EUROPEAN CONFERENCE ON COMPUTER VISION

MILANO 2024

CLR-GAN: Improving GANs Stability and Quality via Consistent Latent Representation and Reconstruction

Shengke Sun[1,*], Ziqian Luan[2,*], Zhanshan Zhao[1,3], Shijie Luo[1], Shuzhen Han[1]

(* Equal contribution)

1 TIANGONG UNIVERSITY ~1912~ 天津工业大学

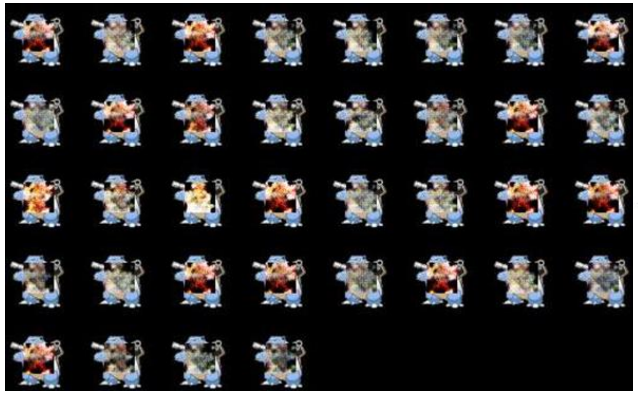
2 XIDIAN UNIVERSITY 1931

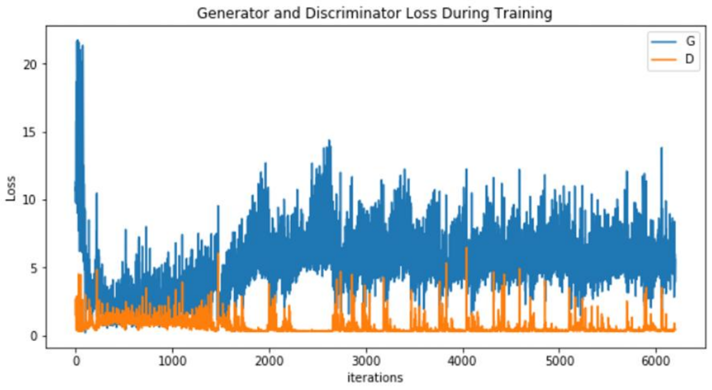
3 University of Victoria

Image Synthesis using GANs

- Pros
  - Fast inference
  - High qualty
  - Easy to edit

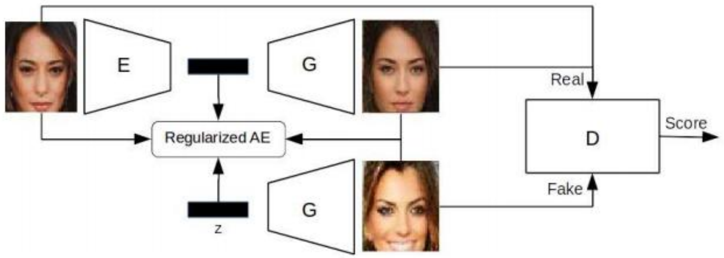
- Cons
  - Hard to train
  - Mode collapse



