Self-Supervised GANs
via Auxiliary Rotation Loss

CVPR2019

2019.07.03

발표자 박성현

DAVIAN
Data and Visual Analytics Lab

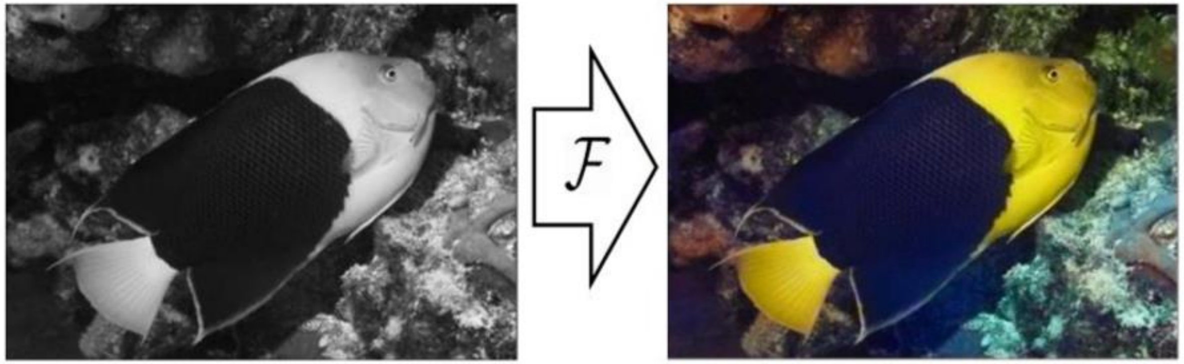
KOREA UNIVERSITY

1 Introduction

Self-Supervision

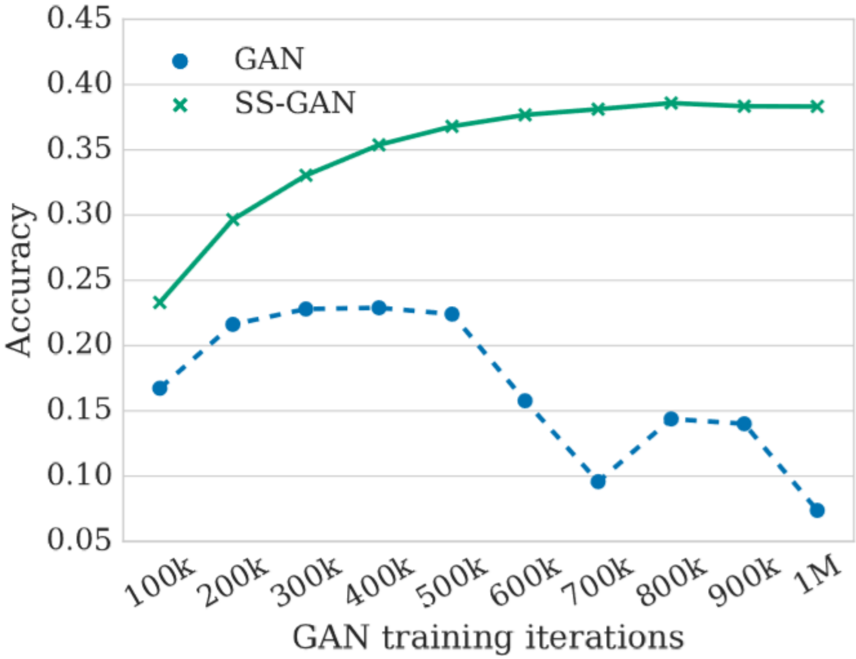
- A form of unsupervised learning where the data provides the supervision
- In general, withhold some part of the data, and task the network with predicting it

