from that distribution. The representations that can be learned by GANs may be used in several applications. However, generating naturalistic images containing ginormous subjects for different tasks like image classification, segmentation, object detection, reconstruction, etc., is continued to be a difficult task. Generative modelling has the potential to learn any kind of data distribution in an unsupervised manner. In this paper, the background of the GAN, theoretic models and extensional variants of GANs are introduced, where the variants can further optimize the original GAN or change the basic structures. Then the typical applications of GANs are explained.*

Keywords- Generative adversarial networks, Supervised learning, Unsupervised learning, adversarial learning, generative models, zero-sum game, machine learning.

I. INTRODUCTION

Generating quality images is a challenging task in the field of computer vision and artificial intelligence, having numerous applications and research scope. Supervised machine learning and deep learning models require large and labelled datasets to generalize the decision making process. However, the availability of large and labelled databases is questioned in many domains like medical diagnosis, fault detection, intrusion detection, etc. Hence, the research community heavily depends on unsupervised learning. In unsupervised learning, the model strives to learn the structure and extracts the useful features of the data. However existing models do not fit the data distributions completely. Goodfellow et al. introduced GANs, an unsupervised generative model, worked on the principle of maximum likelihood, and used adversarial training. Right from the inception of generative adversarial networks (GANs), they have been the most discussed and most researched domains not only in the field of computer science but also in other domains. GANs have gained much popularity in generating high-quality realistic data.

It is now possible using GANs to generate photorealistic object images such as birds and faces, generate indoor or outdoor scenes, translate images from a source domain to the target domain, generate high definition images from low-definition images, and so on [5]. Besides, GANs have been introduced into the study of other artificial intelligence subfields, including speech and language processing [1], [2], malware detection [3], and chess game program [4],[5]. Recently, media interest in generative modelling projects has increased. The StyleGAN introduced by NVIDIA generates an authentic face image. GPT-3 from open artificial intelligence generates a complete sentence by providing a short introduction syntax. As of 2021, GAN and attention-based methods have evolved significantly, generating video, text, speech, and music that even experts cannot distinguish[10].

The goal of this survey is to provide a broad understanding of the progression of GANs and to summarize the current state-of-the-art. In this paper, the author presents an overview of GANs, its different variants, and potential applications in different domains. The paper attempts to identify GANs' advantages, disadvantages and major challenges for successful implementation of GAN in different application areas. Rest of the article is organized as follows. Section II provides basics of GANs, different objective functions of GANs and the variants of GANs. Section III highlights the most significant applications of GAN in real life. Sect. IV concludes the paper at the end[17].

II. BASIC THEORY AND TYPES OF GANs

A. Basic Theory of GAN

The generative models can be thought of as a group of thieves trying to generate counterfeit currency whereas the discriminative model can be thought of as police trying to detect the counterfeit currency. Thus, the entire framework resembles a two-player minimax game where the generator tries minimize its objective function and the discriminator tries to maximize its objective function[14].

Goodfellow et al. [18] introduced the adversarial process to learn generative models. The fundamental aspect of GAN is the min-max two-person zero-sum game. In this game, one player takes the advantages at the equivalent loss of the other player. Here, the players correspond to different networks of GAN called discriminator and generator. The main objective of the discriminator consists of determining whether a sample belongs to a fake distribution or real distribution. Whereas, generator aims to deceive the discriminator by generating fake sample distribution. Discriminator produces the chances or probability of a given sample to be a real sample. A higher value of probability shows that the sample is likely to be a real sample. The value close to zero indicates that the sample is a fake sample. The probability value near 0.5 indicates the generation of an optimal solution, such that discriminator is unable to differentiate fake and real sample[17].

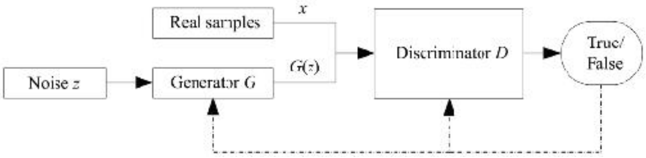

Fig. 1. Computation procedure and structure of GAN

In Fig. 1. Any differentiable function can be used as the generator and the discriminator. Here, differentiable functions D and G is used to represent the discriminator and the generator, and their inputs are real data x and random variables z, respectively. G(z) represents the sample generated by G and obeying the distribution pdata of real data. If the input of discriminator D is from the real data x, D should classify it to be true and label it as 1. If the input is from G(z), D should classify it to be false and label it as 0. The purpose of D is to achieve correct classification of the data source, while the purpose of G is to make performance of the generated data G(z) on D (i.e., D(G(z)) consistent with the performance of real data x on D (i.e., D(x)). The adversarial optimization process improves the performance of D and G gradually. Eventually, when the discrimination ability of D has been improved to a high level but cannot discriminate the data source correctly, it is thought that the generator G has captured the distribution of real data [5].

The loss function that G seeks to minimize and D attempts to maximize is as follows:

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

where x is a sample from the real dataset distribution pdata(x)and z is sampled from a latent space distribution pz(z). Eq.1shows the two networks playing a Mini-Max Game, each trying to improve their own loss function [23].

B. Types of GANs

With the passage of time, several developments have been made to the original architecture of GAN. In this subsection, we will introduce GANs' representative variants.

