Supplementary material for "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications"

Jie Gui, *Senior Member, IEEE*, Zhenan Sun, *Senior Member, IEEE*, Yonggang Wen, *Fellow, IEEE*, Dacheng Tao, *Fellow, IEEE*, Jieping Ye, *Fellow, IEEE*

Certain algorithms and open research problems are discussed in this supplementary material.

1 ALGORITHMS

Due to space limitations, some algorithms [1]–[20] could not be shown in the main part of the paper. In this section, we introduce generative adversarial network (GAN) training, GAN evaluation, and task-driven GANs.

1.1 GAN Training

1.1.1 *Objective function*

1.1.1.1 Hinge loss based GANs:
Hinge loss based GANs were proposed and used in [21]–[23], whose discriminator loss and generator loss are

$$\begin{aligned} & E_{x\sim p_{data}(x)}\left[\min(0,-1+D\left(x\right))\right] \\ & \quad +E_{z\sim p_z(z)}\left[\min(0,-1-D\left(G(z)\right))\right], \end{aligned} \tag{1}$$

$$-E_{z\sim p_z(z)}\left[D\left(G(z)\right)\right]. \tag{2}$$

Softmax cross-entropy loss [24] is also used in GANs.

1.1.1.2 Energy-based generative adversarial network (EBGAN):
The EBGAN discriminator is considered as an energy function that assigns higher energy to fake ("generated") examples and lower energy to real examples. For the energy function, please refer to [25], which includes a corresponding tutorial. Given a positive margin m, the loss functions for EBGAN can be defined as follows:

$$\mathcal{L}_D(x,z) = D(x) + [m - D(G(z))]^+, \tag{3}$$

$$\mathcal{L}_G(z) = D(G(z)), \tag{4}$$

where $[y]^+ = \max(0,y)$ is the rectified linear unit (ReLU) function. Note that in the original GANs, the discriminator D gives a high score to real examples and a low score to the generated ("fake") examples. However, the discriminator in EBGAN attributes lower energy (score) to the real examples and higher energy to the generated examples. EBGAN is more stable than the original GANs during training.

1.1.1.3 Boundary equilibrium generative adversarial networks (BEGAN):
Similar to EBGAN [26], dual-agent GAN (DA-GAN) [27], [28], and margin adaptation for GANs (MAGANs) [29], BEGAN also uses an autoencoder as the discriminator. Using proportional control theory, a novel equilibrium method was proposed for BEGAN to balance the generator and discriminator during training that is fast, stable, and robust to parameter changes.

1.1.1.4 Mode regularized generative adversarial networks (MDGAN) :
Che et al. [30] argued that unstable training and mode collapse in GANs are due to the special functional shapes of the trained discriminators in high-dimensional spaces, which can cause training to become stuck or push the probability mass in the wrong direction—toward a higher concentration than that of the real data distribution. Che et al. [30] introduced several methods to regularize the objective that can stabilize GAN models during training. The key idea of MDGAN is to utilize an encoder $E\left(x\right): x \to z$ to produce the latent variable z for the generator G rather than utilizing noise. This procedure has two advantages:

- The encoder guarantees the correspondence between z ($E(x)$) and x, which makes G capable of covering diverse modes in the data space. Therefore, it prevents the mode collapse problem.
- Because encoder reconstruction can add more information to the generator G, the discriminator D cannot easily distinguish between the real and generated examples.

J. Gui is with the School of Cyber Science and Engineering, Southeast University, 211100, China and Purple Mountain Laboratories, Nanjing 210000, China. He was previously with the Department of Computational Medicine and Bioinformatics, University of Michigan, USA (e-mail: guijie@seu.edu.cn).
Z. Sun is with the Center for Research on Intelligent Perception and Computing, Chinese Academy of Sciences, Beijing 100190, China (e-mail: znsun@nlpr.ia.ac.cn).
Y. Wen is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ygwen@ntu.edu.sg).
D. Tao is with JD Explore Academy, China and also with the School of Computer Science in the University of Sydney, Australia (e-mail: dacheng.tao@gmail.com).
J. Ye is with University of Michigan, Ann Arbor (e-mail: jpye@umich.edu).

The loss functions used for the generator and the encoder of MDGAN are

$$\begin{aligned} \mathcal{L}_G &= -E_{z\sim p_z(z)}\left[\log\left(D\left(G\left(z\right)\right)\right)\right] \\ &+E_{x\sim p_{data}(x)}\left[\begin{array}{c} \lambda_1 d\left(x, G\circ E\left(x\right)\right) \\ +\lambda_2 \log D\left(G\circ E\left(x\right)\right) \end{array}\right], \end{aligned} \tag{5}$$

$$\mathcal{L}_E = E_{x\sim p_{data}(x)}\left[\begin{array}{c} \lambda_1 d\left(x, G\circ E\left(x\right)\right) \\ +\lambda_2 \log D\left(G\circ E\left(x\right)\right) \end{array}\right], \tag{6}$$

where both λ_1 and λ_2 are free-tuning parameters, d is the distance metric such as the Euclidean distance, and $G\circ E\left(x\right) = G\left(E\left(x\right)\right)$.

1.1.1.5 Unrolled GAN:
Metz et al. [31] introduced a GAN stabilization technique by defining the generator objective with regard to an unrolled optimization of the discriminator. This allows training to be adjusted between the current value of the discriminator, which is usually unstable and leads to poor solutions, and the optimal discriminator solution in the generator objective, which is perfect but infeasible in real applications. Let $f\left(\theta_G, \theta_D\right)$ denote the objective function of the original GANs.

Then, a local optimal solution of the discriminator parameters θ_D^* can be expressed as the fixed point of an iterative optimization procedure,

$$\theta_D^0 = \theta_D, \tag{7}$$

$$\theta_D^{k+1} = \theta_D^k + \eta^k \frac{df\left(\theta_G, \theta_D^k\right)}{d\theta_D^k}, \tag{8}$$

$$\theta_D^*\left(\theta_G\right) = \lim_{k\to\infty}\theta_D^k, \tag{9}$$

where η^k is the learning rate. Unrolling for K steps creates a surrogate objective for updating the generator:

$$f_K\left(\theta_G, \theta_D\right) = f\left(\theta_G, \theta_D^K\left(\theta_G, \theta_D\right)\right). \tag{10}$$

When $K = 0$, this objective is the same as the standard GAN objective. When $K \to \infty$, this objective becomes the true generator objective function $f\left(\theta_G, \theta_D^*\left(G\right)\right)$. By adjusting the number of unrolling steps K, we can interpolate between standard GAN training dynamics, with their related pathologies, and the more expensive gradient descent based on the true generator loss. The generator and discriminator parameter updates of the unrolled GAN using this surrogate loss are as follows:

$$\theta_G = \theta_G - \eta\frac{df_K\left(\theta_G, \theta_D\right)}{d\theta_G}, \tag{11}$$

$$\theta_D = \theta_D + \eta\frac{df\left(\theta_G, \theta_D\right)}{d\theta_D}. \tag{12}$$

Metz et al. [31] showed how this method solves mode collapse, stabilizes GAN training, and increases the diversity and coverage of the distribution generated by the generator.

1.1.2 Structure

1.1.2.1 Laplacian generative adversarial networks (LAPGAN) and SinGAN:
LAPGAN [32] was proposed to produce higher resolution images than the original GANs. LAPGAN uses a cascade of CNNs within a Laplacian pyramid framework [33] to generate high-quality images.

SinGAN [34] learns a generative model from a single natural image. SinGAN uses a pyramid of fully convolutional GANs, each of which learns the patch distribution at a different image scale. Similar to SinGAN, internal GAN (InGAN) [35] also learns a generative model from a single natural image.

1.1.2.2 BigGANs and StyleGAN:
Both BigGANs [36] and StyleGAN [37], [38] made substantial advances in GAN quality.

BigGANs [36] is a large-scale tensor processing unit (TPU) implementation of GANs that is similar to self-attention GAN (SAGAN) but greatly upscaled. BigGANs successfully generated images with high resolution (up to 512 by 512 pixels). When insufficient data are available, replicating the BigGAN results from scratch is challenging. Lucic et al. [39] proposed training a BigGAN high-quality model with fewer labels. BiGAN with BigGAN generator (BigBiGAN) [40], extends the idea to representation learning by adding an encoder and modifying the discriminator. BigBiGAN achieved state-of-the-art performances on both unsupervised representation learning on ImageNet and unconditional image generation.

In the original GANs [41], G and D were multilayer perceptron (MLP) models. Karras et al. [37] proposed a StyleGAN architecture for GANs that won the CVPR 2019 best paper honorable mention. The StyleGAN generator is a high-quality generator that has been used for other generation tasks, such as generating faces. This model is particularly exciting because it allows us to separately control different factors, such as hair, age and sex, that are involved in controlling the appearance of the final example. StyleGAN [37] has also been used to generate high-resolution fashion model images wearing custom outfits [42].