(a) Input context (b) Human artist

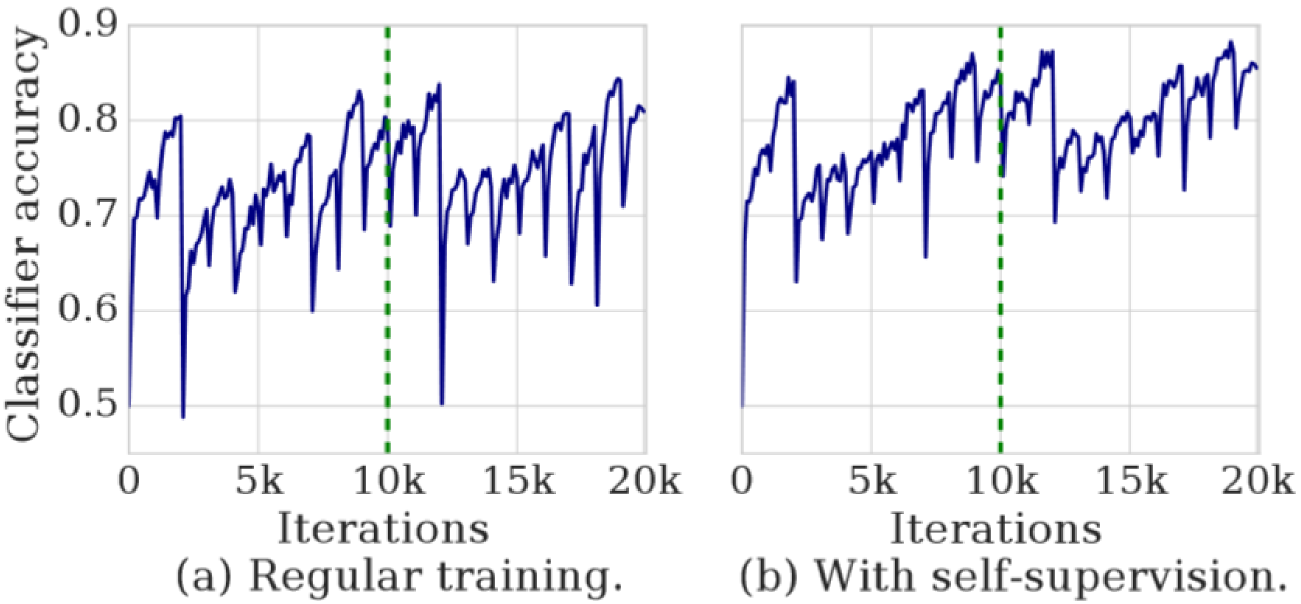


1 Introduction

Discriminator forgetting



[Performance of a linear classification]