1) InfoGAN

Information maximizing GANs (InfoGANs) [15] are an information-theoretic extension of GANs that are able to learn disentangled features in a completely unsupervised manner. Here, InfoGANs modify the objective of GANs to learn meaningful representations by maximizing the mutual information between a fixed small subset of GAN's noise variables and observations. InfoGANs use approach of semantically decomposing a domain according to the semantic features of the data under consideration and thus decompose the input noise vector into two parts: (i)z which is treated as a source of noise, (ii) c called the latent code and targeted at the salient structured semantic features of the data distribution. Thus, the generator network becomes the generator G(z, c). In order to avoid the latent code c being ignored, information-theoretic regularization is done and the information I(c; G(z, c)) is maximized[14].

InfoGAN [22] aims to solve :

$$\min_G \max_D V_I (D, G) = V (D, G) - \lambda I(c, G(z, c)), \quad (2)$$

where V(D, G) is the objective function of original GAN,G(z, c) is the generated sample, I is the mutual information,and λ is the tunable regularization parameter. The final objective function of InfoGAN is

$$\min_G \max_D VI (D, G) = V (D, G) - \lambda L_I(c,Q), \quad (3)$$

where LI (c,Q) is the lower bound of I(c,G(z, c)). InfoGAN has several variants such as causal InfoGAN and semisupervised InfoGAN (ss-InfoGAN) [21].

The uniqueness of the InfoGAN compared to the standard GAN is the introduction of a regularization term (I) that captures the shared information among the interpretable variables (c) and the generator output [24].

Figure 2 shows the structure of InfoGAN[10].

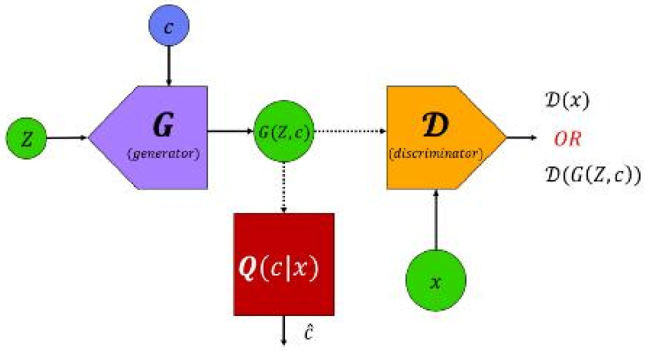

Figure 2. Structure and principle of InfoGAN (Information maximizing Generative Adversarial Networks).

2)Conditional GANs (cGANs)

GANs can be extended to a conditional model if both the G and D networks are conditioned on some extra information yto address the limitation of dependence only on random variables in original model [80]. y could be any kind of auxiliary information, such as class labels or data from other modalities. The conditional information can be added by feeding y into the both the D and G network as an additional input layer as depicted in Fig. 3[17].

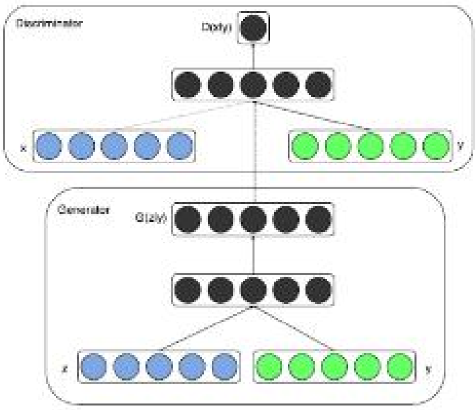

Figure 3. The structure of a Conditional Generative Adversarial Network (CGAN)

In the generator for CGANs, the prior input noise pz(z) and the auxiliary information yare combined in joint hidden representations. In the discriminator, x and y are presented as inputs to the discriminative function. Here, the objective function is similar to that of vanilla GAN except that the data distributions are now conditioned on y. This modified objective function is given as follows:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[log D(x|y)] + E_{z\sim p_z(z)}[log(1 - D(G(z|y)))]$$

(4)

The architectural diagram of CGAN can be seen in Figure 3. Both, the generator and the discriminator are multilayer perceptrons with Rectified Linear Units (ReLU) as the activation for hidden layers and sigmoid for the output layer. [16] also demonstrate the use of CGANs for automated tagging of images with multi-label predictions. This allows them to generate a distribution of tag-vectors conditional on image features [14].

3) WGAN-GP (Wasserstein GAN-Gradient Penalty)

WGAN-GP[11] solves problems such as mode collapse and unstable training, and makes GAN training predictable and reliable. WGAN-GP included a gradient penalty term in the critic loss function [12]. Critics' weights are not clipped. Moreover, the batch normalization layer should not be used for critics. Batch normalization generates a correlation between images in the same batch, so gradient penalty loss has less effect [13]. WGAN-GP suggests another way to enforce the Lipchitz constraint on critics: adding a term to the loss function that penalizes when the gradient norm of critic deviates significantly from "1". As a result, the training process was greatly stabilized. Gradient penalty loss is the squared difference between the gradient norm of the output and one. This model naturally finds weights that minimize the gradient penalty term. In other words, the model is made to follow the Lipchitz constraint. It is difficult to calculate the gradient everywhere during the training process. WGAN-GP only calculates the gradient at some point. In order not to be biased on one side, the real image-synthetic image pair are connected as shown in Figure 4, and the images interpolated using randomly selected points along a straight line are used [10].

