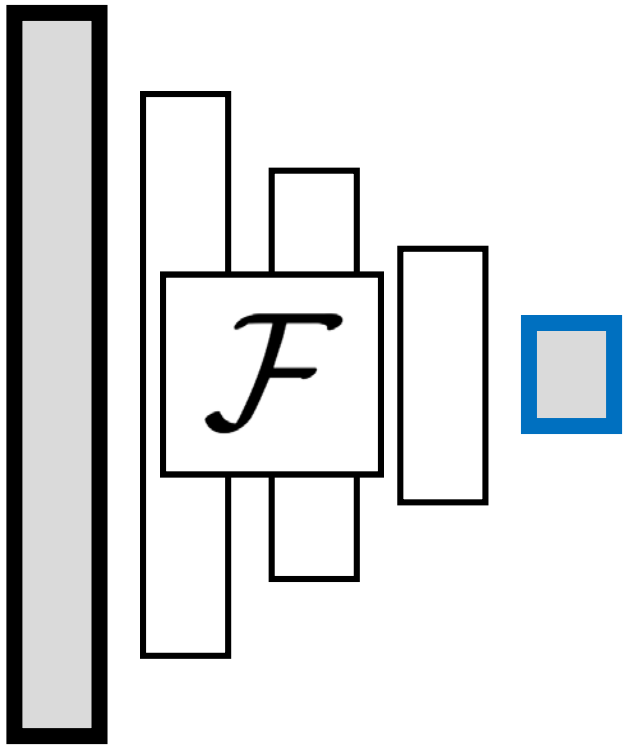
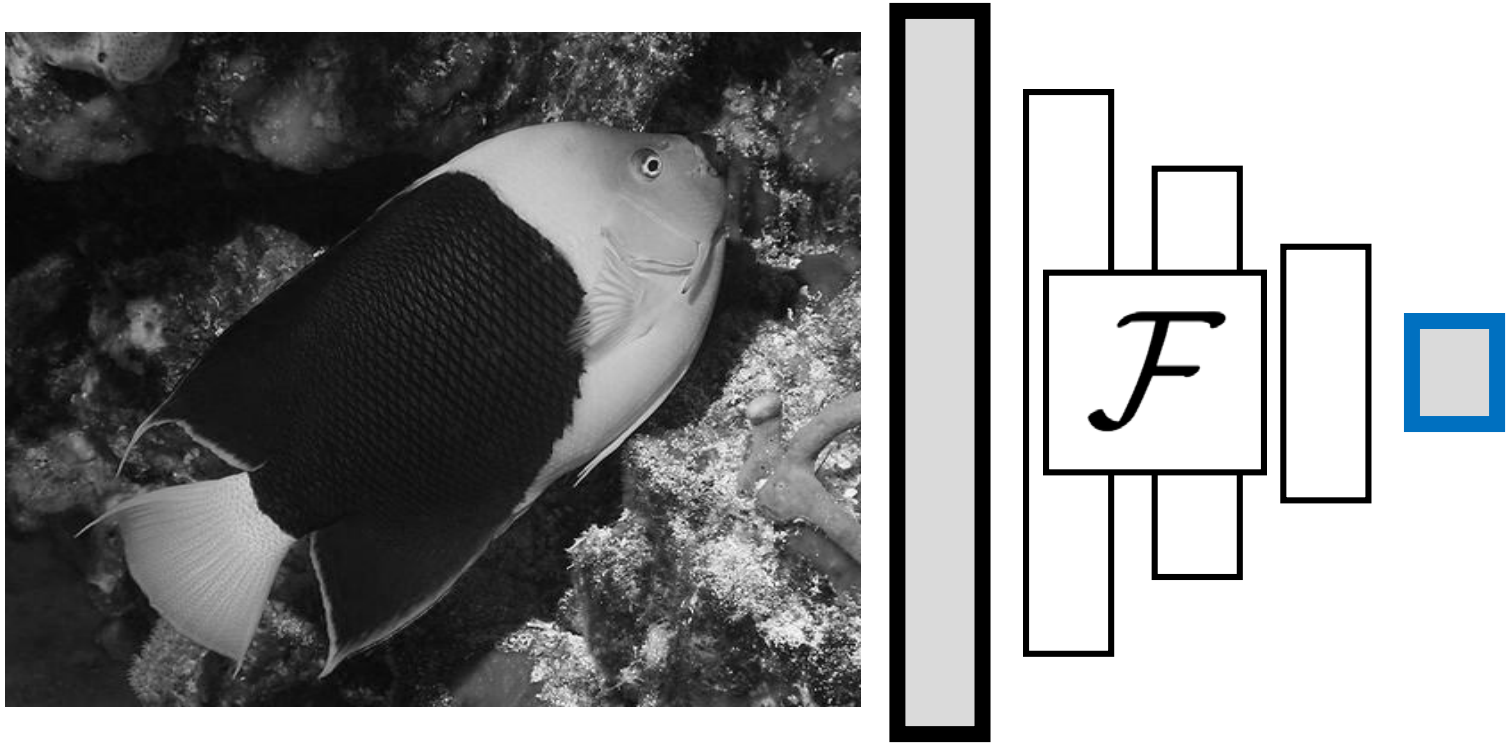
Discriminative Deep Networks