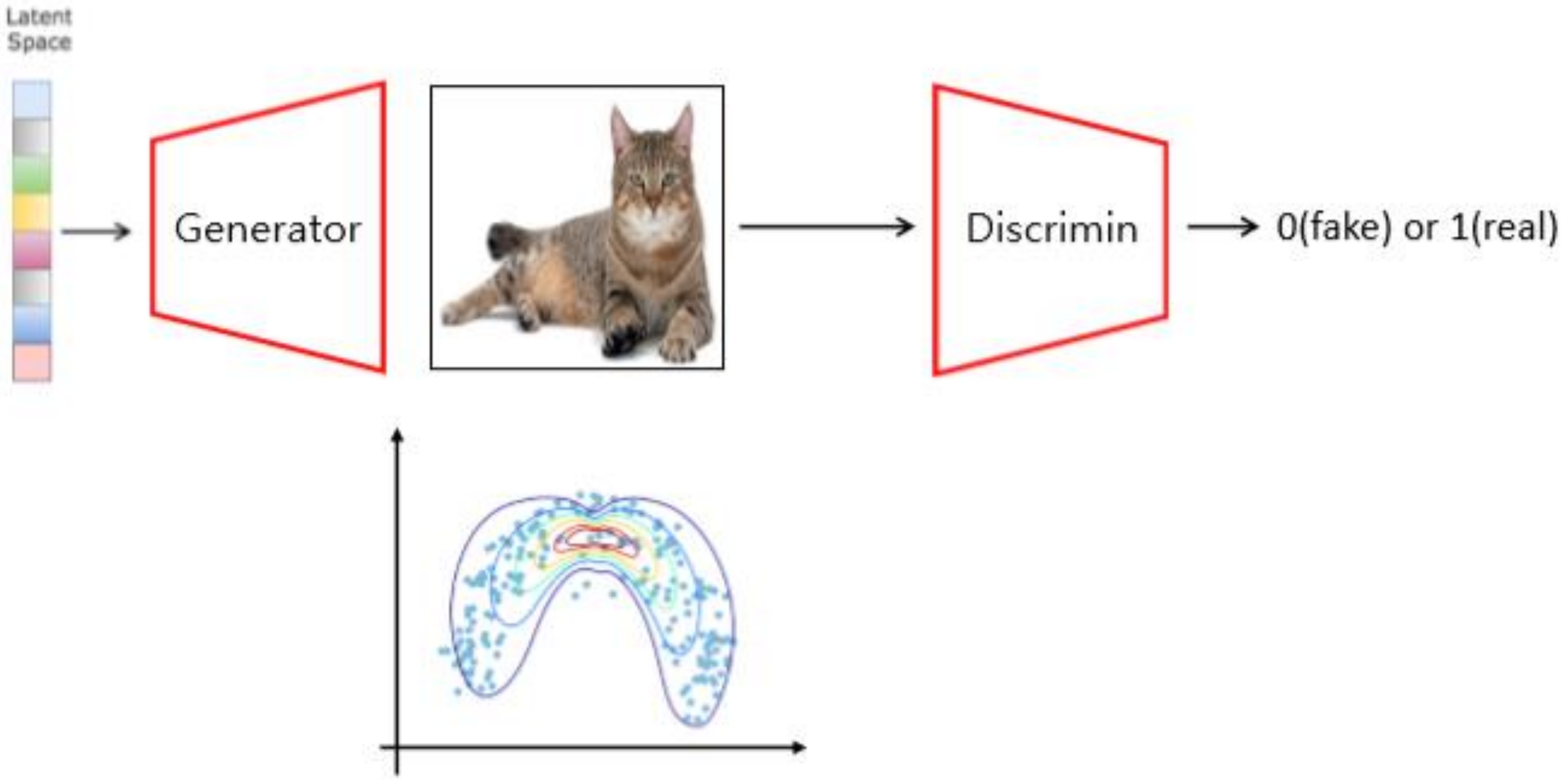
[Image classification accuracy]