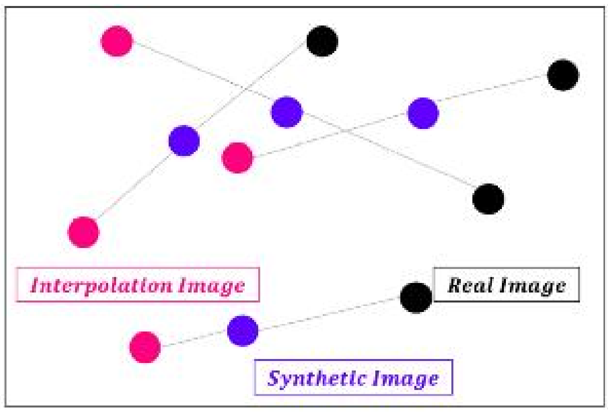

Figure 4. Interpolation between images

4) Laplacian Pyramid of Adversarial Networks(LAPGAN)

A sequential image generation framework Laplacian GAN (LapGAN) proposed by combining the CGAN model with the framework of the Laplacian pyramid (LP) . LapGAN requires the multiscale generation process in which a series of the GAN generates particular levels of details of an image in an LP representation. The GAN at each generation step of the LP can be different. The LP was built from a Gaussian pyramid (GP) using up-sampling u (.) and down-sampling d (.) functions explained as:[24] Let G(I) = [I0, I1,....,IK] be the Gaussian pyramid where I0 = I and IK is k repeated applications of d(:) to I. Then, the coefficient hk at level k of the Laplacian pyramid is given by the difference between the adjacent levels in Gaussian pyramid, upsampling the smaller one with u(:).

$h_k = L_k(I) = G_k(I) - u(G_{k+1}(I) = I_k- (I_{k+1}))$(5)

Modified reconstruction of the Laplacian pyramid coefficients[$h_1,\ldots, h_K$] can be performed through backward recurrenceas follows:

$$\widetilde{I}_k = u(\widetilde{I}_{k+1}) + \widetilde{h}_k = u(\widetilde{I}_{k+1}) + G_k(z_k, u(\widetilde{I}_{k+1}))$$

(6)

LAPGANs also take advantage of the CGAN model by adding a low-pass image l_k to the generator as well as the discriminator[25].

5) Deep Convolutional Generative Adversarial Networks (DCGAN)

A new class of convolutional neural networks (CNN) called Deep Convolutional GAN (DCGAN). DCGAN was the first structure that practiced de-convolutional neural networks (de-CNN) structural design that significantly stabilizes GAN training. These frameworks consist of two networks; one network works as a CNN called the generator, and the other network works as a de-CNN called discriminator. A newly proposed class of architectural constraints included in the CNN architecture is:

- Remove all levels of pooling layers with stride convolutions.
- Both G and D must use Batch Normalization (BN) .
- Use ReLU and Leaky-ReLU in the generator and the discriminator networks, respectively[24].

The DCGAN models performance was evaluated against LSUN, Imagenet1k, CIFAR10 and SVHN datasets. The quality of unsupervised representation learning was evaluated by first using DCGAN as a feature extractor and then the performance accuracy was calculated by fitting a linear model on top of those features. The authors also demonstrated feature learning by the generator show casing how the generator could learn to forget scene components such as bed, windows, lamps and other furniture. They also performed vector arithmetic on face samples leading to good results [17].

6) Adversarial Autoencoders (AAE)

Makhzani et al. [19] proposed adversarial autoencoder which is a probabilistic autoencoder which makes use of GAN to perform variational inference by matching the aggregated posterior of the hidden code vector of the autoencoder with an arbitrary prior distribution. After training, the encoder learns to convert the data distribution to the prior distribution, while the decoder learns a deep generative model that maps the imposed prior to the data distribution. The architectural diagram of an adversarial autoencoder is shown in Fig.5.

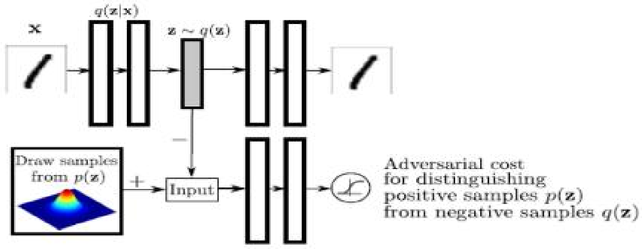

Fig. 5 The architecture of AAE

Let x be the input and z be the latent code vector of an autoencoder. Let p(z) be the prior distribution we want to impose, q(z ∣ x) be the encoding distribution and p(x ∣ z) be the decoding distribution. Also, let $p_d(x)$ be the data distribution and p(x) be the model distribution. The encoding function of the autoencoder $q(z \mid x)$ defines an aggregated posterior distribution of q(z) on the hidden code vector of the autoencoder as follows:

$$q(z) = \int_x q(z \mid x) p_d(x) dx$$

(7)

In adversarial autoencoder, the autoencoder is regularized by matching the aggregated posterior q(z) to an arbitrary priorp(z). The generator of the adversarial network is alsothe encoder of the autoencoder q(z ∣ x) . Both, the adversarial network and the autoencoder are trained jointly with stochastic gradient descent in two phases—the reconstruction phase and the regularization phase. Labels can also be incorporated in AAEs in the adversarial training phase in order to better shape distribution of the hidden code. A one-hot vector is added to the input of the discriminative network to associate the label with the mode of distribution. Here, the one-hot vector acts as a switch that selects the corresponding decision boundary in the discriminative network given the class label. The one hot vector also contains one point corresponding to an extra class which in turn corresponds to unlabelled examples. When an unlabelled example is encountered, the extra class is turned on and the decision boundary for the full mixture of Gaussian distribution is selected [17].

