Lecture 8
Generative Adversarial Networks

Deep Learning for Computer Vision
Valeriya Strizhkova
9 November 2021

About myself

Valeriya Strizhkova

1st year PhD student @ Inria, STARS team

https://scholar.google.ru/citations?user=6n5PrUAAAAAJ&hl

https://github.com/valerystrizh

Lecture Structure

- GANs: Valeriya Strizhkova
- DeepFake Detection: Dr. Antitza Dantcheva

Outline

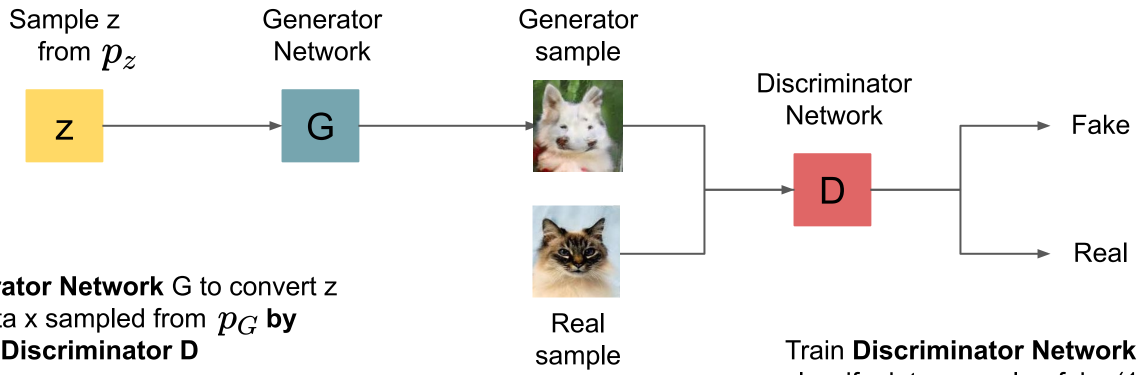
- Basic idea of GAN
- Image generation
  - Conditional GAN
  - Image-to-image translation (Pix2Pix, CycleGAN)
  - StyleGAN
- Video Generation

Generative Adversarial Networks

- **Setup**: Assume we have data x_i drawn from distribution $p_{data}(x)$. Want to sample from p_{data}.
- **Idea**: Introduce a latent variable z with simple prior p(z).
- Sample $z \sim p(z)$ and pass to a Generator Network $x = G(z)$
- Then x is a sample from the Generator distribution p_G. Want $p_G = p_{data}$

Generative Adversarial Networks

- Setup: Assume we have data x_i drawn from distribution $p_{data}(x)$. Want to sample from p_{data}.
- Idea: Introduce a latent variable z with simple prior p(z).
- Sample $z \sim p(z)$ and pass to a Generator Network $x = G(z)$
- Then x is a sample from the Generator distribution p_G. Want $p_G = p_{data}$



Train **Generator Network** G to convert z into fake data x sampled from p_G **by fooling the Discriminator D**

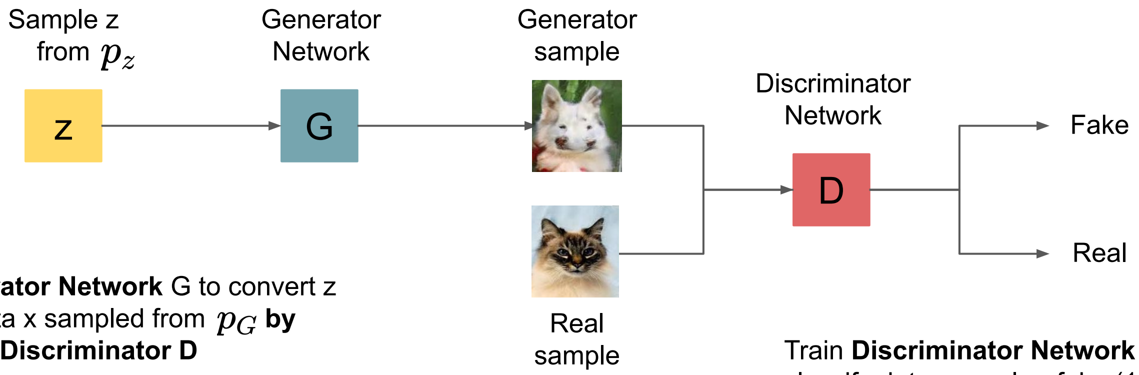
Train **Discriminator Network** D to classify data as real or fake (1/0)

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Training Objective

Jointly train generator G and discriminator D with a minimax game

$$\min_{G} \max_{D} \left(E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \right)$$



Train **Generator Network** G to convert z into fake data x sampled from p_G **by fooling the Discriminator D**

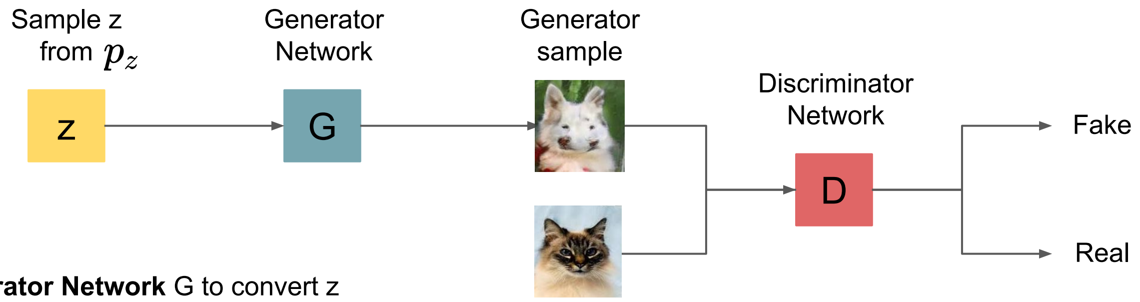
Train **Discriminator Network** D to classify data as real or fake (1/0)

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Training Objective

Jointly train generator G and discriminator D with a minimax game

$$\min_{G} \max_{D} \left(E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \right)$$



Train **Generator Network** G to convert z into fake data x sampled from p_G **by fooling the Discriminator D**

Train **Discriminator Network** D to classify data as real or fake (1/0)