1.1.2.3 Hybrids of autoencoders and GANs:
An autoencoder is a type of neural network used to learn efficient data codings in an unsupervised manner. The autoencoder has an encoder and a decoder. The encoder aims to learn a representation (encoding) for a set of data, $z = E(x)$, typically for dimensionality reduction. The decoder aims to reconstruct the data $\hat{x} = g\left(z\right)$. That is, the decoder tries to generate a representation from the reduced encoding that is as close as possible to the original input x.

GANs with an autoencoder: An adversarial autoencoder (AAE) [43] is a probabilistic autoencoder based on GANs. Adversarial variational Bayes (AVB) [44], [45] was proposed to unify variational autoencoders (VAEs) and GANs. Sun et al. [46] proposed an UNsupervised Image-to-image Translation (UNIT) framework based on GANs and VAEs. Hu et al. [47] aimed to establish formal connections between GANs and VAEs via a new formulation. By combining a VAE with a GAN, Larsen et al. [48] utilized learned feature representations in the GAN discriminator as the basis for the VAE reconstruction. Therefore, [48] replaced

TABLE 1: Different generator losses and discriminator losses of GANs

Method	Generator loss: $\min_G$	Discriminator loss: $\max_D$
Original minimax game	$E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right)\right)\right)\right]$	$E_{x\sim p_{data}(x)}\left[\log D\left(x\right)\right]+E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right)\right)\right)\right]$
Non-saturating game	$E_{z\sim p_z(z)}\left[-\log\left(D\left(G\left(z\right)\right)\right)\right]$	$E_{x\sim p_{data}(x)}\left[\log D\left(x\right)\right]+E_{z\sim p_z(z)}\left[-\log\left(D\left(G\left(z\right)\right)\right)\right]$
Maximum likelihood game	$E_{z\sim p_z(z)}\left[-D\left(G\left(z\right)\right)/\left(1-D\left(G\left(z\right)\right)\right)\right]$	$E_{x\sim p_{data}(x)}\left[\log D\left(x\right)\right]+E_{z\sim p_z(z)}\left[-D\left(G\left(z\right)\right)/\left(1-D\left(G\left(z\right)\right)\right)\right]$
InfoGAN	$E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right)\right)\right)\right]-\lambda L_I(c;Q)$	$E_{x\sim p_{data}(x)}\left[\log D\left(x\right)\right]+E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right)\right)\right)\right]$
Conditional GANs	$E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\mid y\right)\right)\right)\right]$	$E_{x\sim p_{data}(x)}\left[\log D\left(x\mid y\right)\right]+E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\mid y\right)\right)\right)\right]$
LSGANs	$E_{z\sim p_z(z)}\left[\left(D\left(G\left(z\right)\right)-c\right)^2\right]$	$E_{x\sim p_{data}(x)}\left[\left(D\left(x\right)-b\right)^2\right]+E_{z\sim p_z(z)}\left[\left(D\left(G\left(z\right)\right)-a\right)^2\right]$
Hinge loss based GANs	$-E_{z\sim p_z(z)}\left[D\left(G(z)\right)\right]$	$E_{x\sim p_{data}(x)}\left[\min(0,-1+D\left(x\right))\right]+E_{z\sim p_z(z)}\left[\min(0,-1-D\left(G(z)\right))\right]$
EBGAN	$D(G(z))$	$D(x)+\left[m-D(G(z))\right]^+$
WGAN	$-E_{z\sim p_z(z)}\left[D\left(G\left(z\right)\right)\right]$	$E_{x\sim p_{data}(x)}\left[D\left(x\right)\right]-E_{z\sim p_z(z)}\left[D\left(G\left(z\right)\right)\right]$
RSGAN	$-E_{(x_r,x_g)}\left[\log\left(\sigma\left(C\left(x_g\right)-C\left(x_r\right)\right)\right)\right]$	$-E_{(x_r,x_g)}\left[\log\left(\sigma\left(C\left(x_r\right)-C\left(x_g\right)\right)\right)\right]$
BiGAN	$E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right),z\right)\right)\right]$	$E_{x\sim p_{data}(x)}\left[\log D\left(x,E(x)\right)\right]+E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right),z\right)\right)\right]$

element-wise errors with feature-wise errors to better capture the data distribution while offering translation invariance. Rosca et al. [49] also proposed variational approaches for auto-encoding GANs. By adopting an encoder-decoder architecture for the generator, the disentangled representation GAN (DR-GAN) [50] addresses pose-invariant face recognition, which is a challenging problem due to the drastic image changes for each diverse pose.

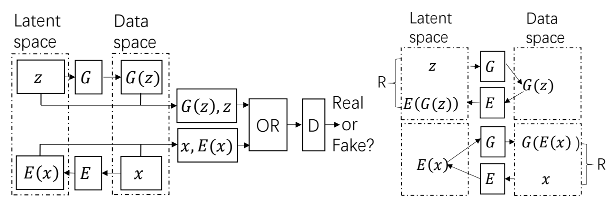
GANs with an encoder: References [51], [52] added only an encoder to GANs. The original GANs [41] cannot learn the inverse mapping—projecting data back into the latent space. To solve this problem, Donahue et al. [51] proposed bidirectional GANs (BiGANs), which learn this inverse mapping through the encoder, and showed that the resulting learned feature representation is useful. Similarly, Dumoulin et al. [53] proposed an adversarially learned inference (ALI) model that also utilizes an encoder to learn the latent feature distribution. The structures of BiGAN and ALI are shown in Fig. 1(a). In addition to the discriminator and generator, BiGAN also has an encoder used to map the data back to the latent space. The input of the discriminator is a data pair composed of the data and their corresponding latent code. For real data x, the pair is $x, E(x)$, where $E(x)$ is obtained from the encoder E. For generated data $G(z)$, the pair is $G(z), z$, where z is the noise vector used to generate $G(z)$ through the generator G. Similar to the objective function of the original GANs, the objective function of BiGAN is

$$\begin{aligned}\min_{G,E}\max_{D}\ V\left(D,E,G\right) &= E_{x\sim p_{data}(x)}\left[\log D\left(x,E(x)\right)\right] \\ &+E_{z\sim p_z(z)}\left[\log\left(1-D\left(G\left(z\right),z\right)\right)\right].\end{aligned} \tag{13}$$

The generator in [51], [52] can be considered a decoder since the generator maps vectors from the latent space to the data space, which is the function performed by a decoder. Different generator losses and discriminator losses of GANs are summarized in Table 1.

Similar to utilizing an encoding process to model the distribution of latent examples, Gurumurthy et al. [54] modeled the latent space as a mixture of Gaussians and learned the mixture components that maximize the likelihood of the generated examples under the data generating distribution.

In an encoding-decoding model, the output (also known as a reconstruction) should be similar to the input in the ideal case. Generally, however, the fidelity of reconstructed examples synthesized by BiGAN/ALI is poor. By imposing an additional adversarial cost on the distribution of data examples and their reconstructions [55], the fidelity of examples can be improved. Arora et al. [56] reported on the theoretical limitations of encoder-decoder GAN architectures such as BiGANs/ALI. Creswell et al. [57] inverted the GAN generator. Other related methods include variational discriminator bottleneck (VDB) [58] and MDGAN (detailed in Paragraph 1.1.1.4).



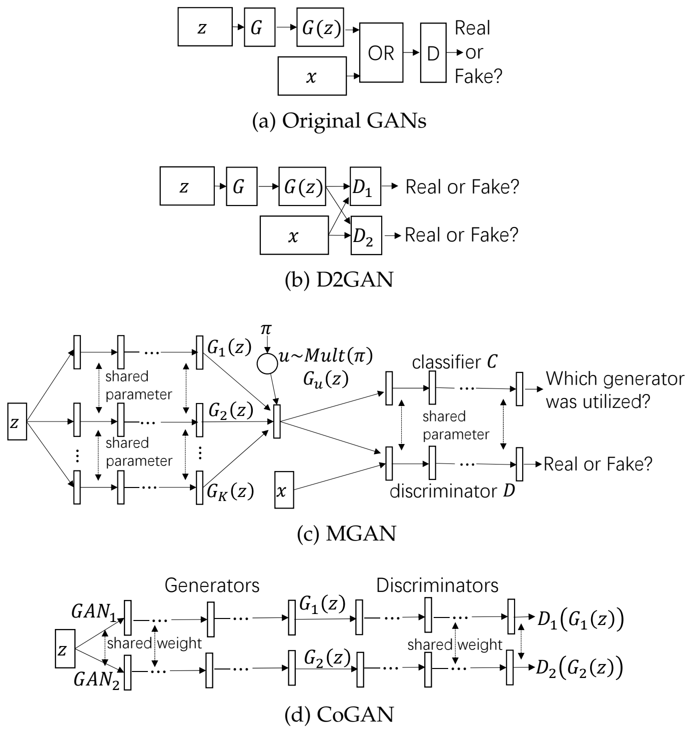
(a) BiGAN/ALI (b) AGE

Fig. 1: Structures of (a) BiGAN and ALI and (b) AGE.