The performance of adversarial autoencoders is evaluated on MNIST and Toronto Face datasets using log likelihood analysis in supervised, semi-supervised and unsupervised settings.

$$p(y) = Cat(y) \qquad p(z) = N(z \mid 0 , I) \qquad (8)$$

Here, the adversarial network and the autoencoder are trained in three phases viz. the reconstruction phase, regularization phase and the semi-supervised classification phase. Moreover, the inference network q(y|x) predicts a one-hot vector whose dimension is the number of categories that the data can be clustered into. It is also showed how adversarial autoencoders can be used for dimensionality reduction [25].

7) Generative Recurrent Adversarial Networks(GRAN)

Im et al. (2016) proposed recurrent generative model showing that unrolling the gradient based optimization yields a recurrent computation that creates images by incrementally adding to a visual “canvas”. Here, the “encoder” convolutional network extracts images of current “canvas”. The resulting code and the code for the reference image get fed to a “decoder” which decides on an update to the “canvas”. Figure 6 depicts an abstraction of how a Generative Recurrent Adversarial Network works. The function f serves as the decoder and the function g serves as encoder in GRAN.

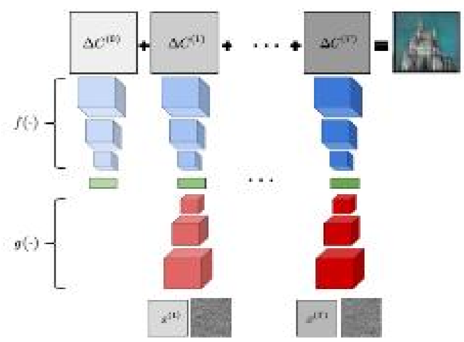

Figure 6. The structure of a Generative Recurrent Adversarial Network (GRAN)

In GRAN, the generator G consists of a recurrent feedback loop that takes a sequence of noise samples drawn from the prior distribution z ~ p(z) and draws the output at different time steps C_1, C_2,….,C_T . At each time step t, a sample z from the prior distribution is passed onto a function f(.)with the hidden state $h_{c,t}$ where $h_{c,t}$ represents the current encoded status of the previous drawing C_{t-1}. C_t is what is drawn to the canvas at time t and it contains the output of the function f(.). Moreover, the function g(.) is used to mimic the inverse of the function f(.). Accumulating the samples at each time step yields the final sample drawn to the canvas C. Ultimately, the function f(.) acts as adecoder and receives the input from the previous hidden state $h_{c,t}$ and noise sample z and the function g(.) acts as anencoder that provides a hidden representation of the outputC_{t-1} for time step t. The authors propose a new evaluation metric for generative models called Generative Adversarial Metric (GAM). The GRAN model’s performance was evaluated against MNIST, CIFAR10 and LSUN datasets with time steps T =1, 3, 5. It was found that GRAN with time steps T = 3and T = 5 performed better than GRAN with time step T = 1. Also, GRAN was compared against other generative models such as denoising VAE and DRAW on the MNIST dataset. It was also found that the samples generated by GRAN were discernible and did not overfit on the training data[14].

8) Bidirectional Generative Adversarial Networks(BiGAN)

Donahue et al. [20] proposed a method for learning the semantics in data distribution as well as its inverse mapping using these learnt feature representations for projecting data back into the latent space. The structure of a Bidirectional Generative Adversarial Network is shown in Fig. 7. As it can be seen from the Fig. 7, in addition to the generator G from the standard GAN framework, BiGAN includes an encoder E which maps the data x to latent representations z. The BiGAN discriminator D discriminates not only in the data space (x versus G(z)), but jointly in data and latent spaces (tuples (x,E(x)) versus (G(z), z)), where the latent component is either the encoder output E(x) or generator input z. Here, according to the objective of GANs, the BiGAN encoder E should learn to invert the generator G. The BiGAN training objective is defined as follows:[17]

$$\underset{G,E}{Min}\ \underset{D,E,G}{Max}\ V(D,E,G) = E_{x\sim p(x)} \underbrace{E_{z\sim PE(.|x)}[\log D(x,z)]}_{\log D(x,E(x))} + E_{z\sim p(z)} \underbrace{E_{z\sim PG(.|z)}[1-\log D(x,z)]}_{(1-\log D(G(z),z))} \qquad (9)$$

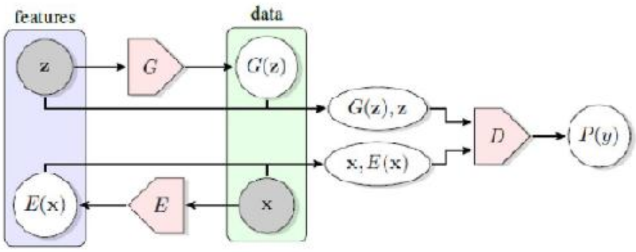

Fig. 7The architecture of BiGAN

9) SAGAN (Self-Attention Generative Adversarial Networks)

Attention is an algorithm used in sequence models, such as transformers [3]. SAGAN is a model that applies the attention algorithm to GAN [9]. The self-attention algorithm is shown in Figure 8..

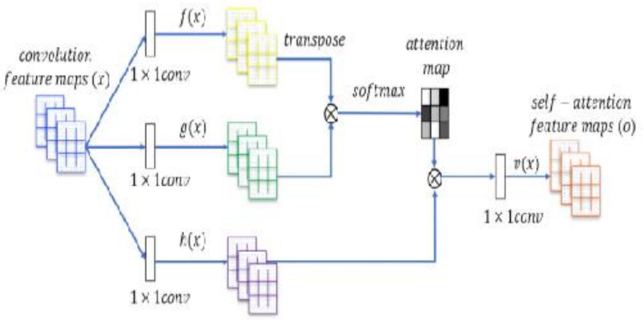

Figure 8. Self-attention algorithm of SAGAN (Self-Attention Generative Adversarial Networks) [9].