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Training Objective

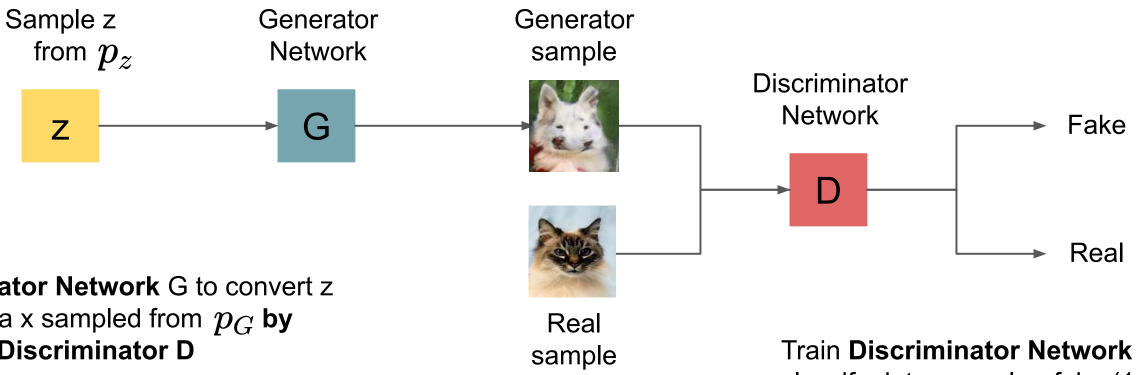
Jointly train generator G and discriminator D with a **minimax game**

Discriminator wants D(x)=1 for real data

Discriminator wants D(x)=0 for fake data

$$\min_{G} \max_{D} \left(E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \right)$$

Generator wants D(x)=1 for fake data



Train **Generator Network** G to convert z into fake data x sampled from p_G **by fooling the Discriminator D**

Train **Discriminator Network** D to classify data as real or fake (1/0)

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Training Objective

Jointly train generator G and discriminator D with a **minimax game**

$$\min_{G} \max_{D} \left(E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \right)$$

$$= \min_{G} \max_{D} V(G, D)$$

Train G and D using alternating gradient updates:

1. Update $D = D + \alpha_D \frac{\delta V}{\delta D}$
2. Update $G = G + \alpha_G \frac{\delta V}{\delta G}$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: vanishing gradient

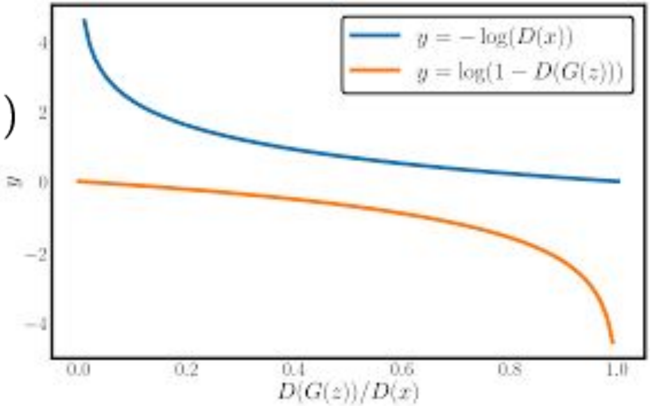
$$\min_G \max_D V(G,D) = \min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

$$\nabla_{\Theta_G} V(G,D) = \nabla_{\Theta_G} E_{z\sim q(z)}[log(1 - D(G(z)))]$$

$$\nabla_a \log(1-\sigma(a)) = \frac{-\nabla_a \sigma(a)}{1-\sigma(a)} = \frac{-\sigma(a)(1-\sigma(a))}{1-\sigma(a)} = -\sigma(a) = -D(G(z))$$

- Gradient goes to 0 if D is confident, i.e. $D(G(z)) \to 0$
- Minimize $-E_{z\sim p(z)}[\log(D(G(z)))]$ for **Generator** instead

  (keep Discriminator as it is)



Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D \left(E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))]\right)$$

$$= \min_G \max_D \left(E_{x\sim p_{data}}[\log D(x)] + E_{x\sim p_G}[\log(1 - D(x))]\right)$$

$$= \min_G \max_D \int_X (p_{data}(x)\log D(x) + p_G(x)\log(1 - D(x)))dx$$

$$= \min_G \int_X \max_D (p_{data}(x)\log D(x) + p_G(x)\log(1 - D(x)))dx$$

$$f(y) = a\log y + b\log(1-y)$$

$$f'(y) = \frac{a}{y} - \frac{b}{1-y}$$

$$f'(y) = 0 \Leftrightarrow y = \frac{a}{a+b}$$

Optimal Discriminator:

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}$$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{x\sim p_G}[\log(1 - D(x))])$$

$$= \min_G \int_X (p_{data}(x) \log D_G^*(x) + p_G(x) \log(1 - D_G^*(x)))dx$$

Optimal Discriminator: $D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1-D(G(z)))])$$

$$= \min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{x\sim p_G}[\log(1-D(x))])$$

$$= \min_G \int_X (p_{data}(x) \log D_G^*(x) + p_G(x) \log(1-D_G^*(x)))dx$$

$$= \min_G \int_X (p_{data}(x) \log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)} + p_G(x) \log \frac{p_G(x)}{p_{data}(x)+p_G(x)})dx$$

Optimal Discriminator: $D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x \sim p_{data}}[\log D(x)] + E_{z \sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G \int_X (p_{data}(x) \log D_G^*(x) + p_G(x) \log(1 - D_G^*(x)))dx$$

$$= \min_G \int_X (p_{data}(x) \log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)} + p_G(x) \log \frac{p_G(x)}{p_{data}(x)+p_G(x)})dx$$

$$= \min_G (E_{x \sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x \sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}])$$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x \sim p_{data}}[\log D(x)] + E_{z \sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G \int_X (p_{data}(x) \log D_G^*(x) + p_G(x) \log(1 - D_G^*(x)))dx$$

$$= \min_G \int_X (p_{data}(x) \log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)} + p_G(x) \log \frac{p_G(x)}{p_{data}(x)+p_G(x)})dx$$

$$= \min_G (E_{x \sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x \sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}])$$

$$= \min_G (E_{x \sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x \sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}] - \log 4)$$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G (E_{x\sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x\sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}] - \log 4)$$

Kullback-Leibler Divergence: $KL(p, q) = E_{x\sim p}[log\frac{p(x)}{q(x)}]$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G (E_{x\sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x\sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}] - \log 4)$$

$$= \min_G (KL(p_{data}, \frac{p_{data}+p_G}{2}) + KL(p_G, \frac{p_{data}+p_G}{2}) - \log 4)$$

Kullback-Leibler Divergence: $KL(p, q) = E_{x\sim p}[log \frac{p(x)}{q(x)}]$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x \sim p_{data}}[\log D(x)] + E_{z \sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G (E_{x \sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x) + p_G(x)}] + E_{x \sim p_G}[\log \frac{p_G(x)}{p_{data}(x) + p_G(x)}] - \log 4)$$

$$= \min_G (KL(p_{data}, \frac{p_{data} + p_G}{2}) + KL(p_G, \frac{p_{data} + p_G}{2}) - \log 4)$$