Combination of a generator and an encoder: In contrast to previous autoencoder and GAN hybrids, the adversarial generator-encoder (AGE) network [52] is set up directly between the generator and the encoder, and no external mappings are learned during the training process. The structure of AGE is shown in Fig. 1(b), where R is the reconstruction loss function. AGE includes two reconstruction losses: between the latent variables z and $E(G(z))$, and between the data x and $G(E(x))$. AGE is similar to CycleGAN. However, there are two differences between them.

- CycleGAN [59] is used for two image modalities such as grayscale and color, while AGE acts between the latent space and the true data space.
- There is a discriminator for each modality in CycleGAN, but no discriminator exists in AGE.

1.1.2.4 Multi-discriminator learning:
GANs include both a discriminator and a generator. Different from GANs, dual discriminator GAN (D2GAN) [60] has **a generator and two binary discriminators**. D2GAN is analogous to a minimax game, wherein one discriminator assigns high scores to examples from the generated distribution while the other discriminator favors data from the true distribution. The generator generates data to fool both discriminators. Reference [60] developed a theoretical analysis to show that when given the optimal discriminators, optimizing the D2GAN generator minimizes both the



(a) Original GANs

(b) D2GAN

(c) MGAN

(d) CoGAN

Fig. 2: Structures of the original GANs [41], D2GAN [60], MGAN [63], and CoGAN [66].

Kullback-Leibler (KL) and the reverse KL divergences between the true and generated distributions, thus effectively overcoming the mode collapse problem. The generative multi-adversarial network (GMAN) [61] further extended GANs to one generator with multiple discriminators. Albuquerque et al. [62] performed multi-objective training of GANs with multiple discriminators.

1.1.2.5 Multi-generator learning:
The multi-generator GAN (MGAN) [63] was proposed to train GANs with a mixture of generators to avoid the mode collapse problem. More specifically, MGAN has **one binary discriminator, K generators, and a multi-class classifier**. The distinguishing feature of MGAN is that the generated examples are produced from multiple generators, but only one is randomly selected as the final output, similar to the mechanism of a probabilistic mixture model. The classifier shows which generator a generated example is from.

The model most closely related to MGAN is multi-agent diverse GAN (MAD-GAN) [64]. A discussion of the difference between MGAN and MAD-GAN can be found in [63]. SentiGAN [65] uses a mixture of generators and a multi-class discriminator to generate sentimental texts.

1.1.2.6 Multi-GAN learning:
Coupled GAN (CoGAN) [66] was proposed to learn a joint distribution of two-domain images. CoGAN is composed of a pair of GANs (GAN_1 and GAN_2), each of which synthesizes images from one domain. Two GANs that consider structure and style were proposed in [67] based on conditional GANs (cGANs). Causal fairness-aware GANs (CFGAN) [68] use two generators and two discriminators to generate fair data. The structures of the original GANs, D2GAN, MGAN, and CoGAN are shown in Fig. 2.

1.2 Evaluating GANs

In this subsection, we show the evaluation metrics [69], [70] that are typically used for GANs.

1.2.1 *Inception score (IS)*

The IS was proposed in [71] and uses the inception model [72] for every generated image to obtain the conditional label distribution $p\left(y|x\right)$. Images that contain meaningful objects should have a conditional label distribution $p\left(y|x\right)$ with low entropy. Furthermore, the model is expected to produce diverse images. Therefore, the marginal $\int p\left(y|x=G\left(z\right)\right)dz$ should have high entropy. In combination with these two requirements, the IS is as following:

$$\exp(E_x KL\left(p\left(y|x\right)||p\left(y\right)\right)), \tag{14}$$

where the exponential results allow easy comparison of the values.

A higher IS indicates that the generative model produces high-quality examples and that the generated examples are diverse. However, the IS metric also has disadvantages. If the generative model falls into mode collapse, the IS might remain high, while the true situation is poor. To address this issue, an independent Wasserstein critic [73] was proposed and trained independently on the validation dataset to assess mode collapse and overfitting.

1.2.2 *Mode score (MS)*

The MS [30], [74] is an improved version of the IS. In contrast to the IS, the MS measures the dissimilarity between the real and generated distributions.

1.2.3 *Fréchet inception distance (FID)*

The FID has also been proposed [75] to evaluate GANs. For a suitable feature function ϕ (the default feature is the inception network convolutional feature), FID models $\phi\left(p_{data}\right)$ and $\phi\left(p_g\right)$ as Gaussian random variables with empirical means $\mu_r,\ \mu_g$ and empirical covariance $C_r,\ C_g$ and computes

$$\begin{aligned} FID(p_{data}, p_g) &= \|\mu_r - \ \mu_g\| \\ &+ tr\left(C_r + \ C_g - 2(C_r C_g)^{1/2}\right), \end{aligned} \tag{15}$$

which is the Fréchet distance (also called the Wasserstein-2 distance) between the two Gaussian distributions. However, the IS and FID do not handle overfitting well. To mitigate this problem, the kernel inception distance (KID) was proposed in [76]. Furthermore, metrics such as FID are insufficient for differentiating between diverse failure cases because they produce only one-dimensional scores. To solve this problem, [77] proposed a new definition of precision and recall for distributions that disentangled the divergence into two separate dimensionalities. An improved precision and recall evaluation metric for assessing generative models was also proposed in [78].

1.2.4 *Multi-scale structural similarity (MS-SSIM)*

SSIM [79] has been used to measure the similarity between two images. Different from the single-scale SSIM measure, the MS-SSIM [80] was proposed for multi-scale image quality assessment. It quantitatively evaluates image similarity by attempting to predict human perceptual similarity judgment. The MS-SSIM values range from 0.0 to 1.0, and lower MS-SSIM values mean perceptually more dissimilar images. References [81], [82] used MS-SSIM to measure the diversity of fake data. Reference [83] suggested that MS-SSIM should be considered only in combination with the FID and IS metrics for testing example diversity.

1.2.5 Human eYe Perceptual Evaluation (HYPE)

The authors of [84] constructed the HYPE, a human benchmark for generative models. HYPE was tested with six popular GANs and two sampling techniques on conditional and unconditional image generation utilizing four datasets. The results showed that HYPE can identify the relative improvements between models, and these measurements were confirmed to be replicable and consistent by bootstrap sampling. Researchers can upload a model and receive a score from an online HYPE deployment.

1.2.6 Generative Information Lower BOund (GILBO)

In [85], a novel metric was defined for evaluating generative models, GILBO, and its value was tested on more than 3,200 models. The advantages and disadvantages of GILBO are discussed in [85], who found that GILBO provides more diverse information than the current metrics based on example quality, such as FID. GILBO can differentiate between GANs with qualitatively diverse representations even when they have equal FID values.

1.2.7 Summary

Overall, selecting a good evaluation metric for GANs is still a difficult problem [86]. Xu et al. [74] proposed an empirical study on the evaluation metrics of GANs. Karol Kurach [83] conducted a large-scale study on regularization and normalization in GANs. Other comparative studies of GANs exist, such as [87]. Reference [88] presented several measures acting as meta-measures that can guide researchers to select appropriate quantitative evaluation metrics. An appropriate evaluation metric should differentiate true examples from fake ones, verify mode drop and mode collapse, and detect overfitting. Hopefully, better methods will be developed to evaluate the quality of GAN models in the future.

1.3 Task-driven GANs

While the focus of this paper is on GANs, there are closely associated fields for specific tasks that provide an enormous volume of literature.

1.3.1 Semi-Supervised Learning

A research field in which GANs have been very successful is the application of generative models to semi-supervised learning [89], [90], as proposed but not shown in the first GAN paper [41].

GANs have been successfully used for semi-supervised learning since categorical GANs (CatGANs) [91]. Feature matching GANs [71] achieved good performances with a small number of labels on datasets such as MNIST, SVHN, and CIFAR-10.

Odena [92] extended GANs to semi-supervised learning by forcing the discriminator network to output class labels. Generally, when training GANs, we do not use the discriminator after training. The discriminator is used only to guide the learning process and is not used to generate data after the generator has been trained. Subsequently, the generator is used to generate the data and the discriminator is abandoned. In traditional GANs, the discriminator is a two-class classifier that outputs one category for real data and another category for generated data. In semi-supervised learning, the discriminator is upgraded to be a multi-class classifier. After training, the classifier is the model we are interested in. For semi-supervised learning, to train an N class classifier, we create GANs with a discriminator that can predict which of N+1 classes the input is from, where the extra class corresponds to the outputs of G. Therefore, suppose that we want to learn to classify two classes, apples and oranges. We can construct a classifier with three labels: the first is the class of real apples, the second is the class of real oranges, and the third is the class of generated data. The system is trained on three different types of data: real labeled data, real unlabeled data, and fake data.