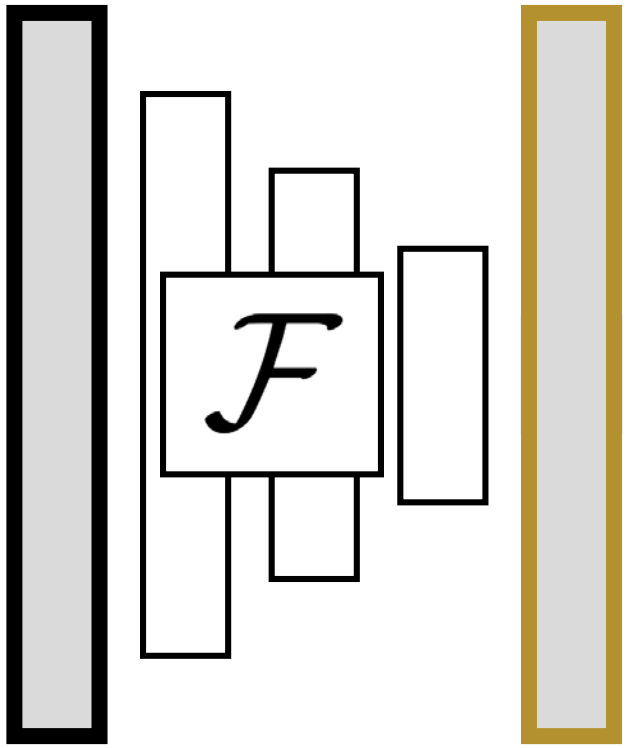
“Rockfish”