Jensen-Shannon Divergence: $JSD(p, q) = \frac{1}{2} KL(p, \frac{p+q}{2}) + \frac{1}{2} KL(q, \frac{p+q}{2})$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G (E_{x\sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x\sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}] - \log 4)$$

$$= \min_G (KL(p_{data}, \frac{p_{data}+p_G}{2}) + KL(p_G, \frac{p_{data}+p_G}{2}) - \log 4)$$

$$= \min_G (2 \times JSD(p_{data}, p_G) - \log 4)$$

Jensen-Shannon Divergence: $JSD(p,q) = \frac{1}{2}KL(p, \frac{p+q}{2}) + \frac{1}{2}KL(q, \frac{p+q}{2})$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

$$= \min_G (E_{x\sim p_{data}}[\log \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}] + E_{x\sim p_G}[\log \frac{p_G(x)}{p_{data}(x)+p_G(x)}] - \log 4)$$

$$= \min_G (KL(p_{data}, \frac{p_{data}+p_G}{2}) + KL(p_G, \frac{p_{data}+p_G}{2}) - \log 4)$$

$$= \min_G (2 \times JSD(p_{data}, p_G) - \log 4)$$

JSD is always nonnegative and zero when the two distributions are equal

=> the global minimum is $p_{data} = p_G$

Jensen-Shannon Divergence: $JSD(p, q) = \frac{1}{2}KL(p, \frac{p+q}{2}) + \frac{1}{2}KL(q, \frac{p+q}{2})$

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: Optimality

$$\min_G \max_D (E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p(z)}[\log(1 - D(G(z)))])$$

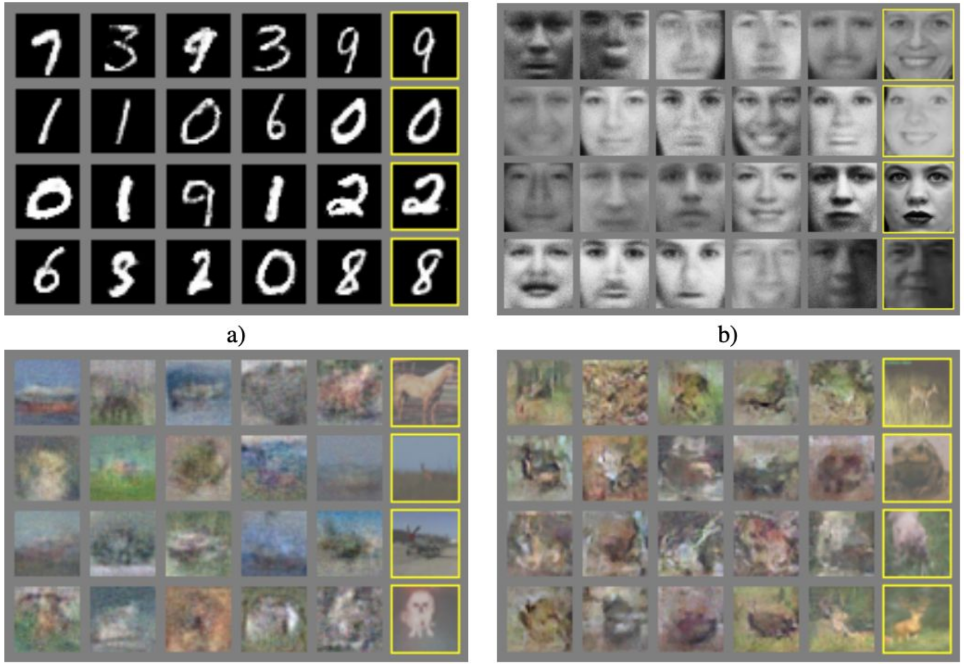
$$= \min_G (2 * JSD(p_{data}, p_G) - \log 4)$$

Summary: The global minimum of the minimax game happens when:

1. $D^*_G(x) = \frac{p_{data}(x)}{p_{data}(x)+p_G(x)}$ (Optimal discriminator for any G)
2. $p_G(x) = p_{data}(x)$ (Optimal generator for optimal D)

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

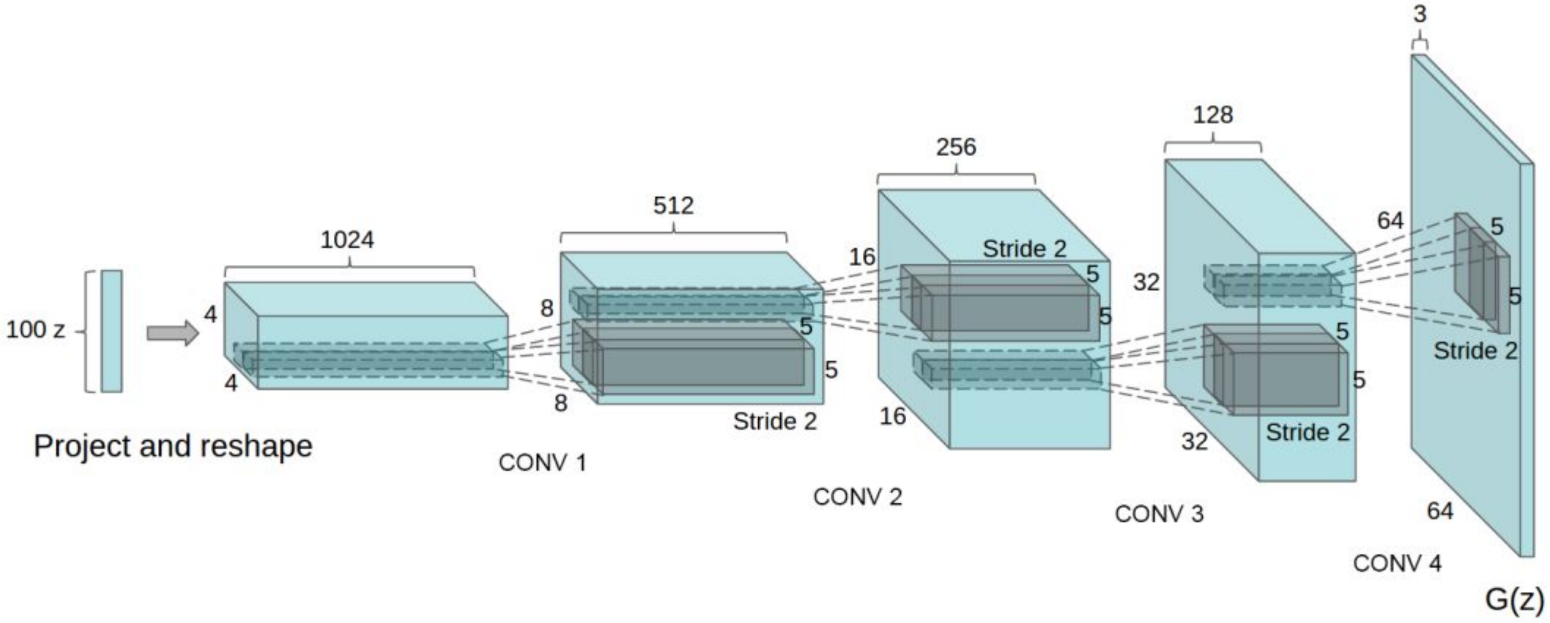
Generative Adversarial Networks: results



a) b) c) d)