Real labeled data: We can tell the discriminator to maximize the probability of the correct class. For example, for an apple photo labeled as an apple, we should maximize the probability of the apple class in the discriminator.

Unlabeled real data: Suppose we have a photo but do not know whether it is an apple or an orange—but we know that it is a real photo. In this situation, we train the discriminator to maximize the sum of the probabilities over all the real classes.

Fake data: When we obtain a generated example from the generator, we train the discriminator to classify it as a fake example.

Miyato et al. [93] proposed virtual adversarial training (VAT), a regularization method for both supervised and semi-supervised learning. Dai et al. [94] showed that given the discriminator objective, good semi-supervised learning requires a bad generator from a theoretical perspective and proposed the definition of a preferred generator. Triangle GANs (Δ-GAN) [95] were proposed for semi-supervised cross-domain joint distribution matching, and Δ-GAN is closely related to Triple-GAN [96]. Madani et al. [97] used semi-supervised learning with GANs to perform chest X-ray classification.

Future improvements to GANs can be expected to simultaneously produce further improvements to both semi-supervised learning and unsupervised learning such as self-supervised learning [98].

1.3.2 Transfer learning

Ganin et al. [99] introduced a domain-adversarial training approach for neural networks to enable domain adaptation, where the training data and the test data are from different but similar distributions. The Professor Forcing algorithm [100] used adversarial domain adaptation to train recurrent networks. Shrivastava et al. [101] used GANs for simulating training data. A novel extension of pixel-level domain adaptation named GraspGAN [102] was proposed for robotic grasping [103], [104]. By using synthetic data and domain adaptation [102], the number of real-world examples needed to achieve a given level of performance was reduced by up to 50 times while utilizing only randomly generated simulated objects.

Recent studies have shown remarkable success in image-to-image translation [105]–[110] between two domains. However, the existing methods, such as CycleGAN [59], learning to discover cross-domain relations with GANs (DiscoGAN) [111], and DualGAN [112], cannot be used directly for more than two domains because different approaches must be built independently for each domain pair. StarGAN [113] solved this problem by conducting image-to-image translations for multiple domains using only a single model. Other related works can be found in [114], [115]. CoGAN [66] can also be used for multiple domains.

Learning fair representations is a problem closely related to domain transfer. Note that different formulations of adversarial objectives [116]–[119] achieve different notations of fairness.

Domain adaptation [120], [121] can be seen as a subset of transfer learning [122]. Recent **visual domain adaptation (VDA)** methods include **visual appearance adaptation, representation adaptation**, and **output adaptation**, which can be considered as using domain adaptation based on the original input, features, and outputs of the domains, respectively.

Visual appearance adaptation: CycleGAN [59] is a representative method in this category. Cycle-consistent adversarial domain adaptation (CyCADA) [123] was proposed to perform visual appearance adaptation and was based on CycleGAN.

Representation adaptation: The key of adversarial discriminative domain adaptation (ADDA) [124], [125] is to learn feature representations such that a discriminator cannot determine the domain to which they belong. Sankaranarayanan et al. [126] focused on adapting the representations learned by segmentation networks across real and synthetic domains based on GANs. Fully convolutional adaptation networks (FCAN) [127] were proposed for semantic segmentation that combined visual appearance adaptation and representation adaptation.

Output adaptation: Tsai [128] made the outputs of the source and target images have a similar structure so that the discriminator cannot differentiate between them.

Other transfer learning based GANs can be found in [129]–[136].

1.3.3 Reinforcement learning

Generative models can be integrated into reinforcement learning (RL) [137] in different ways [138], [139]. Reference [140] already discussed connections between GANs and actor-critic methods. The connections among GANs, inverse reinforcement learning (IRL), and energy-based models were studied in [141]. These connections to RL may be useful for developing both GANs and RL. Furthermore, GANs were combined with RL in programs to synthesize images [142]. The competitive multi-agent learning framework proposed in [143] is also related to GANs and works by learning robust grasping policies from an adversary.

Imitation learning: The connection between imitation learning and EBGAN is discussed in [144]. Ho and Ermon [145] showed that an instantiation of their framework draws an analogy between GANs and imitation learning, from which they derived a model-free imitation learning method that has significant performance advantages over existing model-free algorithms in imitating complex behaviors in large and high-dimensional environments. Song et al. [146] proposed multi-agent generative adversarial imitation learning (GAIL), and Guo et al. [147] proposed generative adversarial self-imitation learning. A multi-agent GAIL framework was used in a deconfounded multi-agent environment reconstruction (DEMER) approach [148] to learn the environment. DEMER was tested on the real Didi Chuxing application and achieved good performances.

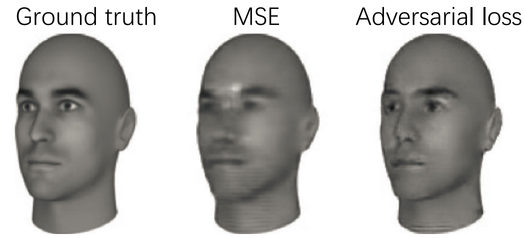


Fig. 3: Lotter et al. [153] provided an excellent description of the significance of being able to model multi-modal data. In this instance, a model is trained to predict the next frame in a video. The video describes a computer rendering of a moving 3D model of a man's head. The image on the left is the ground truth—an instance of an actual frame of a video that the model would like to predict. The image in the middle shows what happens when the model is trained using the MSE between the model predicted next frame and the actual next frame. The model is forced to select only one answer that represents what the next frame will look like. However, because there are multiple possible answers corresponding to slightly diverse head positions, the answer that the model selects is the average of multiple slightly diverse images, which causes the faces to be blurred. Utilizing an additional GAN loss, as shown by the image on the right, the model is capable of knowing that multiple possible outputs exist, each of which is recognizable and clear, and forms a realistic, satisfying image. Images are from [153].

1.3.4 Multi-modal learning

Generative models, especially GANs, allow machine learning techniques to work with multi-modal outputs. In many tasks, an input may correspond to multiple diverse correct outputs, each of which is an acceptable answer. Traditional machine learning methods, such as minimizing the mean squared error (MSE) between the model predicted output and a desired output, are not capable of training models that can produce many different correct outputs. One instance of such a case occurs when attempting to predict the next frame in a video sequence, as shown in Fig. 3. Multi-modal image-to-image translation related works can be found in [149]–[152].

1.3.5 Other task-driven GANs

GANs have been used for feature learning and applied to tasks such as feature selection [154], hashing [155]–[162], and metric learning [163].

MisGAN [164] was proposed to learn from incomplete data with GANs. Evolutionary GANs were proposed in [165]. Ponce et al. [166] combined GANs and genetic algorithms to evolve images for visual neurons. GANs have also been used in other machine learning tasks [167], such as active learning [168], [169], online learning [170], ensemble learning [171], zero-shot learning [172], [173], and multi-task learning [174].

2 OPEN RESEARCH PROBLEMS

Because GANs have become popular throughout the range of deep learning fields, their limitations have recently been improved [175], [176]. However, open research problems for GANs still exist.

GANs for discrete data: GANs rely on the generated examples being completely differentiable with respect to the generative parameters. Therefore, GANs cannot directly produce discrete data, such as hashing code and one-hot words. Solving this problem is highly important because it could unlock the potential of GANs for natural language processing (NLP) and hashing. Goodfellow [138] suggested three ways to solve this problem: using the Gumbel-softmax [177], [178] or the concrete distribution [179], utilizing the REINFORCE algorithm [180], and training the generator to sample continuous values that can be transformed to discrete ones (such as sampling word embeddings directly).

Other methods are applicable to this research direction. Song et al. [155] used a continuous function to approximate the sign function to generate hashing codes. Gulrajani et al. [181] modeled discrete data with a continuous generator. Hjelm et al. [182] introduced an algorithm for training GANs with discrete data that utilizes the estimated difference measure from the discriminator to compute importance weights for the generated examples, thus providing a policy gradient for training the generator. Other related works can be found in [183], [184]. However, more work needs to be done in this interesting area.

New divergences: New families of integral probability metrics (IPMs) for training GANs, such as Fisher GAN [185], [186], mean and covariance feature matching GAN (McGan) [187], and Sobolev GAN [188], have been proposed, which raises the following question: Are there any other interesting classes of divergences? This topic deserves further study.

Estimation uncertainty: Generally, when more data are available, the uncertainty estimation is reduced. GANs do not yield the distribution that generated the training examples, and GANs aim to generate new examples that come from the same distribution as that of the training examples. Therefore, GANs have neither a likelihood nor a well-defined posterior. There have been early attempts in this research direction, such as Bayesian GAN [189]. Although we can use GANs to generate data, how can we measure the uncertainty of the well-trained generator? This is another interesting future issue.