Discriminative Deep Networks

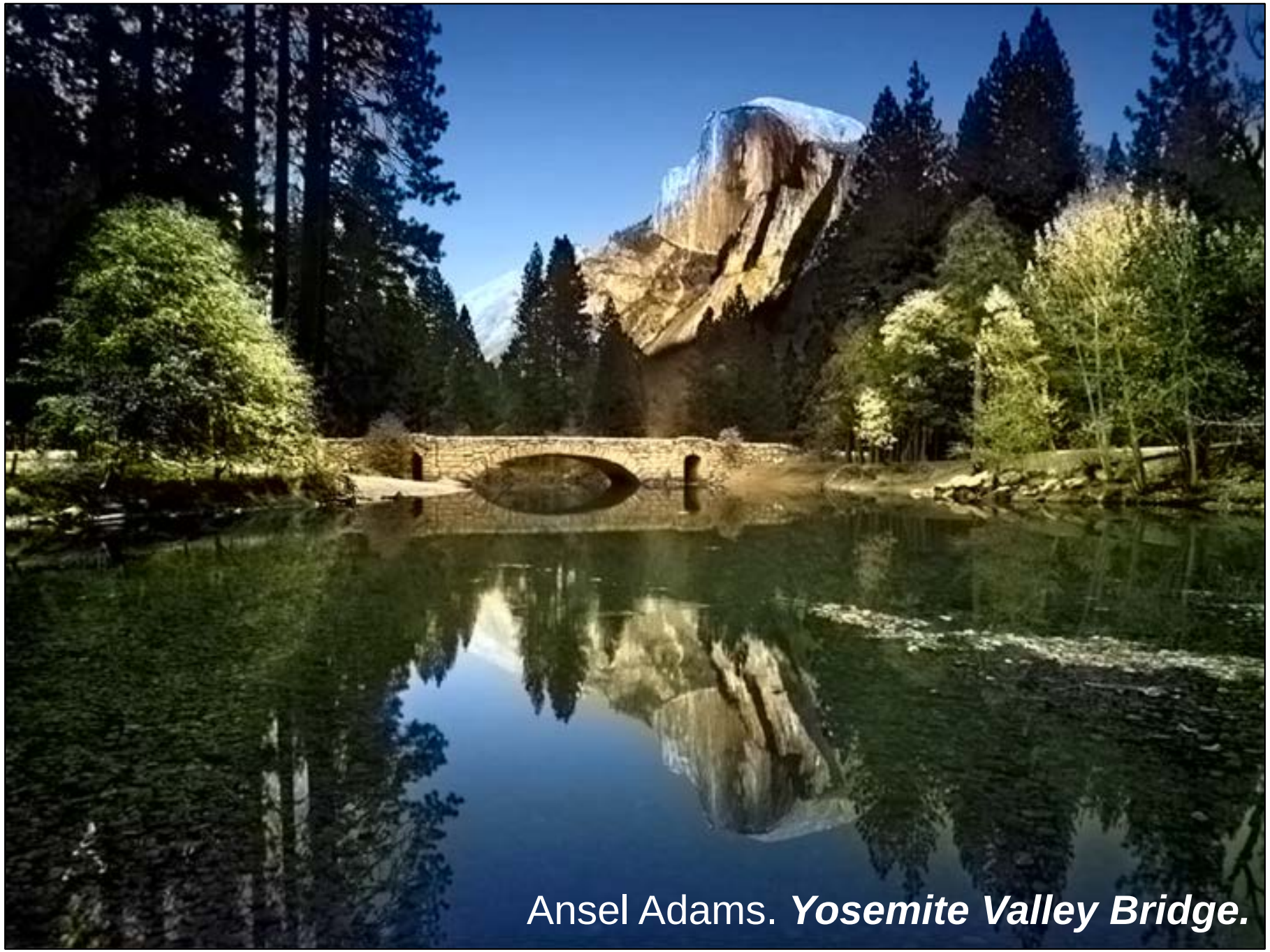


Raw, Unlabeled Pixels

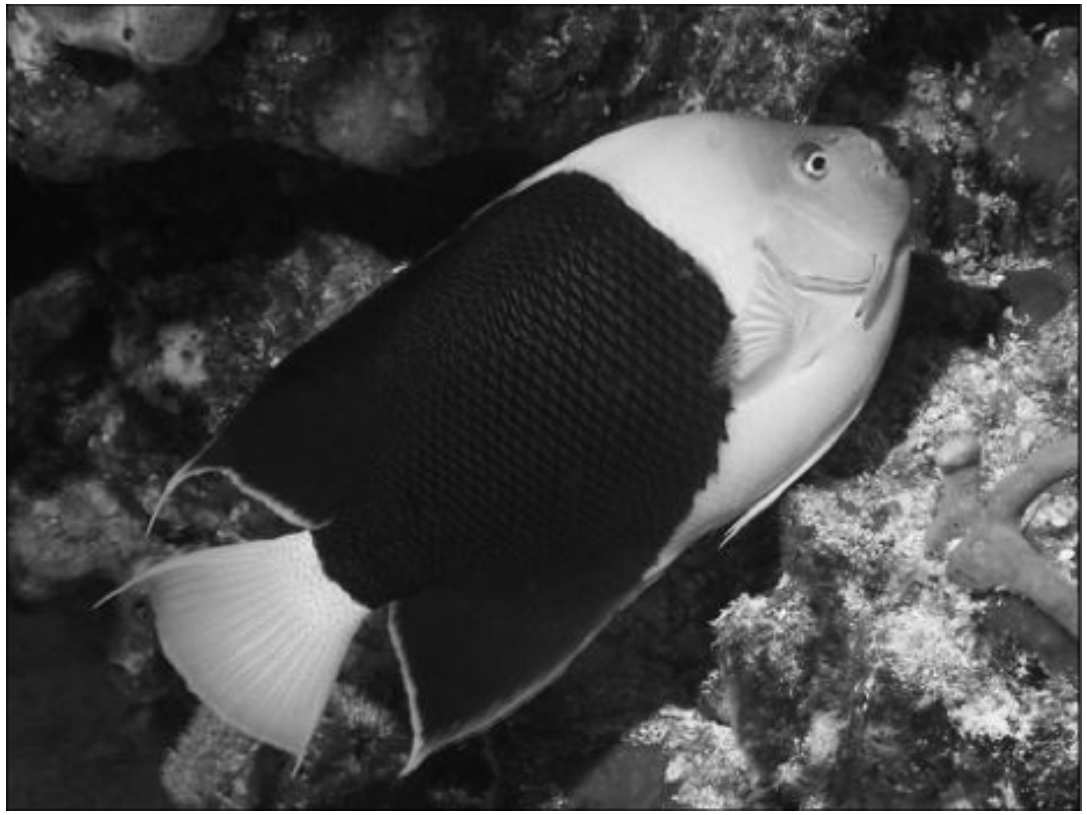
Generative Deep Networks



Raw, Unlabeled Pixels



Ansel Adams. ***Yosemite Valley Bridge.***



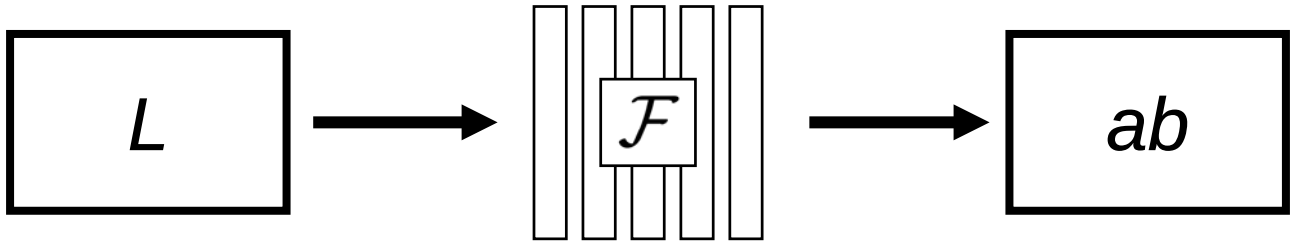
$\mathcal{F}$

Grayscale image: *L* channel

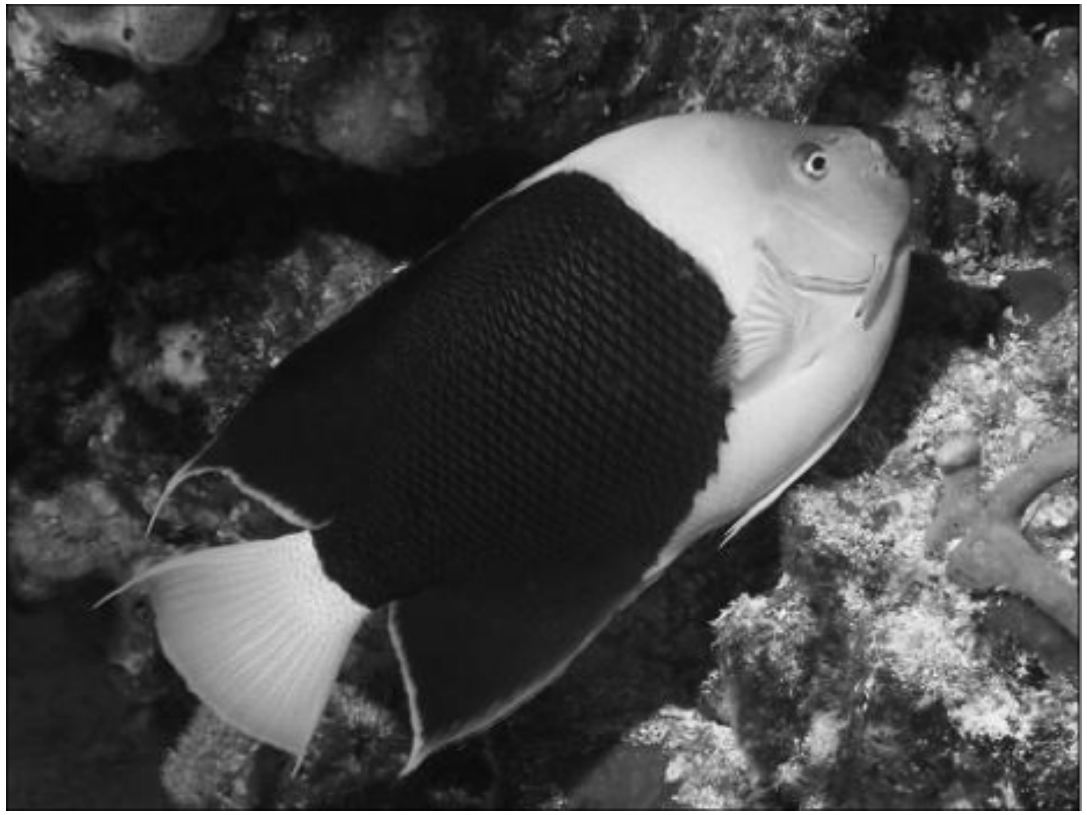
$$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$$

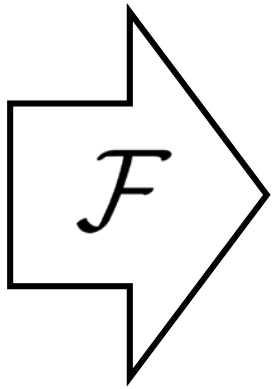