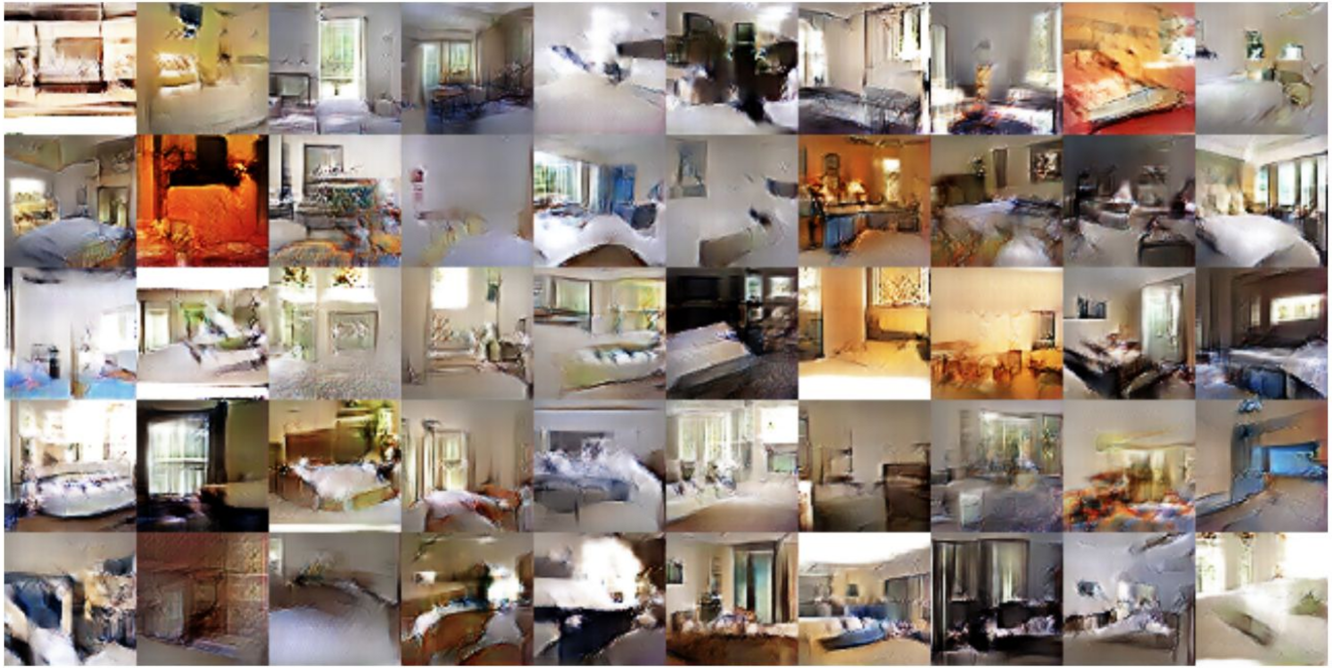
Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS) 2014.

Generative Adversarial Networks: DC-GAN



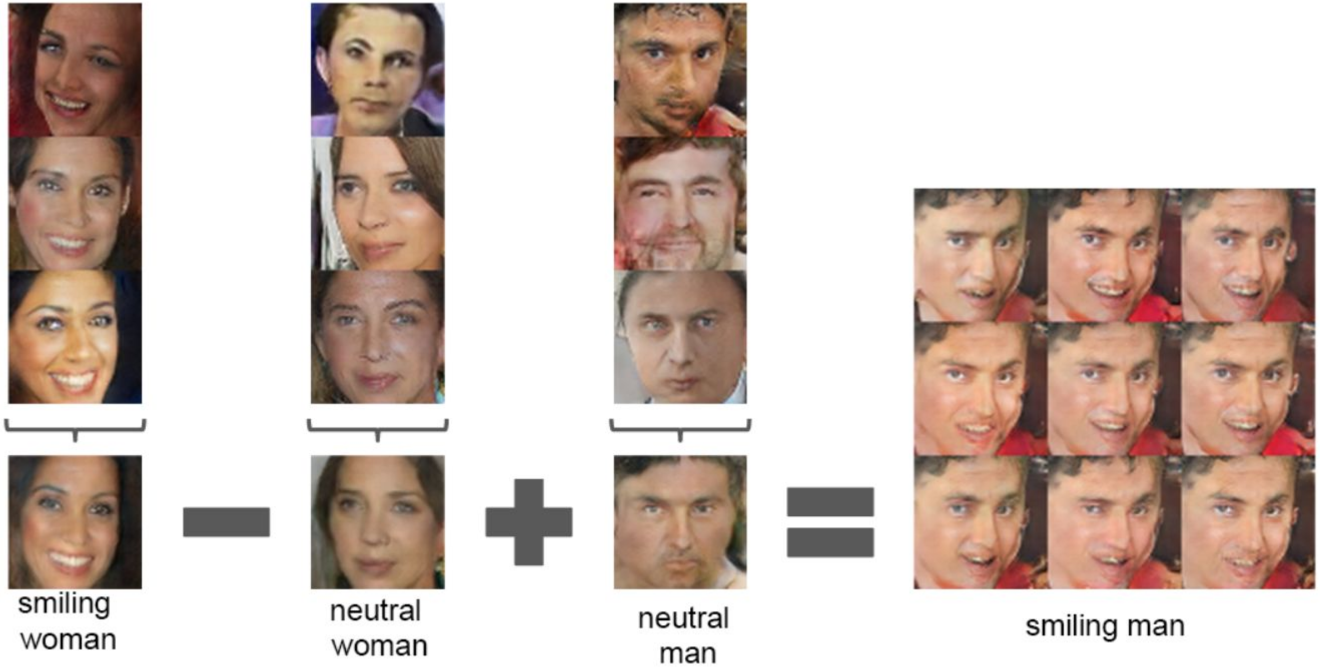
Alec Radford, Luke Metz, Soumith Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. ICLR 2016

