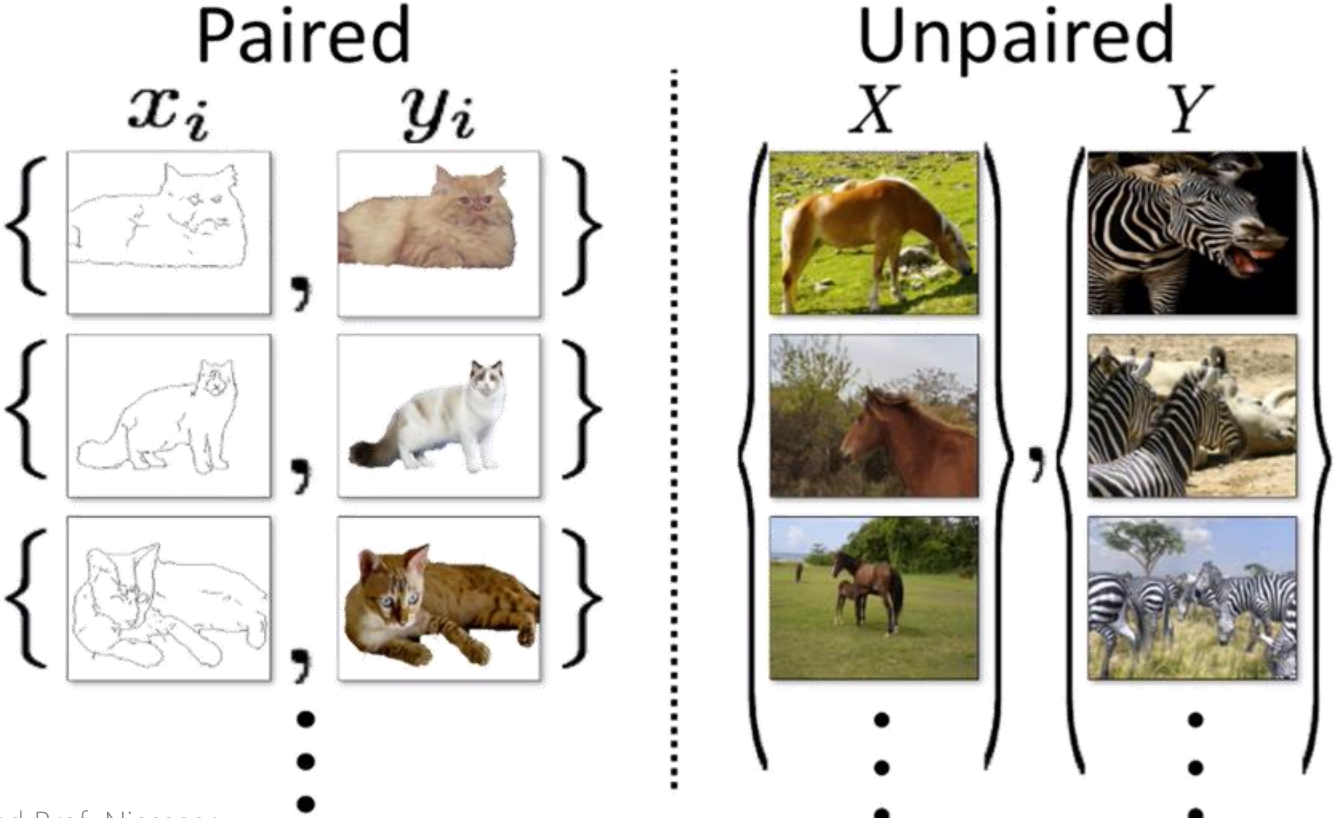
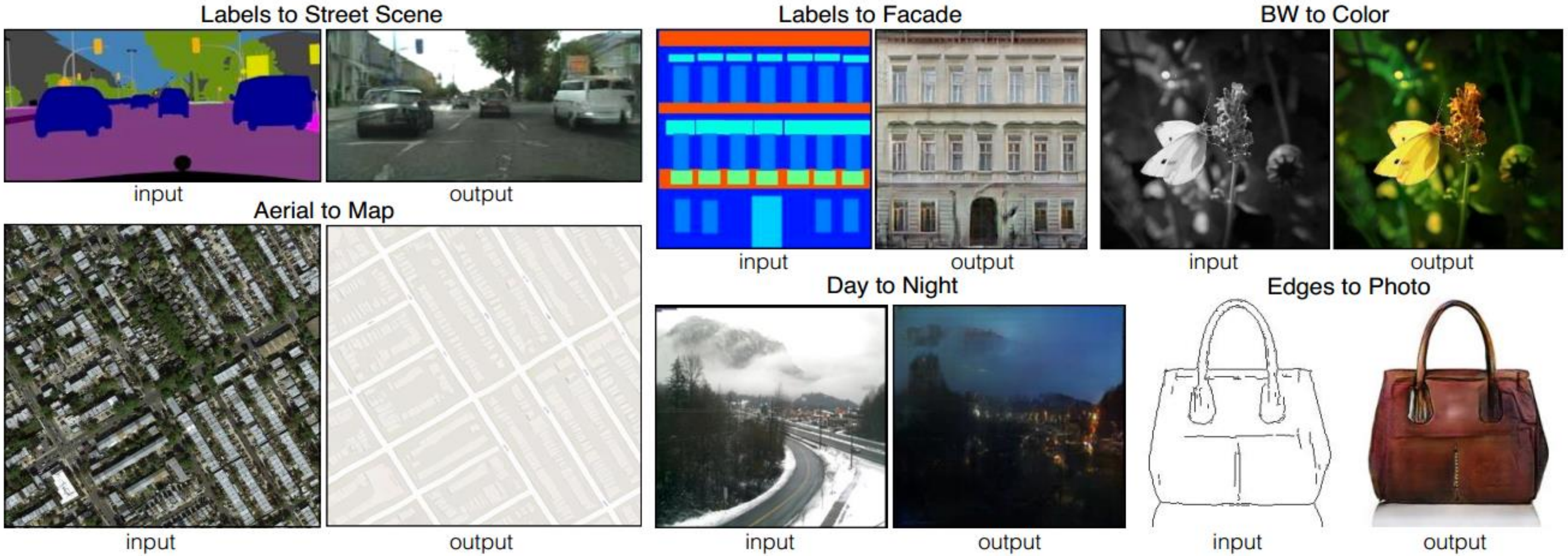
Conditional Generative Adversarial Networks (cGANs) continued!