DAVIAN
Data and Visual Analytics Lab

KOREA UNIVERSITY

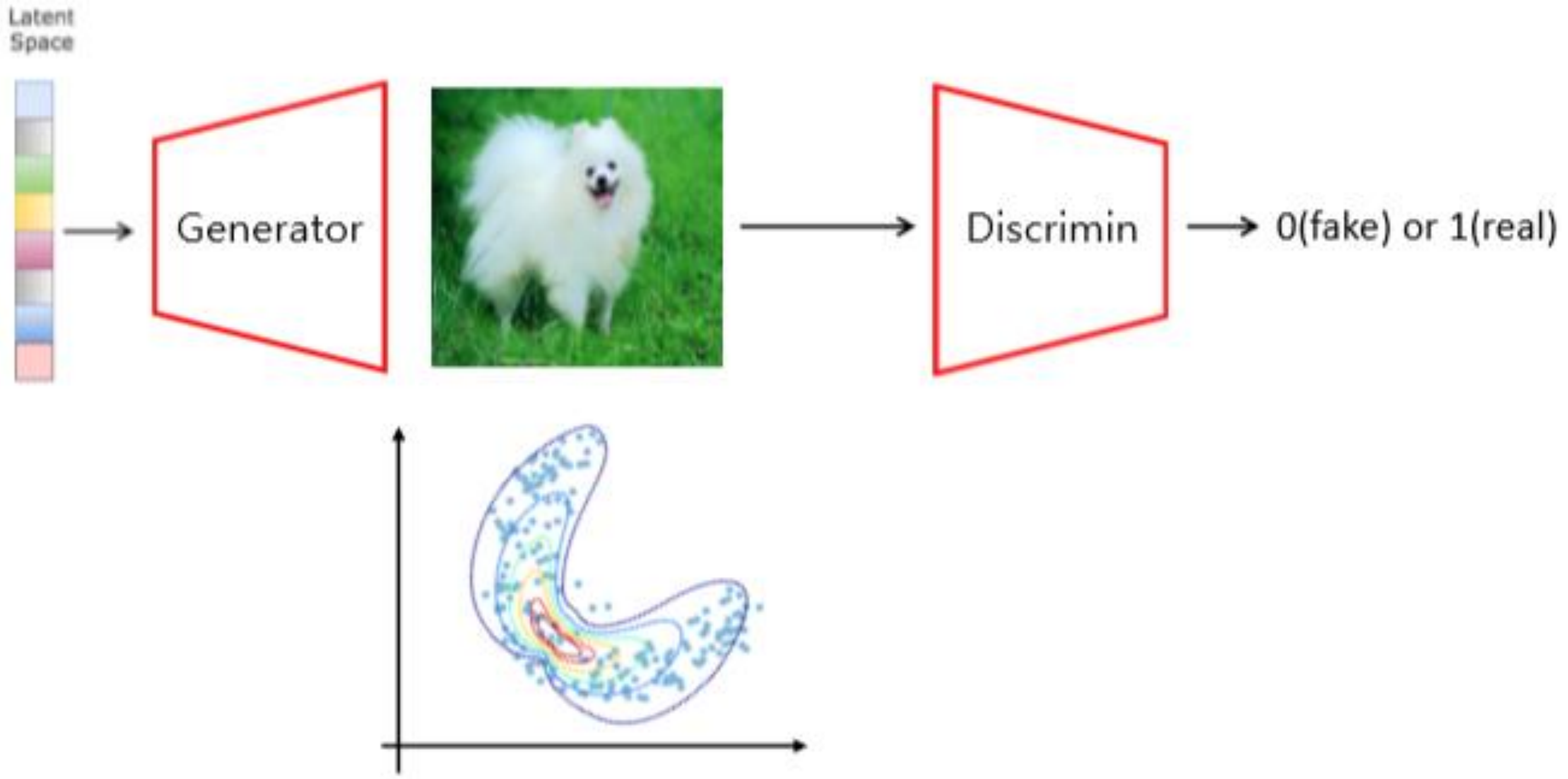
1 Introduction

Why is the GAN unstable?



1 Introduction

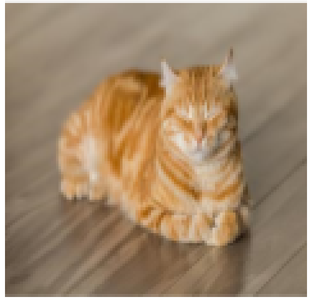
Why is the GAN unstable?



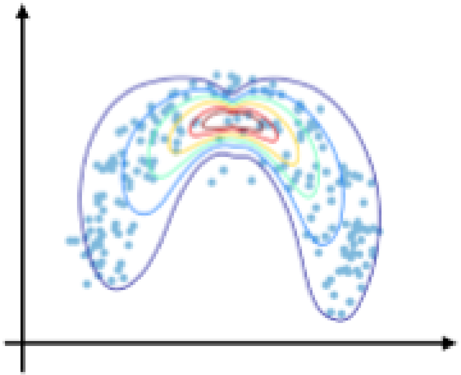
1 Introduction

Why is the GAN unstable?

Training



Learning the conditional model is easier than joint distribution!