Generative Adversarial Networks: Interpolation



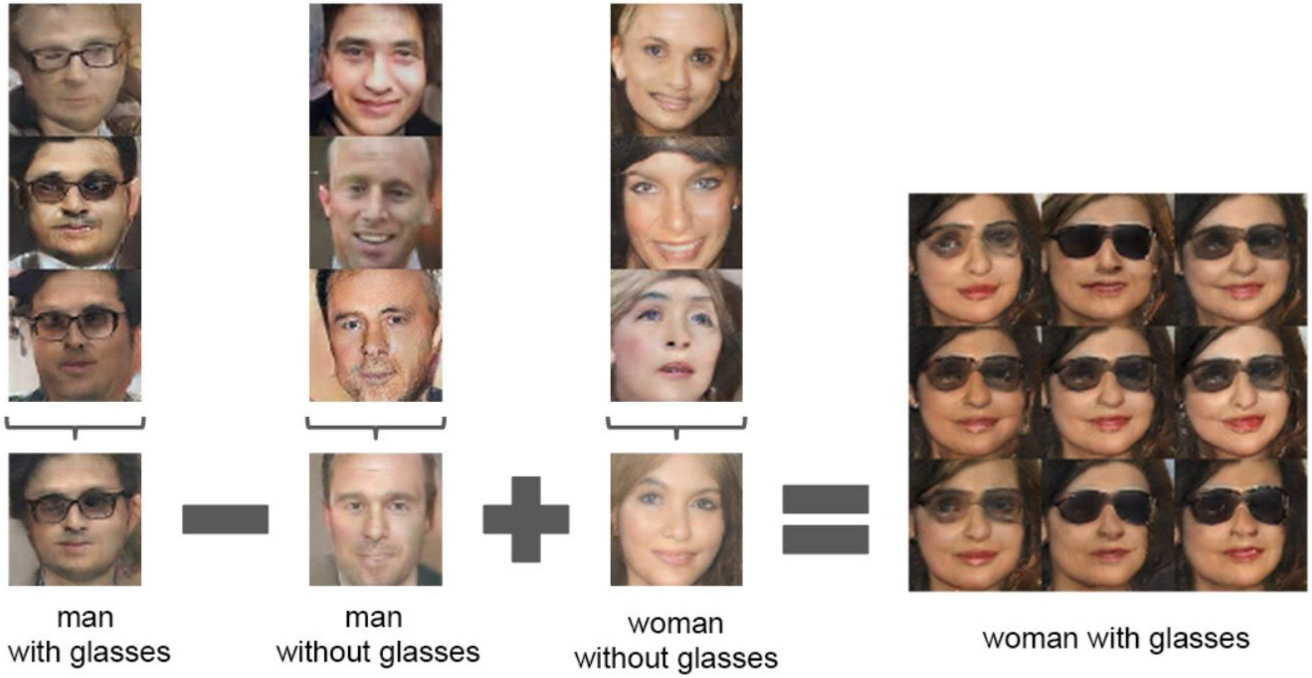
Alec Radford, Luke Metz, Soumith Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. ICLR 2016

Generative Adversarial Networks: Vector Math



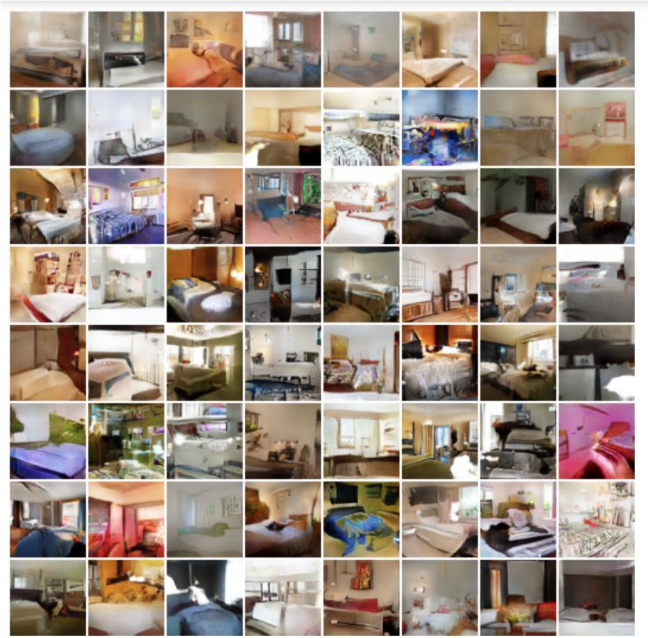
Alec Radford, Luke Metz, Soumith Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. ICLR 2016

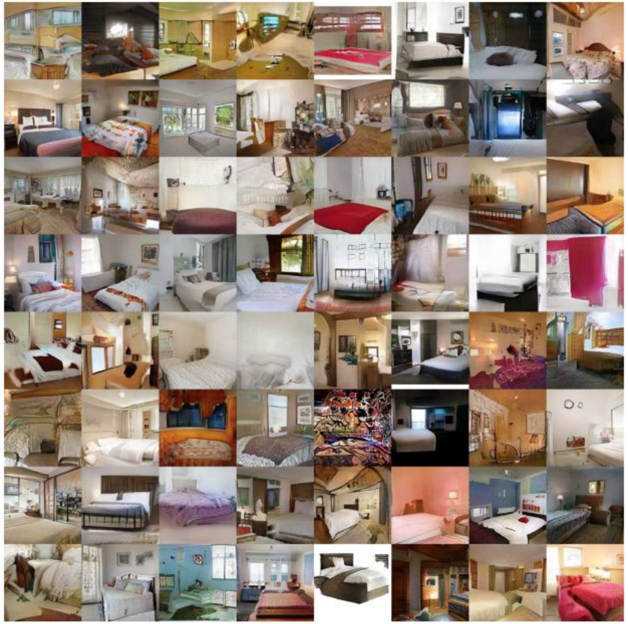
Generative Adversarial Networks: Vector Math



Alec Radford, Luke Metz, Soumith Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. ICLR 2016

GAN: Improved Loss Functions

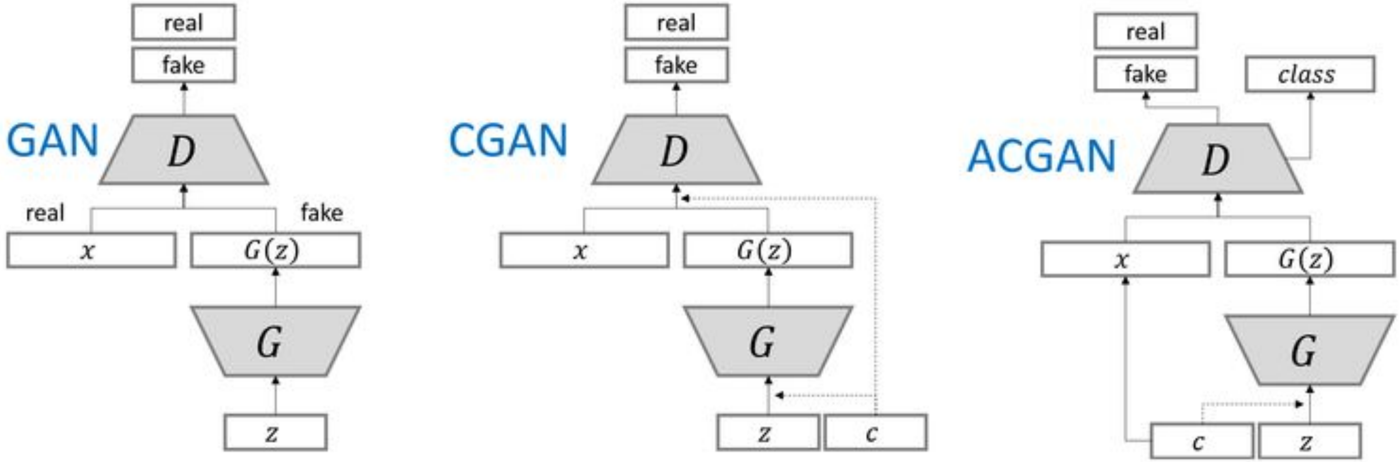




[1] Martin Arjovsky, Soumith Chintala, Léon Bottou. Wasserstein GAN. 2017

[2] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, Aaron Courville. Improved Training of Wasserstein GANs. NeurIPS, 2017.