Color information: *ab* channels

$$\hat{\mathbf{Y}} \in \mathbb{R}^{H \times W \times 2}$$



Zhang, Isola, Efros. ***Colorful Image Colorization.*** In *ECCV*, 2016.



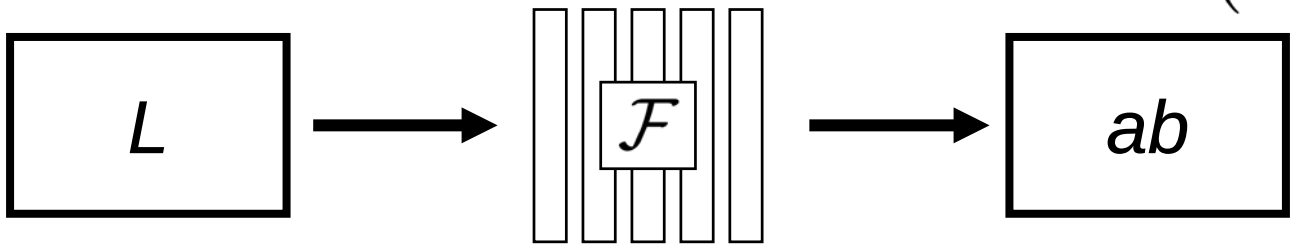


Grayscale image: *L* channel

$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$

Concatenate (*L*,*ab*) channels

$(\mathbf{X}, \widehat{\mathbf{Y}})$



Zhang, Isola, Efros. ***Colorful Image Colorization.*** In *ECCV*, 2016.