In a GAN without attention, the convolution feature map can process only local information. In order to connect pixel information on one side of the image to the other side, the channel must be increased to several convolutional layers, and the dimension of the image space must be reduced. This process, on the other hand, loses accurate location information instead of capturing high-level features. As such, it is inefficient for the model to learn the dependence between distant pixels. SAGAN solved the problem above by applying the attention algorithm to GAN. The outline of the proposed method is shown in Figure 4. As shown in Figure 8, attention focuses on different types of areas. SAGAN made a significant advancement using an attention mechanism that works similar to human perception.

10) Boundary Equilibrium GAN (BEGAN)

Boundary equilibrium GAN (BEGAN) keep-up an equilibrium that manages the trade-off between variety and superiority. The main goal behind BEGAN is to change the loss function. The Wasserstein distance between reconstruction loss of actual and synthesized images gives the real loss. In BEGAN, the discriminator works during training as autoencoder balances the process optimizing of G and D. The idea of making the discriminator as an autoencoder first proposed in BEGAN . BEGAN cost function is:

$$\left.\begin{aligned} L_D(x, z) &= D(x) - k_t\, D(G(z)) \text{ for } \theta_D \\ L_G(z) &= D(G(z)) \text{ for } \theta_{DG} \\ L_{k+1} &= k_t + \alpha\,(\gamma D(x) - D(G(z))) \text{ for k training} \end{aligned}\right\} \quad (10)$$

where LG represents the loss of the generator, LD represents the loss of the discriminator, L(x), L (G (z))represents the auto-encoder L1 loss of real, fake data, and equilibrium hyper-parameter "γ" respectively [24].

The added layer does not freeze but continues training. This algorithm was applied tothe LSUN (large scale scene understanding) dataset image [10].

11) Style GAN

StyleGAN uses a mixture of Progressive Growing of Generative Adversarial Networks and neural-style transfer technologies [8,11,49]. StyleGAN has been in the spotlight by creating full high definition-level results with several steps of control from the details of the image to the whole. StyleGAN solved the problem of latent space entanglement by proposing a method called AdaIN (Adaptive Instance Normalization) , which uses reference style bias and scale are used to adjust the mean and variance of feature map outputted from the layers within the synthesis network. The AdaIN layer prevents style information from leaking between layers. The style vector injected into each layer makes it affect only the features of that layer. This latent vector is better than the original vector.

The synthesis network is based on the PGGAN structure, and the style vector of the front layer of the synthesis network affects features larger than the style vector of the back layer. StyleGAN had full control over the image generated using latent vector and changed the style to various levels by changing the position of the vector in the synthesis network. To merge the two images, a vector is passed through the synthesis network and is converted . When deformation occurs early, styles such as posture, appearance, and glasses are transferred. When deformation occurs later, styles such as color and fine shape of the face are transferred. Both features of the image are maintained.

Finally, the StyleGAN structure adds noise behind each convolutional layer to capture areas such as hair position or face background. The noise injection location determines the fineness and roughness in the image [10].

StyleGAN2 comes with various improvements to image quality, efficiency, diversity, and disentanglement, and the results are incredibly improved. StyleGAN2 simply redesigns the normalization used in the generator of StyleGAN, which removes the artifacts such as blob-shaped artifacts that resemble water droplets. The StyleGAN2 achieves excellent results in face image synthesis and quality than StyleGAN [24].

III. APPLICATIONS OF GANS

In this section diverse applications of GANs are discussed like, medical diagnosis, text generation, hyperspectral image classification, etc., in detail.

A. Clinical diagnosis

In [27], a cycleGAN-based unified framework is discussed to standardize the intensity distribution of MRI images with different parameters coming from multiple groups. The effectiveness of the proposed method is investigated on T2-FLAIR image datasets. Qi et al. [28] developed a model using cascaded conditional GAN (C-cGANs) for automatic bi-ventricle segmentation of magnetic resonance images of the heart. In [29], the authors have addressed two issues while dealing with clinical data for cardiac disease diagnosis using ECG signals. First, they extracted the global features and then increased the stability of training to extract high-quality diverse samples. In [30], the authors used a GAN model as a data augmentation tool to generate synthetic data to tackle the imbalanced classification of multi-class ECG data. In [31] the authors developed a finger veinGAN(FV-GAN) framework that consists of two generators: an image generator that generates vein images from vein patterns and a pattern generator that maps vein images to vein patterns. Another useful resource for clinical diagnosis is ultrasound imaging. A two-stage GAN structure is devised to increase the image quality of hand-held or portable ultrasound devices. MRI and PET images are fused , the authors have proposed an algorithm based on Wasserstein GAN (MWGAN) to surmount the challenges involved in fusing images from multiple sources. A GAN-based unsupervised approach is proposed on the principles of anomaly detection for the diagnosis of lung cancer. The GAN (MDGAN) consists of a generator network and multiple discriminator networks. Ghassemi et al. [32] discussed a GAN-based model as a novel data augmentation method for multi-class classification of MR images. Conditional GAN (CGAN)-based denoising method that removes the noise in reduced radiation chest images and enhances the image quality for clinical diagnosis. Label smoothing GAN (LSGAN) for the classification of optical coherence tomography (OCT) images that can help in detecting and avoiding blindness at early stages. From the above discussion, it is observed that image reconstruction, image synthesis , segmentation, classification, abnormality detection, denoising, data augmentation, etc., are the novel tasks that were solved using GANs [26].