Conditional GANs



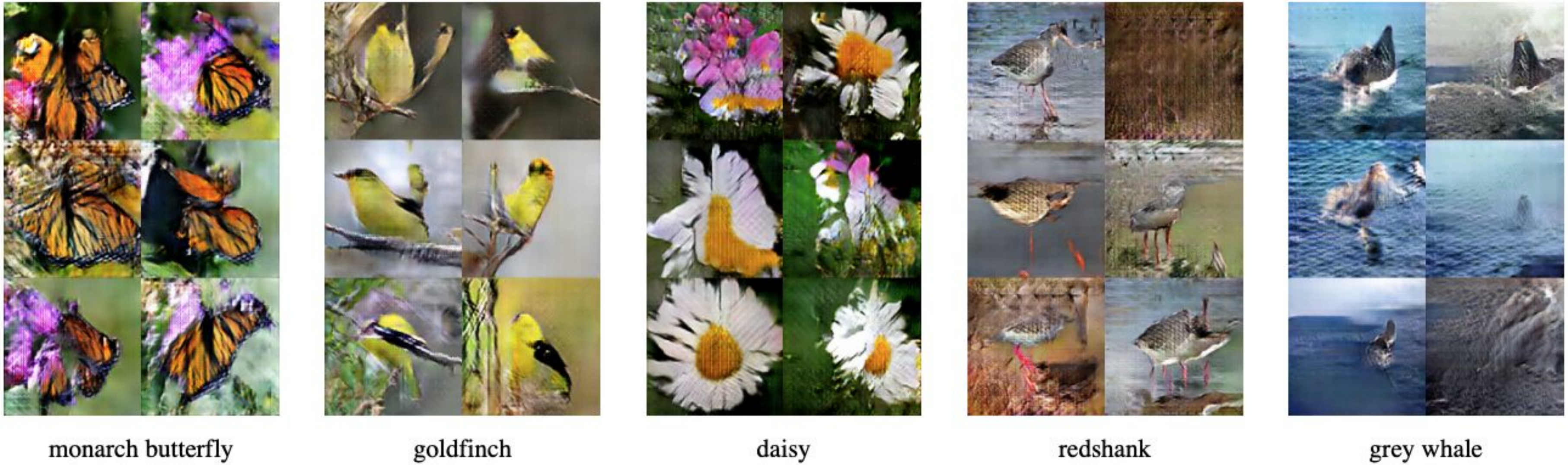
(a) (b) (c)

[1] [2]

[1] Mehdi Mirza, Simon Osindero. Conditional Generative Adversarial Nets. 2014

[2] Augustus Odena, Christopher Olah, Jonathon Shlens. Conditional Image Synthesis With Auxiliary Classifier GANs. ICML 2016

Conditional GANs



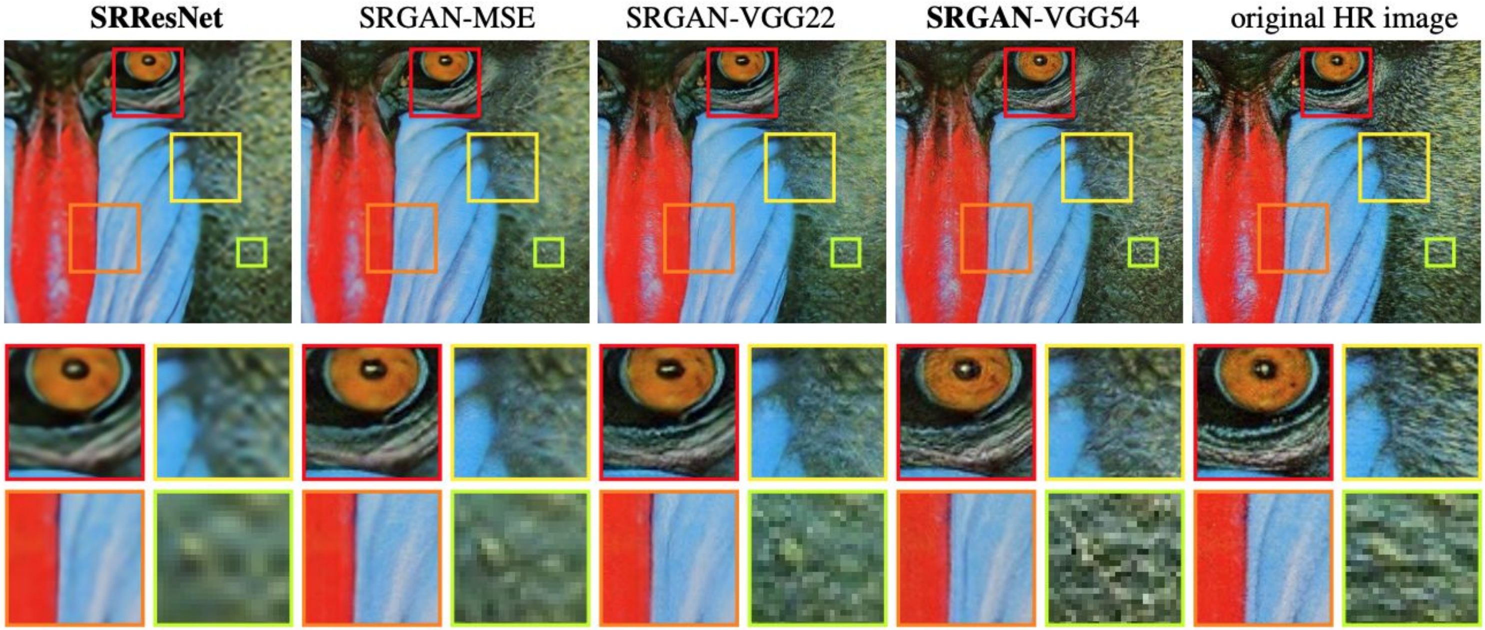
Augustus Odena, Christopher Olah, Jonathon Shlens. Conditional Image Synthesis With Auxiliary Classifier GANs. ICML 2016

Conditional GANs: BigGAN

Figure 6: Samples generated by our BigGAN model at 512×512 resolution.