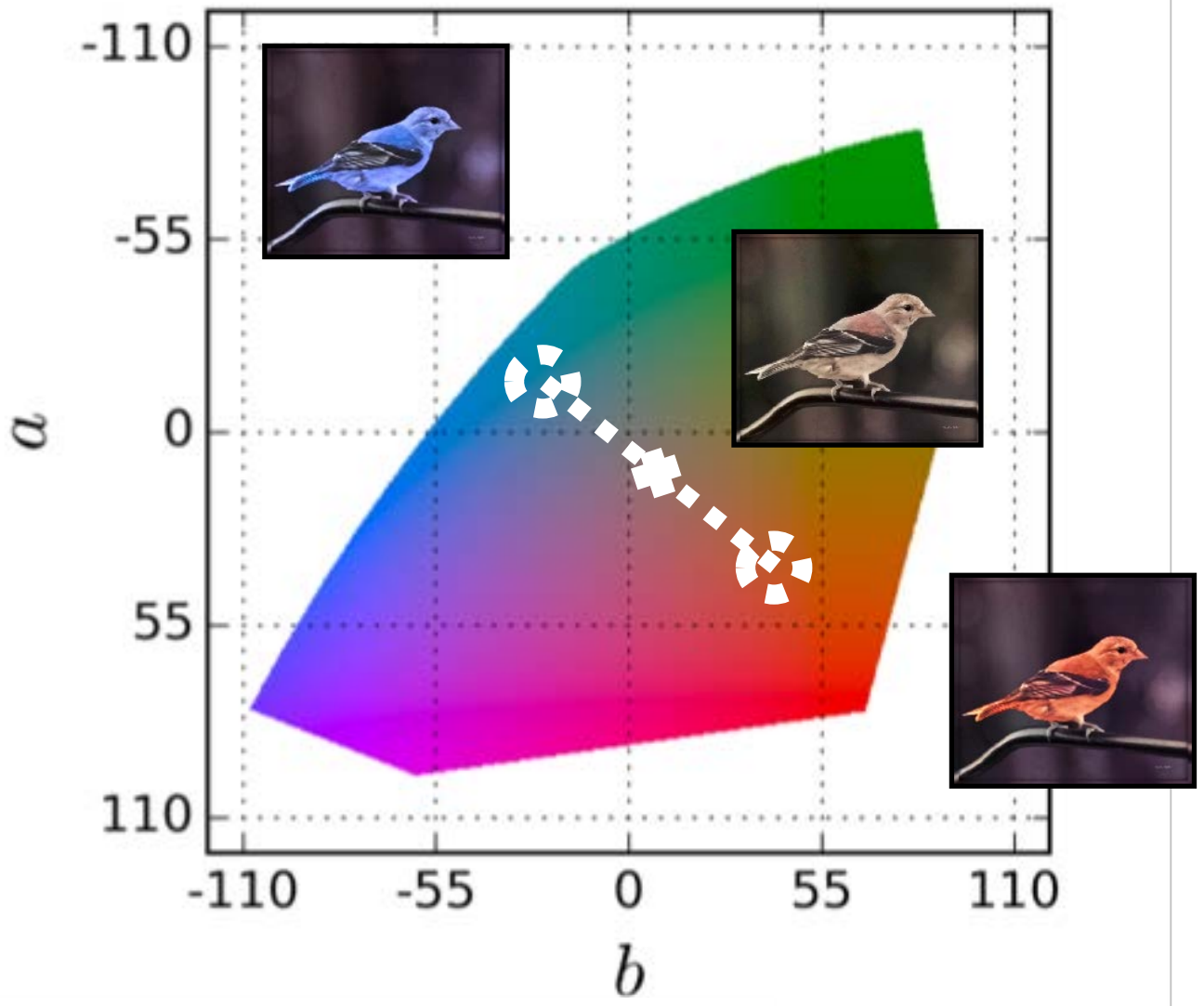
Simple L2 regression doesn’t work ☹

Input

Output

Ground truth

$$\mathrm{L}_2(\widehat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2}\sum_{h,w} \|\mathbf{Y}_{h,w} - \widehat{\mathbf{Y}}_{h,w}\|_2^2$$



$$\mathrm{L}_2(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2}\sum_{h,w} \|\mathbf{Y}_{h,w} - \hat{\mathbf{Y}}_{h,w}\|_2^2$$

Better Loss Function

$$\theta^* = \arg\min_{\theta} \ell(\mathcal{F}_\theta(\mathbf{X}), \mathbf{Y})$$

- Regression with L2 loss inadequate

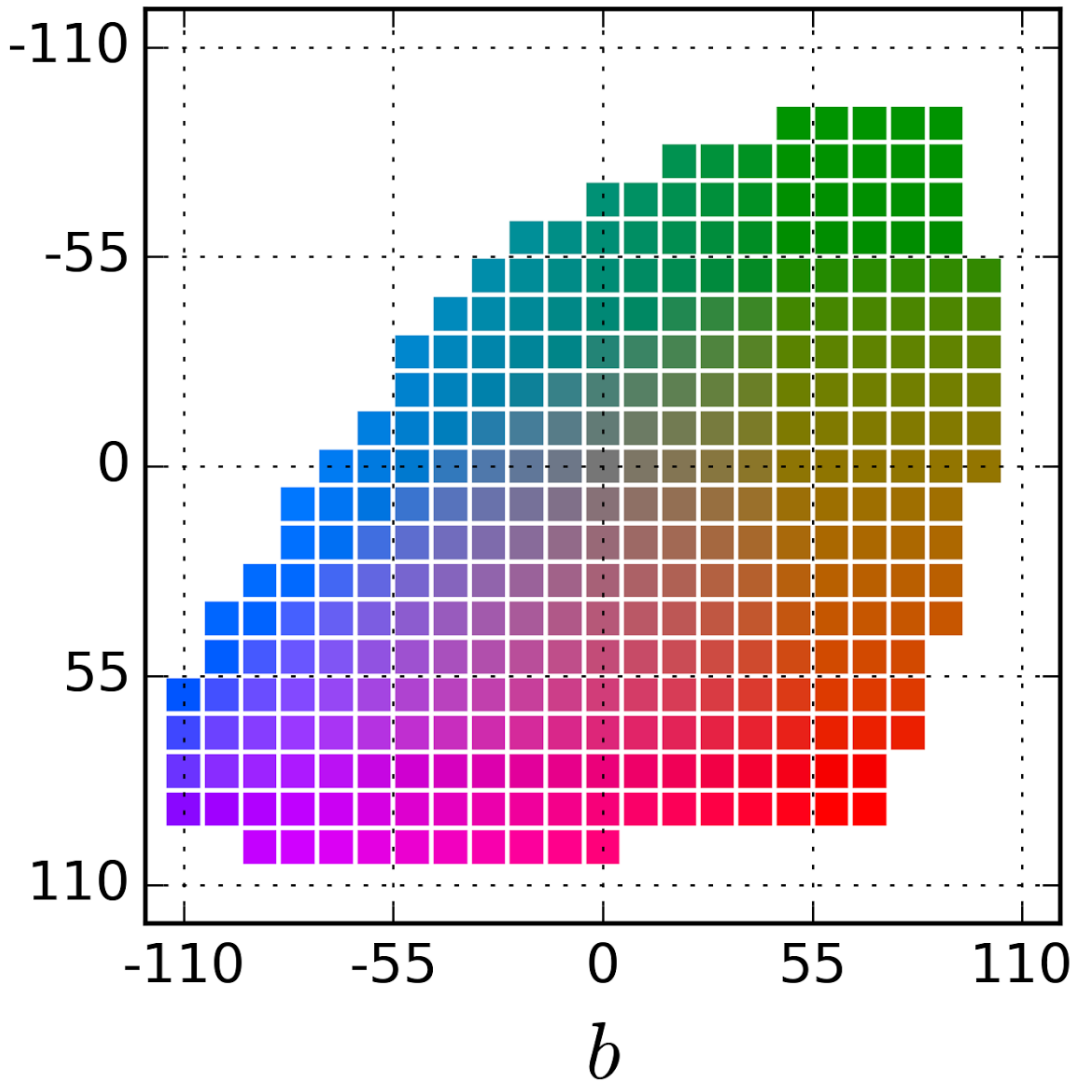
$$\mathrm{L}_2(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2}\sum_{h,w} \|\mathbf{Y}_{h,w} - \hat{\mathbf{Y}}_{h,w}\|_2^2$$