B. Image Generation

Recently, GANs has gained more and more momentum for generating naturalistic images through adversarial training. A novel Geometry-Aware GANs modelcalled GAGAN is proposed which incorporates geometric information into the image generation process. Adversarial image generation model, called LR-GAN is proposed to generate sharp images by considering both scene structure and context. LR-GAN combines foregrounds on the background in a contextually relevant style for generating more realistic images. Results have shown that LR-GAN outperforms DCGAN. Style and Structure GAN (S2-GAN) model is proposed in which a GANs is used to produce the image structure and then output is fed into the second GANs for considering the image style. Generation of realistic images has wide range of practical applications, such as anime character generation, image synthesis, super resolution, image editing and blending, inpainting, etc [38].

C. Intrusion detection

Yuan et al. [33] developed a randomized nonlinear image transformation method to alter and ruin the advanced patterns of attacking noise partly in the adversarial images. They employed a generative cleaning network to retrieve the lost content of the original image during the image transformation phase. The discriminator network is used to defend the classification process and trained not to detect any leftover noise patterns in the images. Extended Monte Carlo tree search (MTCS) algorithm using a GAN model that produces adversarial examples of cross-site scripting (XSS) attack traffic data. They added adversarial examples to an original dataset during the training phase. Also, they assigned a probability value that bypasses the adversarial image from the detector. Imbalance GAN (IGAN) framework to enhance the process of intrusion detection in adhoc networks. The architecture consists of a feed forward network to extract the features, an IGAN with a filter to synthesize the abnormal class samples and a deep neural network to perform the classification task.

D. Fault diagnosis

Fault detection is an important task in the field of control engineering to capture the malfunctioning of machine to avoid machine failure and human loss. A paper discussed in [26] devised a model for monitoring of machine condition and fault diagnosis using sensor data. The model design is based on ACGAN. They also used a novel quantitative method for the evaluation of generated sensor signal data and used time domain and frequency domain characteristics for assessing the diversity of generated samples. A paper discussed in [26] addressed the automated detection and diagnosis (AFDD) of fault training data using an unsupervised framework. A support vector machine (SVM) is trained as a binary classifier on the augmented dataset. In the detection phase, it identifies the faulty state, and in the diagnosis phase, it classifies the type of fault. Another GAN-based framework (CVAE-GAN) using the conditional variational autoencoder (CVAE) for imbalanced fault diagnosis in a planetary gear box. A framework that works in two stages for imbalanced fault diagnosis of rotating machines is also discussed in this paper [26]. Investigations on CWRU and Bogie datasets proved the effectiveness of the proposed model. [26].

E. Applications to Speech and Language Processing

Recently, there are some GANs based applications for speech and language processing. Li et al. [1] use GANs to capture the relevance of dialogue and generate corresponding text. Zhang et al. [6] propose to generate realistic sentence with GANs, by using long short-term memory and convolutional neural networks for adversarial training. SeqGAN [2] employs reinforcement learning to generate speech and language, poem, and music. Pascual et al. [7] propose the use of GANs for speech enhancement, called SEGAN. Reed et al. [8] present a GANs-based method for generating images from the text. Text encoding is used by both the generator and the discriminator. Zhang et al. [9] propose StackGAN for text to photorealistic image synthesis. Given text descriptions, Stage-I of StackGAN sketches rough shapes and basic colors of objects, yielding low resolution images. Stage-II of StackGAN takes Stage-I results and text descriptions as inputs, and generates high resolution images with photorealistic details [5].

F. Geoscience and remote sensing

The remote sensing community suffers from the problem of limited samples. Due to this, the training process end up with over-fitting problem. In Shi et al. automatically generated building footprints from the satellite images using a conditional GAN. Zhu et al. presented two schemes for the classification of 1D and 3D hyperspectral images. First, a spectral classifier is modelled using a 1D GAN. Second, a spectral-spatial classifier is designed using a 3D GAN. The authors have used the GAN model as a regularization unit to alleviate the effect of overfitting due to limited samples. Cloud obstruction is a conventional problem in remote sensing object detection field. To overcome this problem, a deep convolutional-based GAN model is proposed with a novel in painting loss function. spatial-spectral GAN model discussed, that performs a multi-class classification of hyperspectral images. This model addresses two issues of the classification process. First, it addresses the inability of the discriminator in multi-class classification and Second, consideration of spatial and spectral information in the classification of hyperspectral images. Variational GAN is proposed using a semi-supervised method to classify hyperspectral images with limited labels. Multibranch conditional GAN (MCGAN) model is devised to increase data for objection in remote sensing images. The MCGAN architecture consists of one generator, three discriminators, and a classifier that build using deep CNNs [26].

G. Video Generation

Vondrick et al.[35] have made great progress in the video field, 32-frame resolution 64×64 realistic video can be generated, depicting golf courses, beaches, train stations and newborns. After testing, 20 of the markers did not recognize the authenticity of these videos. MD-GAN(Multistage Dynamic GAN) predicts future video frames with the proposed model through a given first frame image. In its two-stage model, the first stage a time-lapse video with realistic content can be generated; the second stage the results of the first stage is optimized. By using this model, realistic time-lapse photographic video with a resolution of 128×128 up to 32 frames can be generated [34].