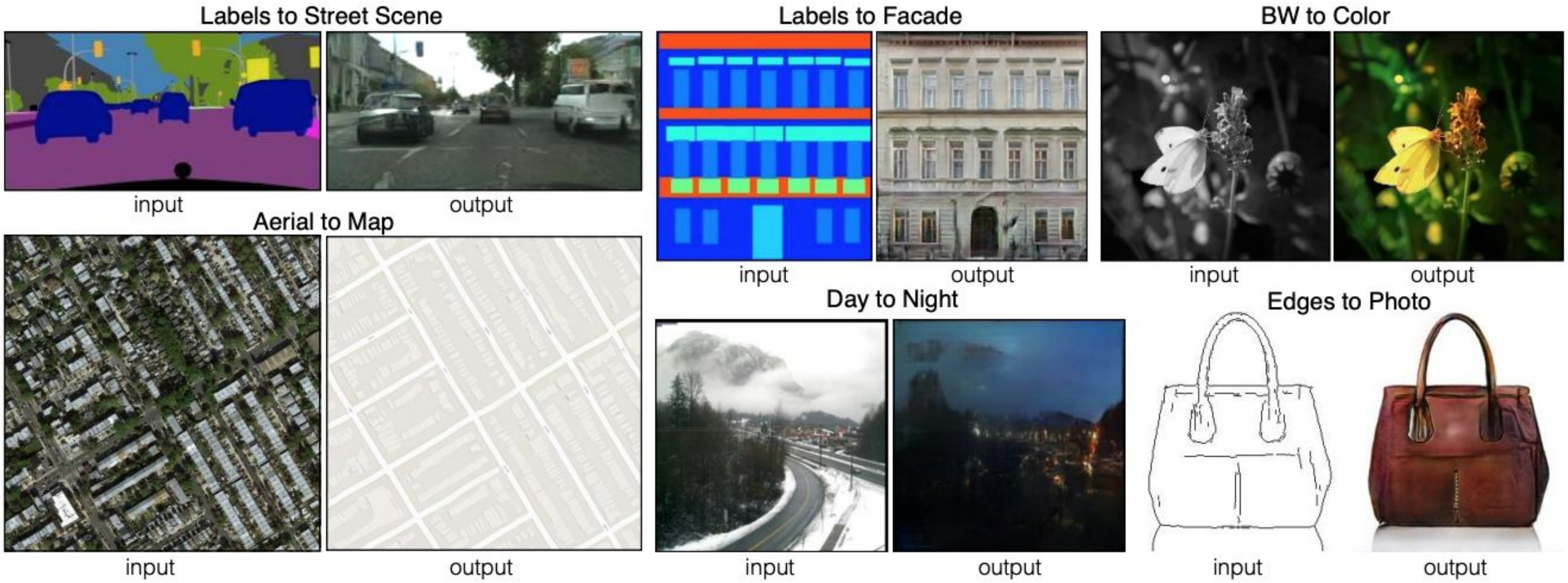
Andrew Brock, Jeff Donahue, Karen Simonyan. Large Scale GAN Training for High Fidelity Natural Image Synthesis. ICLR 2019

Image Super-Resolution



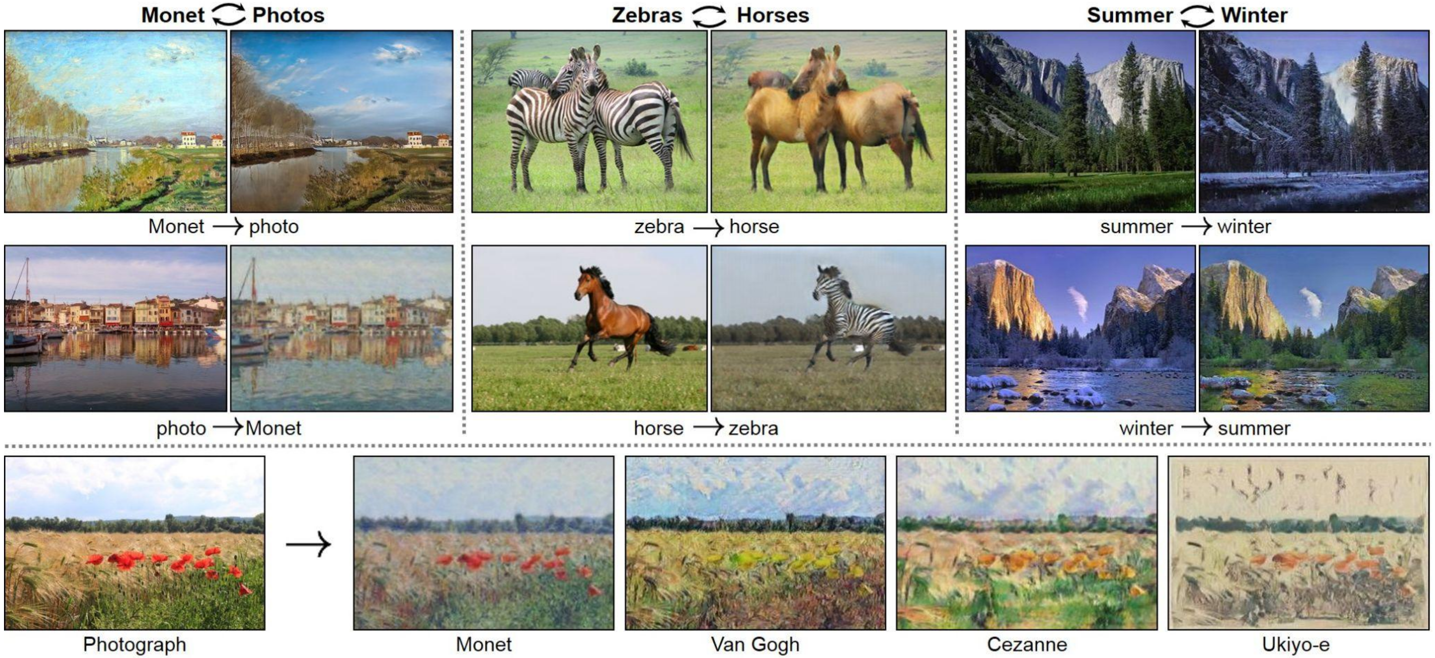
Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. CVPR 2017