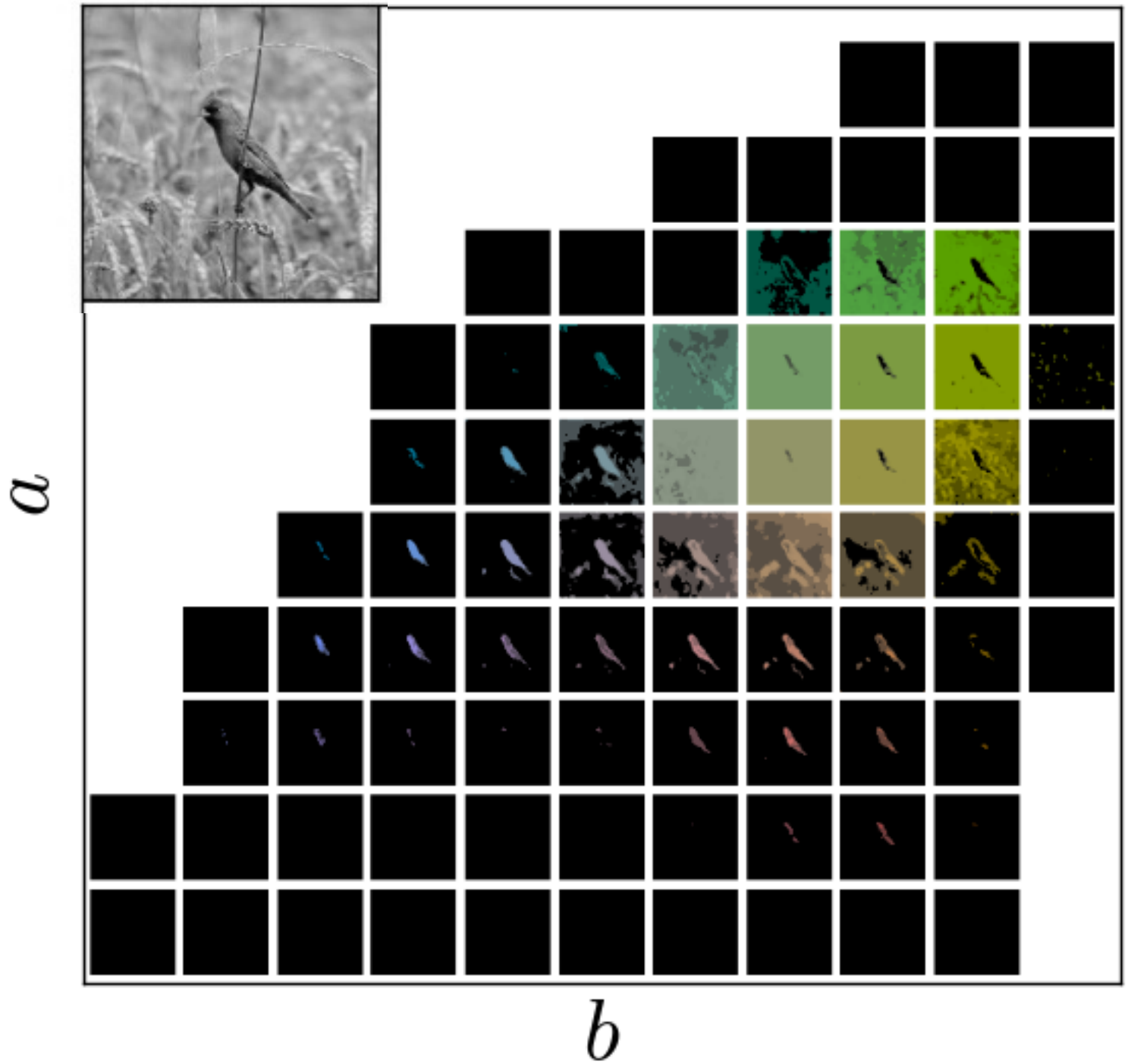
- Use per-pixel multinomial classification

$$\mathrm{L}(\hat{\mathbf{Z}}, \mathbf{Z}) = -\frac{1}{HW}\sum_{h,w}\sum_{q} \mathbf{Z}_{h,w,q} \log(\hat{\mathbf{Z}}_{h,w,q})$$

