Training GANs with Stronger Augmentations via Constructive Discimintor

Paper Summary

Harum Naseem Bhinder and Muhammad Raahim Khan
Lahore University of Management Sciences (LUMS)

Compiled as of 19th April, 2021.

This paper proposes a unique discriminator of Generative Adversarial Networks (GANs). Secondly, it was also showcased that contrastive representation learning such as SimCLR can also benefit once it is jointly trained with GANs. Thirdly, it lays out experimental results showing that GANs with contrastive learning consistently improve Fréchet Inception Distance (FID) and Inception Score (IS) instead of other data augmentation techniques. Most importantly, these measures are improved while maintaining highly discriminative features within the discriminator in terms of the linear evaluation. Lastly, it demonstrates that when GANs are trained in an unsupervised manner, without labels, they can induce many conditional generative models through a brief latent sampling, leveraging the learned features of the contrastive discriminator.

1. Methodology

The authors propose to train discriminators of Generative Adversarial Networks (GANs) using the principle of Constructive Learning (ContraD). This is to develop a training scheme of GANs that can extract, under more robust and strong data augmentation, further useful information despite limited practices such as random translations up to a few pixels.

The methodology is such that the encoder network is trained to minimize two different contrastive losses: the SimCLR Loss on real samples and the Supervised Contrastive Loss on fake samples. Then, this encoder network is jointly trained under the standard framework of GAN.

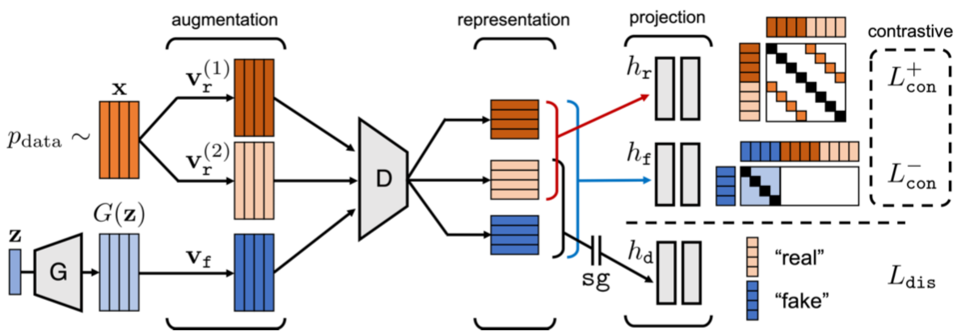

Figure 1: This is an overview of Contrastive Discriminator (ContraD). Overall, the representation of ContraD is learned from the two contrastive losses L+con (real samples) and L−con (fake samples). Here, sg(·) denotes the stop-gradient operation.

Various datasets such as CIFAR-10/100, ImageNet, CelebA-HQ-128, and AFHQ were used throughout the experiment. The effectiveness of ContraD was measured using multiple experiments, including but not limited to ablation study. Fréchet Inception Distance (FID) and Inception Score (IS) were used as quantitative metrics for evaluation purposes.

2. Results

It was observed that ContraD not only improved GAN training but also outperformed many other models (figure 2).

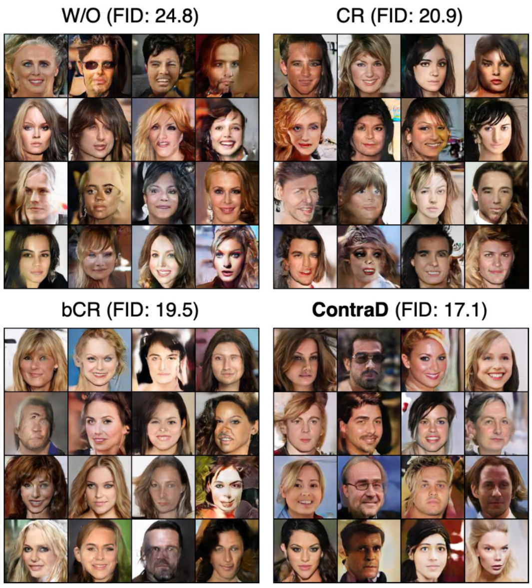

Figure 2: Qualitative comparison of unconditionally generated samples from GANs with different training methods. All the models are trained on CelebA-HQ-128 with an SNDCGAN architecture.

When keeping the hyperparameters constant, it was concluded that ContraD could also further improve its results. However, this was not the case for other models. The other models could not improve their performance under the same setup.

Architecture	Method	Augment.	CIFAR-10		CIFAR-100	
			FID ↓	IS ↑	FID ↓	IS ↑
G: SNDCGAN D: SNDCGAN	-	-	26.6	7.38	28.5	7.25
	CR (Zhang et al., 2020)	HFlip, Trans	19.5	7.87	22.2	7.91
	bCR (Zhao et al., 2020c)	HFlip, Trans	14.0	8.35	19.2	8.46
	DiffAug (Zhao et al., 2020a)	Trans, CutOut	22.9	7.64	27.0	7.47
	ContraD (ours)	**SimCLR**	**10.9**	**8.78**	**15.2**	**9.09**
G: SNDCGAN D: SNResNet-18	-	-	41.3	6.33	52.3	5.24
	CR (Zhang et al., 2020)	HFlip, Trans	32.1	7.08	36.5	6.55
	bCR (Zhao et al., 2020c)	HFlip, Trans	22.8	7.29	28.2	7.30
	DiffAug (Zhao et al., 2020a)	Trans, CutOut	59.5	5.62	58.7	5.39
	ContraD (ours)	**SimCLR**	**9.86**	**9.09**	**15.0**	**9.56**
G: StyleGAN2 D: StyleGAN2	-	-	11.1	9.18	16.5	9.51
	DiffAug* (Zhao et al., 2020a)	Trans, CutOut	9.89	9.40	15.2	**10.0**
	ContraD (ours)	**SimCLR**	**9.80**	**9.47**	**14.1**	**10.0**

Table 1: Here is the comparison of the best FID score and IS on unconditional image generation of CIFAR-10 and CIFAR-100.

The Ablation Study identified that stopping gradients before the discriminator's head in the ContraD design and contrastive losses proved to be essential factors for ContraD architecture. Similarly, maintaining two independent projection headers proved to be more stable than sharing them. Two layered networks for the discriminator head proved to be an efficient choice. Furthermore, even with weaker augmentations proved to be still as good when compared with other models. Finally, as opposed to previous methods, without overly regularizing the discriminator, the proposed ContraD effectively handles the SimCLR augmentations. This could possibly be done by leveraging many components of contrastive learning such as the normalized loss, and the use of separate projection heads.

				D: SNDCGAN		*D*: SNResNet-18	
MLP h_d	$\mathtt{sg}(\cdot)$	L^{+}_{con}	L^{-}_{con}	FID ↓	IS ↑	FID ↓	IS ↑
✓	✓	✓	✓	**11.1**	**8.62**	**10.6**	**8.99**
✗	✓	✓	✓	185	3.43	274	2.09
✓	✗	✓	✓	11.6	8.61	28.0	7.52
✓	✓	✗	✓	11.9	8.45	182	2.02
✓	✓	✓	✗	210	1.93	232	2.56

Table 2: This is a comparison of the best FID score and IS on CIFAR-10 for ablations of proposed components