Theory: For generalization, Zhang et al. [190] developed generalization bounds between the true distribution and learned distribution under different evaluation metrics. When evaluated by neural distance, the bounds in [190] show that generalization is guaranteed as long as the discriminator set is sufficiently small, regardless of the size of the hypothesis set or generator. Arora et al. [191] proposed a novel test for estimating support size using the birthday paradox of discrete probability and showed that GANs do suffer mode collapse even when the images have better visual quality. More deep theoretical research is worth performing. How do we test for generalization empirically? Useful theory should enable the choice of model class, capacity, and architecture. This is another interesting issue that should be investigated in future work.

Moreover, [87] found that no GAN variant outperformed the original GANs, which demonstrates that future GAN research should conduct more systematic experiments and that model comparisons must be conducted neutrally.

REFERENCES

[1] W. Yin, Y. Fu, L. Sigal, and X. Xue, "Semi-latent gan: Learning to generate and modify facial images from attributes," *arXiv preprint arXiv:1704.02166*, 2017.

[2] B. L. Odry, B. Mailhe, M. S. Nadar, and P. Ceccaldi, "Image standardization using generative adversarial networks," Aug. 25 2020. US Patent 10,753,997.

[3] T. Kim, B. Kim, M. Cha, and J. Kim, "Unsupervised visual attribute transfer with reconfigurable generative adversarial networks," *arXiv preprint arXiv:1707.09798*, 2017.

[4] X. Wang, Z. Man, M. You, and C. Shen, "Adversarial generation of training examples: applications to moving vehicle license plate recognition," *arXiv preprint arXiv:1707.03124*, 2017.

[5] J. Xu and X. Han, "Cone-beam ct image enhancement using generative adversarial networks," Oct. 31 2019. US Patent App. 16/044,245.

[6] M. Chidambaram and Y. Qi, "Style transfer generative adversarial networks: Learning to play chess differently," *arXiv preprint arXiv:1702.06762*, 2017.

[7] L. S. Hibbard, "Radiotherapy treatment plan modeling using generative adversarial networks," Oct. 31 2019. US Patent App. 15/966,228.

[8] F. Juefei-Xu, V. N. Boddeti, and M. Savvides, "Gang of gans: Generative adversarial networks with maximum margin ranking," *arXiv preprint arXiv:1704.04865*, 2017.

[9] J. Shor and S. G. Cotado, "Computing systems with modularized infrastructure for training generative adversarial networks," May 9 2019. US Patent App. 16/159,093.

[10] S.-g. Lee, U. Hwang, S. Min, and S. Yoon, "Polyphonic music generation with sequence generative adversarial networks," *arXiv preprint arXiv:1710.11418*, 2018.

[11] M. Gorijala and A. Dukkipati, "Image generation and editing with variational info generative adversarial networks," *arXiv preprint arXiv:1701.04568*, 2017.

[12] S. Hitawala, "Comparative study on generative adversarial networks," *arXiv preprint arXiv:1801.04271*, 2018.

[13] K. J. Liang, C. Li, G. Wang, and L. Carin, "Generative adversarial network training is a continual learning problem," *arXiv preprint arXiv:1811.11083*, 2018.

[14] J.-J. Hwang, S. Azernikov, A. A. Efros, and S. X. Yu, "Learning beyond human expertise with generative models for dental restorations," *arXiv preprint arXiv:1804.00064*, 2018.

[15] A. Clark, J. Donahue, and K. Simonyan, "Efficient video generation on complex datasets," *arXiv preprint arXiv:1907.06571*, 2019.

[16] J. Li, J. Yang, A. Hertzmann, J. Zhang, and T. Xu, "Layoutgan: Generating graphic layouts with wireframe discriminators," in *International Conference on Learning Representations*, 2019.

[17] S. Sinha, H. Zhang, A. Goyal, Y. Bengio, H. Larochelle, and A. Odena, "Small-gan: Speeding up gan training using core-sets," in *International Conference on Machine Learning*, pp. 9005–9015, 2020.

[18] J. Li, J. Yang, A. Hertzmann, J. Zhang, and T. Xu, "Layoutgan: Synthesizing graphic layouts with vector-wireframe adversarial networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 7, pp. 2388–2399, 2021.

[19] M. Niemeyer and A. Geiger, "Giraffe: Representing scenes as compositional generative neural feature fields," in *Conference on Computer Vision and Pattern Recognition*, pp. 11453–11464, 2021.

[20] J. Pan, J. Dong, Y. Liu, J. Zhang, J. Ren, J. Tang, Y.-W. Tai, and M.-H. Yang, "Physics-based generative adversarial models for image restoration and beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 7, pp. 2449–2462, 2021.

[21] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," in *International Conference on Learning Representations*, 2018.

[22] J. H. Lim and J. C. Ye, "Geometric gan," *arXiv preprint arXiv:1705.02894*, 2017.

[23] D. Tran, R. Ranganath, and D. M. Blei, "Hierarchical implicit models and likelihood-free variational inference," in *Neural Information Processing Systems*, pp. 2794–2802, 2017.

[24] M. Lin, "Softmax gan," *arXiv preprint arXiv:1704.06191*, 2017.

[25] Y. LeCun, S. Chopra, R. Hadsell, M. Ranzato, and F. Huang, "A tutorial on energy-based learning," *Predicting structured data*, vol. 1, no. 0, 2006.

[26] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," in *International Conference on Learning Representations*, 2017.

[27] J. Zhao, L. Xiong, P. K. Jayashree, J. Li, F. Zhao, Z. Wang, P. S. Pranata, P. S. Shen, S. Yan, and J. Feng, "Dual-agent gans for photorealistic and identity preserving profile face synthesis," in *Advances in Neural Information Processing Systems*, pp. 66–76, 2017.

[28] J. Zhao, L. Xiong, J. Li, J. Xing, S. Yan, and J. Feng, "3d-aided dual-agent gans for unconstrained face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 10, pp. 2380–2394, 2018.

[29] R. Wang, A. Cully, H. J. Chang, and Y. Demiris, "Magan: Margin adaptation for generative adversarial networks," *arXiv preprint arXiv:1704.03817*, 2017.

[30] T. Che, Y. Li, A. P. Jacob, Y. Bengio, and W. Li, "Mode regularized generative adversarial networks," in *International Conference on Learning Representations*, 2017.

[31] L. Metz, B. Poole, D. Pfau, and J. Sohl-Dickstein, "Unrolled generative adversarial networks," in *International Conference on Learning Representations*, 2017.

[32] E. L. Denton, S. Chintala, R. Fergus, *et al.*, "Deep generative image models using a laplacian pyramid of adversarial networks," in *Neural Information Processing Systems*, pp. 1486–1494, 2015.

[33] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, 1983.

[34] T. R. Shaham, T. Dekel, and T. Michaeli, "Singan: Learning a generative model from a single natural image," in *International Conference on Computer Vision*, pp. 4570–4580, 2019.

[35] A. Shocher, S. Bagon, P. Isola, and M. Irani, "Ingan: Capturing and retargeting the "dna" of a natural image," in *International Conference on Computer Vision*, pp. 4492–4501, 2019.

[36] A. Brock, J. Donahue, and K. Simonyan, "Large scale gan training for high fidelity natural image synthesis," in *International Conference on Learning representations*, 2019.

[37] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Conference on Computer Vision and Pattern Recognition*, pp. 4401–4410, 2019.

[38] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," in *Conference on Computer Vision and Pattern Recognition*, pp. 8107–8116, 2020.

[39] M. Lučić, M. Tschannen, M. Ritter, X. Zhai, O. Bachem, and S. Gelly, "High-fidelity image generation with fewer labels," in *International Conference on Machine Learning*, pp. 4183–4192, 2019.

[40] J. Donahue and K. Simonyan, "Large scale adversarial representation learning," in *Neural Information Processing Systems*, 2019.

[41] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Neural Information Processing Systems*, pp. 2672–2680, 2014.

[42] G. Yildirim, N. Jetchev, R. Vollgraf, and U. Bergmann, "Generating high-resolution fashion model images wearing custom outfits," in *ICCV Workshop on Computer Vision for Fashion, Art and Design*, 2019.

[43] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.

[44] L. Mescheder, S. Nowozin, and A. Geiger, "Adversarial variational bayes: Unifying variational autoencoders and generative adversarial networks," in *International Conference on Machine Learning*, pp. 2391–2400, 2017.

[45] X. Yu, X. Zhang, Y. Cao, and M. Xia, "Vaegan: A collaborative filtering framework based on adversarial variational autoencoders," in *International Joint Conference on Artificial Intelligence*, pp. 4206–4212, 2019.

[46] M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," in *Neural Information Processing Systems*, pp. 700–708, 2017.

[47] Z. Hu, Z. Yang, R. Salakhutdinov, and E. P. Xing, "On unifying deep generative models," in *International Conference on Learning Representations*, 2018.

[48] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," in *International Conference on Machine Learning*, pp. 1558–1566, 2016.

[49] M. Rosca, B. Lakshminarayanan, D. Warde-Farley, and S. Mohamed, "Variational approaches for auto-encoding generative adversarial networks," *arXiv preprint arXiv:1706.04987*, 2017.