Colors in *ab* space

(discrete)




a
b

Designing loss functions

Input

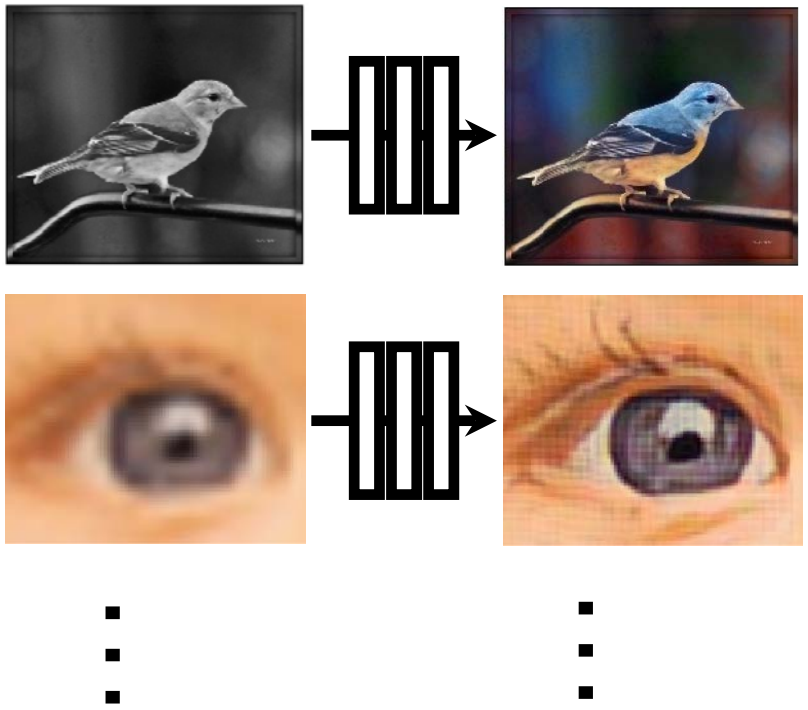
Zhang et al. 2016

Ground truth

Color distribution cross-entropy loss with colorfulness enhancing term.

[Zhang, Isola, Efros, ECCV 2016]

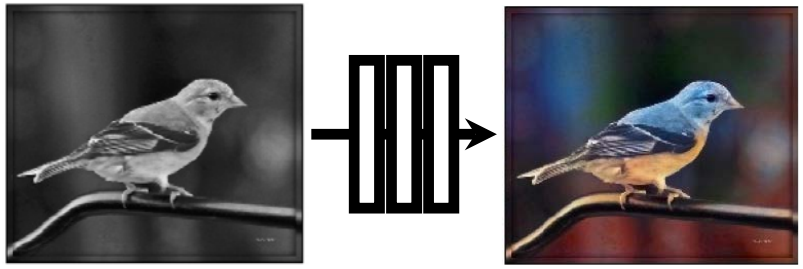

DADS

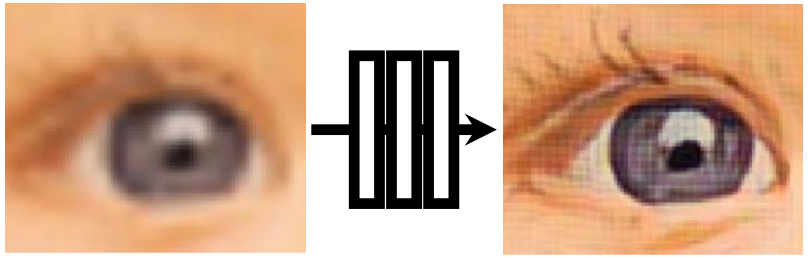


Universal loss?

Generative Adversarial Network (GANs)