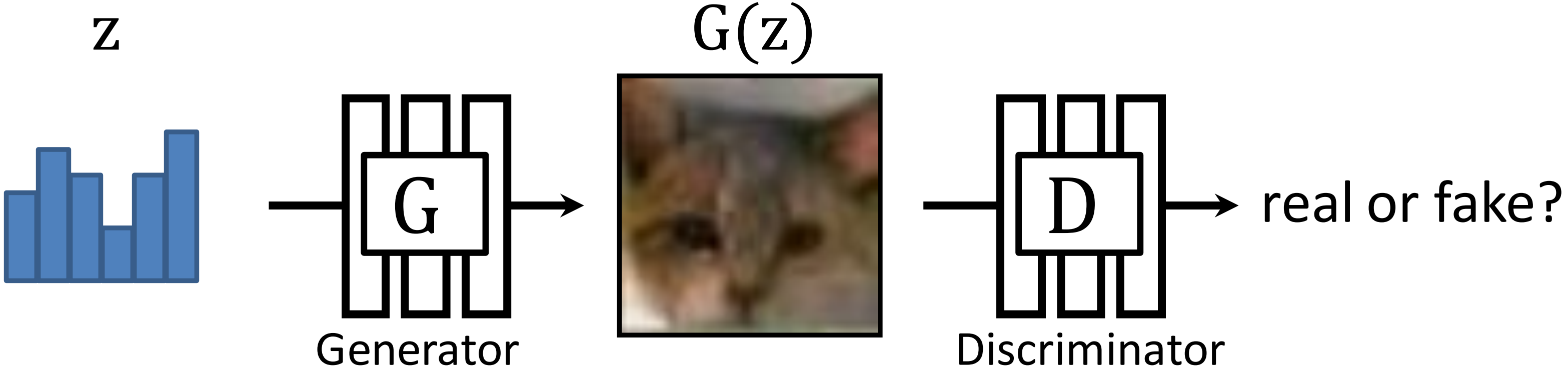
Paired vs Unpaired Setting



pix2pix: Image-to-Image Translation



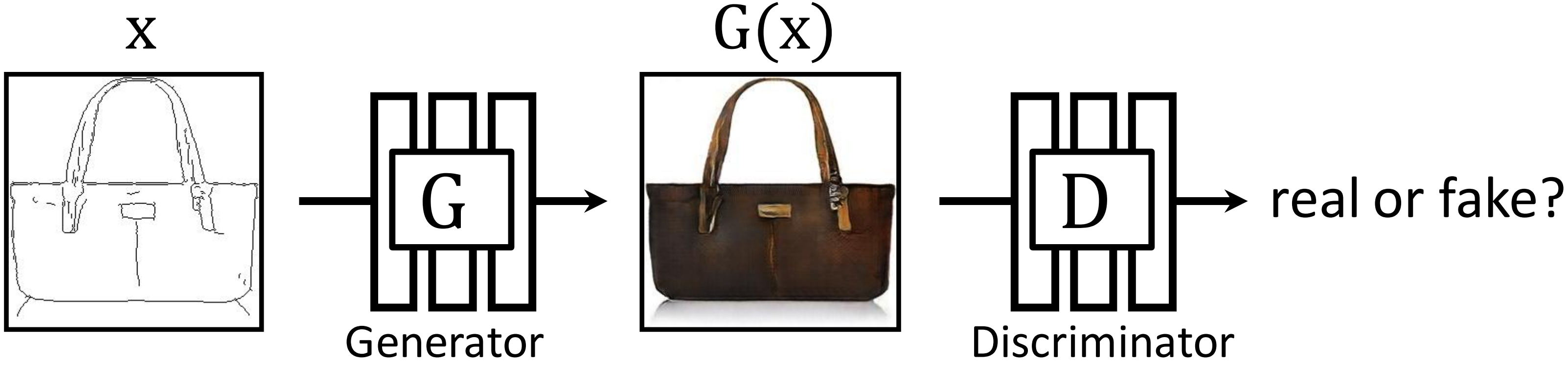
slides credit: Isola / Zhu



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

slides credit: Isola / Zhu

[Goodfellow et al. 2014]

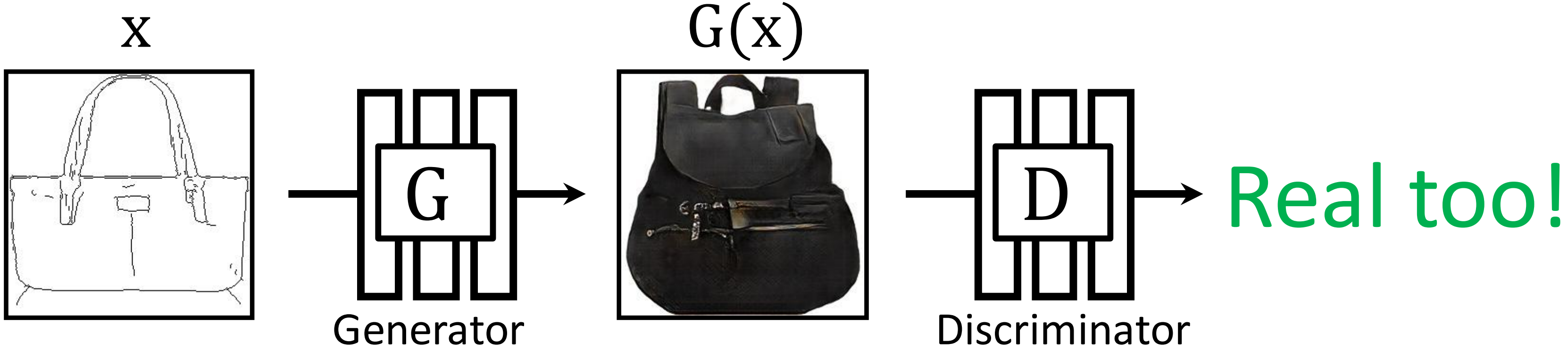


$$\min_{G} \max_{D} \mathbb{E}_{x,y}[\log D(G(x)) + \log(1 - D(y))]$$

slides credit: Isola / Zhu

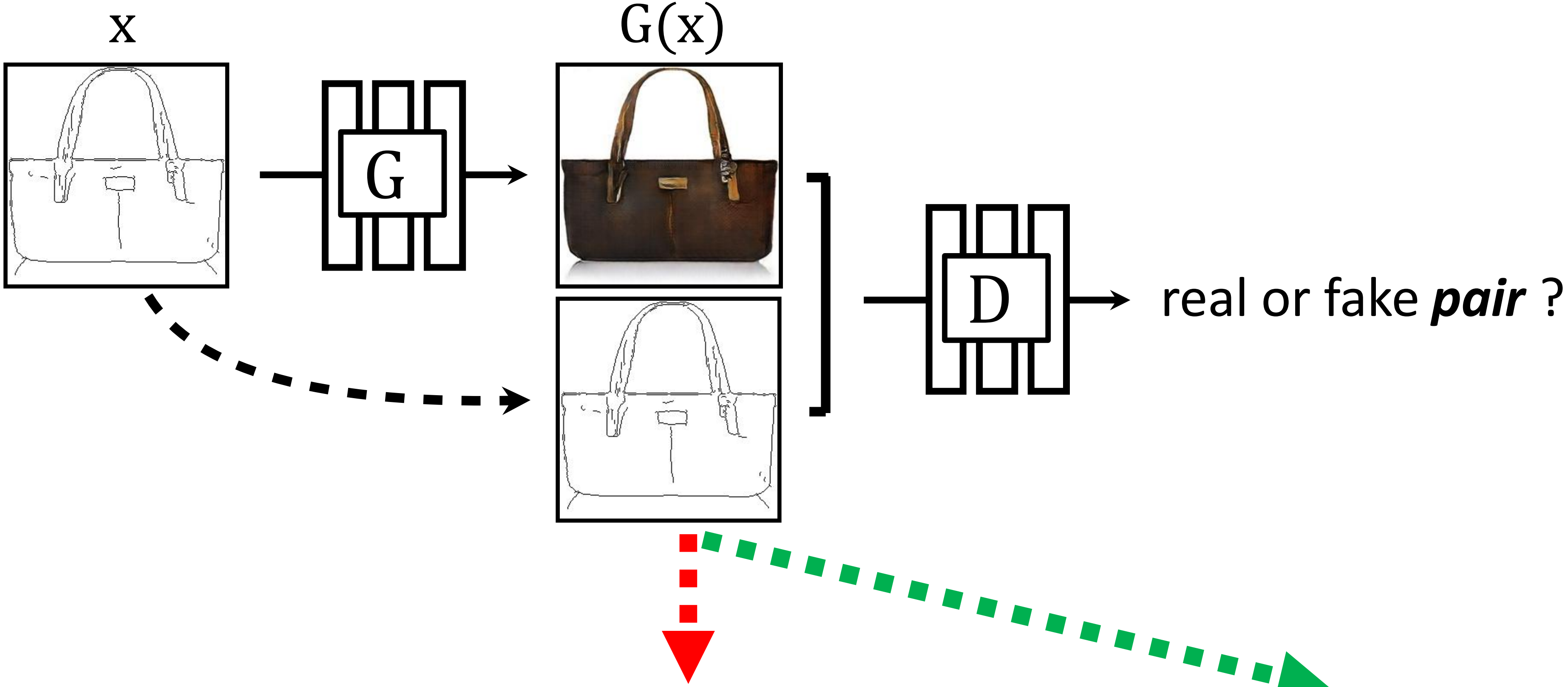
$$\min_G \max_D \mathbb{E}_{x,y}[\log D(G(x)) + \log(1 - D(y))]$$

slides credit: Isola / Zhu



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

slides credit: Isola / Zhu

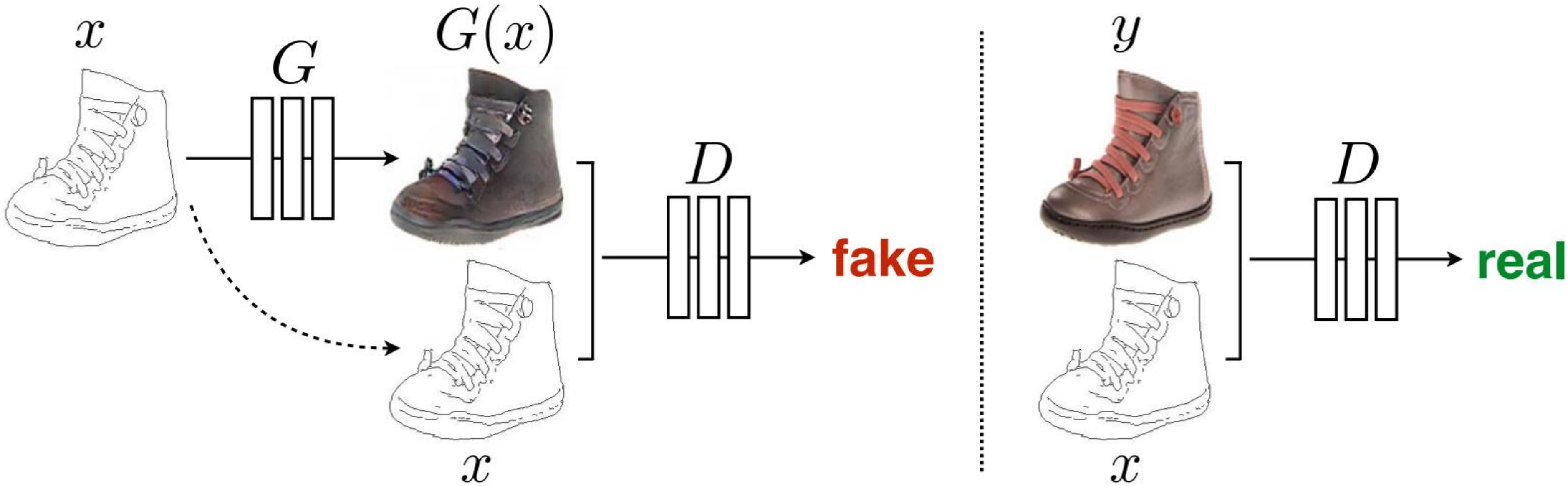


$$\min_G \max_D \mathbb{E}_{x,y}[\log \underbrace{D(x, G(x))}_{\text{fake pair}} + \log(1 - \underbrace{D(x, y)}_{\text{real pair}})]$$

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

slides credit: Isola / Zhu

pix2pix



pix2pix: Paired Setting

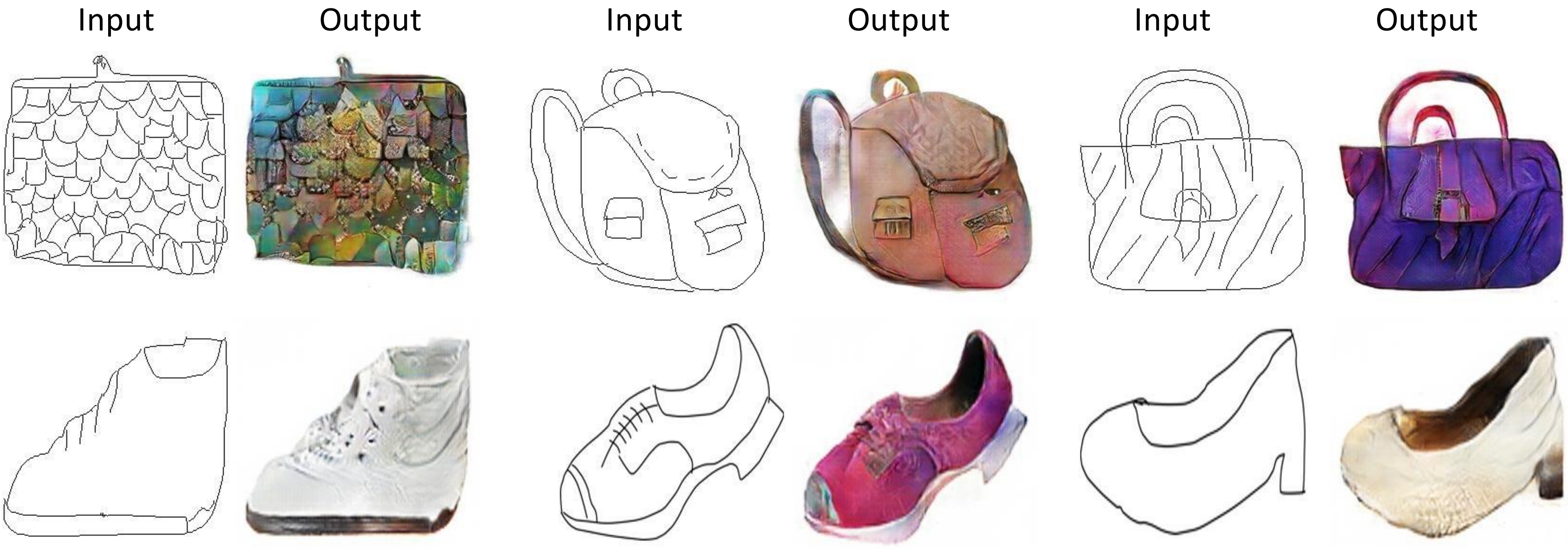
- Great when we have ‘free’ training data
- Often called self-supervised
- Think about these settings ☺

Edges → Images

Edges from [Xie & Tu, 2015]

slides credit: Isola / Zhu

Sketches → Images

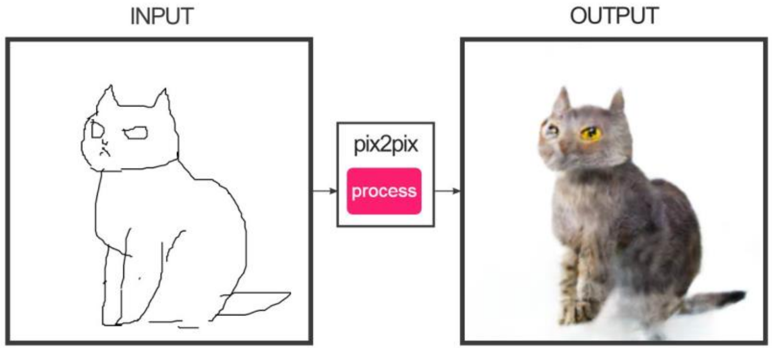


Trained on Edges → Images

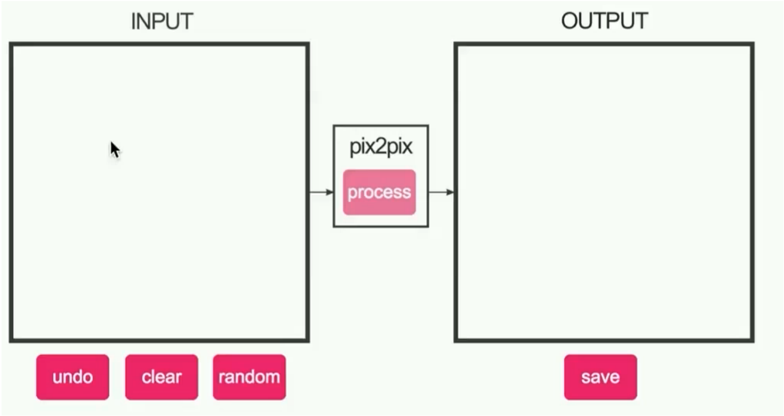
Data from [Eitz, Hays, Alexa, 2012]

slides credit: Isola / Zhu

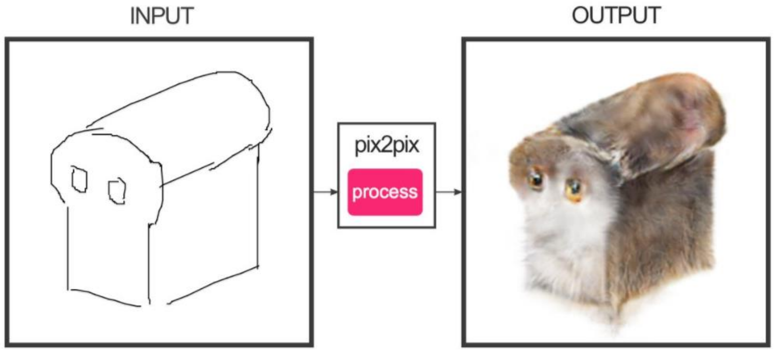
#edges2cats [Christopher Hesse]



@gods_tail



@matthematician



Ivy Tasi @ivymyt

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

Ideas behind Pix2Pix

- $L = L_{GAN} + \lambda L_1$ (makes it more constrained)
- Unet / skip connections for preserving structure
- Noise only through dropout
  - cGANs tend to learn to ignore the random vector z
  - Still want probabilistic model

Ideas behind Pix2Pix

- L1 or L2 loss for low frequency details

- GAN discriminator for high frequency details

-> PatchGAN

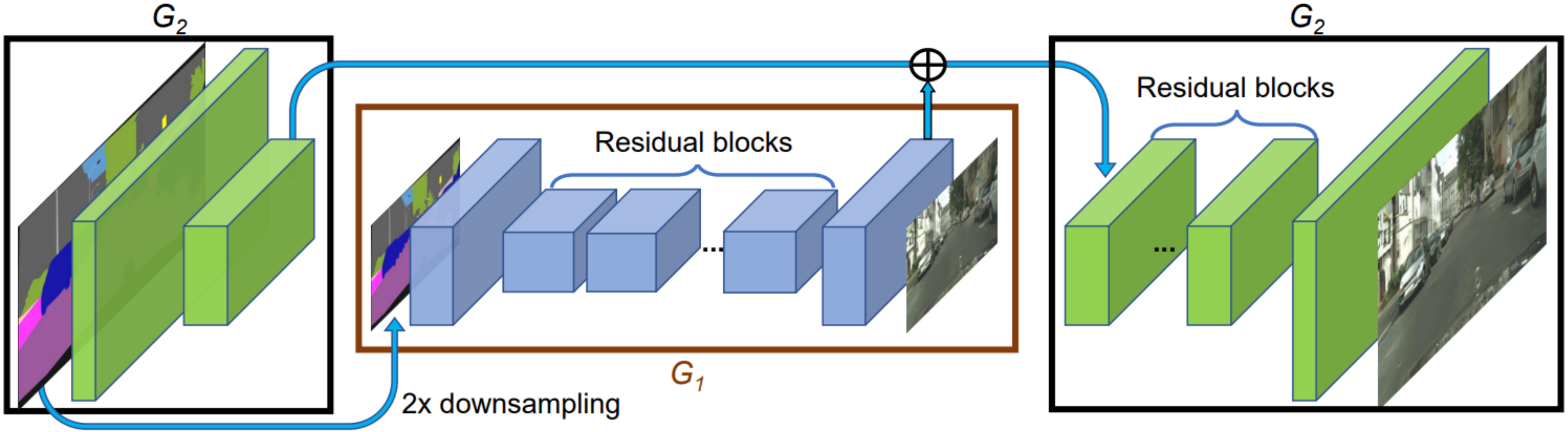
  – GAN discriminator applied only to local patches
  – It’ s fully-convolutional; i.e., can run on arbitrary image sizes

Pix2PixHD

- Expand the pix2pix idea to multi-scale
- Coarse-to-fine generator + discriminator
- G’s and D’s are the same but since they operate on different resolutions, they have effectively a larger receptive field

[Wang et al. 18]

Pix2PixHD



[Wang et al. 18]

Pix2PixHD

- Use of multi-scale discriminators

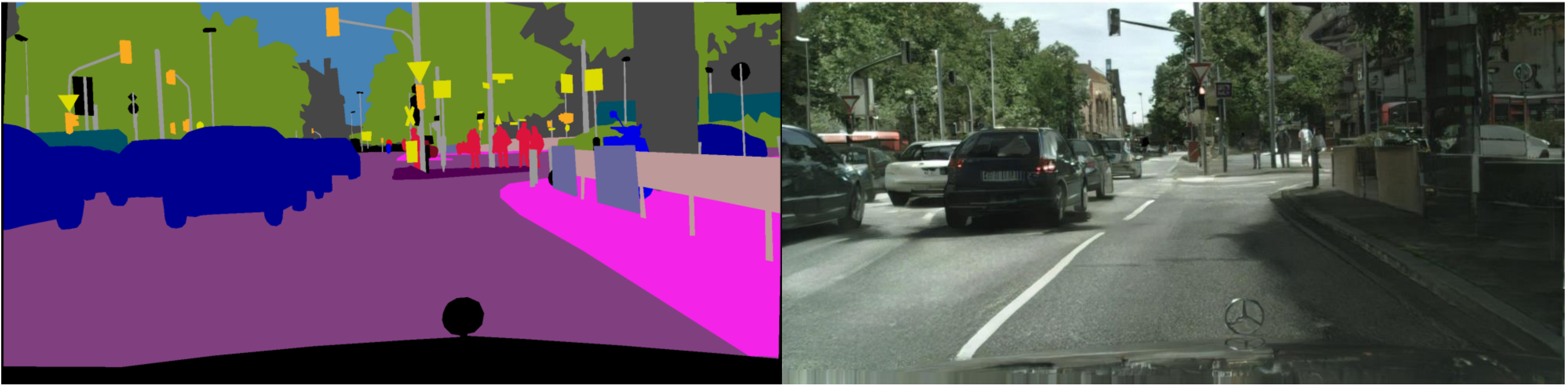
- $\min_{G} \max_{D_1,D_2,D_3} \sum_{k=1,2,3} L_{GAN}(G, D_k)$

- Can make various combinations of stacking discriminator and generator
  - E.g., have a single G and downsample generated and real images – or have intermediate real images (cf. ProGAN)

[Wang et al. 18]

Pix2PixHD

Input labels　　　　　　　　　　Synthesized image



[Wang et al. 18]

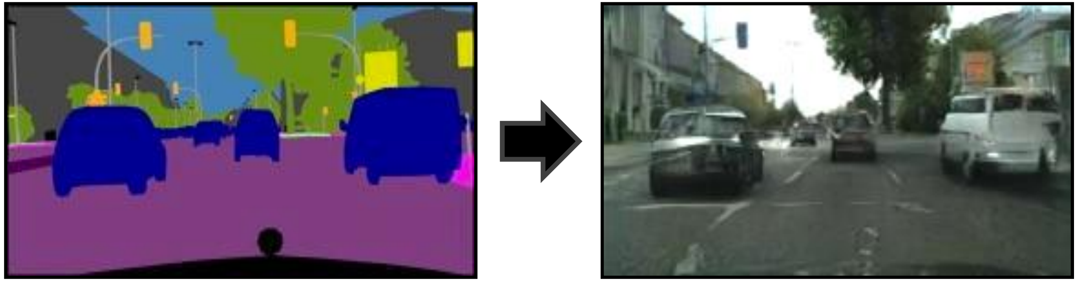
Pix2PixHD

[Wang et al. 18]

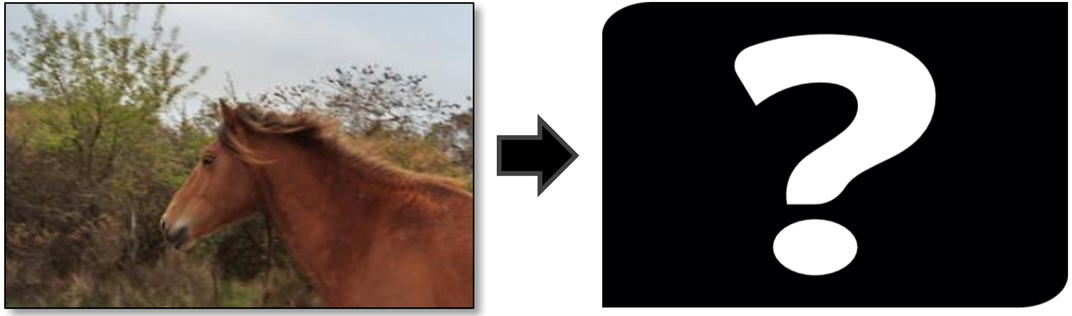
Pix2PixHD (interactive results)

[Wang et al. 18]

Paired



Label ↔ photo: per-pixel labeling



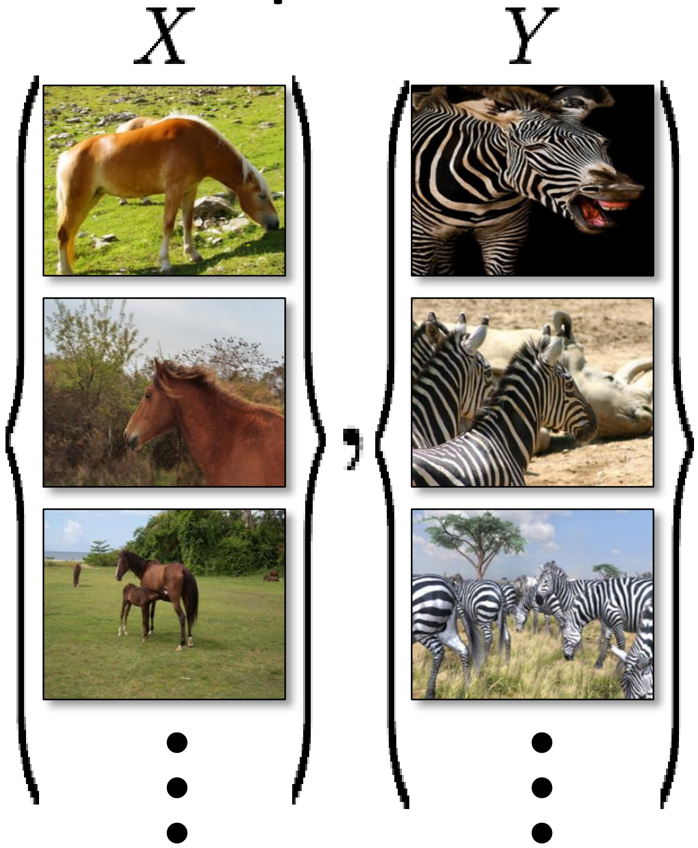
Horse ↔ zebra: how to get zebras?

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

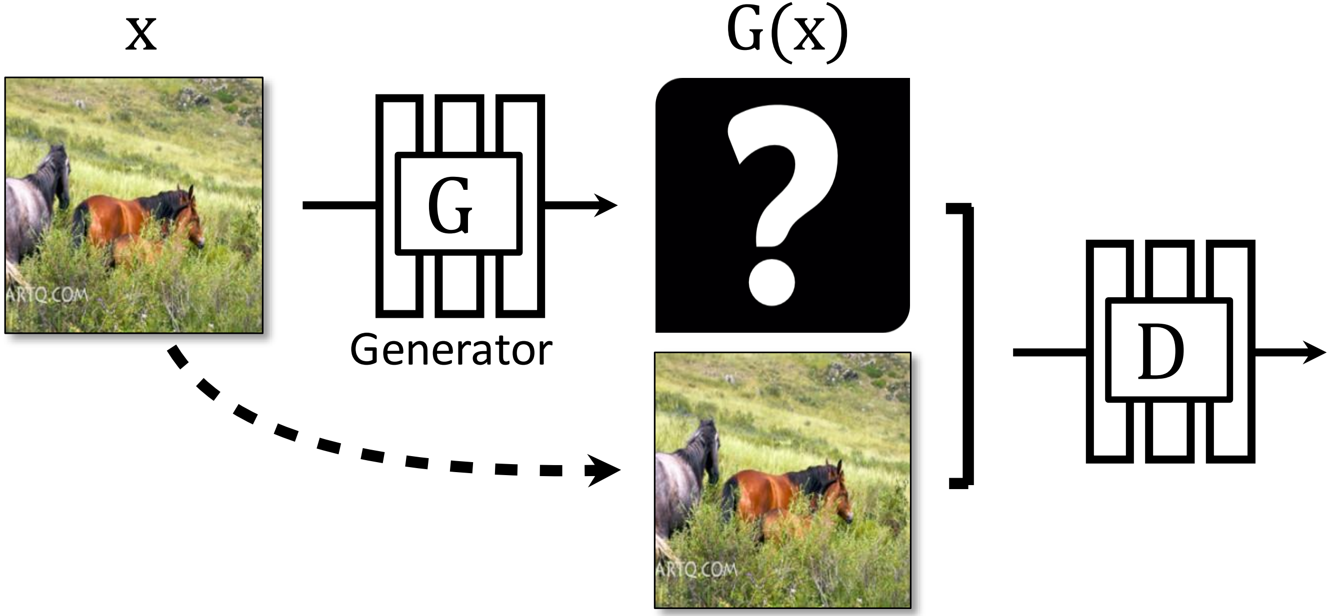
slides credit: Isola / Zhu

Paired

Unpaired

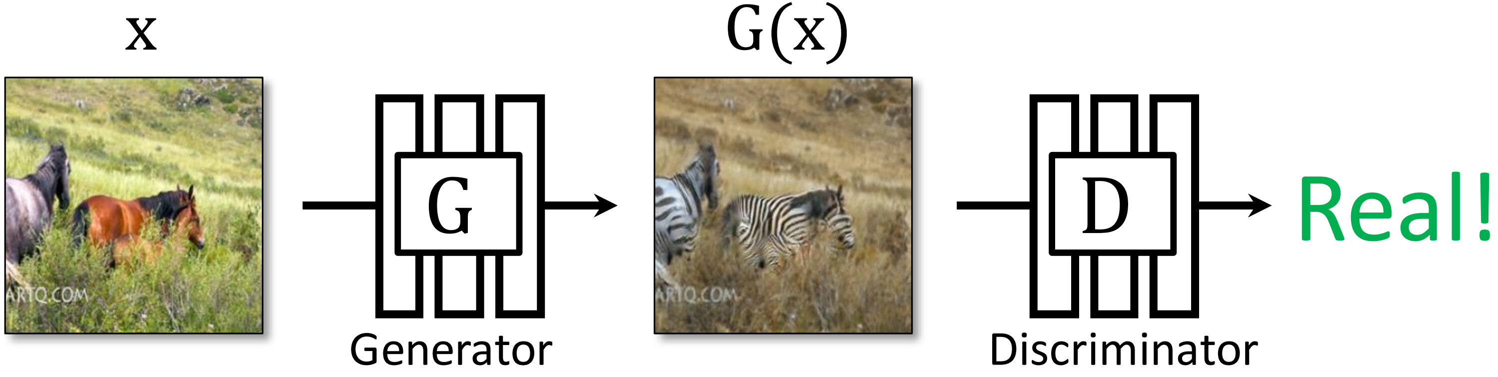


slides credit: Isola / Zhu

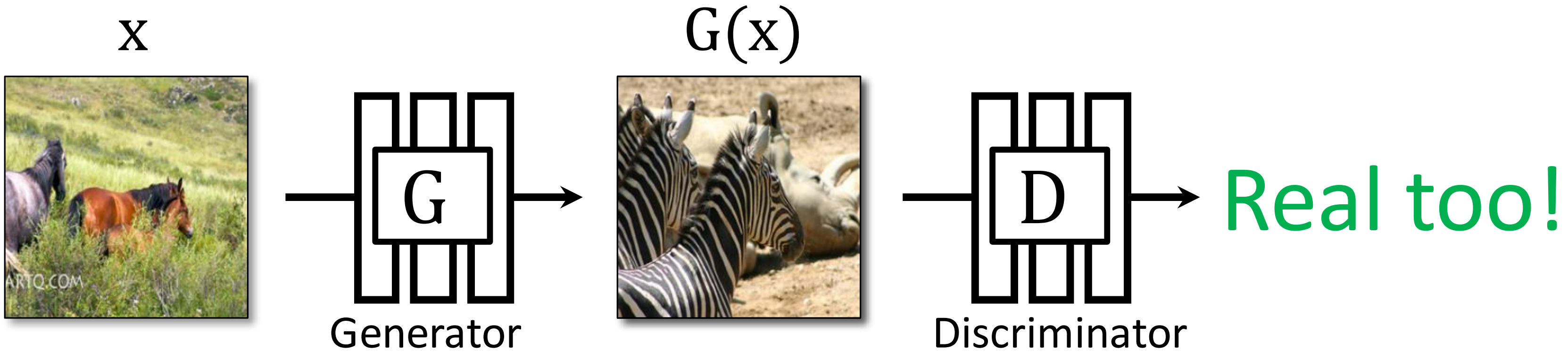


No input-output pairs!

slides credit: Isola / Zhu

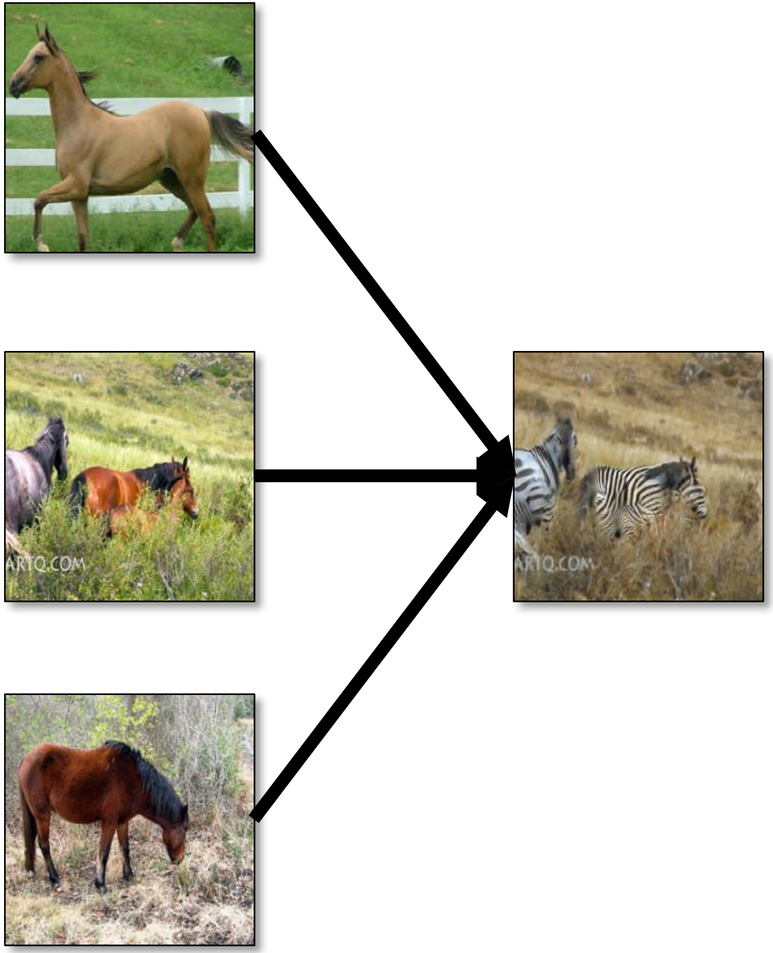


slides credit: Isola / Zhu



GANs doesn’t force output to correspond to input

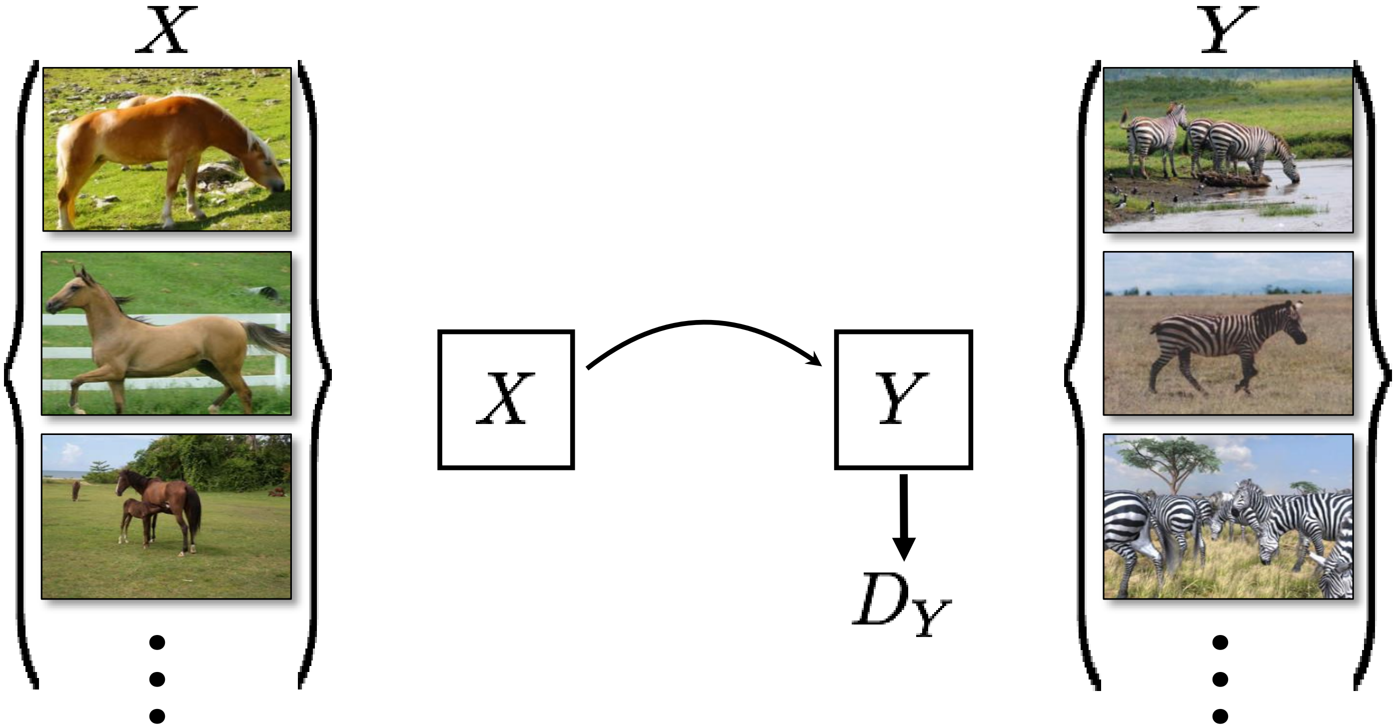
slides credit: Isola / Zhu



mode collapse!

slides credit: Isola / Zhu

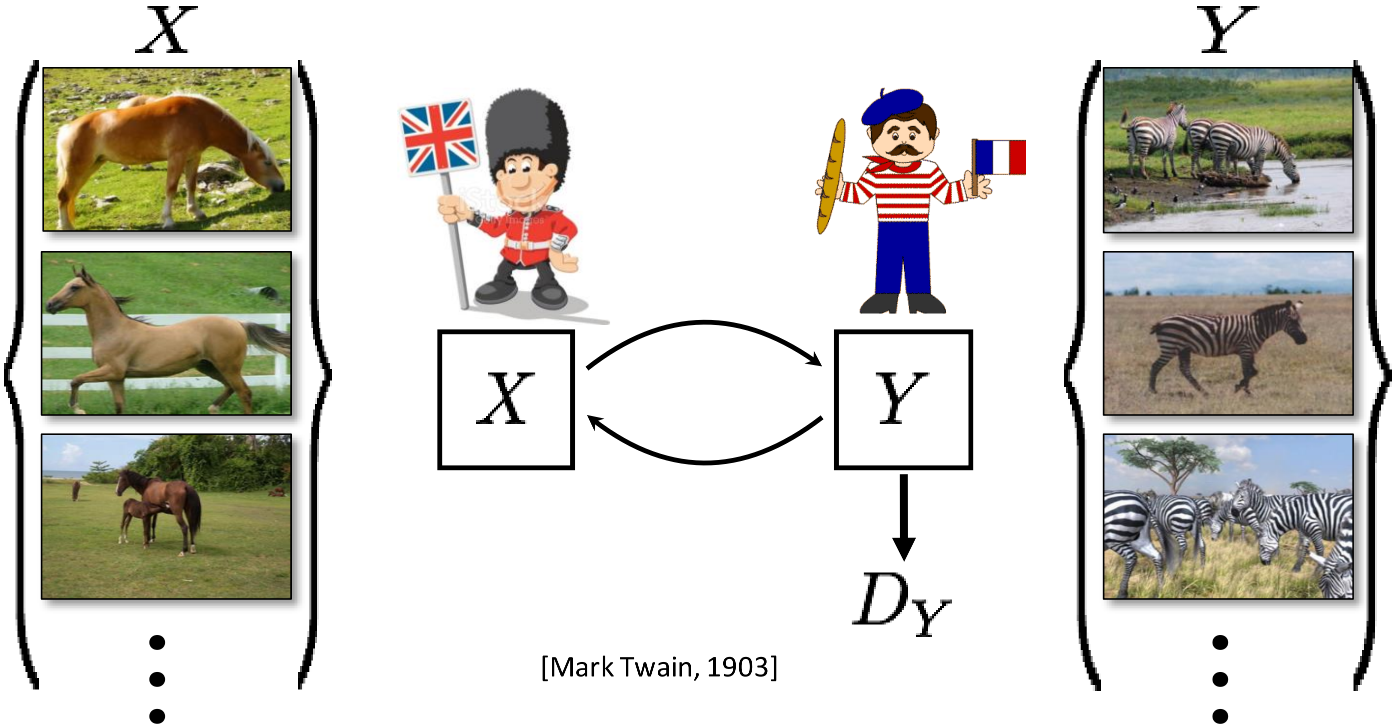
Cycle-Consistent Adversarial Networks



slides credit: Isola / Zhu

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

Cycle-Consistent Adversarial Networks

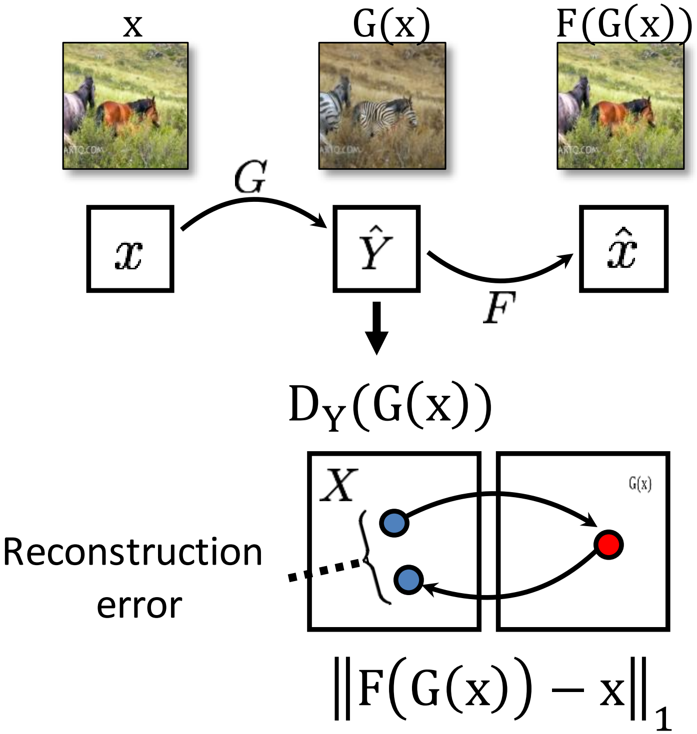


[Mark Twain, 1903]

slides credit: Isola / Zhu

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

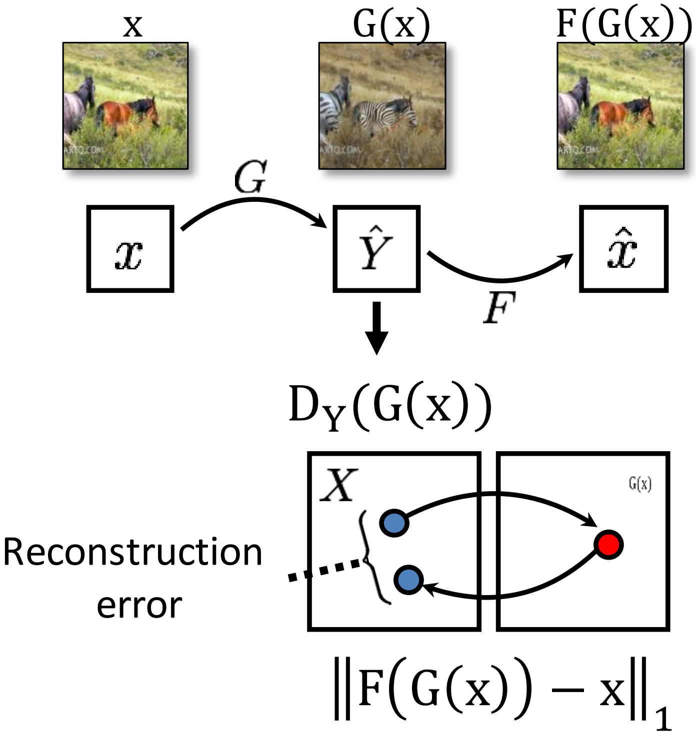
Cycle Consistency Loss

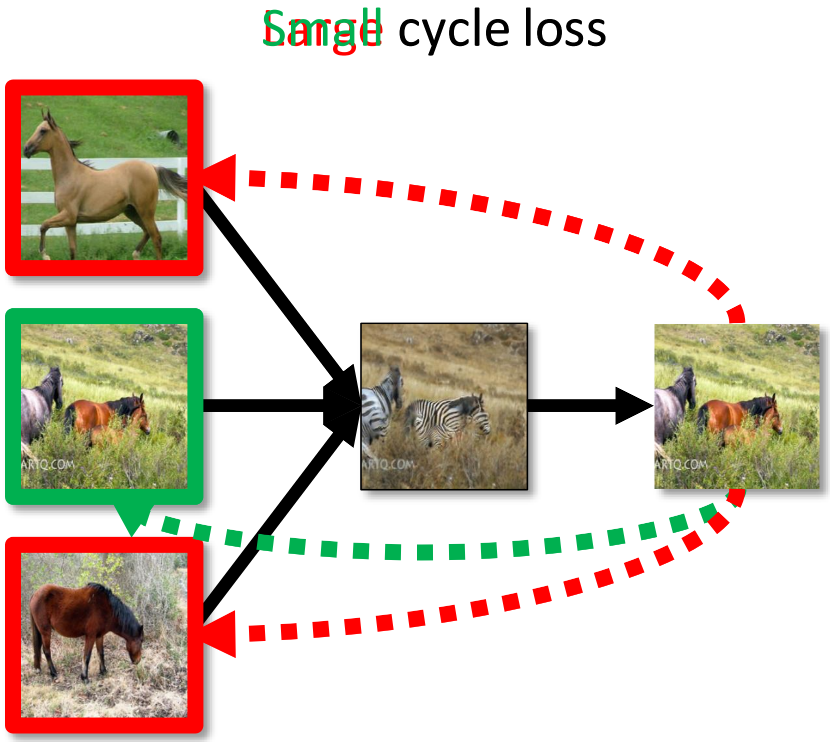


slides credit: Isola / Zhu

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

Cycle Consistency Loss

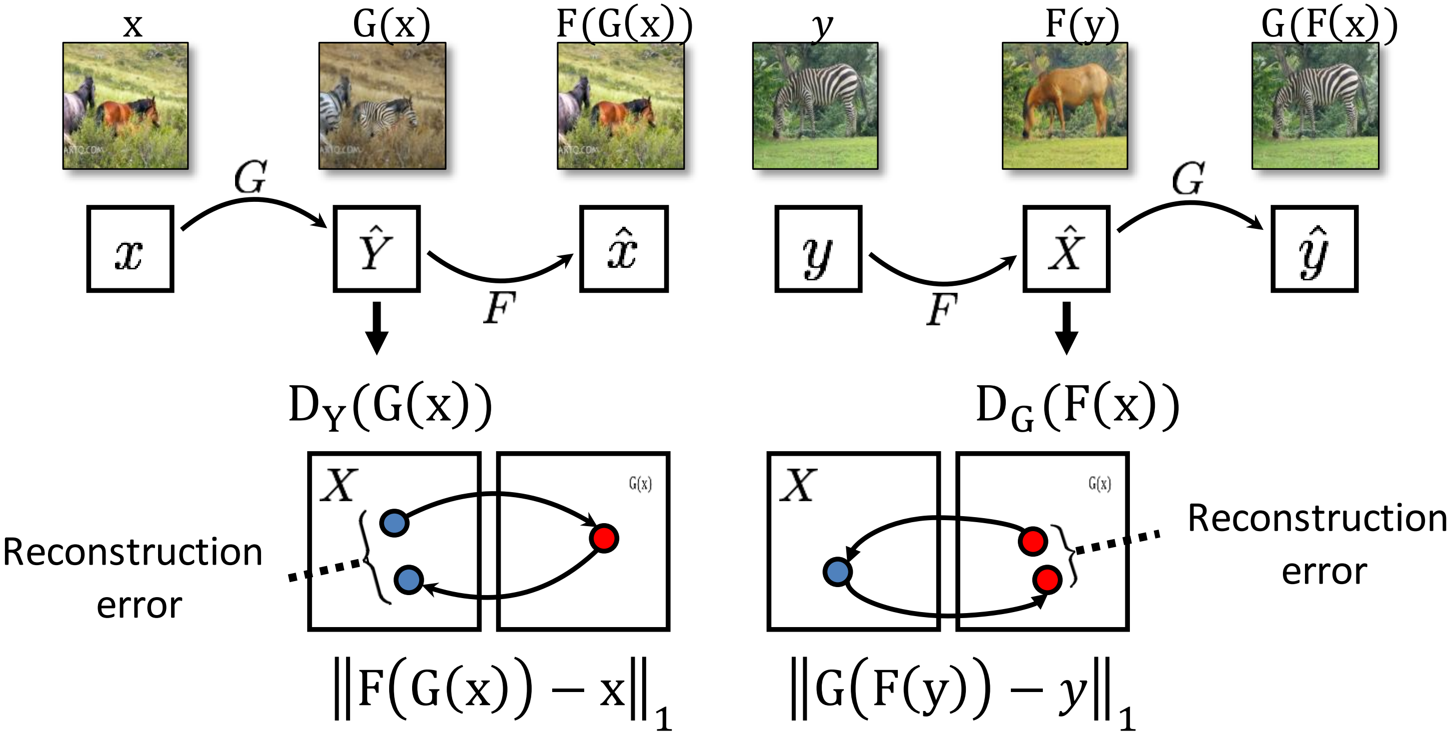




slides credit: Isola / Zhu

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

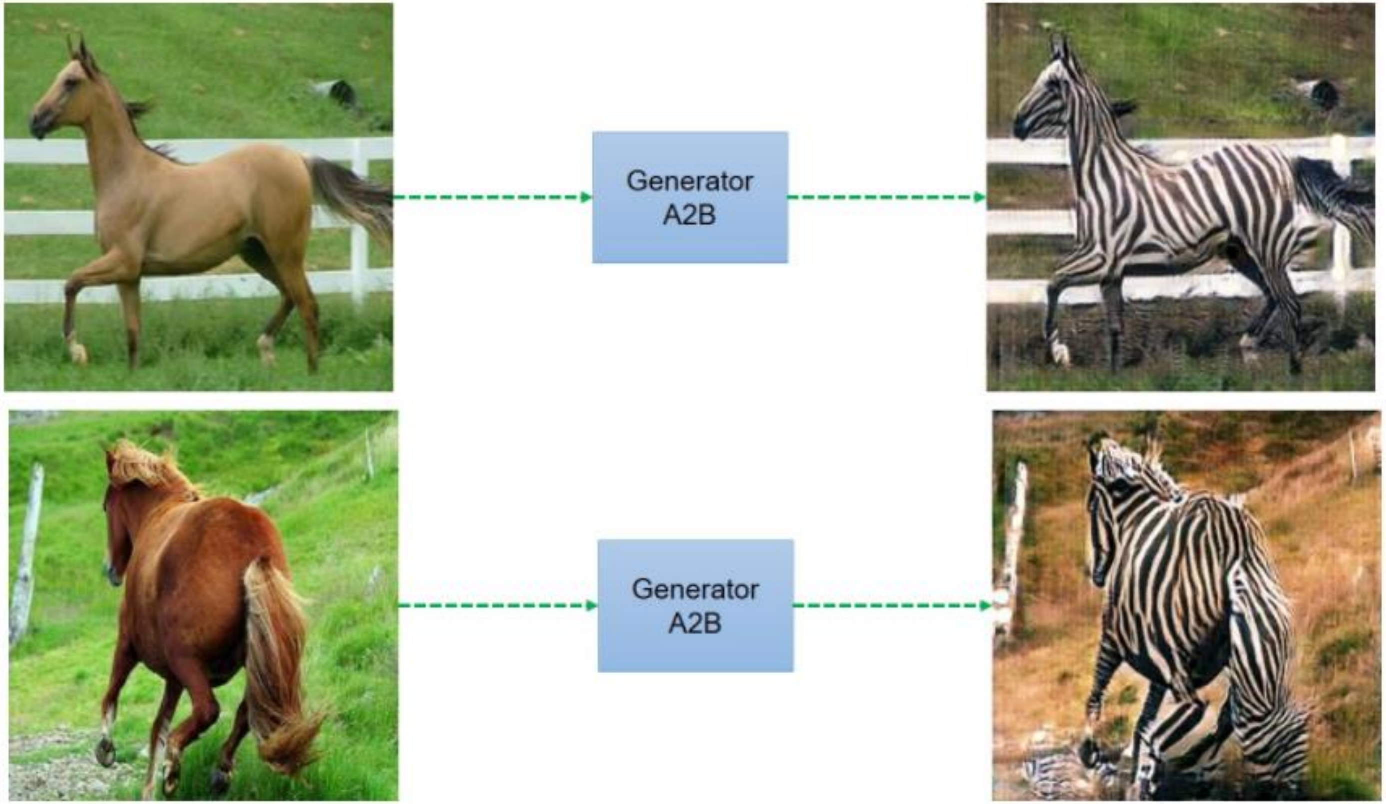
Cycle Consistency Loss



slides credit: Isola / Zhu

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

Cycle GAN - Overview



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

Cycle GAN: Objective

$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x)))],$$

Domain X   Domain Y

$$\underbrace{\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1].}_{\text{Cycle consistency}}$$

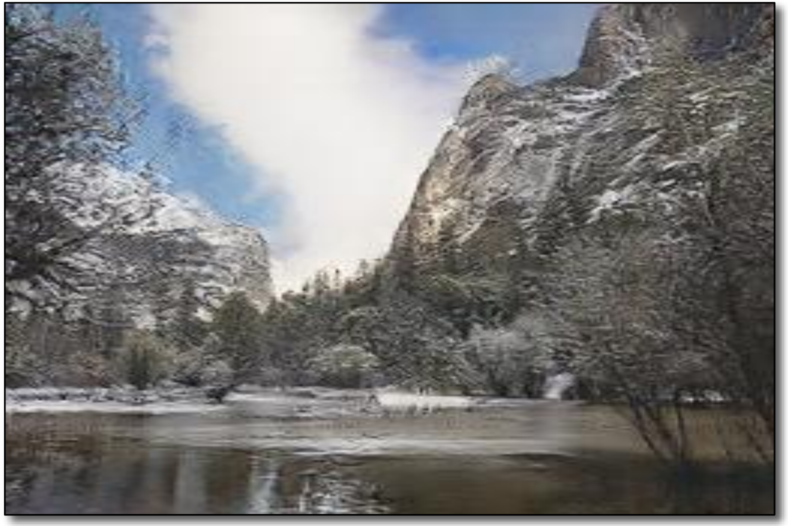
Full Loss: $$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) + \lambda \mathcal{L}_{\text{cyc}}(G, F),$$

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

Monet's paintings → photos

slides credit: Isola / Zhu

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



slides credit: Isola / Zhu

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

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

Administrative

Administrative

- Deadline for final projects
  - Wed Feb 6th, 11:59pm
  - Submission via moodle
  - Submission must contain
    - Code (results must be replicable)
    - 2-3 pages of final report (at most 1 page of text, rest results; i.e., images and tables)
    - Use CVPR templates: http://cvpr2019.thecvf.com/submission/main_conference/author_guidelines

Administrative

- Poster presentation
  - **Friday Feb 8th, 1pm-3pm**
  - Location:
    - Magistrale (preliminary – will update if it changes)
    - In the area next to the back entrance (parking lot direction)
  - Poster stands will be provided
  - You need to print posters yourself (poster@in.tum.de)
  - Hang posters 15 mins before presentation session starts

Guest Speakers

- Oriol Vinyals:
  - https://ai.google/research/people/OriolVinyals
  - Time: **January 31st, 6pm – 8pm**
  - Location: HS-1 (CS building – the big one)

Next Lectures

- Next Lecture -> Jan 21st
- Keep working on the projects!

Conditional Generative Adversarial Networks (cGANs) continued!