H. Auxiliary Automatic Driving

Santana et al.[36] implemented the assisted autonomous driving with GAN. First, an image is generated, that is consistent with the distribution of the official traffic scene image, and then a transition model is trained based on the cyclic neural network to predict the next traffic scene[34].

I. Drug Discovery

Researchers from Insilico Medicine [37] proposed an approach of artificially intelligent drug discovery using GANs. They attempt to train the Generator to sample drug candidates for a given disease as precisely as possible to existing drugs from a Drug Database. After training, it’s possible to generate a drug for a previously incurable disease using the Generator, and using the Discriminator to determine whether the sampled drug actually cures the given disease.

J. Molecule Development in Oncology

Insilico Medicine showed the pipeline of generating new anticancer molecules with a defined set of parameters. Researchers proposed an Adversarial Autoencoder (AAE) model for identification and generation of new compounds based on available biochemical data. AAE was trained using Growth Inhibition percentage data, drug concentrations, and fingerprints as inputs [17].

K. Deepfakes

GANs generates high-quality data (text, image, video, speech) which can be used to generate fake information, i.e., deepfakes. Deepfakes algorithms generate fake images and videos whose authenticity cannot be distinguished from the real data. Deepfakes mainly have been explored for facial manipulation which can be classified into three categories, (1) face synthesis is about creating non-existent realistic faces using GANs ; (2) face swap is about swapping faces; (3) facial attributes and expressionis about manipulating attributes of the face, such as color tone, age, gender, etc. There is a need of algorithms that can detect fake content automatically.A novel solution is proposed to detect deepfakes by identifying subtle visual artifacts in the image.

L. Handling privacy issues in data generation

GANs has also been used for generating synthetic data which can be made available publicly instead of real data. However, an adversary can get the training set membership through the GANs model and generated synthetic data. A model memorizes the training samples which leads the issue of privacy and makes the GANs model vulnerable. It is easy to identify the training samples through the observation that D is more likely to classify training samples as real rather than samples not present in the training data. To achieve the privacy, records in the data can be de-identified, but now de-identified records can be re-identified by relating them to other identifiable datasets. Differentially Private Generative Adversarial Network (DP-GAN) is proposed to guarantee differential privacy of D by introducing noise during the model optimization. But, proposed model does not maintain trade-off between sample quality and diversity while supporting differential privacy. Hence, model is not useful for practical applications. An architecture design is proposed maintaining trade-off between privacy and sample quality.

M. Fairness

Recently, achieving fairness in learning models have gained momentum in machine learning for several applications, such as granting loan, shortlisting candidate for interview, etc. A GANs-based approach is proposed to generate fair dataset w.r.to sensitive attributes in allocative decision making from the real dataset for model training. Reinforcement learning based race balance network is proposed to handle the bias in the data. Existing deep learning models are pointed out for face recognition encode gender information implicitly. A new direction introduced for handling the issue of missing fairness in the outcome. It is identified that existing deep learning classifiers trained to generate diagnostic labels from X-ray images are biased w.r.to sensitive attributes[38].

IV. CONCLUSION

This paper summarizes the background of GANs, expounds its basic principles, and introduces its derived model and its application in various fields. Despite being a relatively newly-proposed approach, GANs have been widely accepted in the machine learning community, with the quantity of research carried out in respect of them increasing at a significant pace. The ability of GAN to generate “infinite” new samples from potential distribution has great application value in many fields such as image and vision computing, speech and language processing, and information security. The author also investigate the development trends of GANs, there are many opportunities for the developments in both theory and algorithms, and by using deep networks, there are vast opportunities for new applications.

REFERENCES

[1] J. W. Li, W. Monroe, T. L. Shi, S. Jean, A. Ritter, and D. Jurafsky, “Adversariallearning for neural dialogue generation,” arXiv: 1701.06547,2017.
[2] L. T. Yu, W. N. Zhang, J. Wang, and Y. Yu, “SeqGAN: sequencegenerative adversarial nets with policy gradient,” arXiv: 1609.05473,2016.
[3] W. W. Hu and Y. Tan, “Generating adversarial malware examples forblack-box attacks based on GAN,” arXiv: 1702.05983, 2017.
[4] M. Chidambaram and Y. J. Qi, “Style transfer generative adversarialnetworks: Learning to play chess differently,” arXiv: 1702.06762, 2017.
[5] X. Kunfeng Wang, Chao Gou, Yanjie Duan, Yilun Lin, Xinhu Zheng,and Fei-Yue Wang,. ieee/caa Journal Of Automatica Sinica, vol. 4, no. 4, 2017
[6] Y. Z. Zhang, Z. Gan, and L. Carin, “Generating text via adversarialtraining,” Proc. Workshop on Adversarial Training, Barcelona, Spain,2016.

[7] S. Pascual, A. Bonafonte, and J. Serr´a, "SEGAN: speech enhancementgenerative adversarial network," arXiv: 1703.09452, 2017.

[8] S. Reed, Z. Akata, X. C. Yan, L. Logeswaran, B. Schiele, and H. Lee,"Generative adversarial text to image synthesis," in Proc. 33rd Int. Conf.Machine Learning, New York, NY, USA, vol. 48, pp. 1060¡1069, 2016.

