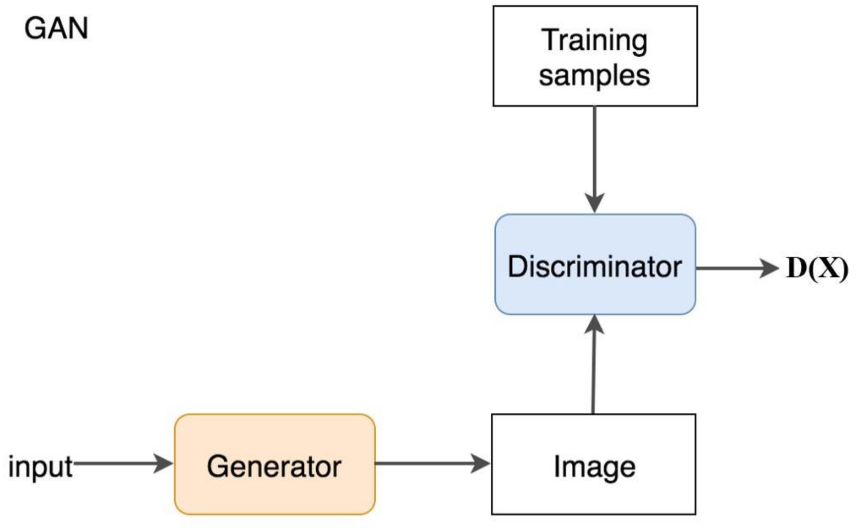
Generative Multi-Adversarial Network

Ishan Durugkar, Ian Gemp, Sridhar Mahadevan

ICLR 2017

발표자 : 김용규

Introduction



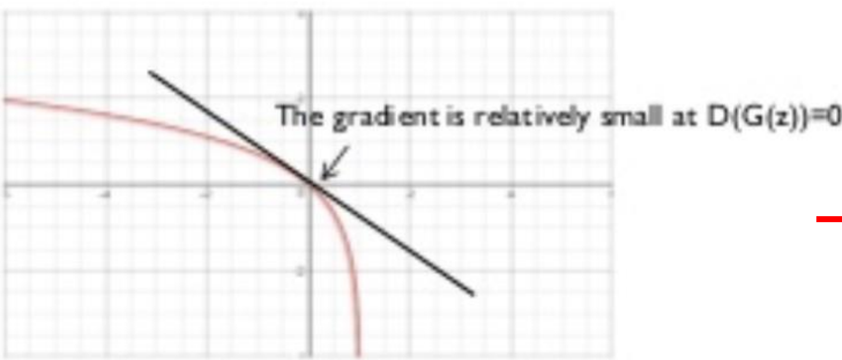
- GANs have proven useful in a variety of application domains including learning censored representations, imitating expert policies and domain transfer
- GANs are reputably difficult to train.

$$\min_{G} \max_{D \in \mathcal{D}} V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} \Big[\log(D(x)) \Big] + \mathbb{E}_{z \sim p_z(z)} \Big[\log(1 - D(G(z))) \Big]$$

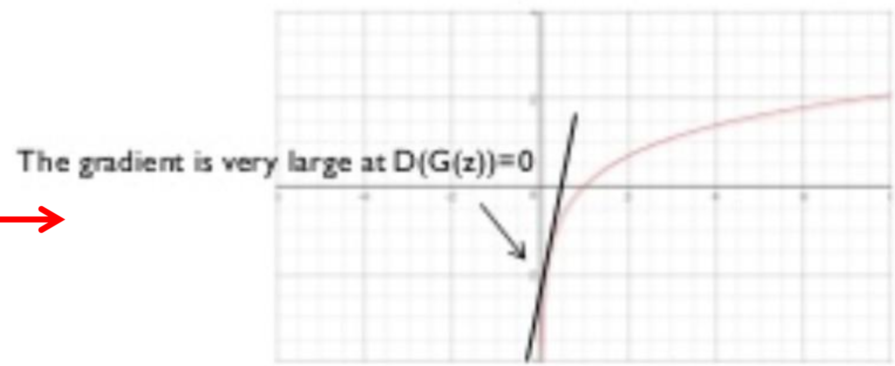
Introduction

$$\boxed{\min_{G}}\max_{D\in\mathcal{D}} V(D,G) = \cancel{\mathbb{E}_{x\sim p_{data}(x)}\left[\log(D(x))\right]} + \mathbb{E}_{z\sim p_z(z)}\left[\log(1-D(G(z)))\right]$$

$$-\log(D(G(z)))$$



$y = \log(1-x)$



$y = \log(x)$

Contributions

- A multi-discriminator GAN framework, GMAN, that allows training with the original, untampered minimax objective
- A generative multi-adversarial metric (GMAM) to perform pairwise evaluation of separately trained frameworks
- A particular instance of GMAN that allows the generator to automatically regulate training and reach higher performance in a fraction of the training time required for the standard GAN model.

Contributions

1. 여러 개 다양한 D 사용

2. A generative multi-adversarial metric (GMAM) 제시

3. G가 스스로 학습 규제

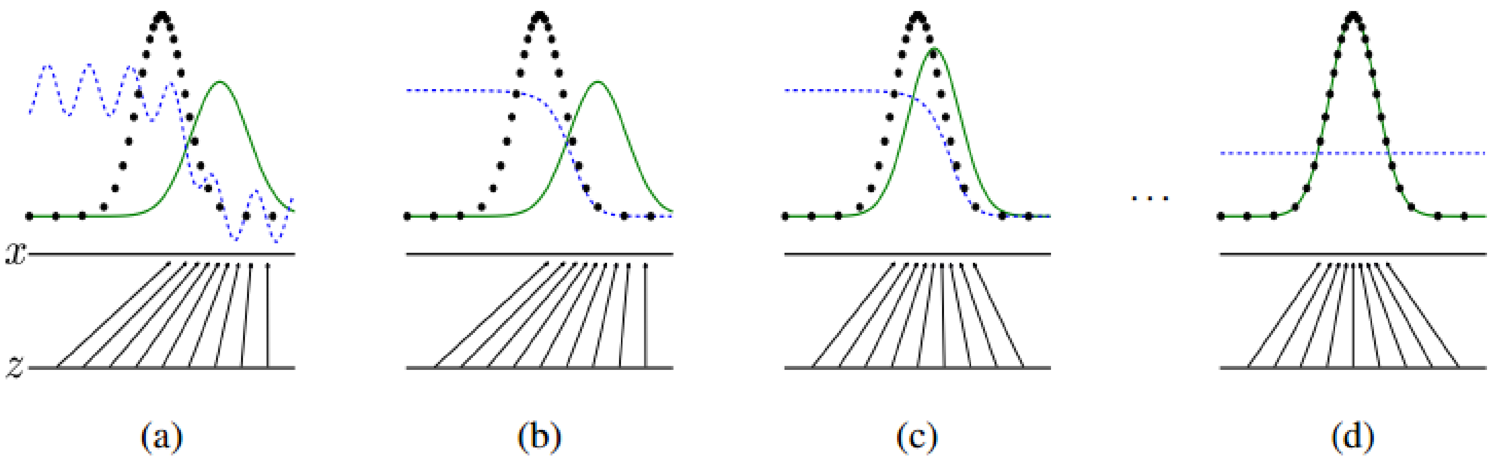
Methods

여러 개 다양한 D 사용

1) A more discriminating D better approximating $\max_{\mathcal{D}} V(D,G)$

$$\min_{G} \boxed{\max_{D \in \mathcal{D}} V(D,G)} = \mathbb{E}_{x \sim p_{data}(x)}\left[\log(D(x))\right] + \mathbb{E}_{z \sim p_z(z)}\left[\log(1 - D(G(z)))\right]$$

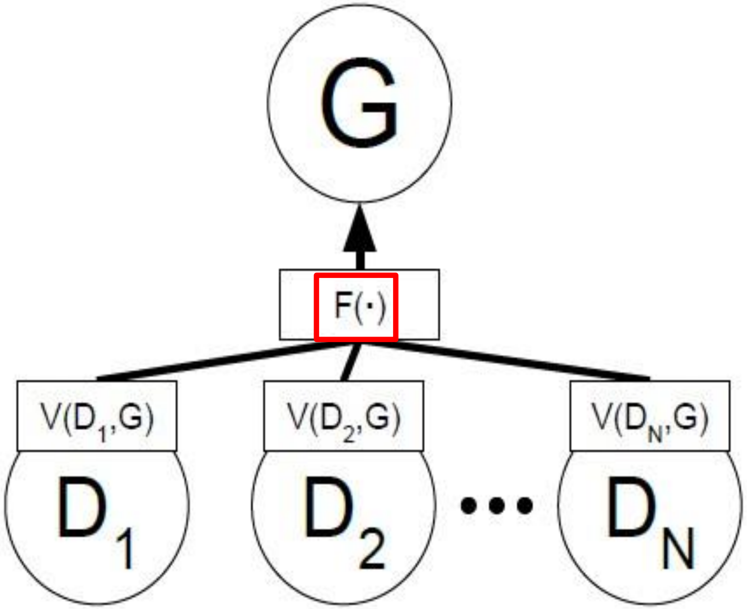
2) A D better matched to the generator’s capabilities.



Methods

여러 개 다양한 D 제시

$$\min_G \max_D V(D,G) \rightarrow \min_G \max \boxed{F}(V(D_1,G),\ldots,V(D_N,G))$$

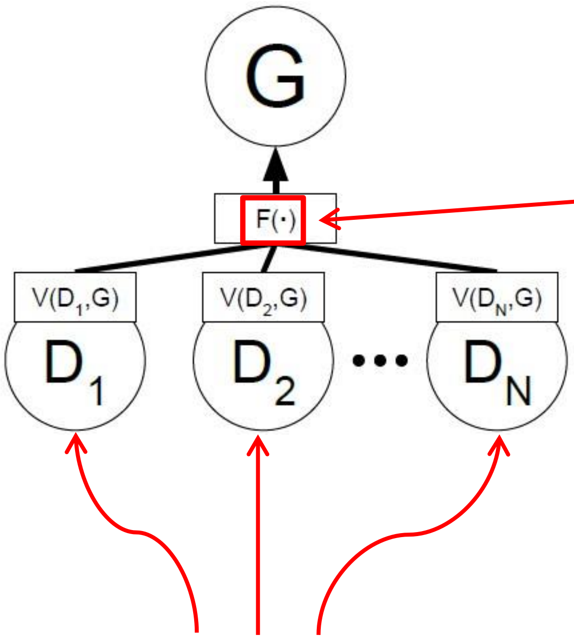


$$V(D_i,G) \rightarrow V_i$$

$$F(V_1,\ldots,V_N) \rightarrow F_G(V_i)$$

Methods

Formidable Adversary : Maximizing V(D, G)



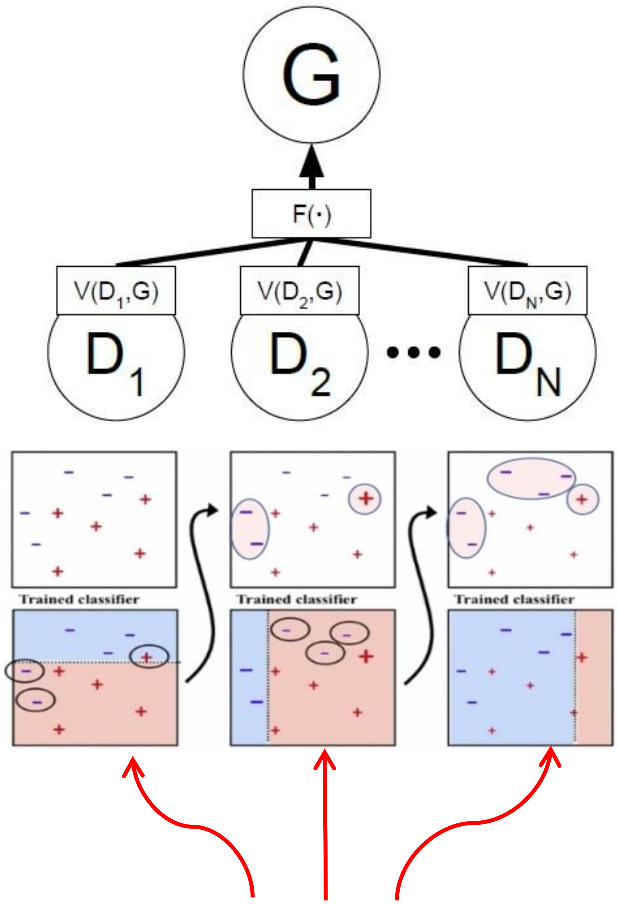
$$\max_{D_i \in \mathcal{D}} V(D_i, G)$$

$$\max\{V_1(t), \ldots, V_N(t)\} > \max\{V_1'(t)\}\,\forall t$$

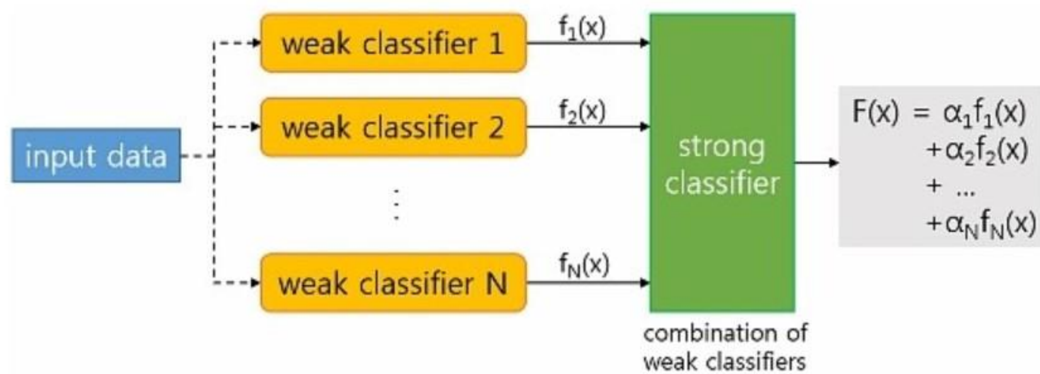
generated image
by G

Methods

Boosting



generated image
by G



Methods

Soft-Discriminator : Forgiving Teacher

$$V(D_i, G)$$

$$\mathrm{AM}_{soft}(V, \lambda) = \sum_i^N w_i V_i$$

$$\mathrm{GM}_{soft}(V, \lambda) = -\exp\Big(\sum_i^N w_i \log(-V_i)\Big)$$

$$\mathrm{HM}_{soft}(V, \lambda) = \Big(\sum_i^N w_i V_i^{-1}\Big)^{-1}$$

$w_i = e^{\lambda V_i}/\Sigma_j e^{\lambda V_j}$ with $\lambda \geq 0, V_i < 0$

$\lambda = 0$ corresponds to the mean

the max is recovered as $\lambda \to \infty$

Methods

Soft-Discriminator : Forgiving Teacher

soft versions of the three classical Pythagorean means

$$\mathrm{AM}(x_1,\ \ldots,\ x_n) = \frac{1}{n}\left(x_1 + \ \cdots\ + x_n\right)$$

$$\mathrm{GM}(x_1,\ \ldots,\ x_n) = \sqrt[n]{|x_1 \times\ \cdots\ \times x_n|}$$

$$\mathrm{HM}(x_1,\ \ldots,\ x_n) = \frac{n}{\frac{1}{x_1} + \ \cdots\ + \frac{1}{x_n}}$$

Using the original minimax objective

- G training phase

$$\frac{1}{N}\sum_{i}^{N}\mathbb{E}_{x\sim p_G(x)}\Big[\log(1-D_i(x))\Big] = \frac{1}{N}\mathbb{E}_{x\sim p_G(x)}\Big[\log(z)\Big].$$

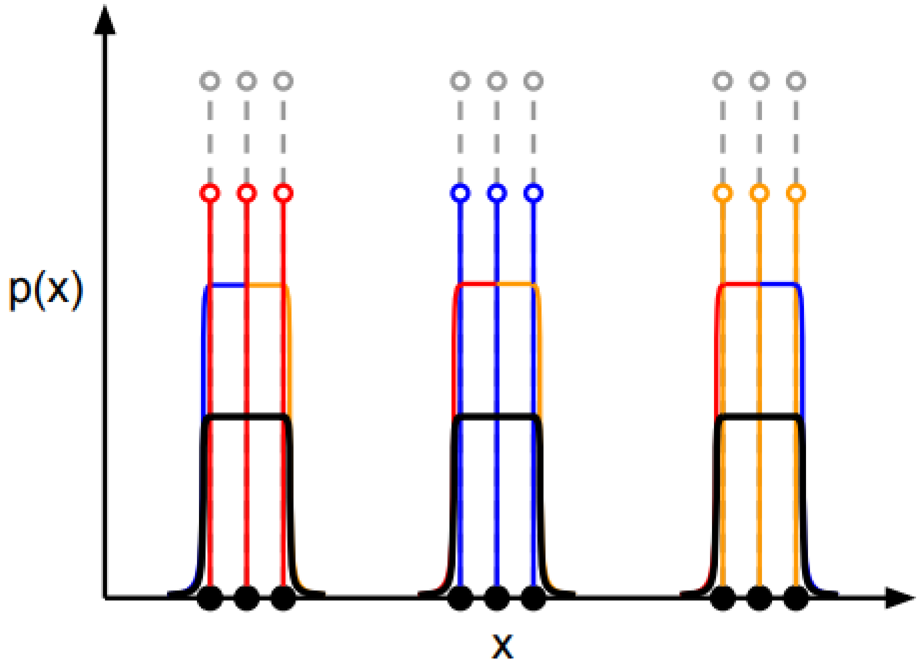
$$z = \prod_i^N (1-D_i(x)).$$

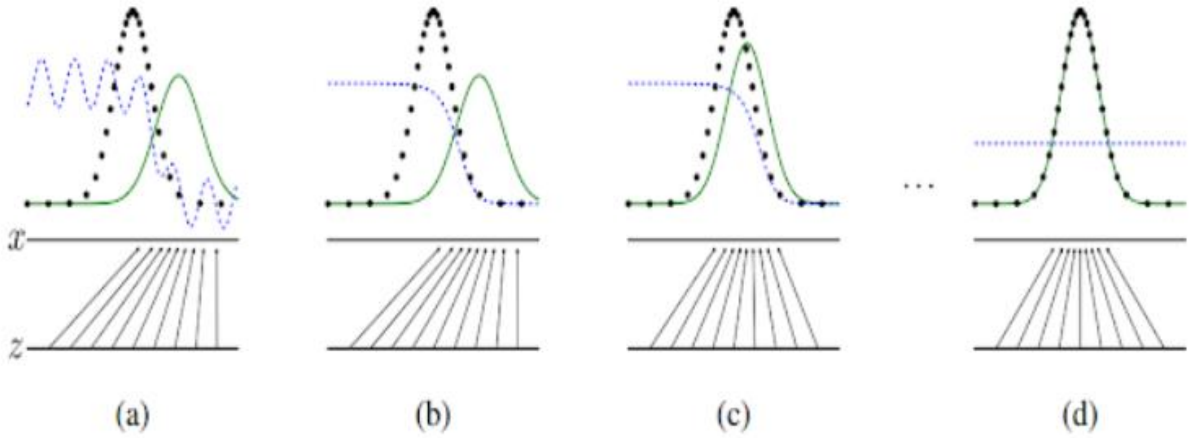
possible z = 0~1.
z = 1 일때, f(x)=0 => 학습이 안됨
(모든 D_i 가 generated image x 에 대해 fake로 판별한 상황에 발생)

⇒ unlikely for large N.
⇒ original generator objective 사용해도 무방 ($-\log(D)$를 사용 안해도 학습에 문제 없음)

Justification of multi-discriminators

- Each discriminator may specialize in discriminating a region of the data space





original GAN paper

- Averaging over these multiple locally optimal discriminators increases the entropy of a generated dist.

Automatic Regulation

- 처음에는 평균으로, 학습이 진행되면서 하나의 discriminator에 의존하도록. => 이것을 adaptive하게 해보자

$$\min_{G,\lambda>0} F_G(V_i) - f(\lambda)$$

$f(\lambda) = c\lambda$ with c a constant (e.g., 0.001).

- generator training phase에서 적용,
- 네트워크에 lambda를 커지게 해야하는 제약을 걸어줌

Experiments

- metric

$$\mathrm{GMAM} = \log\left(\frac{F^a_{G_b}(V^a_i)}{F^a_{G_a}(V^a_i)} \bigg/ \frac{F^a_{G_a}(V^b_i)}{F^b_{G_b}(V^b_i)}\right).$$

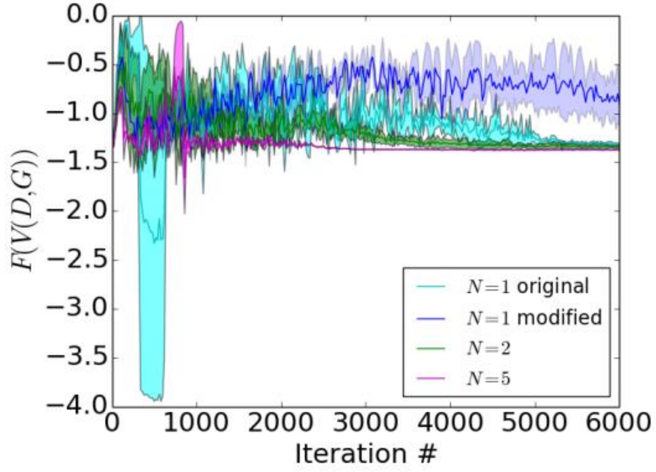
- 상대 평가 메트릭
- 서로 다른 두 셋 (G_a, D_a1, D_a2,…)
- G_b 가 D_a와 D_b를 더 잘 속이면? >0
- G_a 가 D_a와 D_b를 더 잘 속이면? <0

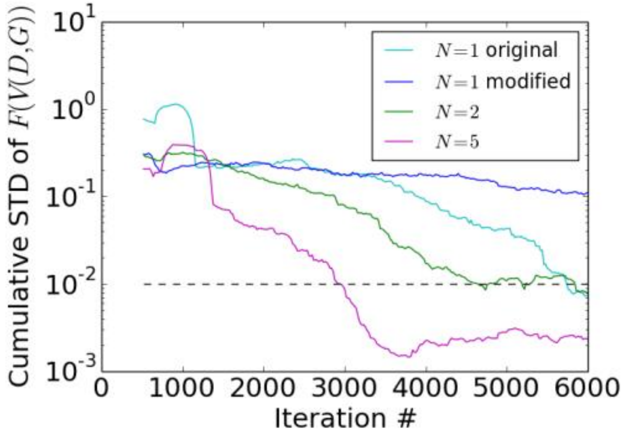
- Baseline

- GMAN-max: $\max\{V_i\}$ is presented to the generator.
- GAN: Standard GAN with a single discriminator (see Appendix A.2).
- mod-GAN: GAN with modified objective (generator minimizes $-\log(D(G(z)))$).
- GMAN-λ: GMAN with F :=arithmetic *softmax* with parameter λ.
- GMAN*: The arithmetic *softmax* is controlled by the generator through λ.

Experiments (MNITST)

	Score	Variant	GMAN*	GMAN-0	GMAN-max	mod-GAN
Better→	**0.127**	GMAN*	-	-0.020 ± 0.009	-0.028 ± 0.019	-0.089 ± 0.036
	0.007	GMAN-0	0.020 ± 0.009	-	-0.013 ± 0.015	-0.018 ± 0.027
	-0.034	GMAN-max	0.028 ± 0.019	0.013 ± 0.015	-	-0.011 ± 0.024
	-0.122	mod-GAN	0.089 ± 0.036	0.018 ± 0.027	0.011 ± 0.024	-

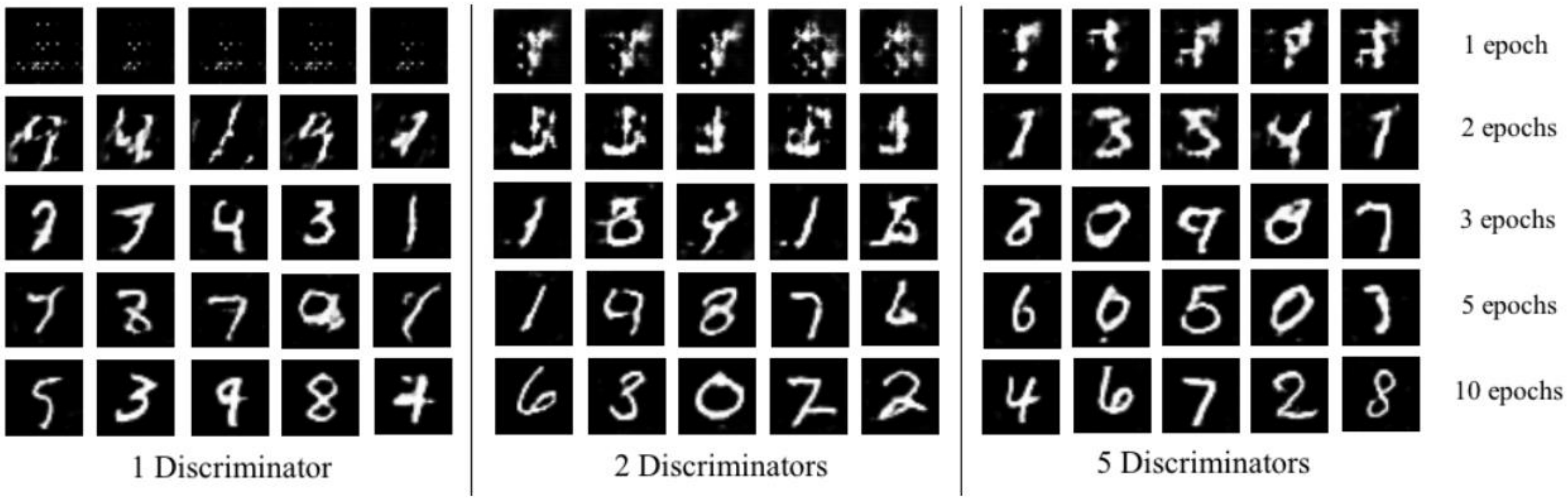


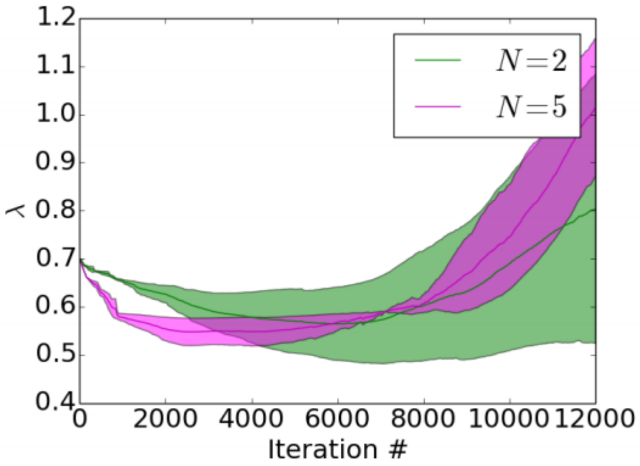


- (left figure) The more discriminators, the faster convergence.
- (right figure) Lower values indicate a more steady-state. (represented as shadows in the left fig.)

Experiments (MNITST)

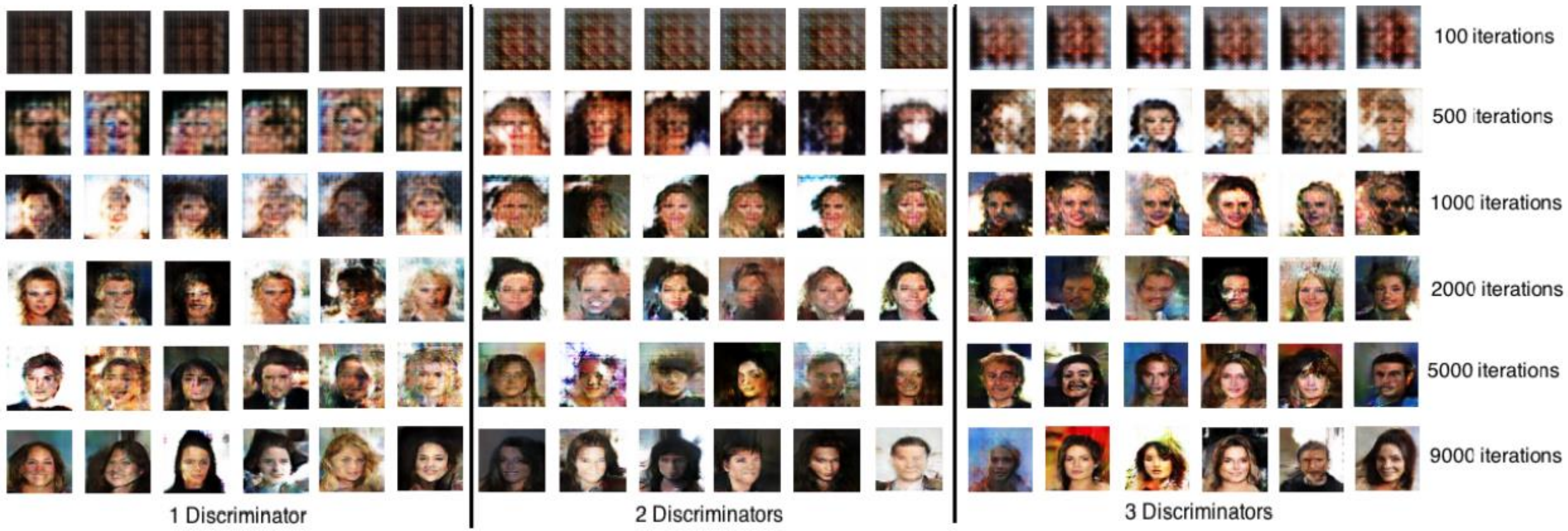
- digits at steady-state appear slightly sharper as well





	Score	λ $(N=5)$	λ^*	$\lambda=1$	$\lambda=0$
Better→	**0.028**	λ^*	-	$\frac{-0.008}{\pm 0.009}$	$\frac{-0.019}{\pm 0.010}$
	0.001	$\lambda=1$	$\frac{0.008}{\pm 0.009}$	-	$\frac{-0.008}{\pm 0.010}$
	-0.025	$\lambda=0$	$\frac{0.019}{\pm 0.010}$	$\frac{0.008}{\pm 0.010}$	-

Experiments (CelebA & CIFAR10)



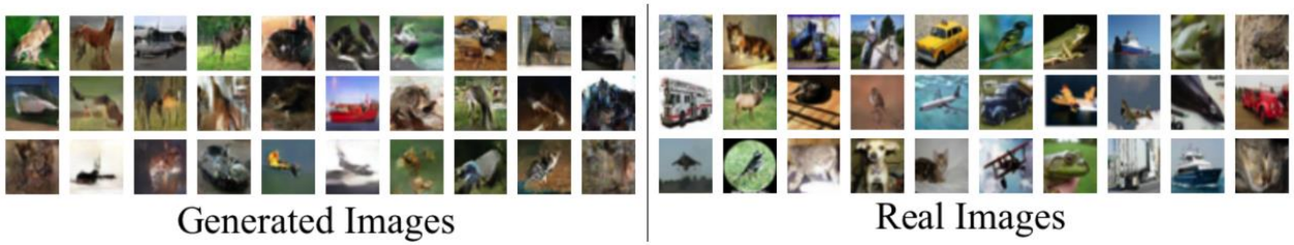


Figure 9: Images generated by GMAN-0 on the CIFAR-10 dataset.