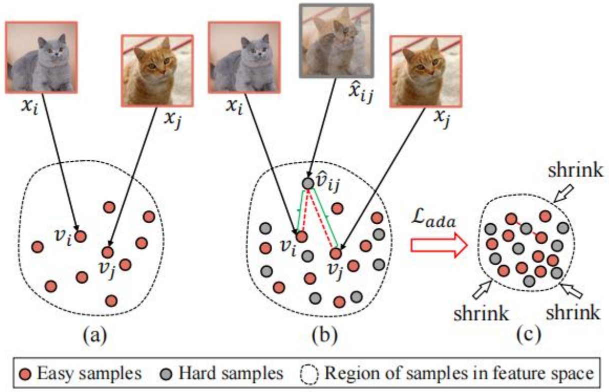


Related Works

- Discriminator based

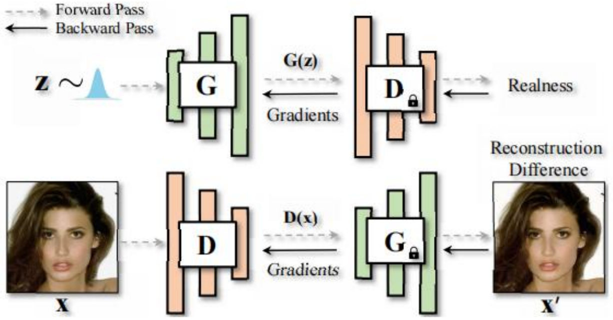


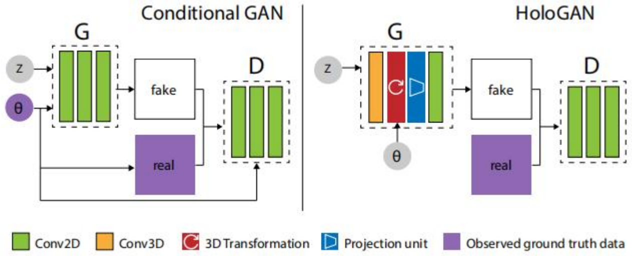
Regulize Discriminator



Build Hard Samples

- Reconstruction based



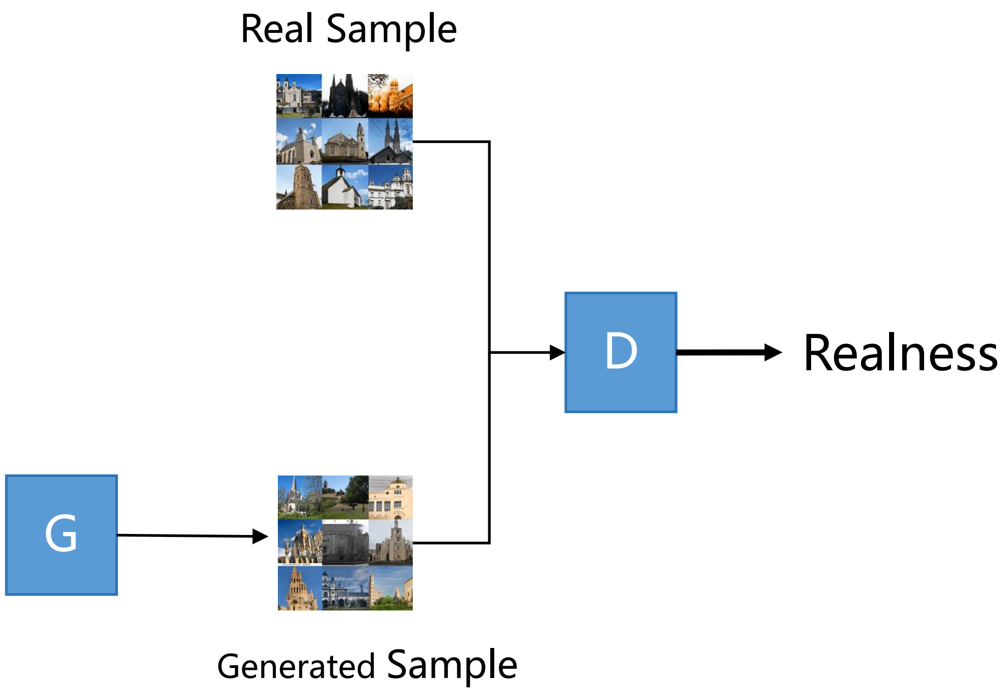


Constrain Discriminator

Delve into the paradigm

- Gradient Perspective
    - All gradient comes from D
    - Training D first

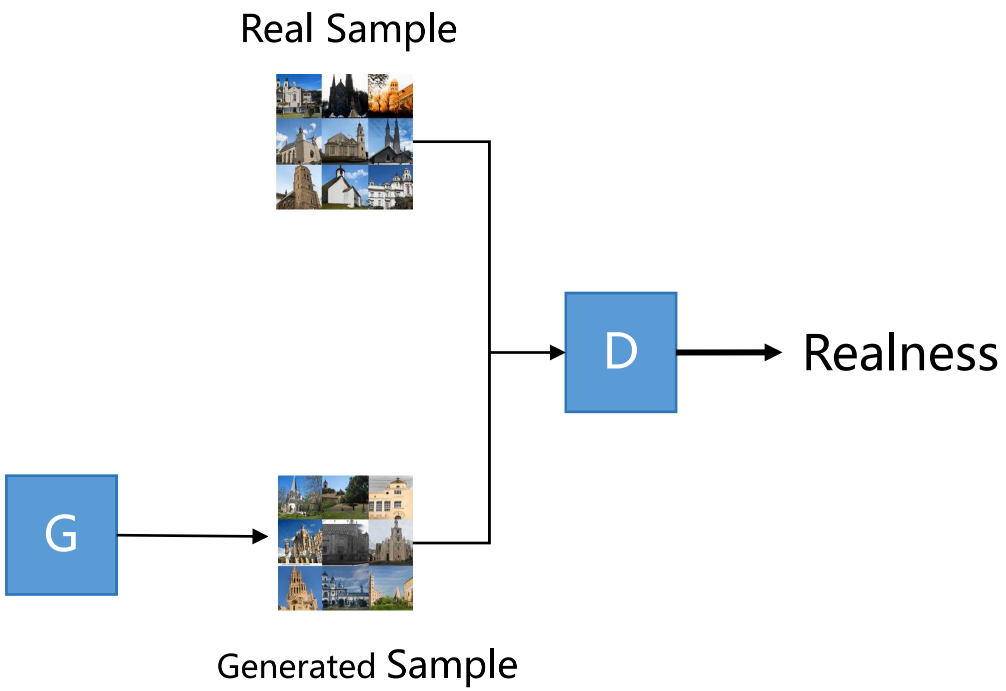
    - Makes D a natural dominant



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

Training Objective

Delve into the paradigm



- Data Perspective
    - Only D has access to data
    - G only update according to D
    - Gradient space more abstract