[9] H. Zhang, T. Xu, H. S. Li, S. T. Zhang, X. L. Huang, X. G. Wang, andD. Metaxas, "StackGAN: text to photo-realistic image synthesis withstacked generative adversarial networks," arXiv: 1612.03242, 2016

[10] Sung-Wook Park, Jae-Sub Ko, Jun-Ho Huh, and Jong-Chan Kim, Review on Generative Adversarial Networks: Focusing on Computer Vision and Its Applications, https://www .mdpi.com/ journal/electronics Electronics 2021, 10, 1216

[11] Gouk, H.; Frank, E.; Pfahringer, B.; Cree, M. Regularisation of Neural Networks by Enforcing Lipschitz Continuity. arXiv 2020,arXiv:1804.04368. Available online: https://arxiv.org/abs/1804.04368 (accessed on 26 May 2020).

[12] Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A.C. Improved training of wasserstein gans. In Proceedings of theAdvances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY,USA, 2018; pp. 5767–5777.

[13] Ioffe, S.; Szegedy, C. Batch normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv 2020,arXiv:1502.03167. Available online: https://arxiv.org/abs/1502.03167 (accessed on 28 May 2020).

[14] Saifuddin Hitawala,Comparative Study on Generative Adversarial Networks, arXiv:1801.04271v1 [cs.LG]

[15] Chen, Xi, Duan, Yan, Houthooft, Rein, Schulman, John, Sutskever, Ilya, and Abbeel, Pieter. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In Advances inNeural Information Processing Systems, pp. 2172–2180,2016.

[16] Mirza, Mehdi and Osindero, Simon. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.

[17] Hamed Alqahtani, Manolya Kavakli Thorne, Gulshan Kumar, Applications of Generative Adversarial Networks (GANs): An UpdatedReview, Archives of Computational Methods in Engineering, https://doi.org/ 10.1007/ s11831-019-09388-y, 2019

[18] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-FarleyD, Ozair S, Courville A, Bengio Y (2014) Generative adversarialnets. In: Advances in neural information processing systems, pp2672–2680

[19] Makhzani A, Shlens J, Jaitly N, Goodfellow I, Frey B (2015)Adversarial autoencoders. arXiv preprint arXiv :1511.05644

[20] Donahue J, Krahenbuhl P, Darrell T (2016) Adversarial featurelearning. arXiv preprint arXiv :1605.09782

[21] Jie Gui, Zhenan Sun, Yonggang Wen, Dacheng Tao, Jieping Ye, A Review on Generative Adversarial Networks:
Algorithms, Theory, and Applications, Journal Of Latex Class Files, Vol. 14, NO. 8, 2015, arXiv:2001.06937v1 [cs.LG] 2020

[22] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I.Sutskever, andP. Abbeel, "Infogan: Interpretable representation learning byinformation maximizing generative adversarial nets," in NeuralInformation Processing Systems, pp. 2172–2180, 2016.

[23] Bhaskar Ghosh, Indira Kalyan Dutta, Michael Totaro, Magdy Bayoumi, A Survey on the Progression and Performance ofGenerative Adversarial Networks, DOI: 10.1109/ICCCNT49239.2020.9225510, 2020

[24] Abdul Jabbar, Xi Li*, and Bourahla Omar, A Survey on Generative Adversarial Networks:Variants, Applications, and Training

[25] Comparative Study on Generative Adversarial Networks, Saifuddin Hitawala

[26] M. R. Pavan Kumar, Prabhu Jayagopal, Generative adversarial networks: a survey on applicationsand challenges, Springer, International Journal of Multimedia Information Retrievalhttps://doi.org/10.1007/s13735-020-00196-w, 2020

[27] Gao Y, Liu Y, Wang Y, Shi Z, Yu J (2019) A universal intensity standardization method based on a many-to-one weak-paired cycle generative adversarial network for magnetic resonance images. IEEE Trans Med Imaging 38(9):2059–2069

[28] Qi L, Zhang H, Tan W, Qi S, Xu L, Yao Y, Qian W (2019) Cascaded conditional generative adversarial networks with multiscale attention fusion for automated bi-ventricle segmentation in cardiac mri. IEEE Access 7:172305–172320

[29] Ye F, Zhu F, Fu Y, Shen B (2019) Ecg generation with sequencegenerative adversarial nets optimized by policy gradient. IEEEAccess 7:159369–159378

[30] Shaker AM, Tantawi M, Shedeed HA, Tolba MF (2020) Generalization of convolutional neural networks for ecg classification using generative adversarial networks. IEEE Access 8:35592– 35605

[31] Yang W, Hui C, Chen Z, Xue JH, Liao Q (2019a) Fv-gan: fingervein representation using generative adversarialnetworks. IEEETrans Inf Forensics Secur 14(9):2512–2524

[32] He K, Zhang X, Ren S, Sun J (2016) Deep residuallearning forimage recognition. In: Proceedings of the IEEE conference oncomputer vision and pattern recognition, pp 770–778
[33] Yuan J, He Z (2020) Adversarial dual network learning with randomizedimage transform for restoring attackedimages. IEEEAccess 8:22617–22624
[34] Liang Gonog1,2 and Yimin Zhou, A Review: Generative Adversarial Networks, 978-1-5386-9490-9/19/$31.00_c 2019 IEEE
[35] C. Vondrick, H. Pirsiavash, and A. Torralba, "Generating videos withscene dynamics," in Advances In NeuralInformation Processing Systems,2016, pp. 613–621.
[36] E. Santana and G. Hotz, "Learning a driving simulator," arXiv preprintarXiv:1608.01230, 2016.
[37] Medicine I (2019) Artificial intelligence for drug discovery, biomarkerdevelopment and aging research. URL https ://insil ico.com/
[38] Divya Saxena, Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions