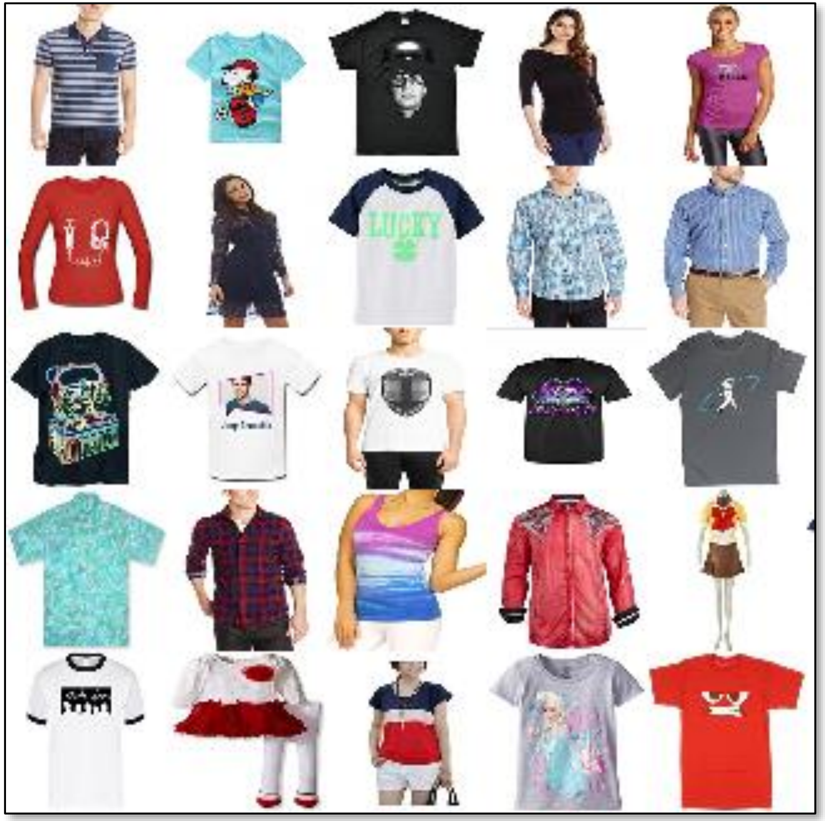
Conditional Generative Adversarial Networks (cGANs)

Conditional GANs (cGANs)

- Gain control of output
- Modeling (e.g., sketch-based modeling, etc.)
  - Add semantic meaning to latent space manifold
- Domain transfer
  - Labels on A -> transfer to B, train network on ‘B’, test on B
  - More later

GAN Manifold

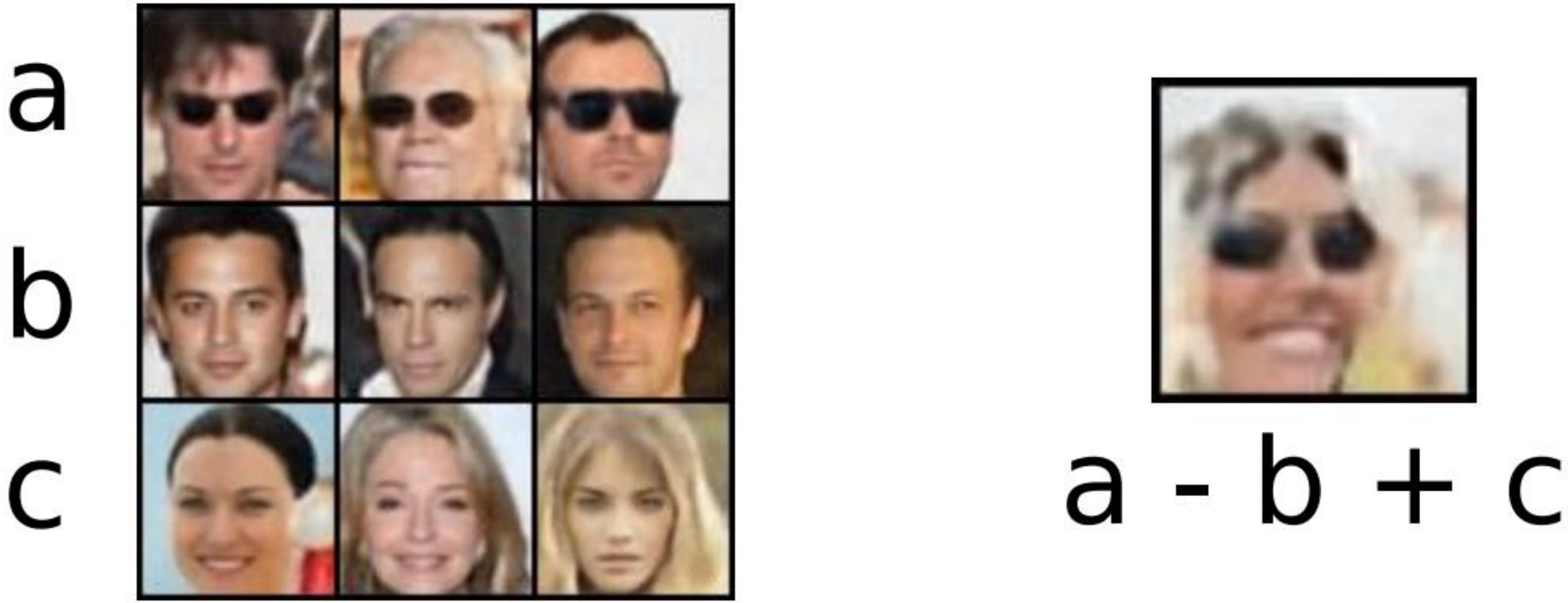


Train Data

Sampled Data -> G(z)

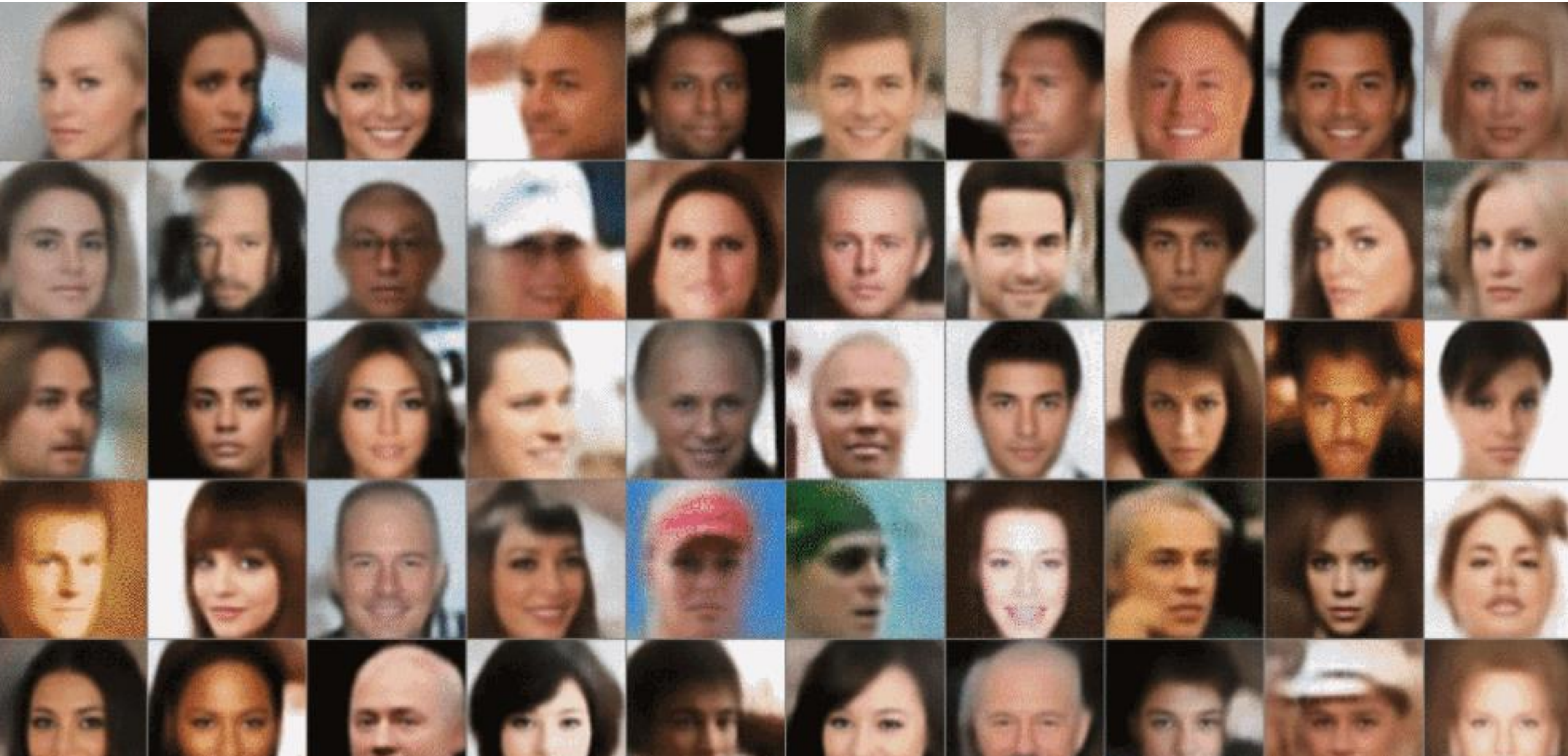
[Radford et al. 15]

GAN Manifold



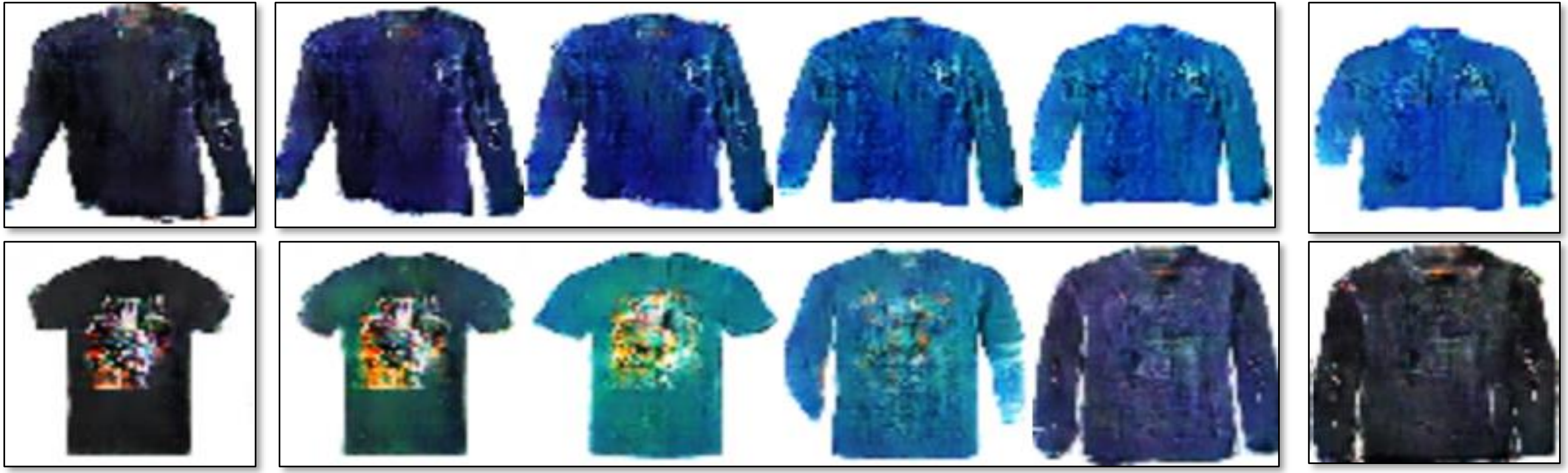
[Bojanowski et al 17]

GAN Manifold



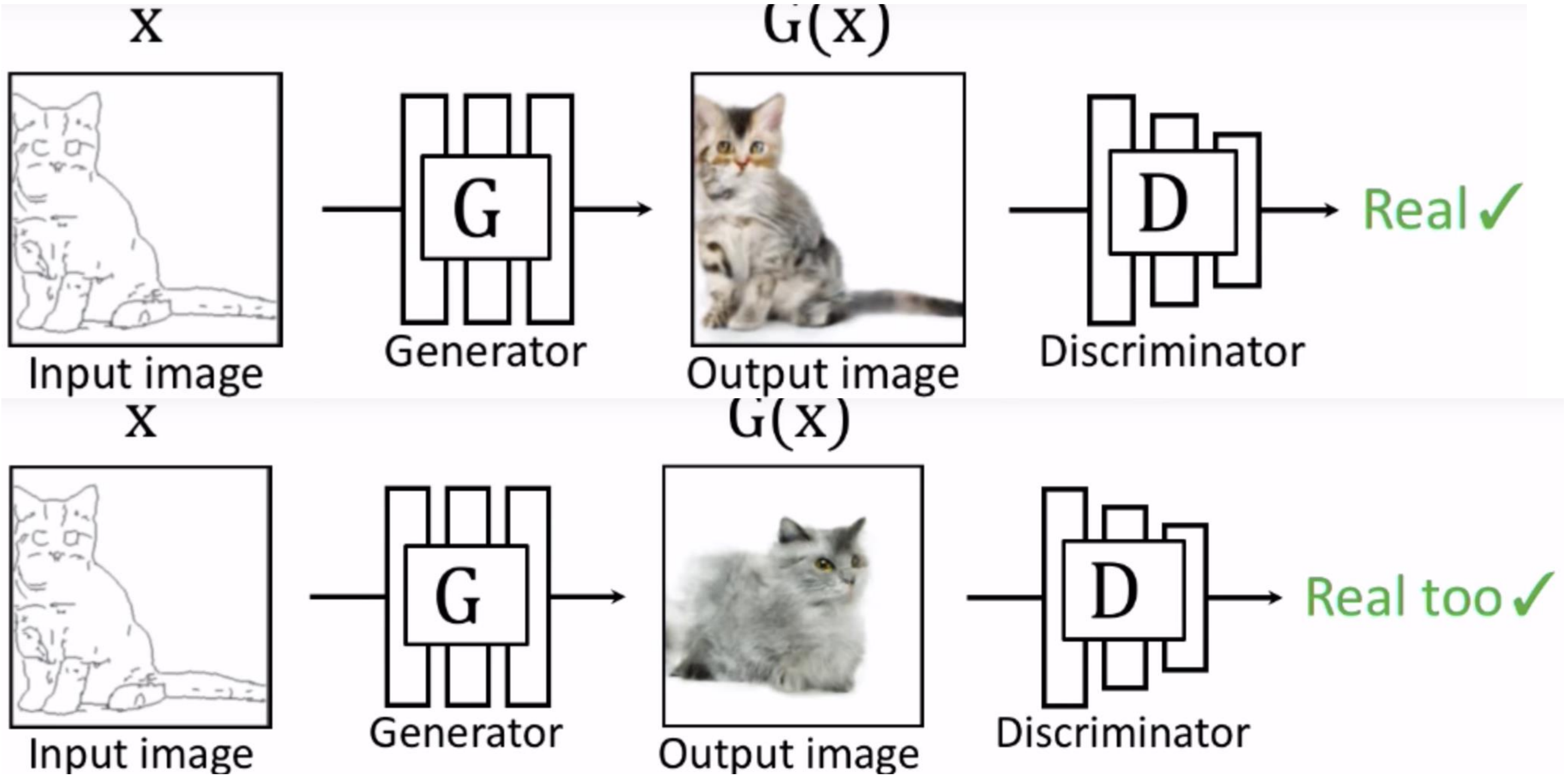
GAN Manifold

$G(z_0)$ Linear interpolation in z space: $G(z_0 + t \cdot (z_1 - z_0))$ $G(z_1)$



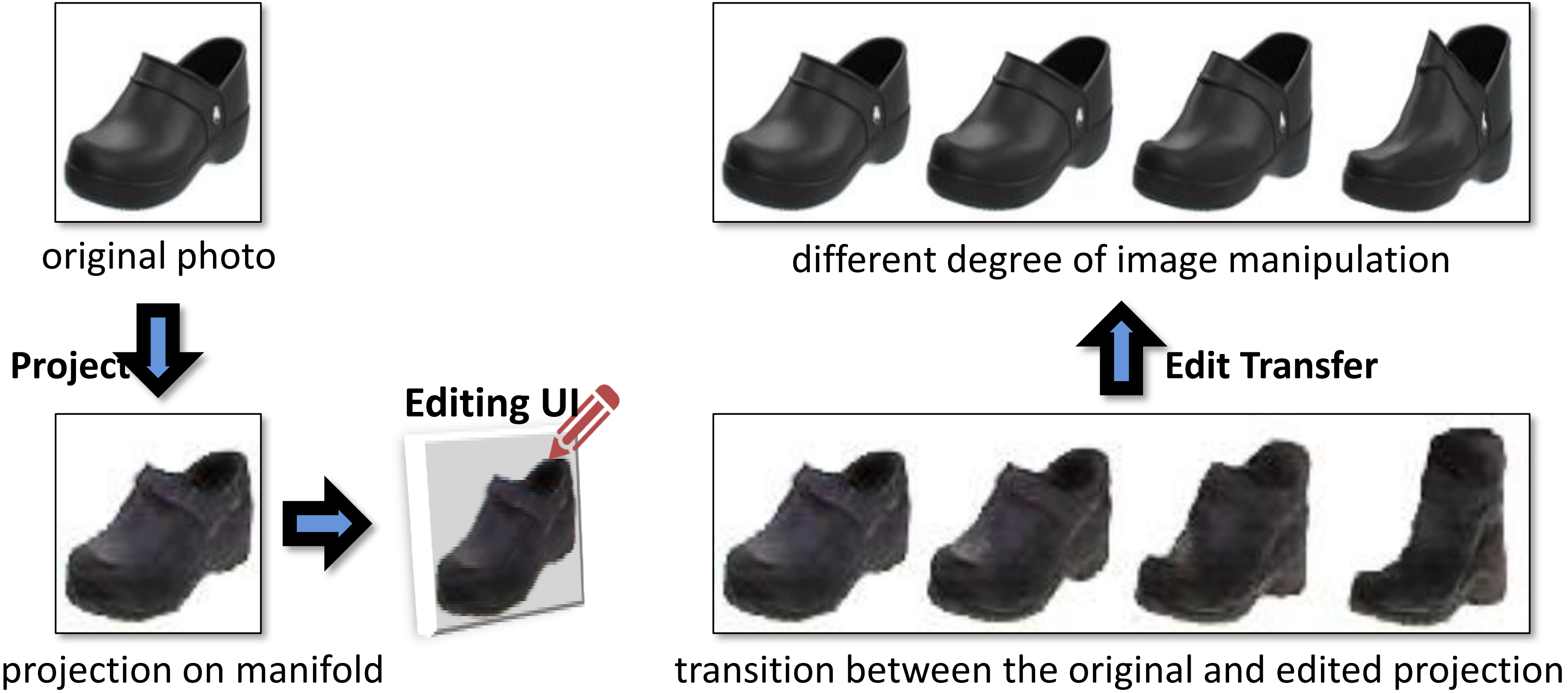
[Radford et al. 15]

Conditional GANs (cGANs)



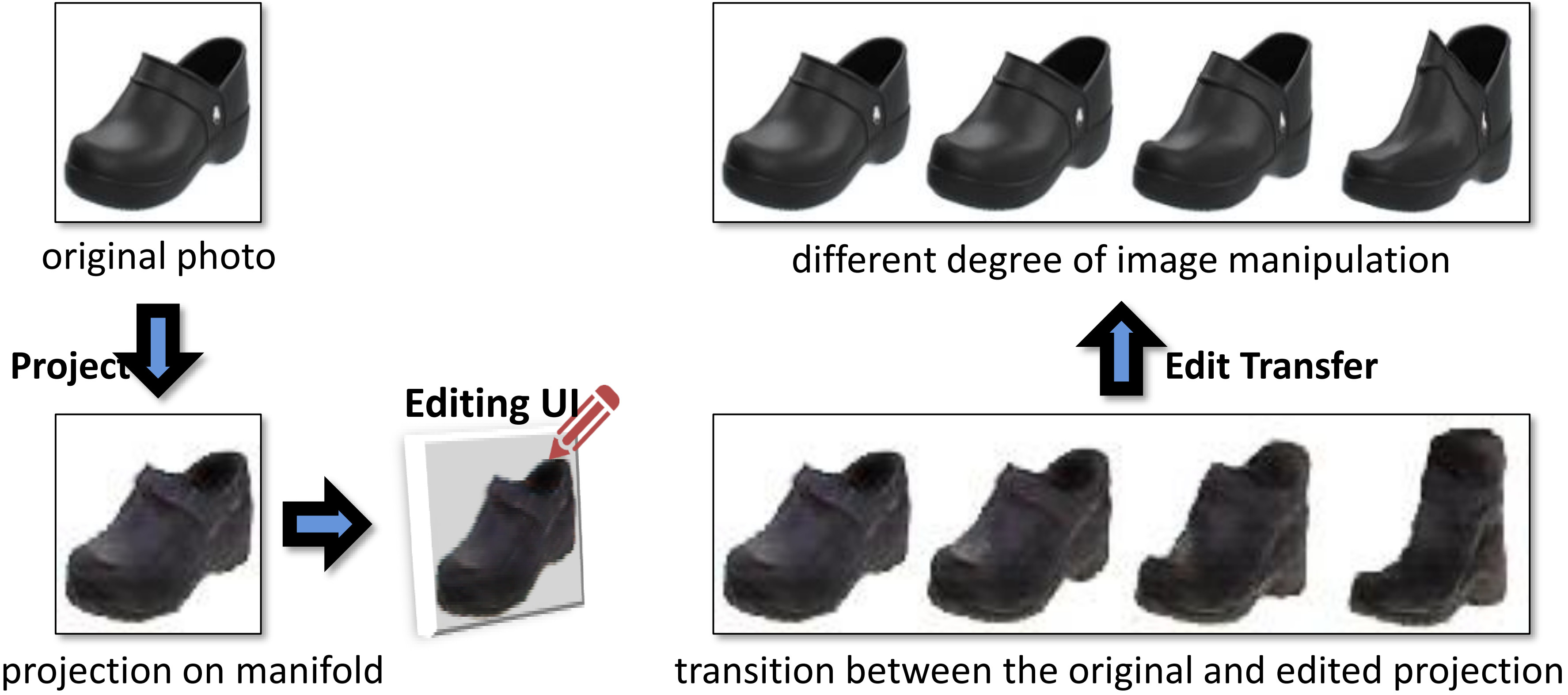
Slide credit Zhu

iGANs: Overview



Slide credit Zhu / [Zhu et al. 16]

iGANs: Overview



Slide credit Zhu / [Zhu et al. 16]

iGANs: Projecting an Image onto the Manifold

Input: real image x^R

Output: latent vector z

Optimization

$$z^* = \arg\min \mathcal{L}(G(z), x^R)$$

Reconstruction loss L

Generative model $G(z)$

0.196 0.238 0.332

iGANs: Projecting an Image onto the Manifold

Input: real image x^R

Output: latent vector z

Optimization

$$z^* = \arg\min \mathcal{L}(G(z), x^R)$$

Inverting Network $z = P(x)$

$$\theta_P^* = \arg\min_{\theta_P} \sum_{x_n^R} \mathcal{L}(\underbrace{G(P(x^R;\theta_P))}_{\text{Auto-encoder}}, x^R)$$

with a fixed decoder G

iGANs: Projecting an Image onto the Manifold

Input: real image x^R
Output: latent vector z

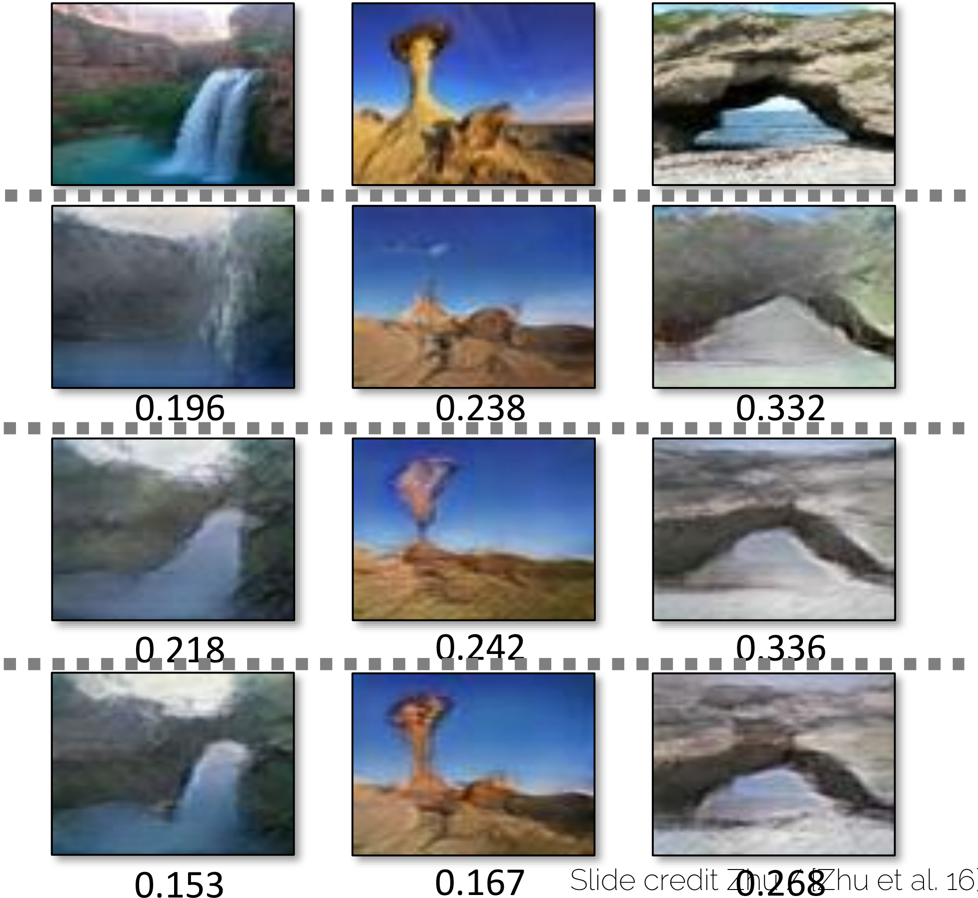
Optimization

$$z^* = \arg\min \mathcal{L}(G(z), x^R)$$

Inverting Network $\mathrm{z} = P(x)$

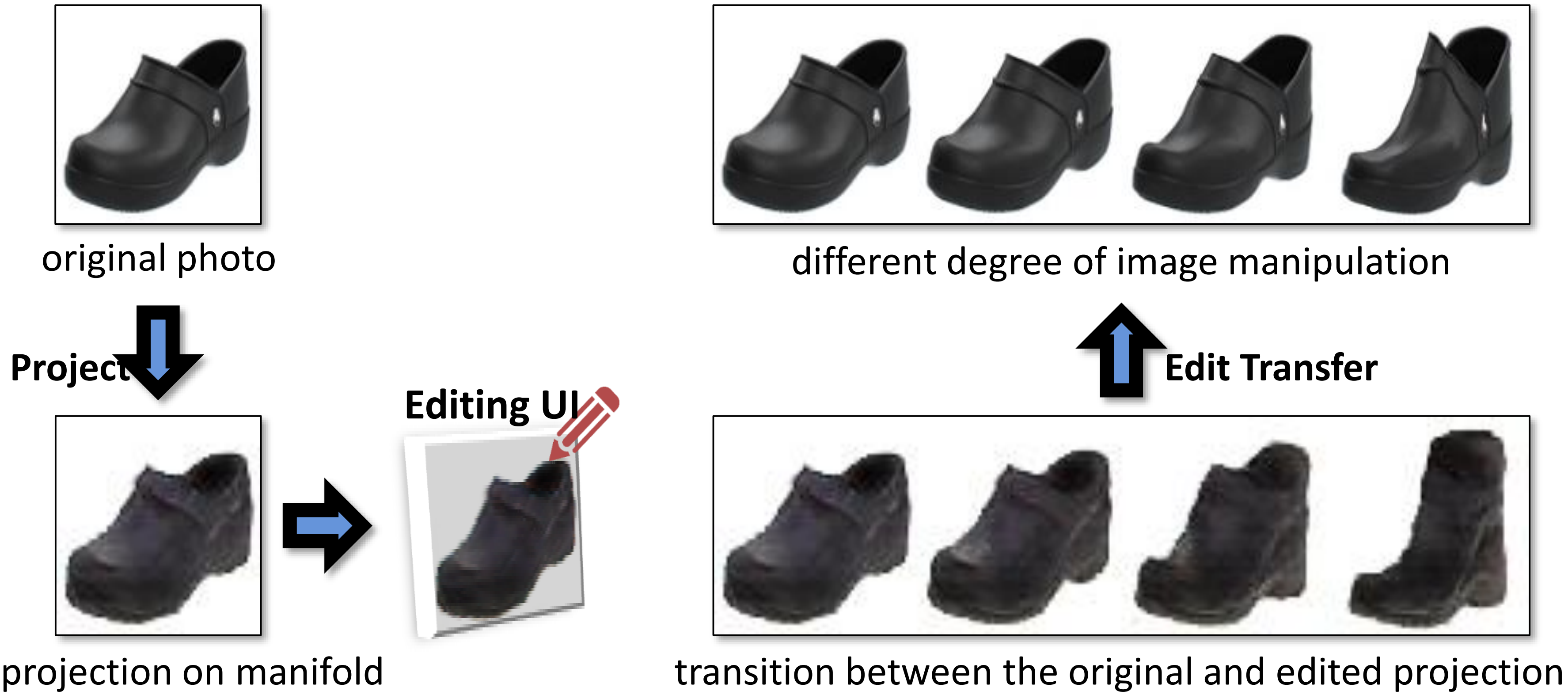
$$\theta_P^* = \arg\min_{\theta_P} \sum_{x_n^R} \mathcal{L}(G(P(x^R; \theta_P)), x^R)$$

Hybrid Method
Use the network as initialization
for the optimization problem



Slide credit Zhu [Zhu et al. 16]

iGANs: Overview



Slide credit Zhu / [Zhu et al. 16]

iGANs: Manipulating the Latent Vector

constraint violation loss L_g

user guidance image

Objective: $$z^* = \arg\min_{z \in \mathbb{Z}} \Big\{ \underbrace{\sum_g (\mathcal{L}_g(G(z), v_g)}_{\text{data term}} + \underbrace{\lambda_s \cdot \|z - z_0\|_2^2}_{\text{manifold smoothness}} \Big\}.$$

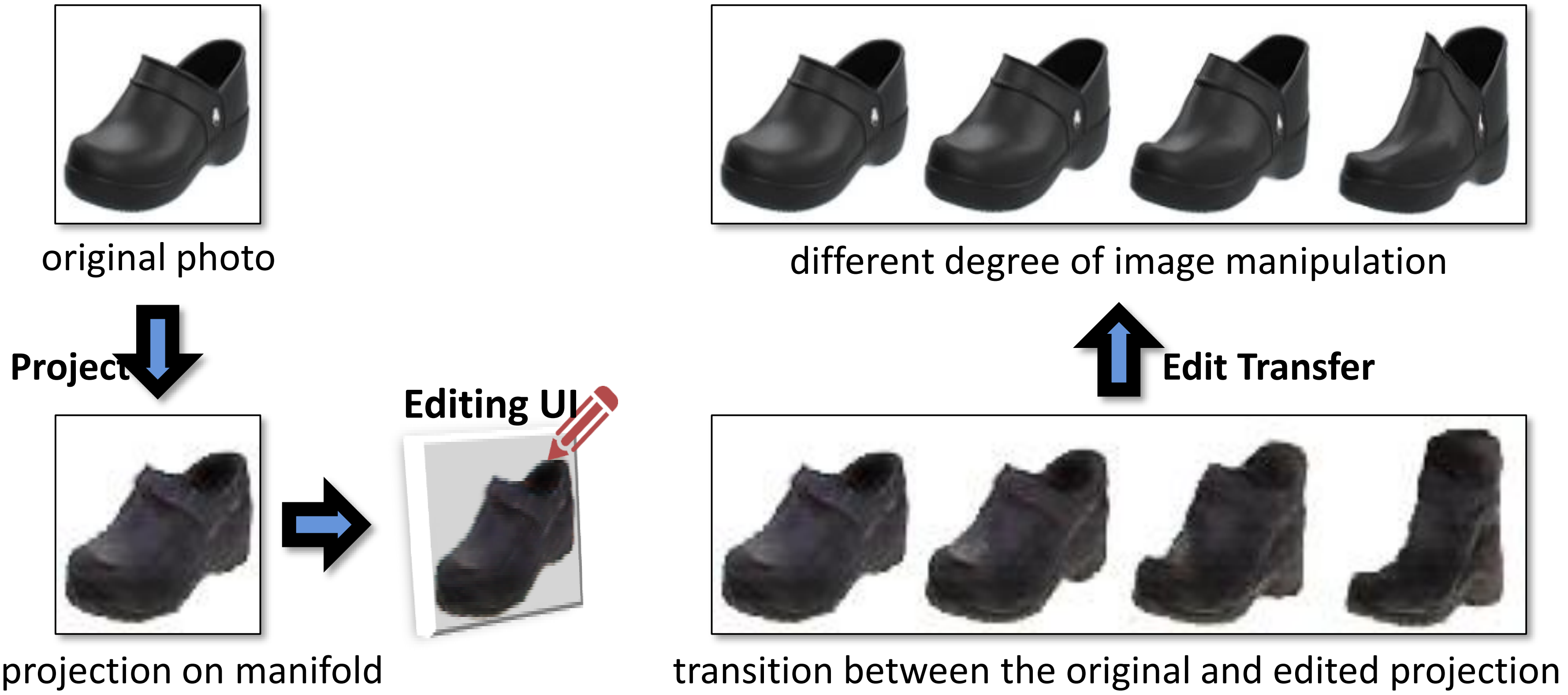
Guidance v_g

$G(z)$

z_0

Slide credit Zhu / [Zhu et al. 16]

iGANs: Overview

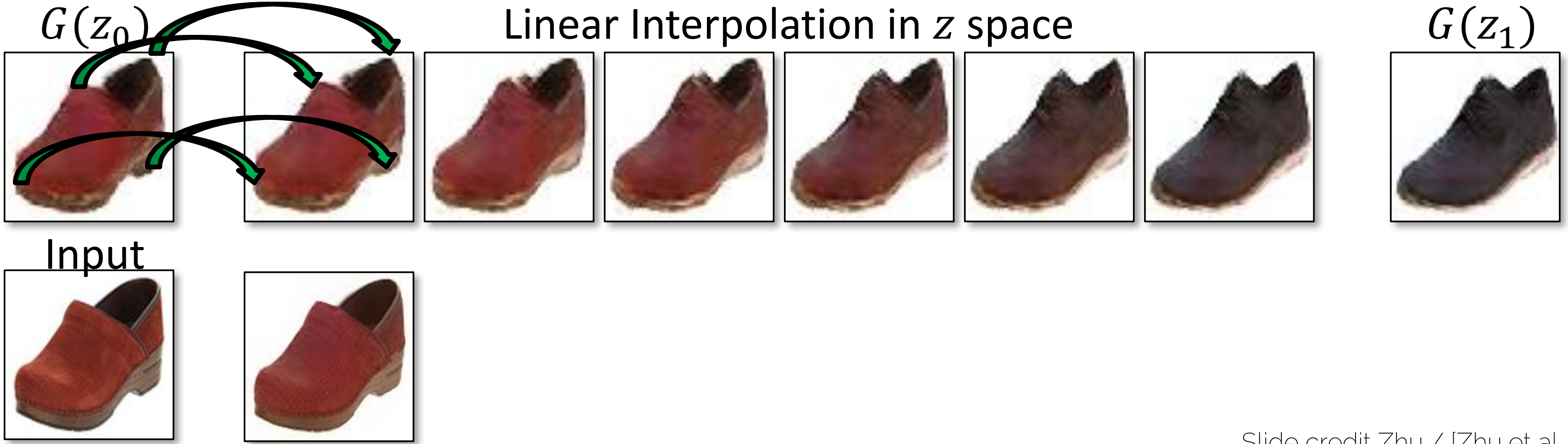


Slide credit Zhu / [Zhu et al. 16]

iGANs: Edit Transfer

Motion (u, v)+ Color ($A_{3\times4}$): estimate per-pixel geometric and color variation

$$\iint \underbrace{\|I(x,y,t) - \boxed{A}\cdot I(x+\boxed{u}, y+\boxed{v}, t+1)\|^2}_{\text{data term}} + \underbrace{\sigma_s(\|\nabla u\|^2 + \|\nabla v\|^2)}_{\text{spatial reg}} + \underbrace{\sigma_c\|\nabla A\|^2}_{\text{color reg}} dxdy$$



iGANs: Edit Transfer

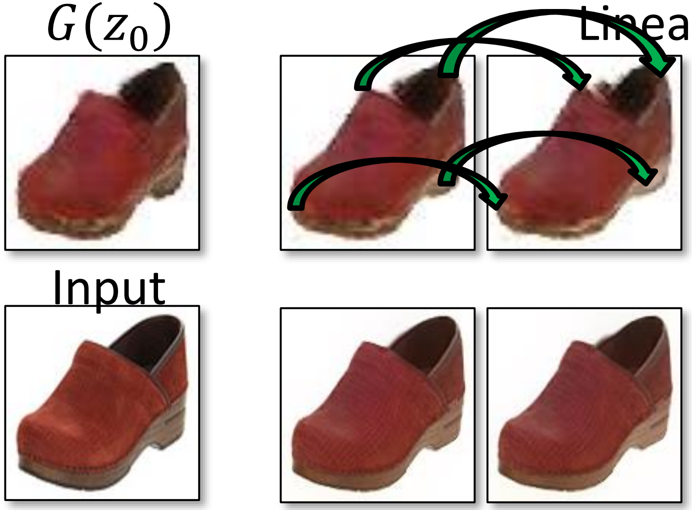
Motion (u, v)+ Color ($A_{3\times4}$): estimate per-pixel geometric and color variation

$$\iint \underbrace{\|I(x,y,t) - \boxed{A}\cdot I(x+\boxed{u}, y+\boxed{v}, t+1)\|^2}_{\text{data term}} + \underbrace{\sigma_s(\|\nabla u\|^2+\|\nabla v\|^2)}_{\text{spatial reg}} + \underbrace{\sigma_c\|\nabla A\|^2}_{\text{color reg}} dxdy$$

$G(z_0)$

Linear Interpolation in z space

$G(z_1)$

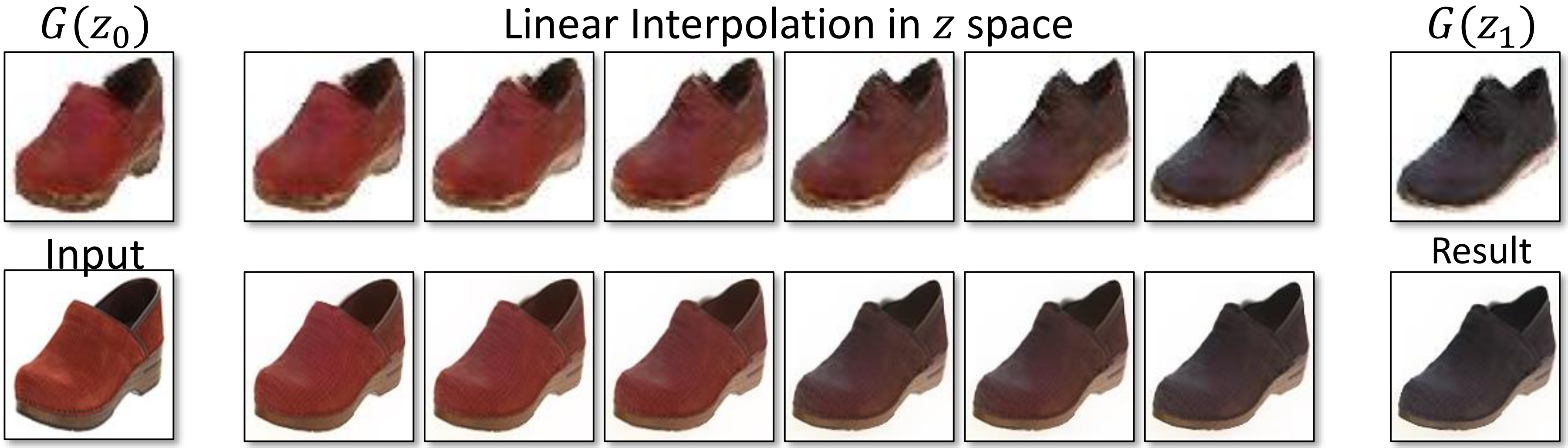


Input

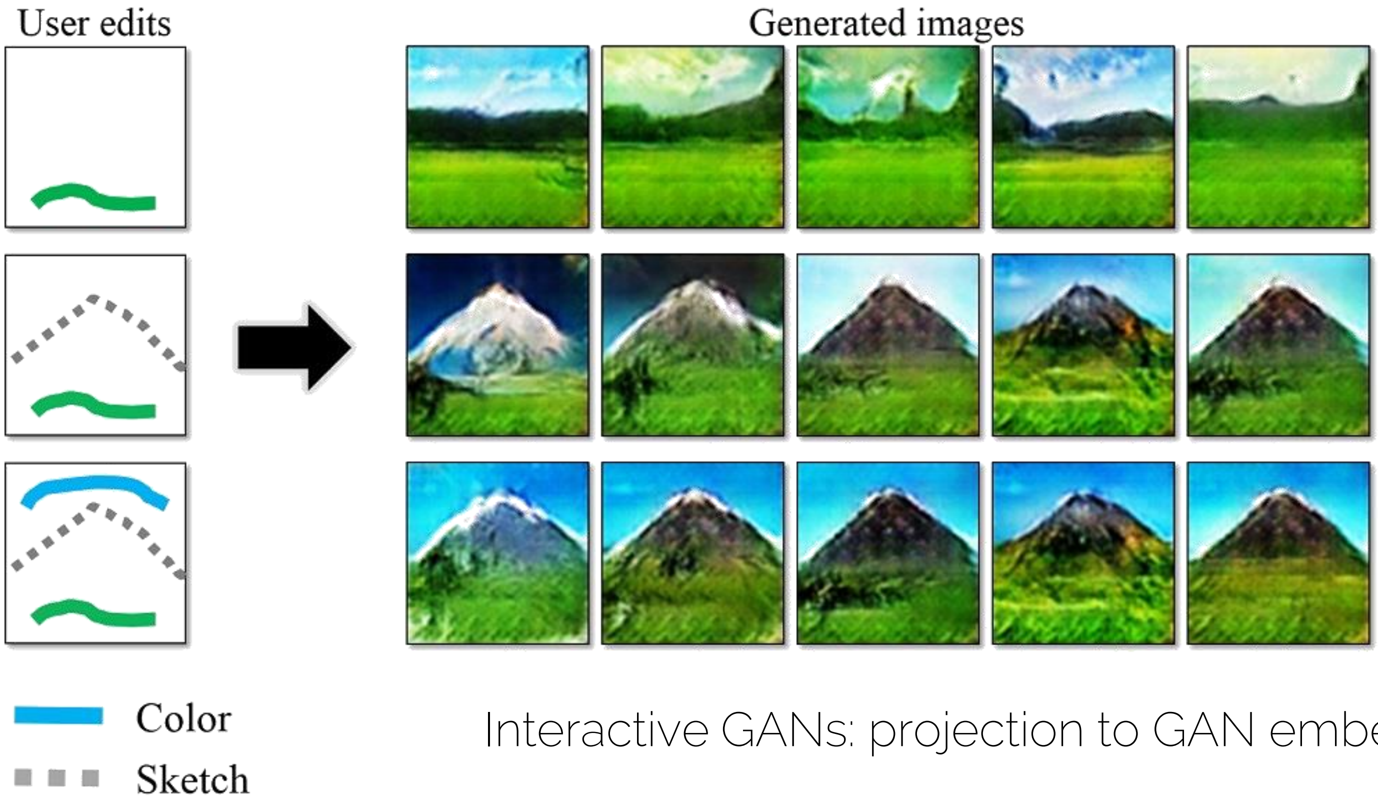
iGANs: Edit Transfer

Motion **(u, v)+** **Color** **($A_{3\times4}$):** estimate per-pixel geometric and color variation

$$\iint \underbrace{\|I(x,y,t) - \boxed{A}\cdot I(x+\boxed{u}, y+\boxed{v}, t+1)\|^2}_{\text{data term}} + \underbrace{\sigma_s(\|\nabla u\|^2 + \|\nabla v\|^2)}_{\text{spatial reg}} + \underbrace{\sigma_c\|\nabla A\|^2}_{\text{color reg}} dxdy$$



cGANs: Interactive GANs



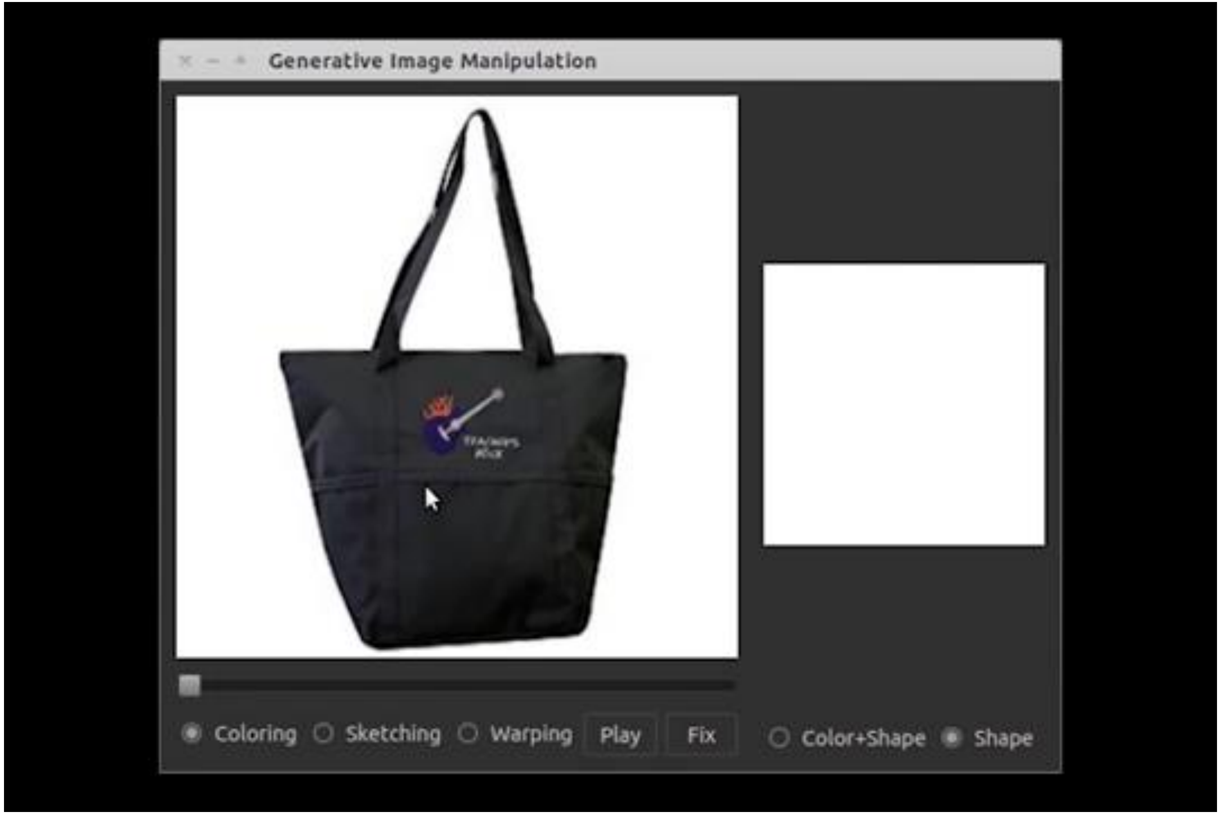
Interactive GANs: projection to GAN embedding

https://github.com/junyanz/iGAN [Zhu et al. 16.]

cGANs: Interactive GANs

https://github.com/junyanz/iGAN [Zhu et al. 16.]

cGANs: Interactive GANs



https://github.com/junyanz/iGAN [Zhu et al. 16.]

Mapping in Latent Space is Difficult!

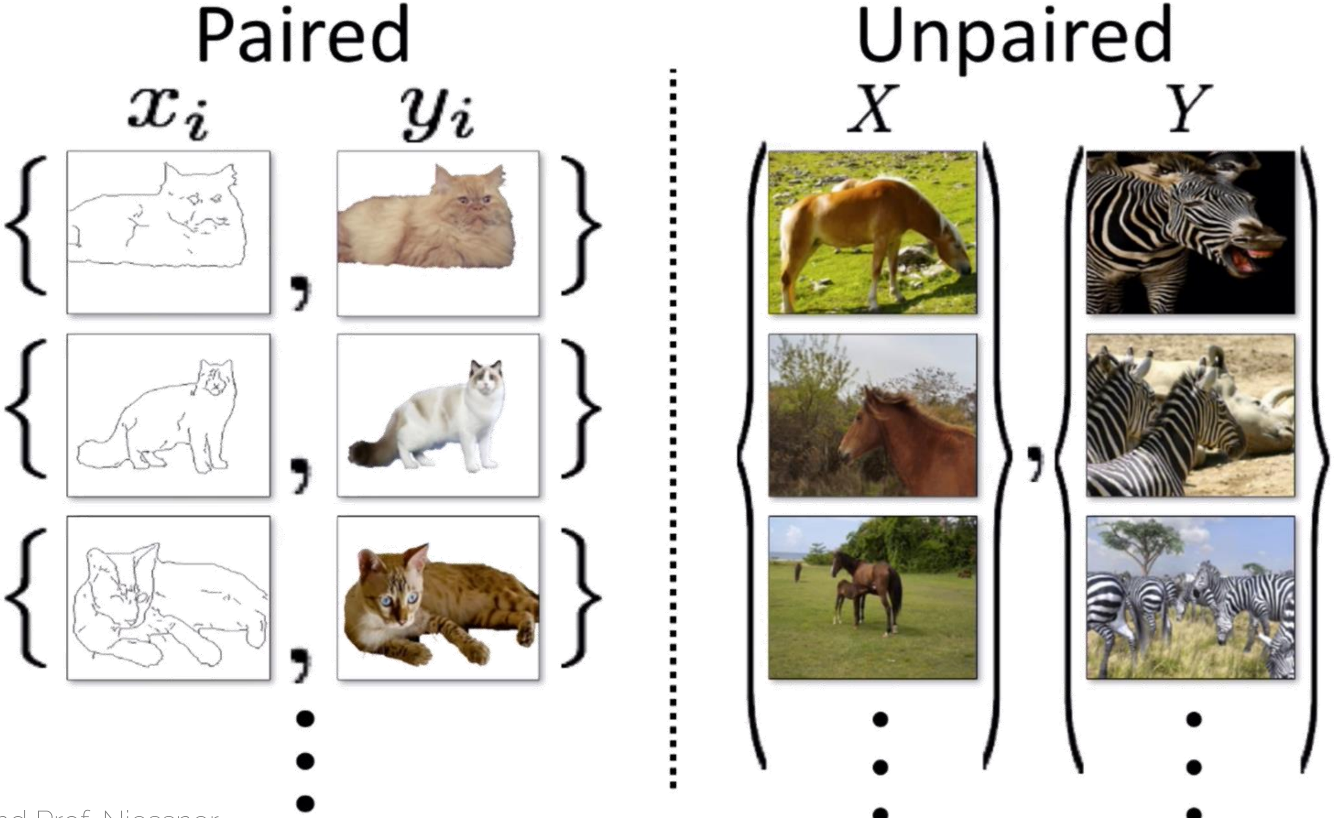
- Semantics are missing
- In most cases, no labels available
- Ideally, need some unsupervised disentangled rep.

(a) Azimuth (pose)

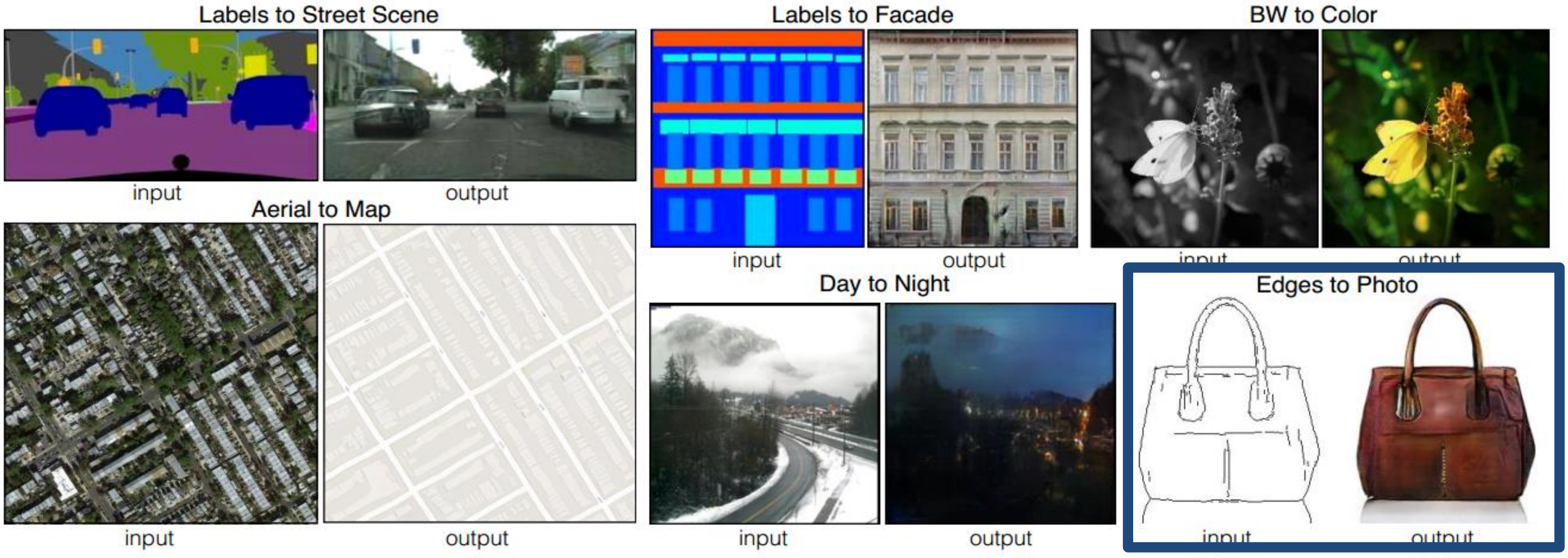
(b) Presence or absence of glasses

InfoGAN [Chen et al. 16]

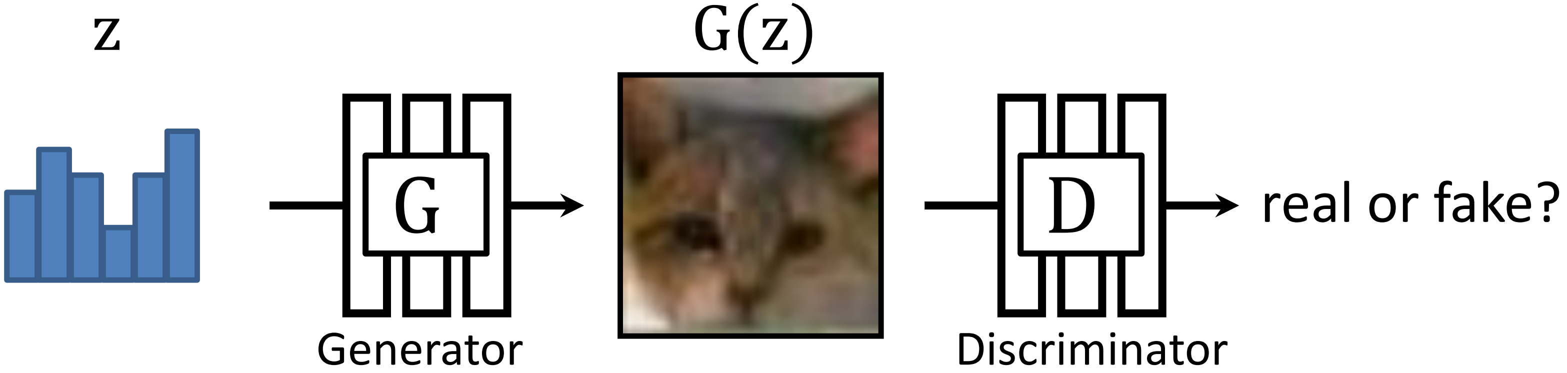
Paired vs Unpaired Setting



pix2pix: Image-to-Image Translation



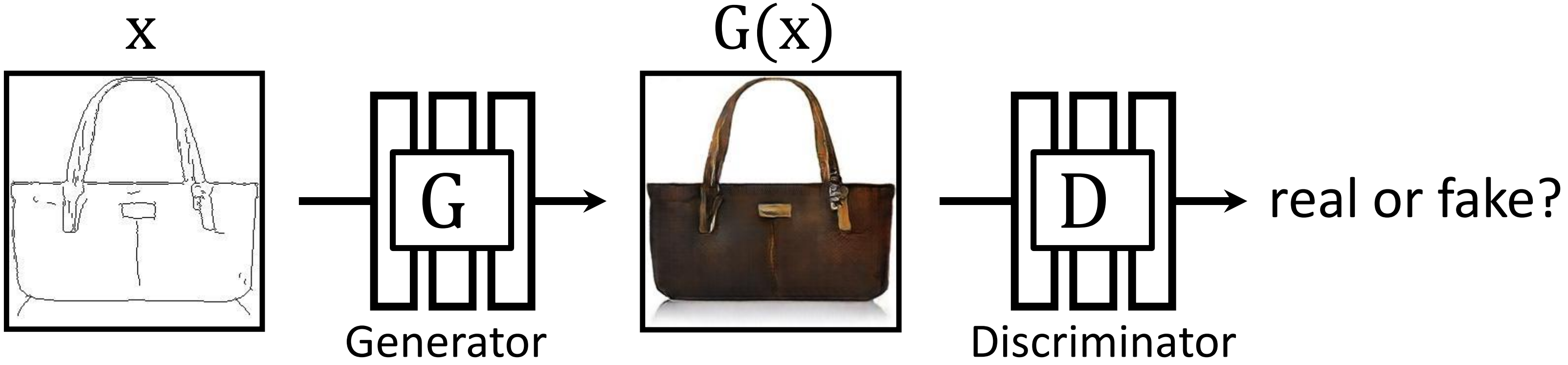
slides credit: Isola / Zhu



$$\min_G \max_D \mathbb{E}_{z,x}[\log D(G(z)) + \log(1 - D(x))]$$

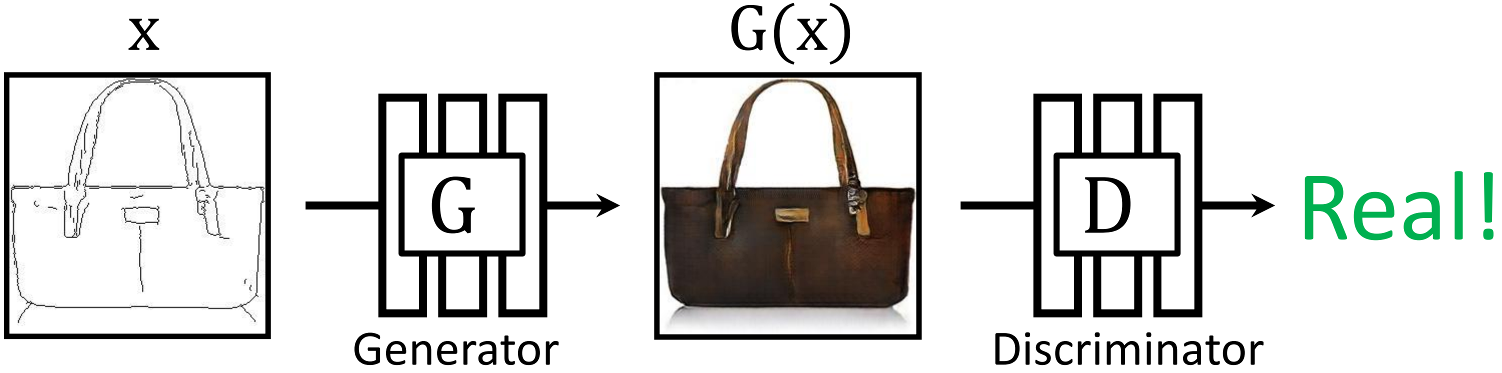
slides credit: Isola / Zhu

[Goodfellow et al. 2014]



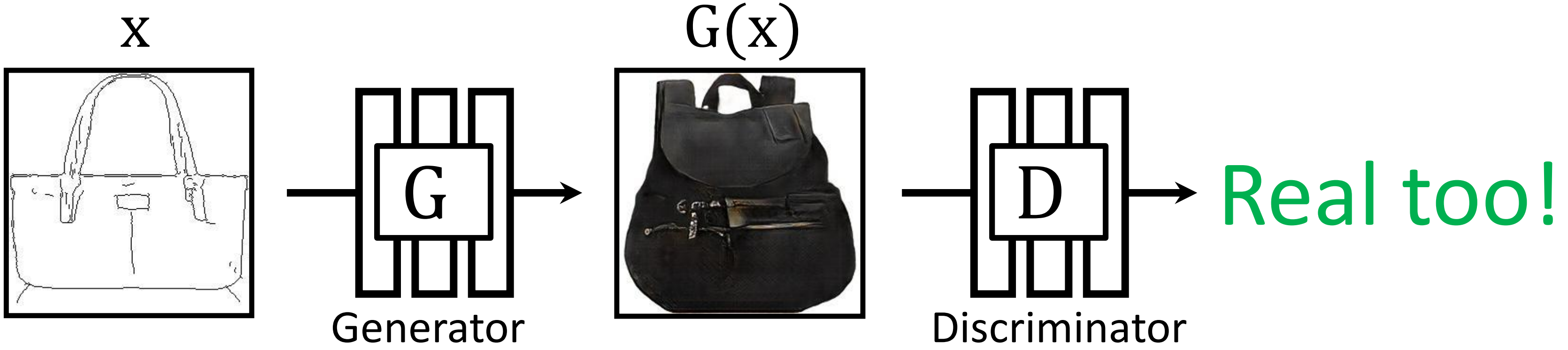
$$\min_{G} \max_{D} \mathbb{E}_{x,y}[\log D(G(x)) + \log(1 - D(y))]$$

slides credit: Isola / Zhu



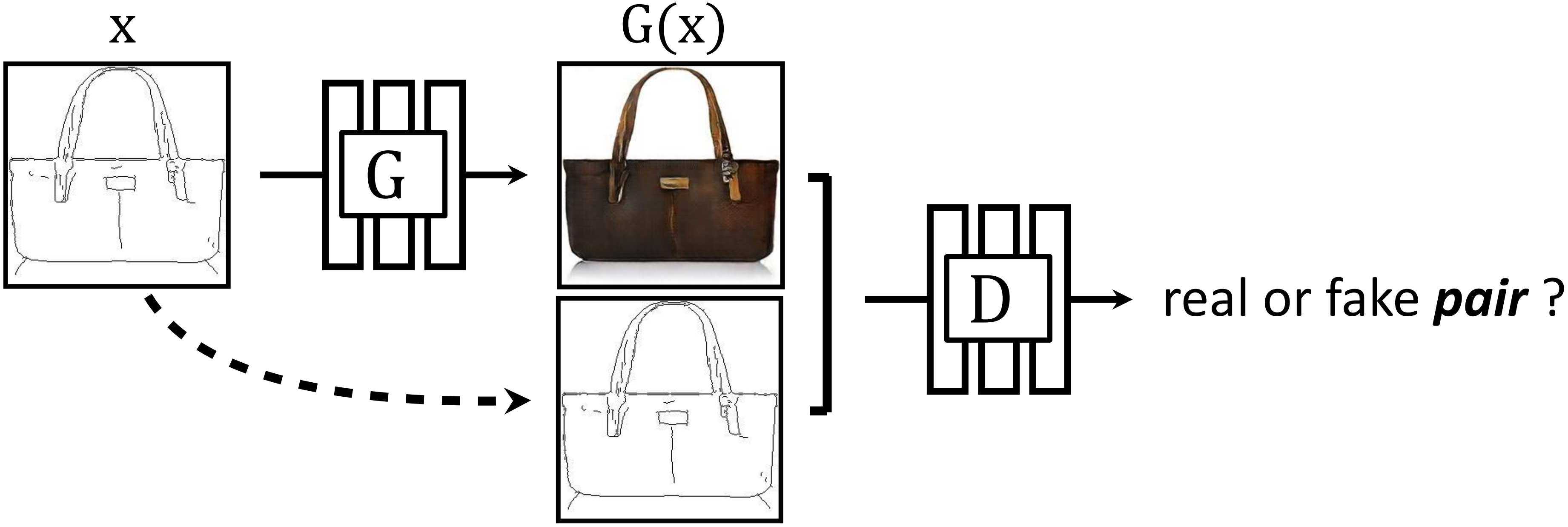
$$\min_G \max_D \mathbb{E}_{x,y}[\log D(G(x)) + \log(1 - D(y))]$$

slides credit: Isola / Zhu



$$\min_{G} \max_{D} \mathbb{E}_{x,y}[\log D(G(x)) + \log(1 - D(y))]$$

slides credit: Isola / Zhu



$$\min_G \max_D \mathbb{E}_{x,y}[\log \underbrace{D(x, G(x))}_{\text{fake pair}} + \log(1 - \underbrace{D(x, y)}_{\text{real pair}})]$$

match joint distribution $\mathrm{p}(\mathrm{G}(\mathrm{x}), \mathrm{y}) \sim \mathrm{p}(\mathrm{x}, \mathrm{y})$

slides credit: Isola / Zhu

Edges → Images

Edges from [Xie & Tu, 2015]

slides credit: Isola / Zhu

Pix2Pix: Paired Setting

- Great when we have ‘free’ training data
- Often called self-supervised
- Think about these settings ☺

Sketches → Images

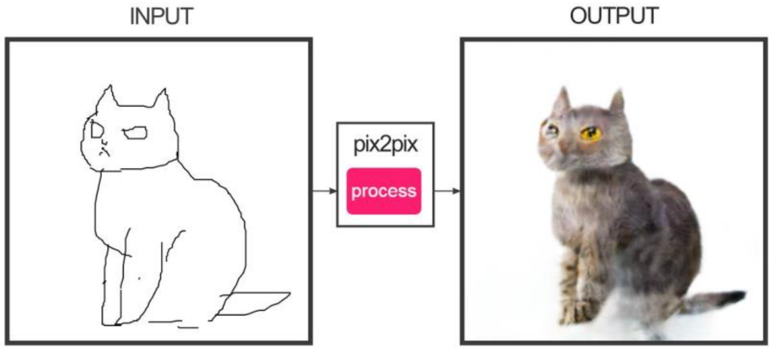
Input Output Input Output Input Output

Trained on Edges → Images

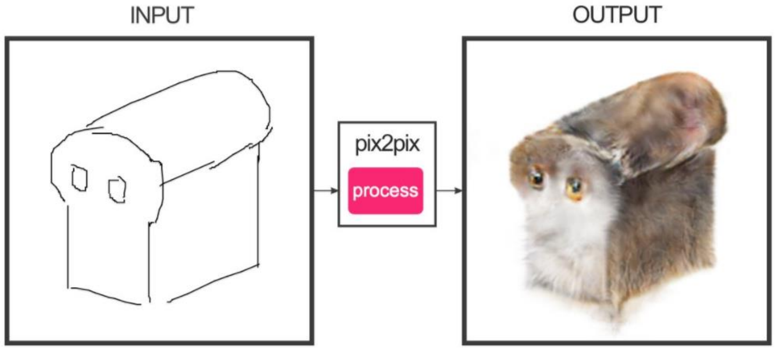
Data from [Eitz, Hays, Alexa, 2012]

slides credit: Isola / Zhu

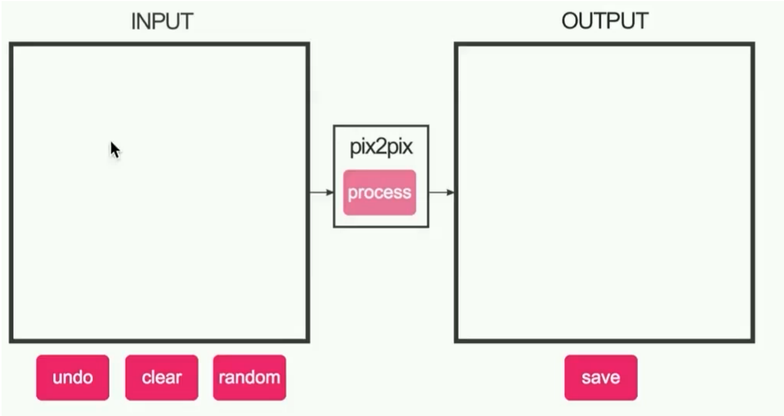
#edges2cats [Christopher Hesse]



@gods_tail



Ivy Tasi @ivymyt



@matthematician

Vitaly Vidmirov @vvid

https://affinelayer.com/pixsrv/

slides credit: Isola / Zhu

Input Output Groundtruth

Data from
[maps.google.com]

slides credit: Isola / Zhu

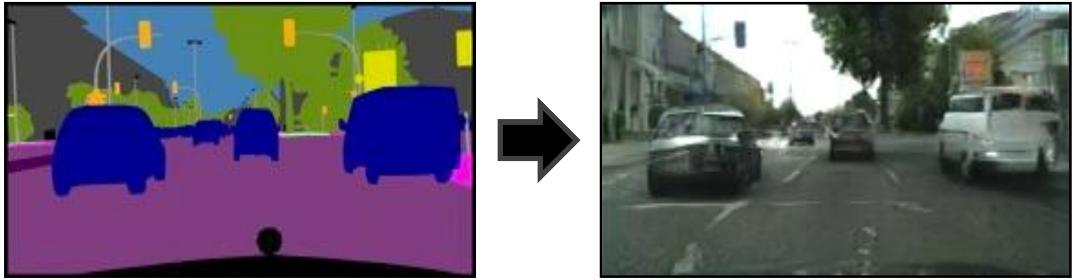
BW → Color

Input Output Input Output Input Output

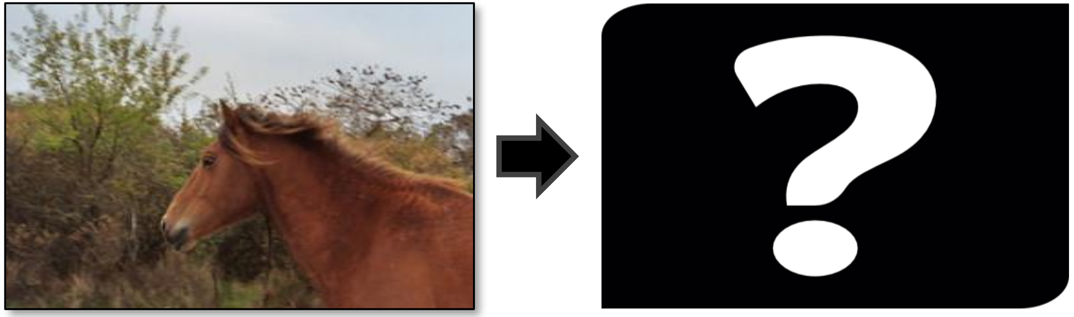
Data from [Russakovsky et al. 2015]

slides credit: Isola / Zhu

Paired



Label ↔ photo: per-pixel labeling



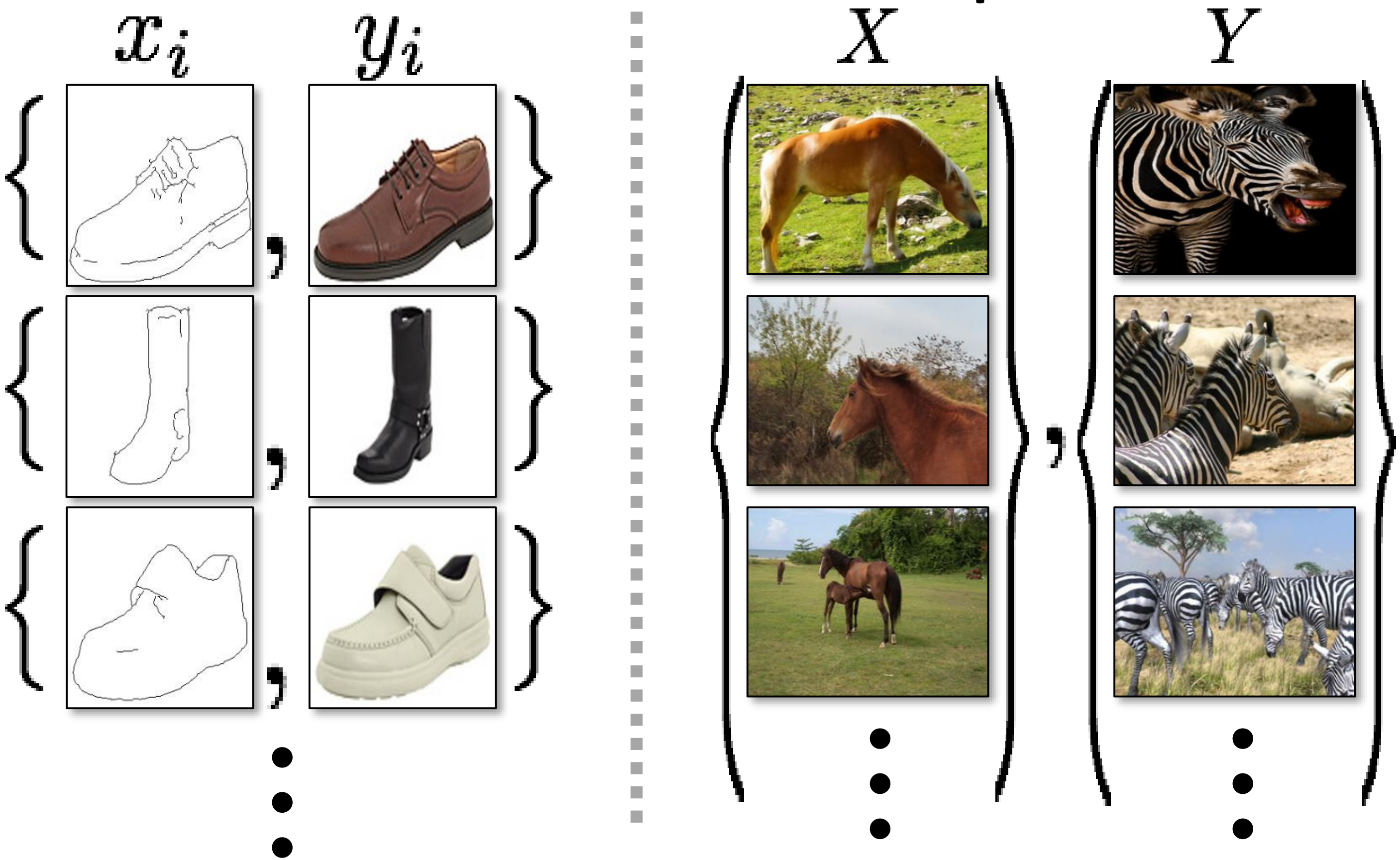
Horse ↔ zebra: how to get zebras?

- Expensive to collect pairs.
- Impossible in many scenarios.