2 Model

SSGAN

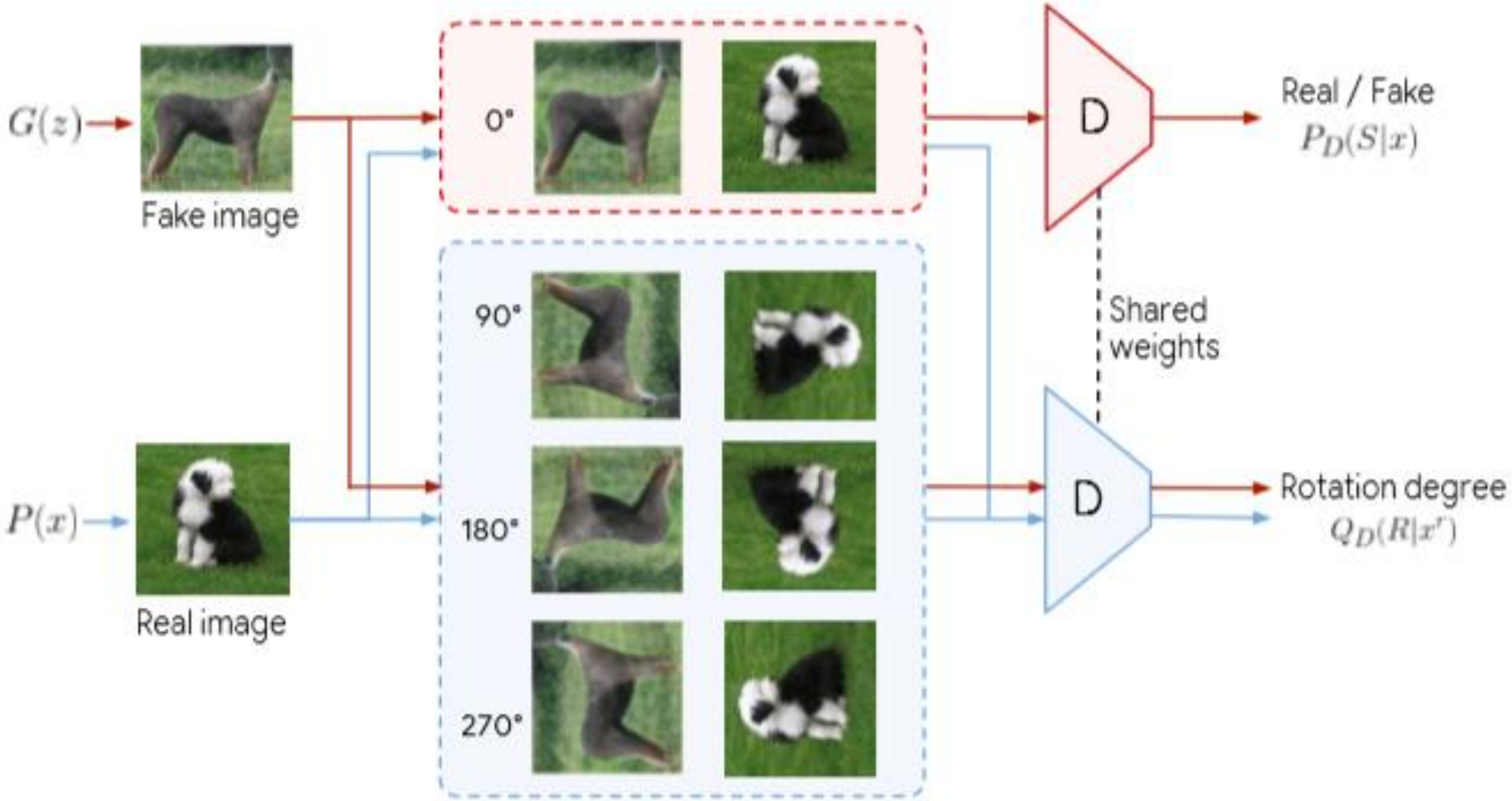


Figure 1: Discriminator with rotation-based self-supervision. The discriminator, D, performs two tasks: true vs. fake binary classification, and rotation degree classification. Both the fake and real images are rotated by 0, 90, 180, and 270 degrees. The colored arrows indicate that only the upright images are considered for true vs. fake classification loss task. For the rotation loss, all images are classified by the discriminator according to their rotation degree.