[50] L. Q. Tran, X. Yin, and X. Liu, "Representation learning by rotating your faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 3007–3021, 2019.

[51] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," in *International Conference on Learning Representations*, 2017.

[52] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "It takes (only) two: Adversarial generator-encoder networks," in *AAAI Conference on Artificial Intelligence*, pp. 1250–1257, 2018.

[53] V. Dumoulin, I. Belghazi, B. Poole, O. Mastropietro, A. Lamb, M. Arjovsky, and A. Courville, "Adversarially learned inference," in *International Conference on Learning Representations*, 2017.

[54] S. Gurumurthy, R. Kiran Sarvadevabhatla, and R. Venkatesh Babu, "Deligan: Generative adversarial networks for diverse and limited data," in *Conference on Computer Vision and Pattern Recognition*, pp. 166–174, 2017.

[55] C. Li, H. Liu, C. Chen, Y. Pu, L. Chen, R. Henao, and L. Carin, "Alice: Towards understanding adversarial learning for joint distribution matching," in *Neural Information Processing Systems*, pp. 5495–5503, 2017.

[56] S. Arora, A. Risteski, and Y. Zhang, "Theoretical limitations of encoder-decoder gan architectures," *arXiv preprint arXiv:1711.02651*, 2017.

[57] A. Creswell and A. A. Bharath, "Inverting the generator of a generative adversarial network," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 7, pp. 1967–1974, 2019.

[58] X. B. Peng, A. Kanazawa, S. Toyer, P. Abbeel, and S. Levine, "Variational discriminator bottleneck: Improving imitation learning, inverse rl, and gans by constraining information flow," in *International Conference on Learning Representations*, 2019.

[59] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *International Conference on Computer Vision*, pp. 2223–2232, 2017.

[60] T. Nguyen, T. Le, H. Vu, and D. Phung, "Dual discriminator generative adversarial nets," in *Neural Information Processing Systems*, pp. 2670–2680, 2017.

[61] I. Durugkar, I. Gemp, and S. Mahadevan, "Generative multi-adversarial networks," in *International Conference on Learning Representations*, 2017.

[62] I. Albuquerque, J. Monteiro, T. Doan, B. Considine, T. Falk, and I. Mitliagkas, "Multi-objective training of generative adversarial networks with multiple discriminators," in *International Conference on Machine Learning*, 2019.

[63] Q. Hoang, T. D. Nguyen, T. Le, and D. Phung, "Mgan: Training generative adversarial nets with multiple generators," in *International Conference on Learning Representations*, 2018.

[64] A. Ghosh, V. Kulharia, V. P. Namboodiri, P. H. Torr, and P. K. Dokania, "Multi-agent diverse generative adversarial networks," in *Conference on Computer Vision and Pattern Recognition*, pp. 8513–8521, 2018.

[65] K. Wang and X. Wan, "Sentigan: Generating sentimental texts via mixture adversarial networks.," in *International Joint Conference on Artificial Intelligence*, pp. 4446–4452, 2018.

[66] M.-Y. Liu and O. Tuzel, "Coupled generative adversarial networks," in *Neural Information Processing Systems*, pp. 469–477, 2016.

[67] X. Wang and A. Gupta, "Generative image modeling using style and structure adversarial networks," in *European Conference on Computer Vision*, pp. 318–335, 2016.

[68] D. Xu, Y. Wu, S. Yuan, L. Zhang, and X. Wu, "Achieving causal fairness through generative adversarial networks," in *International Joint Conference on Artificial Intelligence*, pp. 1452–1458, 2019.

[69] Z. Wang, G. Healy, A. F. Smeaton, and T. E. Ward, "Use of neural signals to evaluate the quality of generative adversarial network performance in facial image generation," *Cognitive Computation*, vol. 12, no. 1, pp. 13–24, 2020.

[70] Z. Wang, Q. She, A. F. Smeaton, T. E. Ward, and G. Healy, "Neuroscore: A brain-inspired evaluation metric for generative adversarial networks," *arXiv preprint arXiv:1905.04243*, 2019.

[71] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," in *Neural Information Processing Systems*, pp. 2234–2242, 2016.

[72] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.

[73] I. Danihelka, B. Lakshminarayanan, B. Uria, D. Wierstra, and P. Dayan, "Comparison of maximum likelihood and gan-based training of real nvps," *arXiv preprint arXiv:1705.05263*, 2017.

[74] Q. Xu, G. Huang, Y. Yuan, C. Guo, Y. Sun, F. Wu, and K. Weinberger, "An empirical study on evaluation metrics of generative adversarial networks," *arXiv preprint arXiv:1806.07755*, 2018.

[75] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," in *Neural Information Processing Systems*, pp. 6626–6637, 2017.

[76] M. Bińkowski, D. J. Sutherland, M. Arbel, and A. Gretton, "Demystifying mmd gans," *arXiv preprint arXiv:1801.01401*, 2021.

[77] M. S. Sajjadi, O. Bachem, M. Lucic, O. Bousquet, and S. Gelly, "Assessing generative models via precision and recall," in *Advances in Neural Information Processing Systems*, pp. 5228–5237, 2018.

[78] T. Kynkäänniemi, T. Karras, S. Laine, J. Lehtinen, and T. Aila, "Improved precision and recall metric for assessing generative models," in *Advances in Neural Information Processing Systems*, pp. 3929–3938, 2019.

[79] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[80] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Asilomar Conference on Signals, Systems & Computers*, vol. 2, pp. 1398–1402, 2003.

[81] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in *International Conference on Machine Learning*, pp. 2642–2651, 2017.

[82] W. Fedus, M. Rosca, B. Lakshminarayanan, A. M. Dai, S. Mohamed, and I. Goodfellow, "Many paths to equilibrium: Gans do not need to decrease a divergence at every step," in *International Conference on Learning Representations*, 2018.

[83] K. Kurach, M. Lučić, X. Zhai, M. Michalski, and S. Gelly, "A large-scale study on regularization and normalization in gans," in *International Conference on Machine Learning*, pp. 3581–3590, 2019.

[84] S. Zhou, M. Gordon, R. Krishna, A. Narcomey, L. F. Fei-Fei, and M. Bernstein, "Hype: A benchmark for human eye perceptual evaluation of generative models," in *Advances in Neural Information Processing Systems*, pp. 3444–3456, 2019.

[85] A. A. Alemi and I. Fischer, "Gilbo: one metric to measure them all," in *Advances in Neural Information Processing Systems*, pp. 7037–7046, 2018.

[86] L. Theis, A. v. d. Oord, and M. Bethge, "A note on the evaluation of generative models," in *International Conference on Learning Representations*, 2015.

[87] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, "Are gans created equal? a large-scale study," in *Neural Information Processing Systems*, pp. 700–709, 2018.

[88] A. Borji, "Pros and cons of gan evaluation measures," *Computer Vision and Image Understanding*, vol. 179, pp. 41–65, 2019.

[89] E. Denton, S. Gross, and R. Fergus, "Semi-supervised learning with context-conditional generative adversarial networks," *arXiv preprint arXiv:1611.06430*, 2016.

[90] M. Ding, J. Tang, and J. Zhang, "Semi-supervised learning on graphs with generative adversarial nets," in *CM International Conference on Information and Knowledge Management*, pp. 913–922, ACM, 2018.

[91] J. T. Springenberg, "Unsupervised and semi-supervised learning with categorical generative adversarial networks," in *International Conference on Learning Representations*, 2016.

[92] A. Odena, "Semi-supervised learning with generative adversarial networks," *arXiv preprint arXiv:1606.01583*, 2016.

[93] T. Miyato, S.-i. Maeda, S. Ishii, and M. Koyama, "Virtual adversarial training: a regularization method for supervised and semi-supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 1979–1993, 2019.

[94] Z. Dai, Z. Yang, F. Yang, W. W. Cohen, and R. R. Salakhutdinov, "Good semi-supervised learning that requires a bad gan," in *Neural Information Processing Systems*, pp. 6510–6520, 2017.

[95] Z. Gan, L. Chen, W. Wang, Y. Pu, Y. Zhang, H. Liu, C. Li, and L. Carin, "Triangle generative adversarial networks," in *Neural Information Processing Systems*, pp. 5247–5256, 2017.

[96] L. Chongxuan, T. Xu, J. Zhu, and B. Zhang, "Triple generative adversarial nets," in *Neural Information Processing Systems*, pp. 4088–4098, 2017.

[97] A. Madani, M. Moradi, A. Karargyris, and T. Syeda-Mahmood, "Semi-supervised learning with generative adversarial networks for chest x-ray classification with ability of data domain adaptation," in *International Symposium on Biomedical Imaging*, pp. 1038–1042, 2018.