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

Training Objective

- G lacks of priors

How to stablize training from the above perspectives?

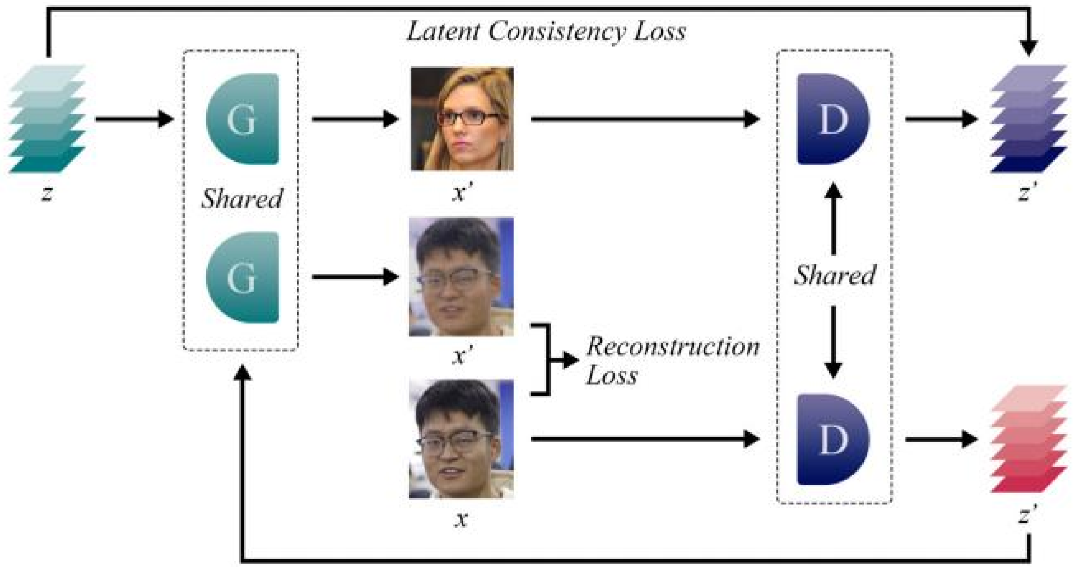
Proposed Method—CLR-GAN

- To constrain the D using gradient from G
  - Reconstruction Loss

$$\mathcal{L}_{rec} = \frac{1}{n}\sum_{i=1}^{n} \|I(x) - I_{rec}\|_1$$

- To make G has access to real data distribution
  - Latent Consistency Loss

$$\mathcal{L}_{CLR} = \frac{1}{n}\sum_{i=1}^{n} \|z_i - \phi(x_i)\|_1$$



Theoretical Analysis

- Training D with fixed G

$$\nabla_{\theta_d} V(D) = \mathbb{E}_{x \sim p_{data}}[\nabla_{\theta_d} log D(x)] - \mathbb{E}_{x,z} \text{sign}(D(x) - z)$$

- Reduce the gradient
- Alleviate gradient explosion

- Training G with fixed D

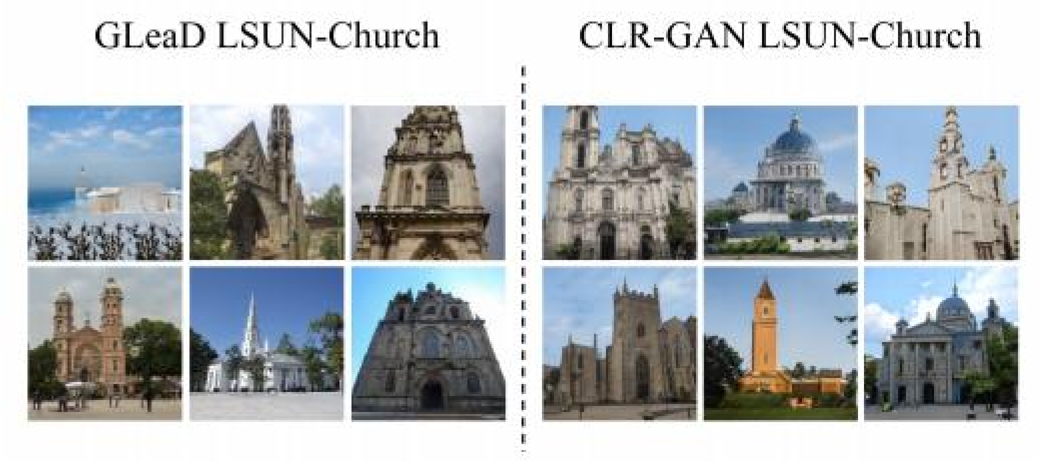
$$\nabla_{\theta_g} V(G) = \mathbb{E}_{z \sim p_z}[\nabla_{\theta_d, \theta_g} log D(1 - D(G(z)))] + \mathbb{E}_{x \sim p_{data}} \text{sign}(G(D(x)) - x)$$