3 Experiments

Experimental Setting

- **Dataset** : CIFAR10, IMAGENET,
  LSUN-BEDROOM, CELEBA-HQ
- **GAN** : Spectral Normalization GAN
- **Hyperparameters**
  - learning rate : 0.0002
  - alpha : 0.2
  - beta : 1
  - Adam Optimizer
  - Batch Size : 16 x 4 = 64

DATASET	METHOD	FID
CIFAR10	Uncond-GAN	19.73
	Cond-GAN	**15.60**
	SS-GAN	17.11
	SS-GAN (sBN)	15.65
IMAGENET	Uncond-GAN	56.67
	Cond-GAN	**42.07**
	SS-GAN	47.56
	SS-GAN (sBN)	43.87
LSUN-BEDROOM	Uncond-GAN	16.02
	SS-GAN	13.66
	SS-GAN (sBN)	**13.30**
CELEBA-HQ	Uncond-GAN	**23.77**
	SS-GAN	26.11
	SS-GAN (sBN)	24.36

Table 1: Best FID attained across three random seeds. In this setting the proposed approach recovers most of the benefits of conditioning.

3 Experiments

Experimental Setting

Type	λ	β_1	β_2	D Iters	CIFAR10 Uncond-GAN	CIFAR10 SS-GAN	ImageNet Uncond-GAN	ImageNet SS-GAN
Gradient Penalty	1	0.0	0.900	1	121.05 ± 31.44	**25.8 ± 0.71**	183.36 ± 77.21	**80.67 ± 0.43**
				2	28.11 ± 0.66	**26.98 ± 0.54**	85.13 ± 2.88	**83.08 ± 0.38**
		0.5	0.999	1	78.54 ± 6.23	**25.89 ± 0.33**	104.73 ± 2.71	**91.63 ± 2.78**
	10	0.0	0.900	1	188.52 ± 64.54	**28.48 ± 0.68**	227.04 ± 31.45	**85.38 ± 2.7**
				2	29.11 ± 0.85	**27.74 ± 0.73**	227.74 ± 16.82	**80.82 ± 0.64**
		0.5	0.999	1	117.67 ± 17.46	**25.22 ± 0.38**	242.71 ± 13.62	**144.35 ± 91.4**
Spectral Norm	0	0.0	0.900	1	87.86 ± 3.44	**19.65 ± 0.9**	129.96 ± 6.6	**86.09 ± 7.66**
				2	20.24 ± 0.62	**17.88 ± 0.64**	80.05 ± 1.33	**70.64 ± 0.31**
		0.5	0.999	1	86.87 ± 8.03	**18.23 ± 0.56**	201.94 ± 27.28	**99.97 ± 2.75**

Table 2: FID for unconditional GANs under different hyperparameter settings. Mean and standard deviations are computed across three random seeds. Adding the self-supervision loss reduces the sensitivity of GAN training to hyperparameters.