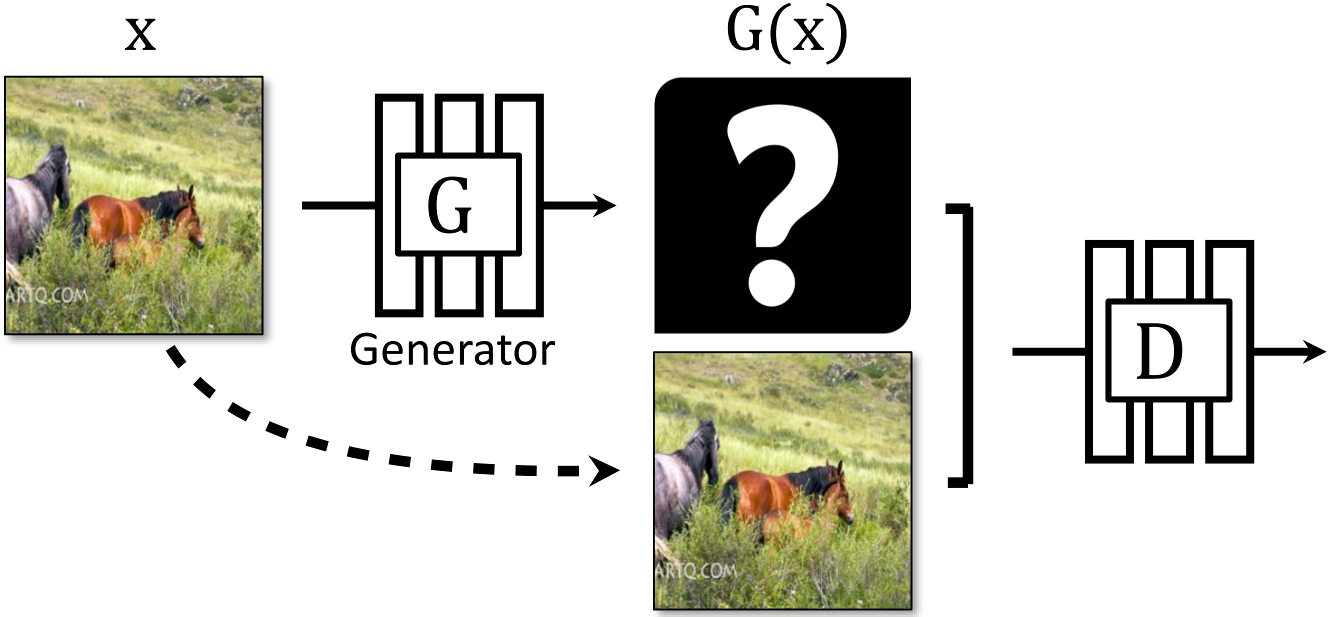
slides credit: Isola / Zhu

Paired

Unpaired

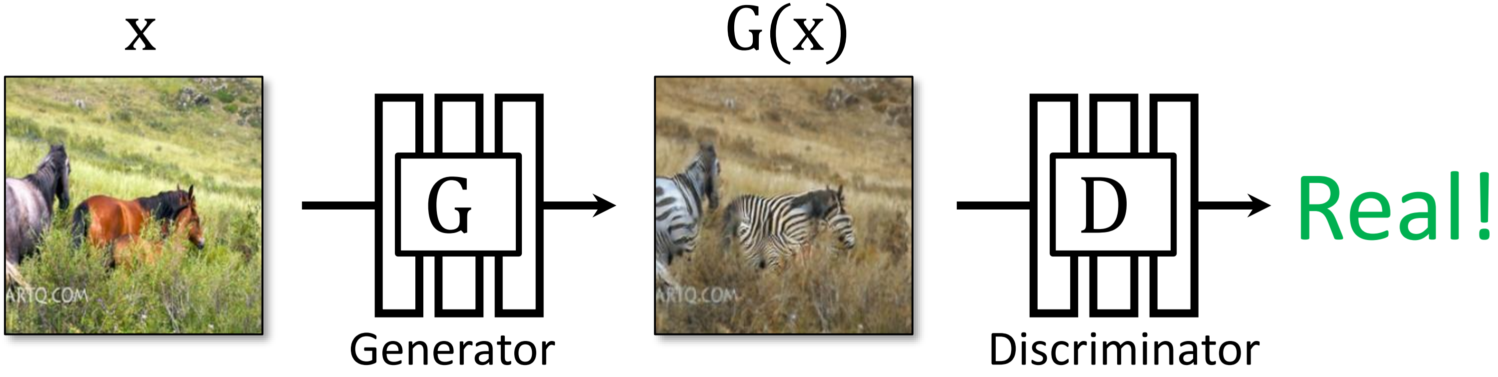


slides credit: Isola / Zhu

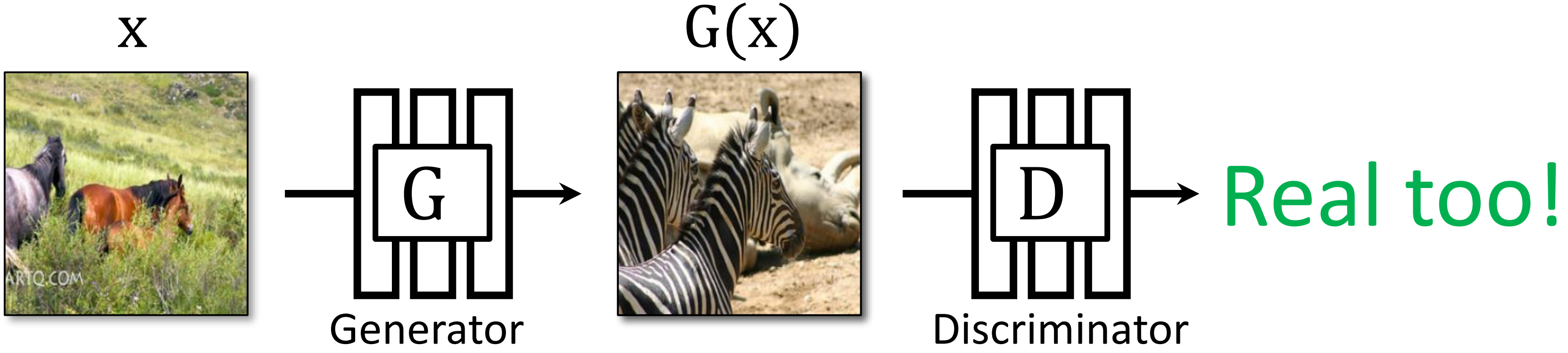


No input-output pairs!

slides credit: Isola / Zhu

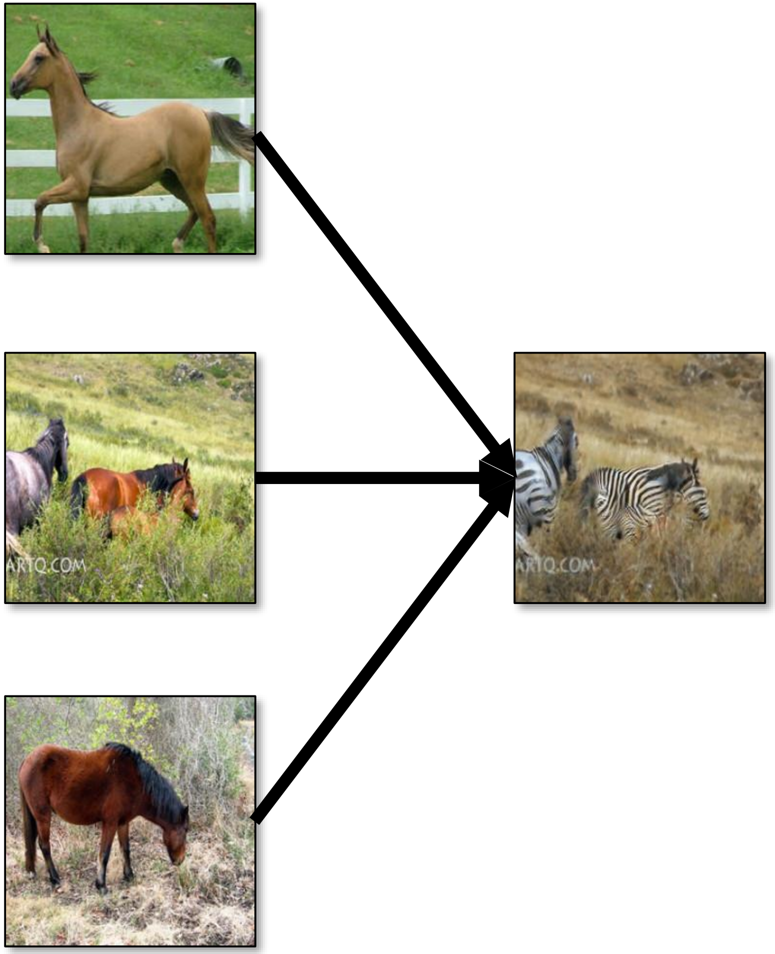


slides credit: Isola / Zhu



GANs doesn’t force output to correspond to input

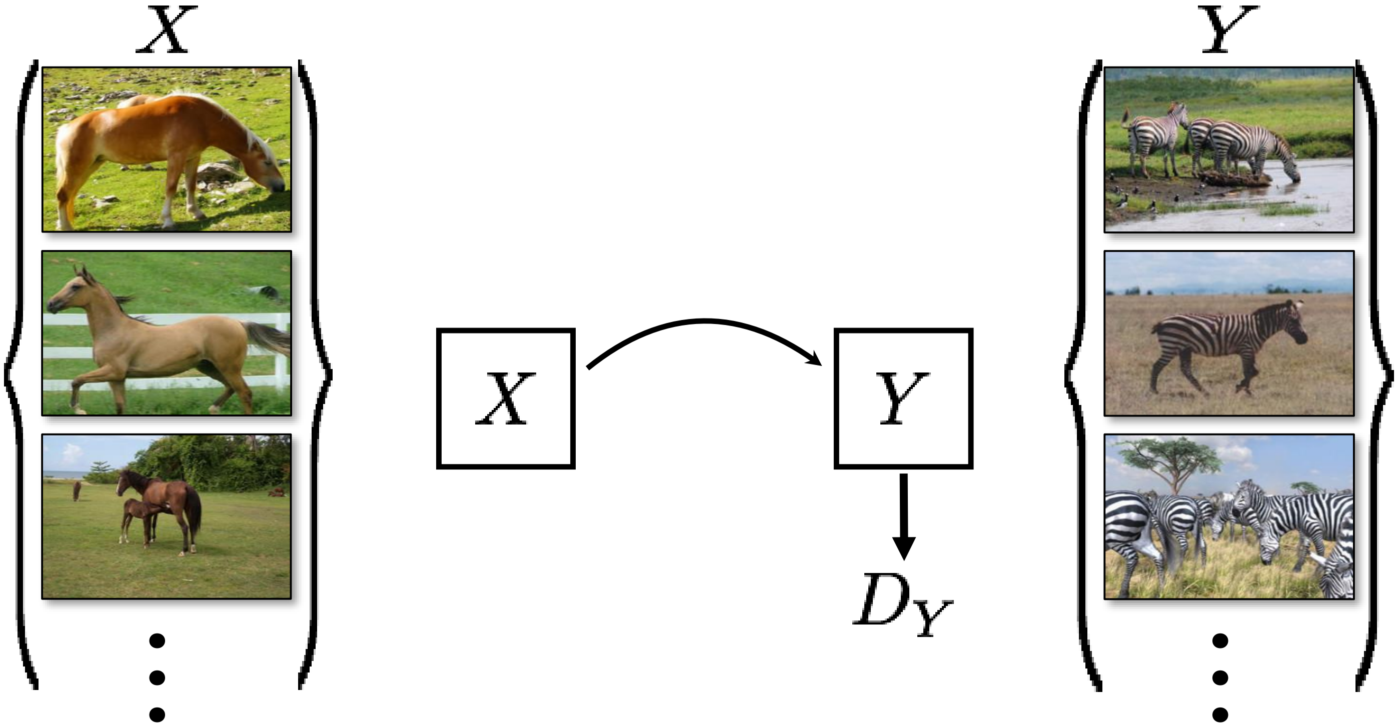
slides credit: Isola / Zhu



mode collapse!

slides credit: Isola / Zhu

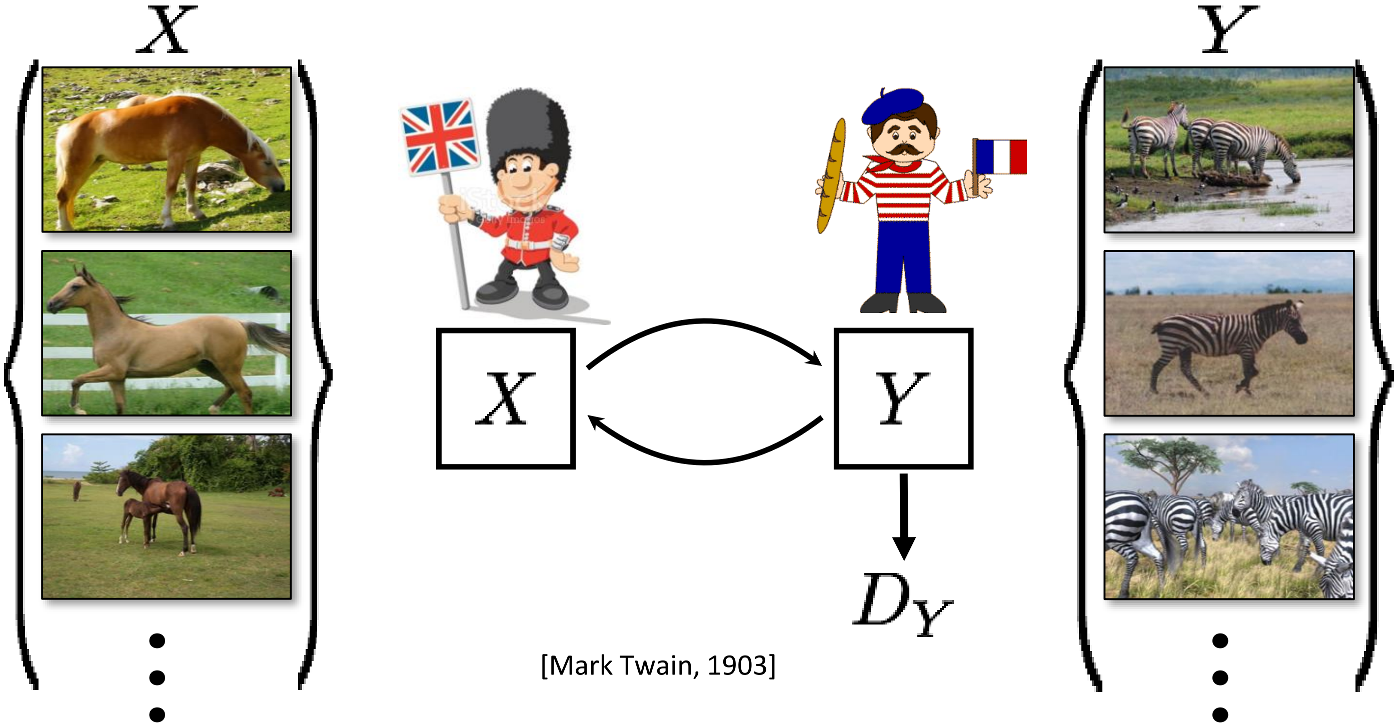
Cycle-Consistent Adversarial Networks



slides credit: Isola / Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

Cycle-Consistent Adversarial Networks

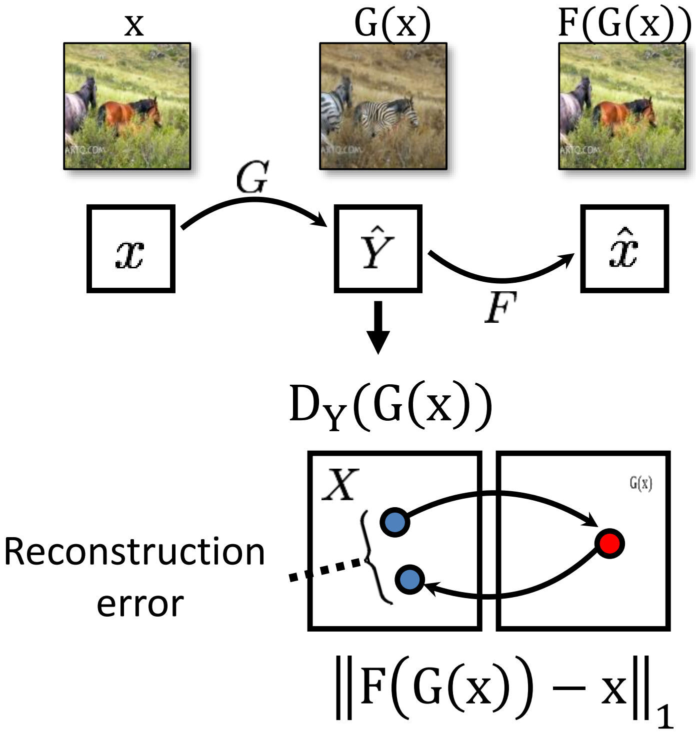


[Mark Twain, 1903]