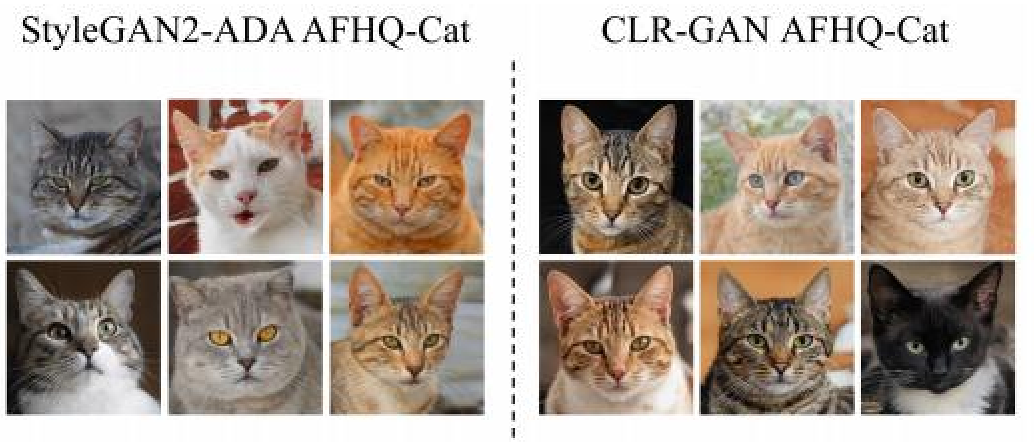
- Enhance the gradient
- Alleviate gradient vanish

Experimental & Visualization Results

Learning Objective	CelebA	CIFAR-10
WGAN [1]	36.47	55.96
HingeGAN [54]	25.57	42.4
LSGAN [36]	30.76	42.01
DCGAN [40]	27.02	38.56
WGAN-GP [14]	70.28	41.86
Realness GAN-Obj.1 [49]	-	36.73
Realness GAN-Obj.2 [49]	23.51	34.59
Realness GAN-Obj.3 [49]	-	36.21
AdaptiveMix [32]	**12.43**	30.85
CLR-GAN(Ours)	13.63	**23.3**

Method	AFHQ-Cat			FFHQ			LSUN Church		
	FID ↓	P ↑	R ↑	FID ↓	P ↑	R ↑	FID ↓	P ↑	R ↑
StyleGAN-V2 [22]	7.92	0.68	0.27	3.86	0.68	0.25	4.04	0.58	0.40
LC-Reg [46]	6.70	-	-	3.93	-	-	4.07	-	-
StyleGAN-V2-ADA [19]	6.05	0.66	0.25	4.01	0.66	0.26	4.01	0.61	0.43
StyleGAN-V2-APA [18]	4.88	0.65	0.30	3.75	0.67	0.29	3.92	0.60	0.43
StyleGAN-V2 + Ours	4.79(-3.13)	**0.76**	0.28	3.44(-0.42)	0.70	0.41	3.52(-0.52)	**0.63**	0.46
StyleGAN-V2-ADA + Ours	**4.45(-0.43)**	0.74	**0.33**	**3.37(-0.64)**	**0.71**	**0.44**	**3.43(-0.58)**	0.61	**0.48**





EUROPEAN CONFERENCE ON COMPUTER VISION
MILANO 2024

CLR-GAN: Improving GANs Stability and Quality via Consistent Latent Representation and Reconstruction

Shengke Sun[1,*], Ziqian Luan[2,*], Zhanshan Zhao[13], Shijie Luo[1], Shuzhen Han[1]

(* Equal contribution)

sunshengke@tiangong.edu.cn

Project Page

Oct 1, 10:30-12:30(GMT+2)

Zoom:
https://us05web.zoom.us/j/6682329594?pwd=eo6TpuMzST2WR5KcMhumaHPA0nlQiV.1&omn=87392099087

1 TIANGONG UNIVERSITY

2 XIDIAN UNIVERSITY

3 University of Victoria