3 Experiments

Sample Quality

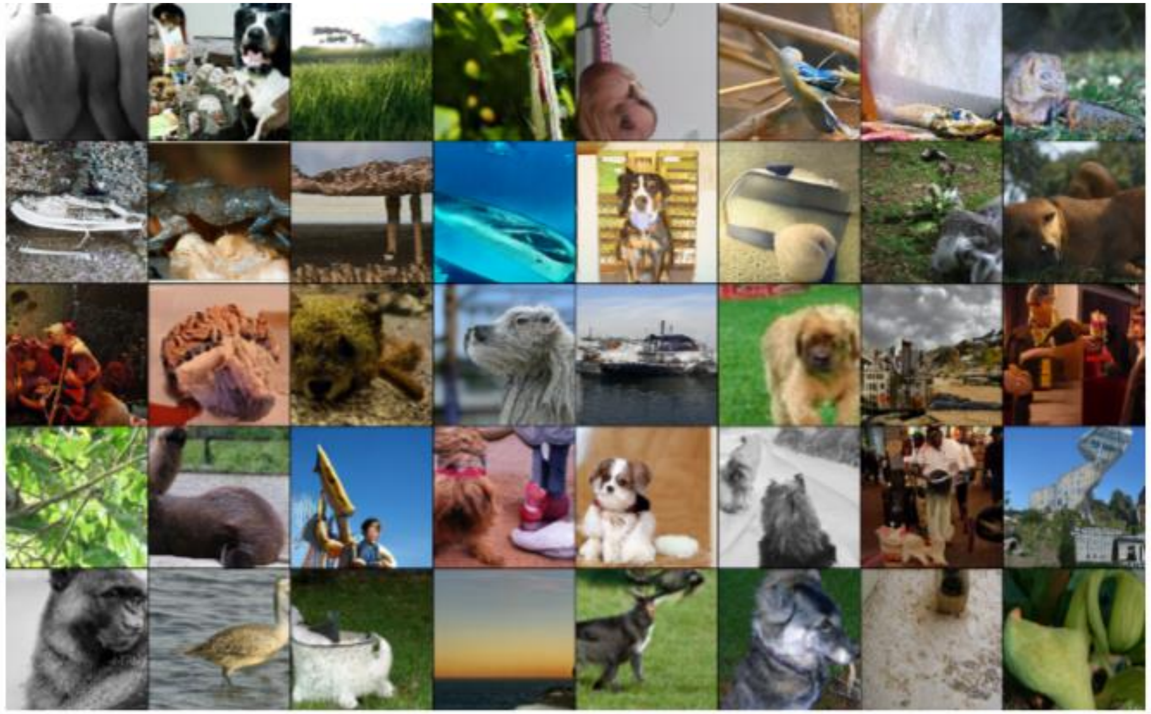
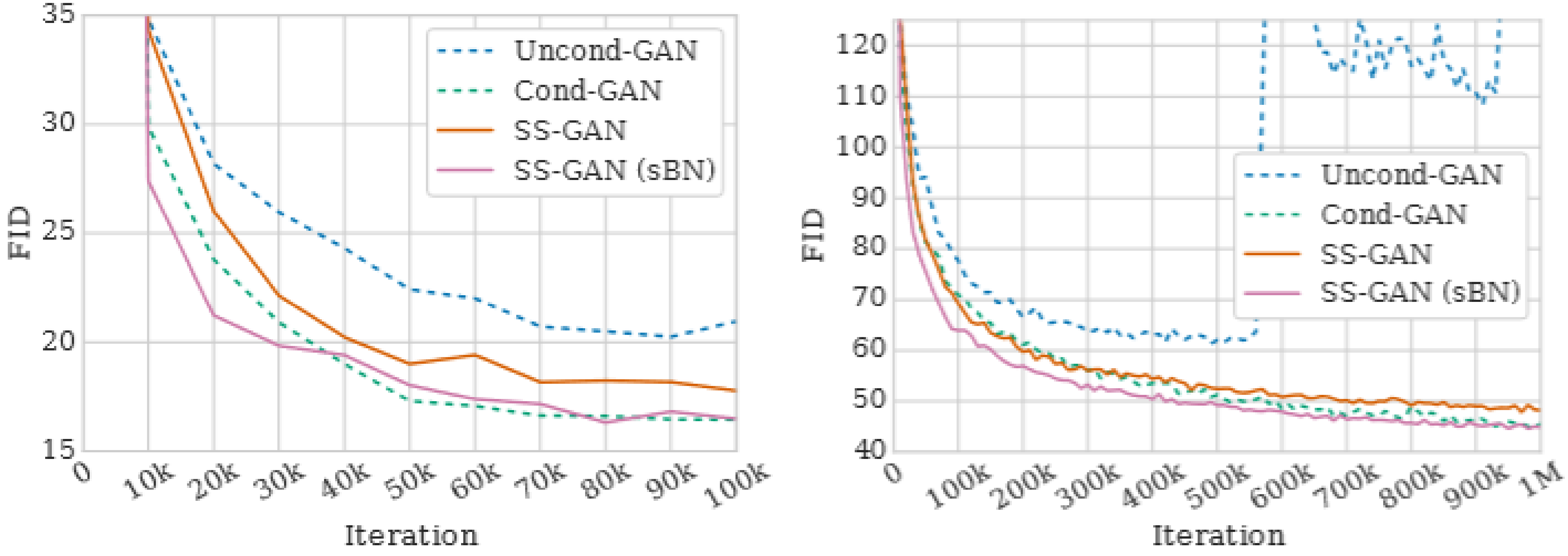


Figure 5: A random sample of unconditionally generated images from the self-supervised model. To our knowledge, this is the best results attained training unconditionally on IMAGENET.