slides credit: Isola / Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

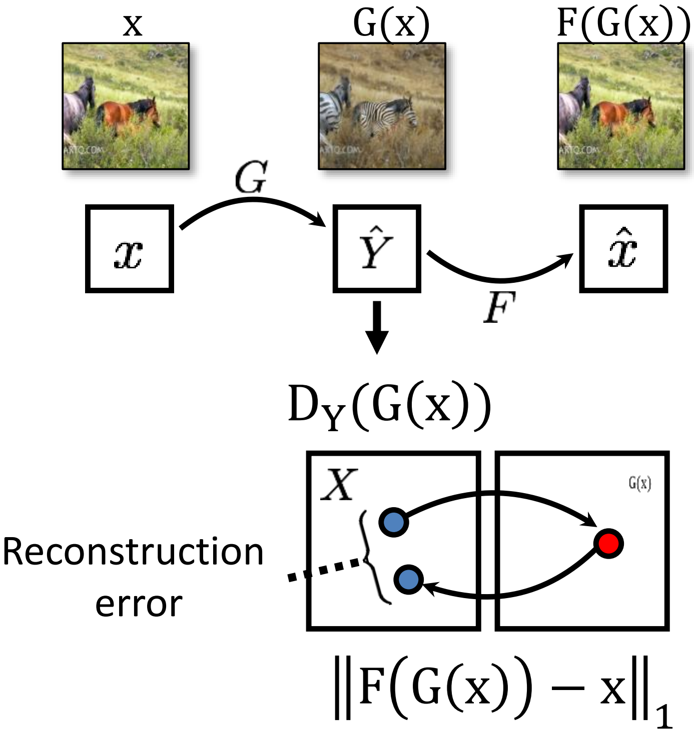
Cycle Consistency Loss

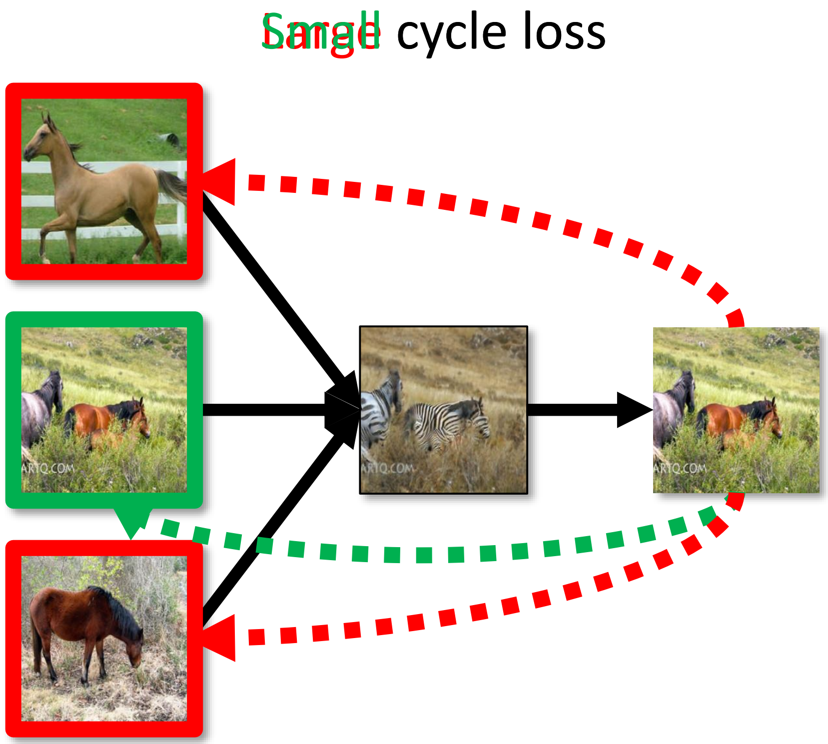


slides credit: Isola / Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

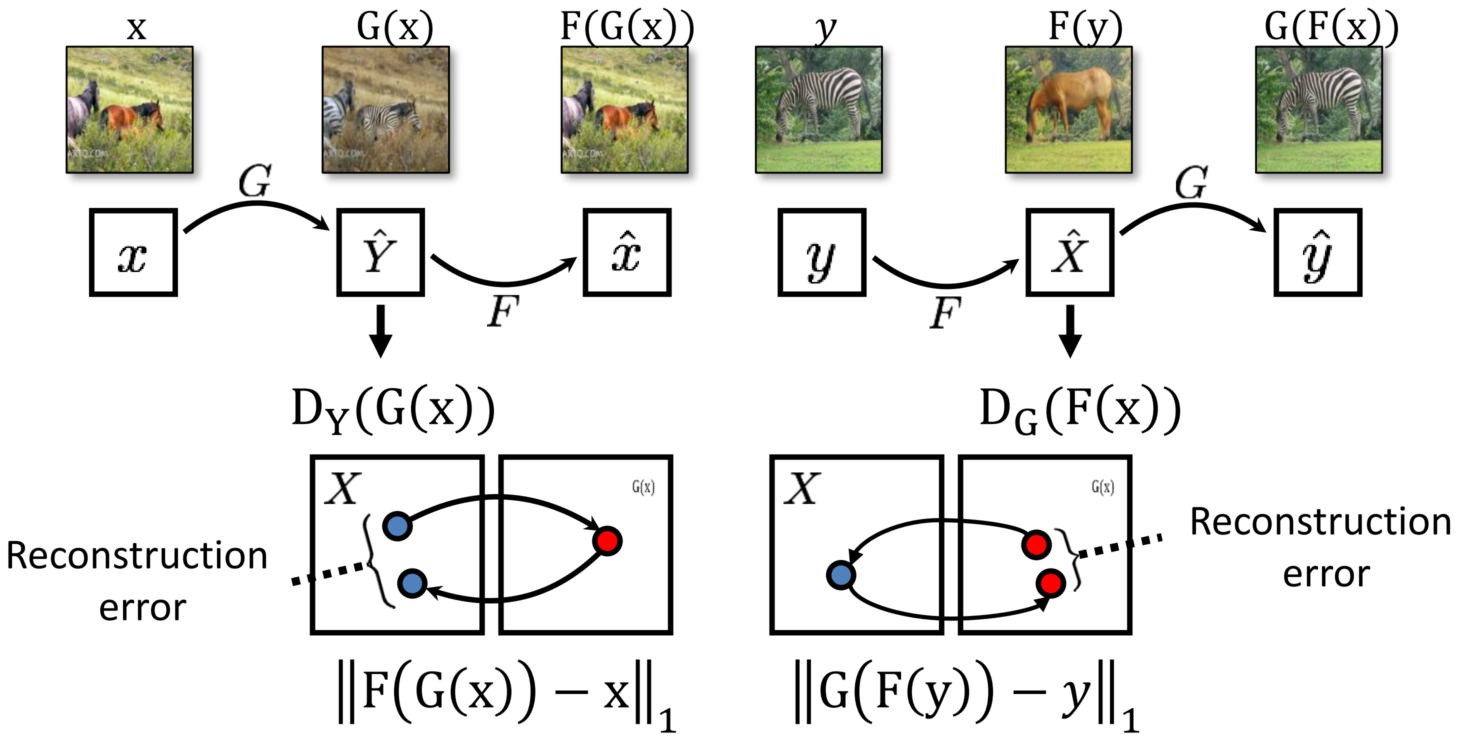
Cycle Consistency Loss





[Zhu*, Park*, Isola, and Efros, ICCV 2017]

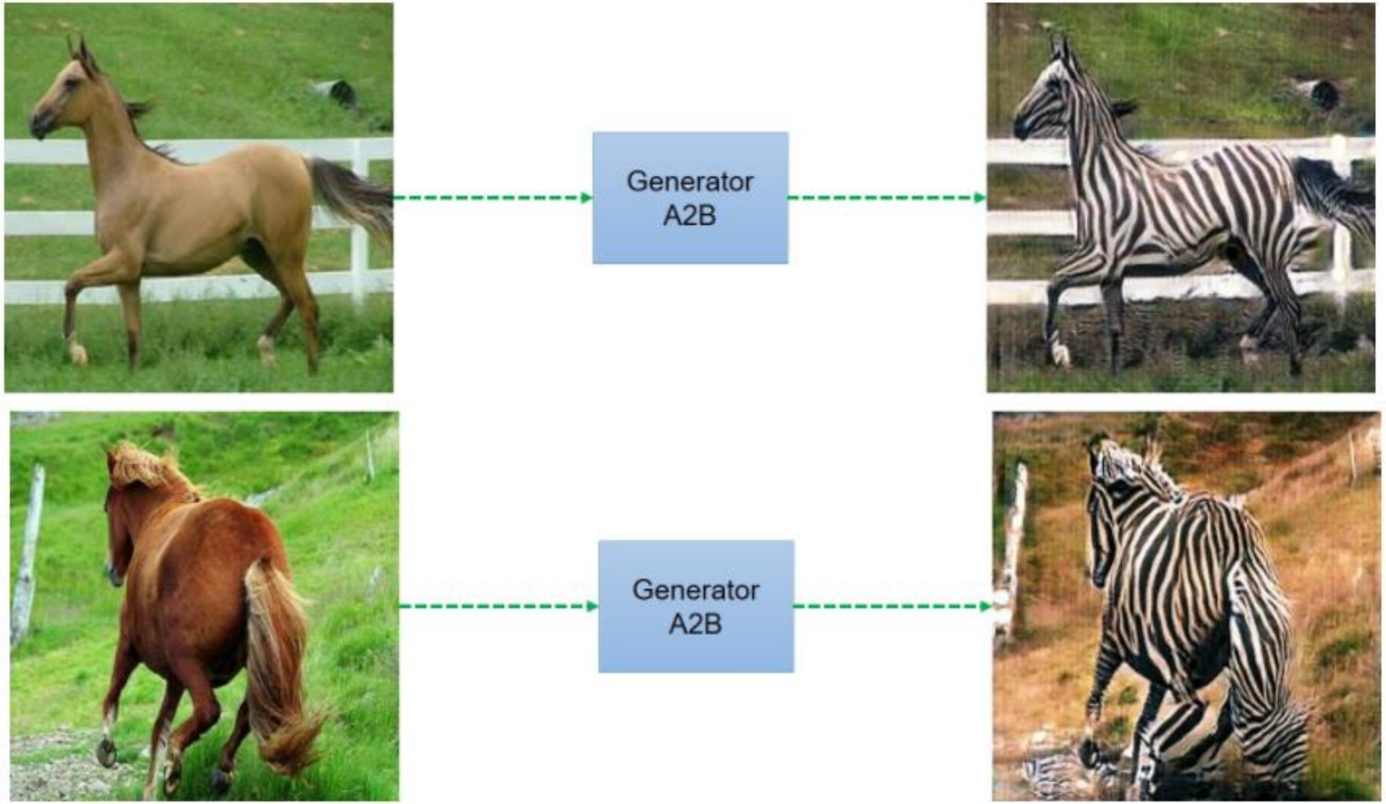
Cycle Consistency Loss



slides credit: Isola / Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

Cycle GAN - Overview



https://junyanz.github.io/CycleGAN/ [Zhu et al. 17.]

Monet's paintings → photos

slides credit: Isola / Zhu

slides credit: Isola / Zhu

slides credit: Isola / Zhu

Next Lectures

- Next Monday 24^{th},
  - Xmas GANs
  - No Lecture
- Jan 14^{th} -> No lecture, but office hours

- Next Lecture -> Jan 14^{th}

- We are still working on feedback for presentations – will send around asap…

- Keep working on the projects!

See you next year ☺