[98] T. Chen, X. Zhai, M. Ritter, M. Lucic, and N. Houlsby, "Self-supervised gans via auxiliary rotation loss," in *Conference on Computer Vision and Pattern Recognition*, pp. 12154–12163, 2019.

[99] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2096–2030, 2016.

[100] A. M. Lamb, A. G. A. P. Goyal, Y. Zhang, S. Zhang, A. C. Courville, and Y. Bengio, "Professor forcing: A new algorithm for training recurrent networks," in *Neural Information Processing Systems*, pp. 4601–4609, 2016.

[101] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," in *Conference on Computer Vision and Pattern Recognition*, pp. 2107–2116, 2017.

[102] K. Bousmalis, A. Irpan, P. Wohlhart, Y. Bai, M. Kelcey, M. Kalakrishnan, L. Downs, J. Ibarz, P. Pastor, K. Konolige, *et al.*, "Using simulation and domain adaptation to improve efficiency of deep robotic grasping," in *International Conference on Robotics and Automation*, pp. 4243–4250, 2018.

[103] S. James, P. Wohlhart, M. Kalakrishnan, D. Kalashnikov, A. Irpan, J. Ibarz, S. Levine, R. Hadsell, and K. Bousmalis, "Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks," in *Conference on Computer Vision and Pattern Recognition*, pp. 12627–12637, 2019.

[104] L. Pinto, J. Davidson, and A. Gupta, "Supervision via competition: Robot adversaries for learning tasks," in *International Conference on Robotics and Automation*, pp. 1601–1608, 2017.

[105] A. Anoosheh, E. Agustsson, R. Timofte, and L. Van Gool, "Combogan: Unrestrained scalability for image domain translation," in *CVPR Workshops*, pp. 783–790, 2018.

[106] X. Chen, C. Xu, X. Yang, and D. Tao, "Attention-gan for object transfiguration in wild images," in *European Conference on Computer Vision*, pp. 164–180, 2018.

[107] C. Wang, C. Xu, C. Wang, and D. Tao, "Perceptual adversarial networks for image-to-image transformation," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4066–4079, 2018.

[108] J. Cao, H. Huang, Y. Li, J. Liu, R. He, and Z. Sun, "Biphasic learning of gans for high-resolution image-to-image translation," *arXiv preprint arXiv:1904.06624*, 2019.

[109] M. Amodio and S. Krishnaswamy, "Travelgan: Image-to-image translation by transformation vector learning," in *Conference on Computer Vision and Pattern Recognition*, pp. 8983–8992, 2019.

[110] M.-Y. Liu, X. Huang, A. Mallya, T. Karras, T. Aila, J. Lehtinen, and J. Kautz, "Few-shot unsupervised image-to-image translation," in *International Conference on Computer Vision*, pp. 10550–10559, 2019.

[111] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," in *International Conference on Machine Learning*, pp. 1857–1865, 2017.

[112] Z. Yi, H. Zhang, P. Tan, and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," in *International Conference on Computer Vision*, pp. 2849–2857, 2017.

[113] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "Stargan: Unified generative adversarial networks for multi-domain image-to-image translation," in *Conference on Computer Vision and Pattern Recognition*, pp. 8789–8797, 2018.

[114] Z. He, W. Zuo, M. Kan, S. Shan, and X. Chen, "Attgan: Facial attribute editing by only changing what you want," *IEEE Transactions on Image Processing*, 2019.

[115] M. Liu, Y. Ding, M. Xia, X. Liu, E. Ding, W. Zuo, and S. Wen, "Stgan: A unified selective transfer network for arbitrary image attribute editing," in *Conference on Computer Vision and Pattern Recognition*, pp. 3673–3682, 2019.

[116] H. Edwards and A. Storkey, "Censoring representations with an adversary," in *International Conference on Learning Representations*, 2016.

[117] A. Beutel, J. Chen, Z. Zhao, and E. H. Chi, "Data decisions and theoretical implications when adversarially learning fair representations," in *Workshop on Fairness, Accountability, and Transparency in Machine Learning*, 2017.

[118] B. H. Zhang, B. Lemoine, and M. Mitchell, "Mitigating unwanted biases with adversarial learning," in *AAAI/ACM Conference on AI, Ethics, and Society*, pp. 335–340, 2018.

[119] D. Madras, E. Creager, T. Pitassi, and R. Zemel, "Learning adversarially fair and transferable representations," in *International Conference on Machine Learning*, 2018.

[120] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krishnan, "Unsupervised pixel-level domain adaptation with generative adversarial networks," in *Conference on Computer Vision and Pattern Recognition*, pp. 3722–3731, 2017.

[121] L. Hu, M. Kan, S. Shan, and X. Chen, "Duplex generative adversarial network for unsupervised domain adaptation," in *Conference on Computer Vision and Pattern Recognition*, pp. 1498–1507, 2018.

[122] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[123] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. A. Efros, and T. Darrell, "Cycada: Cycle-consistent adversarial domain adaptation," in *International Conference on Machine Learning*, 2018.

[124] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Conference on Computer Vision and Pattern Recognition*, pp. 7167–7176, 2017.

[125] S. Wang and L. Zhang, "Catgan: Coupled adversarial transfer for domain generation," *arXiv preprint arXiv:1711.08904*, 2017.

[126] S. Sankaranarayanan, Y. Balaji, A. Jain, S. Nam Lim, and R. Chellappa, "Learning from synthetic data: Addressing domain shift for semantic segmentation," in *Conference on Computer Vision and Pattern Recognition*, pp. 3752–3761, 2018.

[127] Y. Zhang, Z. Qiu, T. Yao, D. Liu, and T. Mei, "Fully convolutional adaptation networks for semantic segmentation," in *Conference on Computer Vision and Pattern Recognition*, pp. 6810–6818, 2018.

[128] Y.-H. Tsai, W.-C. Hung, S. Schulter, K. Sohn, M.-H. Yang, and M. Chandraker, "Learning to adapt structured output space for semantic segmentation," in *Conference on Computer Vision and Pattern Recognition*, pp. 7472–7481, 2018.

[129] H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, and M. Marchand, "Domain-adversarial neural networks," in *NIPS Workshop on Transfer and Multi-Task Learning: Theory meets Practice*, 2014.

[130] J. Shen, Y. Qu, W. Zhang, and Y. Yu, "Wasserstein distance guided representation learning for domain adaptation," in *AAAI Conference on Artificial Intelligence*, pp. 4058–4065, 2018.

[131] S. Benaim and L. Wolf, "One-sided unsupervised domain mapping," in *Neural Information Processing Systems*, pp. 752–762, 2017.

[132] H. Zhao, S. Zhang, G. Wu, J. M. Moura, J. P. Costeira, and G. J. Gordon, "Adversarial multiple source domain adaptation," in *Neural Information Processing Systems*, pp. 8559–8570, 2018.

[133] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in *Neural Information Processing Systems*, pp. 1640–1650, 2018.

[134] W. Hong, Z. Wang, M. Yang, and J. Yuan, "Conditional generative adversarial network for structured domain adaptation," in *Conference on Computer Vision and Pattern Recognition*, pp. 1335–1344, 2018.

[135] K. Saito, K. Watanabe, Y. Ushiku, and T. Harada, "Maximum classifier discrepancy for unsupervised domain adaptation," in *Conference on Computer Vision and Pattern Recognition*, pp. 3723–3732, 2018.

[136] R. Volpi, P. Morerio, S. Savarese, and V. Murino, "Adversarial feature augmentation for unsupervised domain adaptation," in *Conference on Computer Vision and Pattern Recognition*, pp. 5495–5504, 2018.

[137] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, *et al.*, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[138] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," *arXiv preprint arXiv:1701.00160*, 2017.

[139] X. Chen, S. Li, H. Li, S. Jiang, Y. Qi, and L. Song, "Generative adversarial user model for reinforcement learning based recommendation system," in *International Conference on Machine Learning*, pp. 1052–1061, 2019.

[140] D. Pfau and O. Vinyals, "Connecting generative adversarial networks and actor-critic methods," *arXiv preprint arXiv:1610.01945*, 2016.

[141] C. Finn, P. Christiano, P. Abbeel, and S. Levine, "A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models," in *NIPS Workshop on Adversarial Training*, 2016.

[142] Y. Ganin, T. Kulkarni, I. Babuschkin, S. Eslami, and O. Vinyals, "Synthesizing programs for images using reinforced adversarial learning," in *International Conference on Machine Learning*, pp. 1666–1675, 2018.

[143] T. Bansal, J. Pachocki, S. Sidor, I. Sutskever, and I. Mordatch, "Emergent complexity via multi-agent competition," in *International Conference on Learning Representations*, 2018.

[144] J. Yoo, H. Ha, J. Yi, J. Ryu, C. Kim, J.-W. Ha, Y.-H. Kim, and S. Yoon, "Energy-based sequence gans for recommendation and their connection to imitation learning," *arXiv preprint arXiv:1706.09200*, 2017.