Image-to-Image Translation: Pix2Pix



Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. CVPR 2017

Unpaired Image-to-Image Translation: CycleGAN

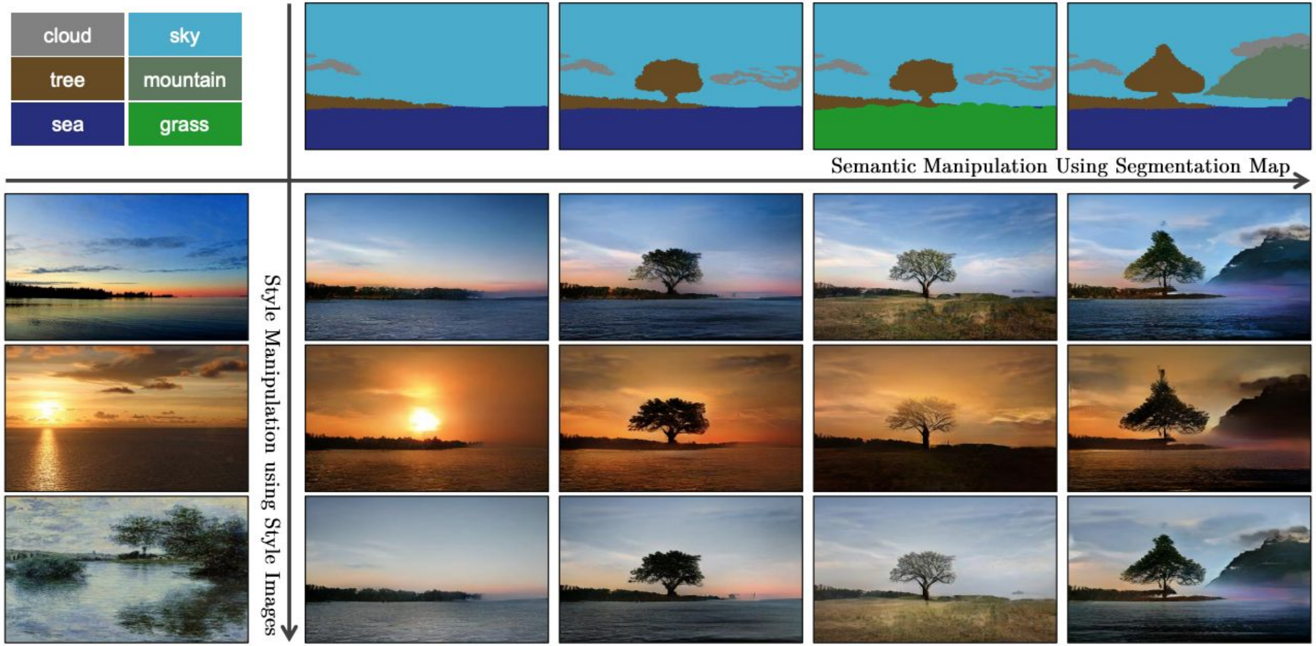


Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV 2017

Unpaired Image-to-Image Translation: CycleGAN

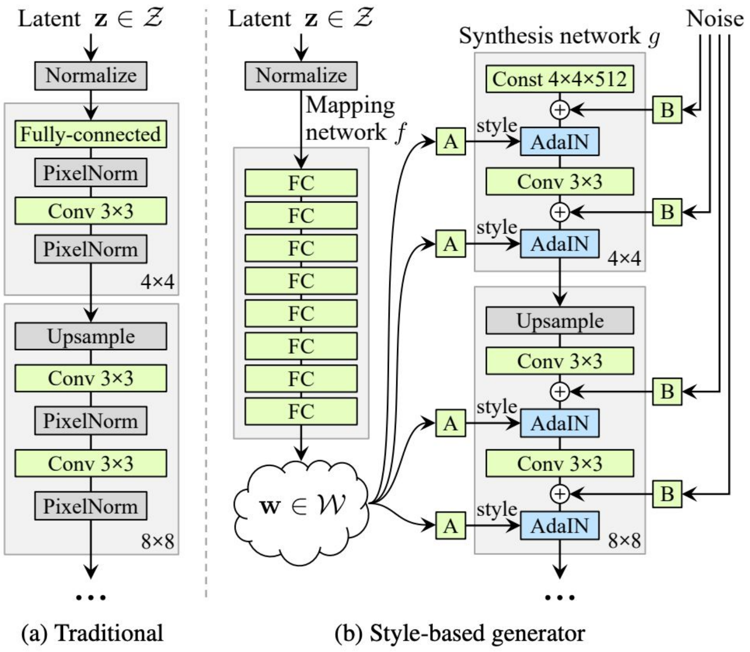
Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV 2017

Label Map to Image



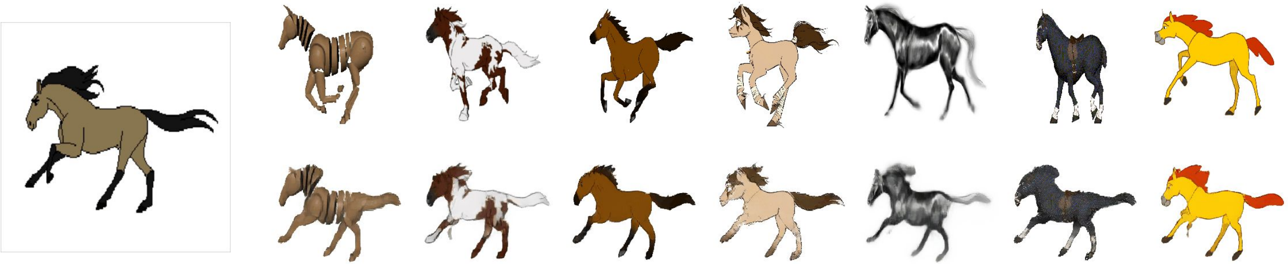
Taesung Park, Ming-Yu Liu, Ting-Chun Wang, Jun-Yan Zhu. Semantic Image Synthesis with Spatially-Adaptive Normalization. CVPR 2019

StyleGAN



Tero Karras, Samuli Laine, Timo Aila. A Style-Based Generator Architecture for Generative Adversarial Networks. CVPR 2019

Video Generation



Aliaksandr Siarohin, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, Nicu Sebe. First Order Motion Model for Image Animation. NeurIPS 2019

Video Generation

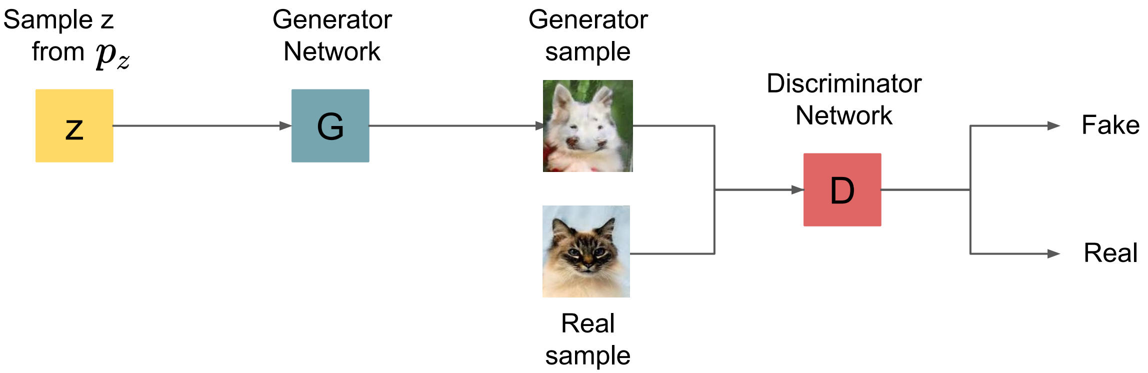
Caroline Chan, Shiry Ginosar, Tinghui Zhou, Alexei A. Efros. Everybody Dance Now. ICCV 2019

GAN Summary

Jointly train two networks:

Discriminator classifies data as real or fake

Generator generates data that fools the discriminator



Q&A