3 Experiments

Large Scale Self-supervised GAN



(a) CIFAR10

(b) IMAGENET

DAVIAN Data and Visual Analytics Lab

KOREA UNIVERSITY

3 Experiments

Representation Quality

	Uncond.	Cond.	Rot-only	SS-GAN (sBN)
Block0	0.719	0.719	0.710	**0.721**
Block1	0.762	0.759	0.749	**0.774**
Block2	0.778	0.776	0.762	**0.796**
Block3	0.776	0.780	0.752	**0.799**
Best	0.778	0.780	0.762	**0.799**

Table 3: Top-1 accuracy on CIFAR10. Mean score across three training runs of the original model. All standard deviations are smaller than 0.01 and are reported in the appendix.

Method	Uncond.	Cond.	Rot-only	SS-GAN (sBN)
Block0	0.074	0.156	0.147	**0.158**
Block1	0.063	0.187	0.134	**0.222**
Block2	0.073	0.217	0.158	**0.250**
Block3	0.083	0.272	0.202	**0.327**
Block4	0.077	0.253	0.196	**0.358**
Block5	0.074	0.337	0.195	**0.383**
Best	0.083	0.337	0.202	**0.383**

Table 4: Top-1 accuracy on IMAGENET. Mean score across three training runs of the original model. All standard deviations are smaller than 0.01, except for Uncond-GAN whose results exhibit high variance due to training instability. All standard deviations are reported in the appendix.

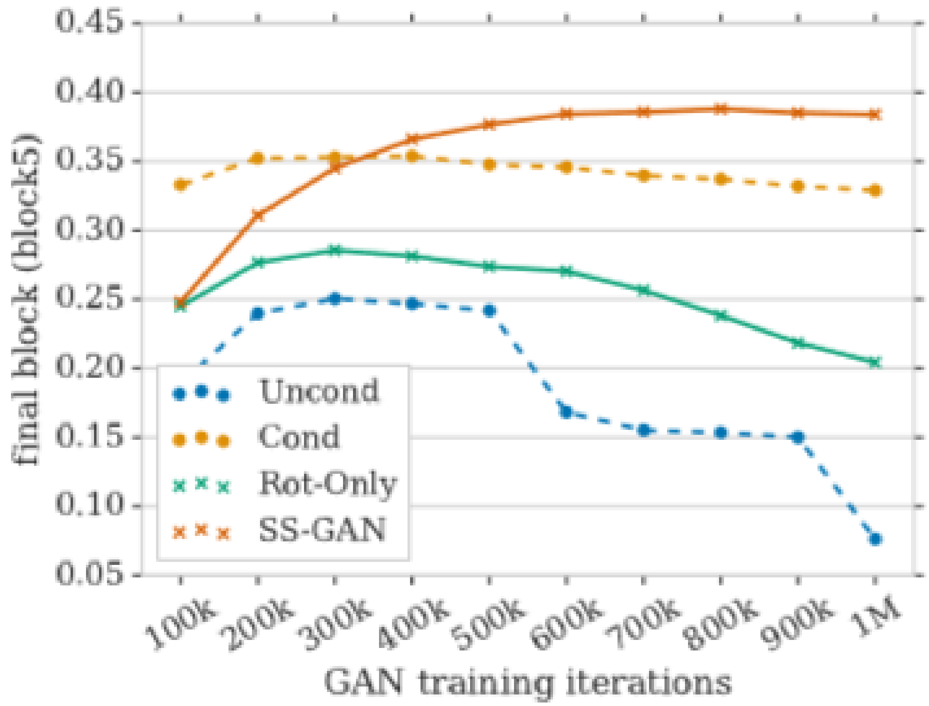


Figure 6: IMAGENET Top 1 accuracy (mean across three seeds) to predict labels from discriminator representations. X-axis gives the number of GAN training iterations.