[145] J. Ho and S. Ermon, "Generative adversarial imitation learning," in *Neural Information Processing Systems*, pp. 4565–4573, 2016.

[146] J. Song, H. Ren, D. Sadigh, and S. Ermon, "Multi-agent generative adversarial imitation learning," in *Neural Information Processing Systems*, pp. 7461–7472, 2018.

[147] Y. Guo, J. Oh, S. Singh, and H. Lee, "Generative adversarial self-imitation learning," in *NeurIPS Deep Reinforcement Learning Workshop*, 2018.

[148] W. Shang, Y. Yu, Q. Li, Z. Qin, Y. Meng, and J. Ye, "Environment reconstruction with hidden confounders for reinforcement learning based recommendation," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 566–576, 2019.

[149] J.-Y. Zhu, R. Zhang, D. Pathak, T. Darrell, A. A. Efros, O. Wang, and E. Shechtman, "Toward multimodal image-to-image translation," in *Neural Information Processing Systems*, pp. 465–476, 2017.

[150] A. Almahairi, S. Rajeswar, A. Sordoni, P. Bachman, and A. Courville, "Augmented cyclegan: Learning many-to-many mappings from unpaired data," in *International Conference on Machine Learning*, pp. 195–204, 2018.

[151] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, "Multimodal unsupervised image-to-image translation," in *European Conference on Computer Vision (ECCV)*, pp. 172–189, 2018.

[152] H.-Y. Lee, H.-Y. Tseng, J.-B. Huang, M. Singh, and M.-H. Yang, "Diverse image-to-image translation via disentangled representations," in *European Conference on Computer Vision*, pp. 35–51, 2018.

[153] W. Lotter, G. Kreiman, and D. Cox, "Unsupervised learning of visual structure using predictive generative networks," *arXiv preprint arXiv:1511.06380*, 2015.

[154] J. Jordon, J. Yoon, and M. van der Schaar, "Knockoffgan: Generating knockoffs for feature selection using generative adversarial networks," in *International Conference on Learning Representations*, 2019.

[155] J. Song, T. He, L. Gao, X. Xu, A. Hanjalic, and H. T. Shen, "Binary generative adversarial networks for image retrieval," in *AAAI Conference on Artificial Intelligence*, pp. 394–401, 2018.

[156] J. Zhang, Y. Peng, and M. Yuan, "Unsupervised generative adversarial cross-modal hashing," in *AAAI Conference on Artificial Intelligence*, pp. 539–546, 2018.

[157] J. Zhang, Y. Peng, and M. Yuan, "Sch-gan: Semi-supervised cross-modal hashing by generative adversarial network," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 489–502, 2020.

[158] K. Ghasedi Dizaji, F. Zheng, N. Sadoughi, Y. Yang, C. Deng, and H. Huang, "Unsupervised deep generative adversarial hashing network," in *Conference on Computer Vision and Pattern Recognition*, pp. 3664–3673, 2018.

[159] G. Wang, Q. Hu, J. Cheng, and Z. Hou, "Semi-supervised generative adversarial hashing for image retrieval," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 469–485, 2018.

[160] Z. Qiu, Y. Pan, T. Yao, and T. Mei, "Deep semantic hashing with generative adversarial networks," in *ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 225–234, 2017.

[161] J. Zhang, Z. Wei, I. C. Duta, F. Shen, L. Liu, F. Zhu, X. Xu, L. Shao, and H. T. Shen, "Generative reconstructive hashing for incomplete video analysis," in *ACM International Conference on Multimedia*, pp. 845–854, ACM, 2019.

[162] Y. Wang, L. Zhang, F. Nie, X. Li, Z. Chen, and F. Wang, "Wegan: Deep image hashing with weighted generative adversarial networks," *IEEE Transactions on Multimedia*, 2019.

[163] G. Dai, J. Xie, and Y. Fang, "Metric-based generative adversarial network," in *ACM International Conference on Multimedia*, pp. 672–680, 2017.

[164] S. C.-X. Li, B. Jiang, and B. Marlin, "Misgan: Learning from incomplete data with generative adversarial networks," in *International Conference on Learning Representations*, 2019.

[165] C. Wang, C. Xu, X. Yao, and D. Tao, "Evolutionary generative adversarial networks," *IEEE Transactions on Evolutionary Computation*, 2019.

[166] C. R. Ponce, W. Xiao, P. F. Schade, T. S. Hartmann, G. Kreiman, and M. S. Livingstone, "Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences," *Cell*, vol. 177, no. 4, pp. 999–1009, 2019.

[167] Y. Wang, Y. Xia, T. He, F. Tian, T. Qin, C. Zhai, and T.-Y. Liu, "Multi-agent dual learning," in *International Conference on Learning Representations*, 2019.

[168] J.-J. Zhu and J. Bento, "Generative adversarial active learning," *arXiv preprint arXiv:1702.07956*, 2017.

[169] M.-K. Xie and S.-J. Huang, "Learning class-conditional gans with active sampling," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 998–1006, ACM, 2019.

[170] P. Grnarova, K. Y. Levy, A. Lucchi, T. Hofmann, and A. Krause, "An online learning approach to generative adversarial networks," in *International Conference on Learning Representations*, 2018.

[171] I. O. Tolstikhin, S. Gelly, O. Bousquet, C.-J. Simon-Gabriel, and B. Schölkopf, "Adagan: Boosting generative models," in *Neural Information Processing Systems*, pp. 5424–5433, 2017.

[172] Y. Zhu, M. Elhoseiny, B. Liu, X. Peng, and A. Elgammal, "A generative adversarial approach for zero-shot learning from noisy texts," in *Conference on Computer Vision and Pattern Recognition*, pp. 1004–1013, 2018.

[173] P. Qin, X. Wang, W. Chen, C. Zhang, W. Xu, and W. Y. Wang, "Generative adversarial zero-shot relational learning for knowledge graphs," in *AAAI Conference on Artificial Intelligence*, 2020.

[174] P. Yang, Q. Tan, H. Tong, and J. He, "Task-adversarial co-generative nets," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1596–1604, ACM, 2019.

[175] C. Daskalakis, A. Ilyas, V. Syrgkanis, and H. Zeng, "Training gans with optimism," in *International Conference on Learning Representations*, 2018.

[176] A. Bora, E. Price, and A. G. Dimakis, "Ambientgan: Generative models from lossy measurements," in *International Conference on Learning Representations*, 2018.

[177] M. J. Kusner and J. M. Hernández-Lobato, "Gans for sequences of discrete elements with the gumbel-softmax distribution," *arXiv preprint arXiv:1611.04051*, 2016.

[178] E. Jang, S. Gu, and B. Poole, "Categorical reparameterization with gumbel-softmax," in *International Conference on Learning Representations*, 2017.

[179] C. J. Maddison, A. Mnih, and Y. W. Teh, "The concrete distribution: A continuous relaxation of discrete random variables," in *International Conference on Learning Representations*, 2017.

[180] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine learning*, vol. 8, no. 3-4, pp. 229–256, 1992.

[181] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in *Neural Information Processing Systems*, pp. 5767–5777, 2017.

[182] R. D. Hjelm, A. P. Jacob, T. Che, A. Trischler, K. Cho, and Y. Bengio, "Boundary-seeking generative adversarial networks," in *International Conference on Learning Representations*, 2018.

[183] T. Che, Y. Li, R. Zhang, R. D. Hjelm, W. Li, Y. Song, and Y. Bengio, "Maximum-likelihood augmented discrete generative adversarial networks," *arXiv preprint arXiv:1702.07983*, 2017.

[184] Z. Junbo, Y. Kim, K. Zhang, A. Rush, and Y. LeCun, "Adversarially regularized autoencoders," in *International Conference on Machine Learning*, pp. 5902–5911, 2018.

[185] Y. Mroueh and T. Sercu, "Fisher gan," in *Neural Information Processing Systems*, pp. 2513–2523, 2017.

[186] J. Yoo, Y. Hong, Y. Noh, and S. Yoon, "Domain adaptation using adversarial learning for autonomous navigation," *arXiv preprint arXiv:1712.03742*, 2018.

[187] Y. Mroueh, T. Sercu, and V. Goel, "Mcgan: Mean and covariance feature matching gan," in *International Conference on Machine Learning*, pp. 2527–2535, 2017.

[188] Y. Mroueh, C.-L. Li, T. Sercu, A. Raj, and Y. Cheng, "Sobolev gan," in *International Conference on Learning Representations*, 2018.

[189] Y. Saatci and A. G. Wilson, "Bayesian gan," in *Neural Information Processing Systems*, pp. 3622–3631, 2017.

[190] P. Zhang, Q. Liu, D. Zhou, T. Xu, and X. He, "On the discrimination-generalization tradeoff in gans," in *International Conference on Learning Representations*, 2018.

[191] S. Arora, A. Risteski, and Y. Zhang, "Do gans learn the distribution? some theory and empirics," in *International Conference on Learning Representations*, 2018.