Generated images



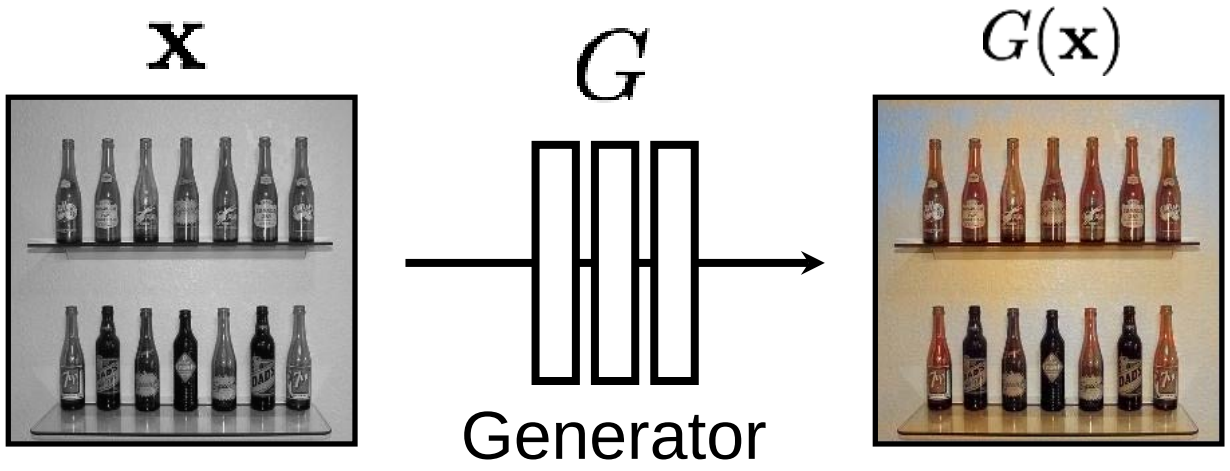


Generated vs Real (classifier)

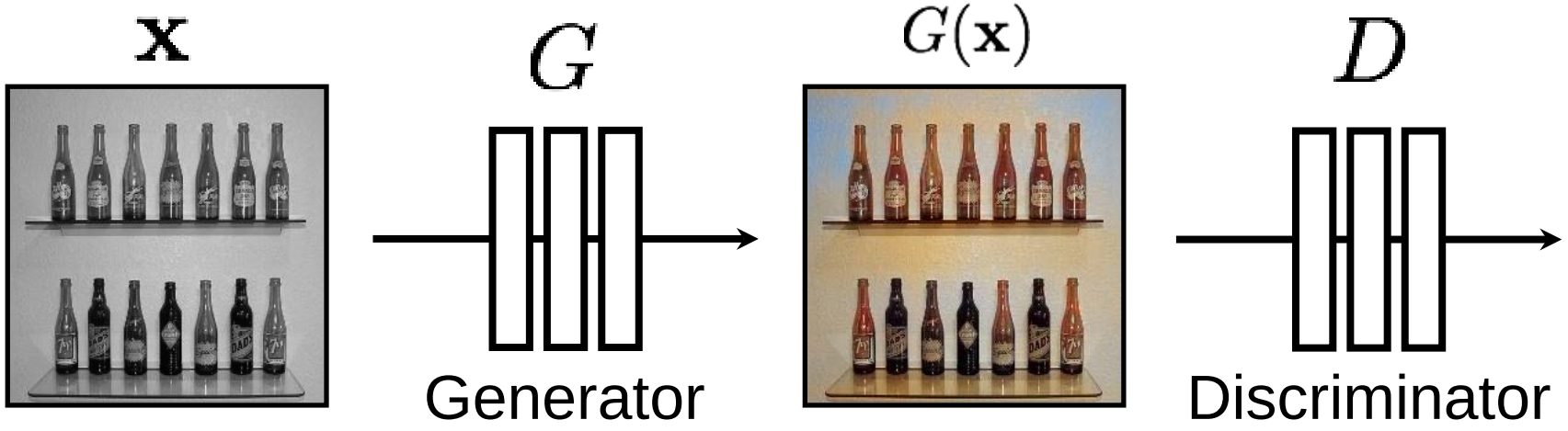
Real photos

…

[Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, Bengio 2014]



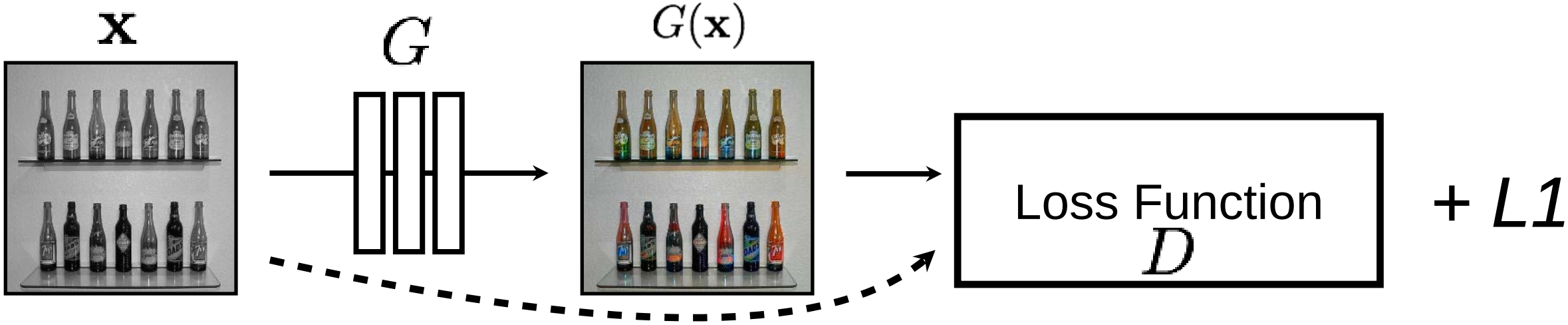
[Goodfellow et al., 2014]



G tries to synthesize fake images that fool D

D tries to identify the fakes

[Goodfellow et al., 2014]



G's perspective: **D** is a loss function.

Rather than being hand-designed, it is *learned*.

[Goodfellow et al., 2014]
[Isola et al., 2017]

BW → Color






Data from [Russakovsky et al. 2015]

BW → Color






Data from [Russakovsky et al. 2015]

Input

Output

Groundtruth

Data from
[maps.google.com]

Input

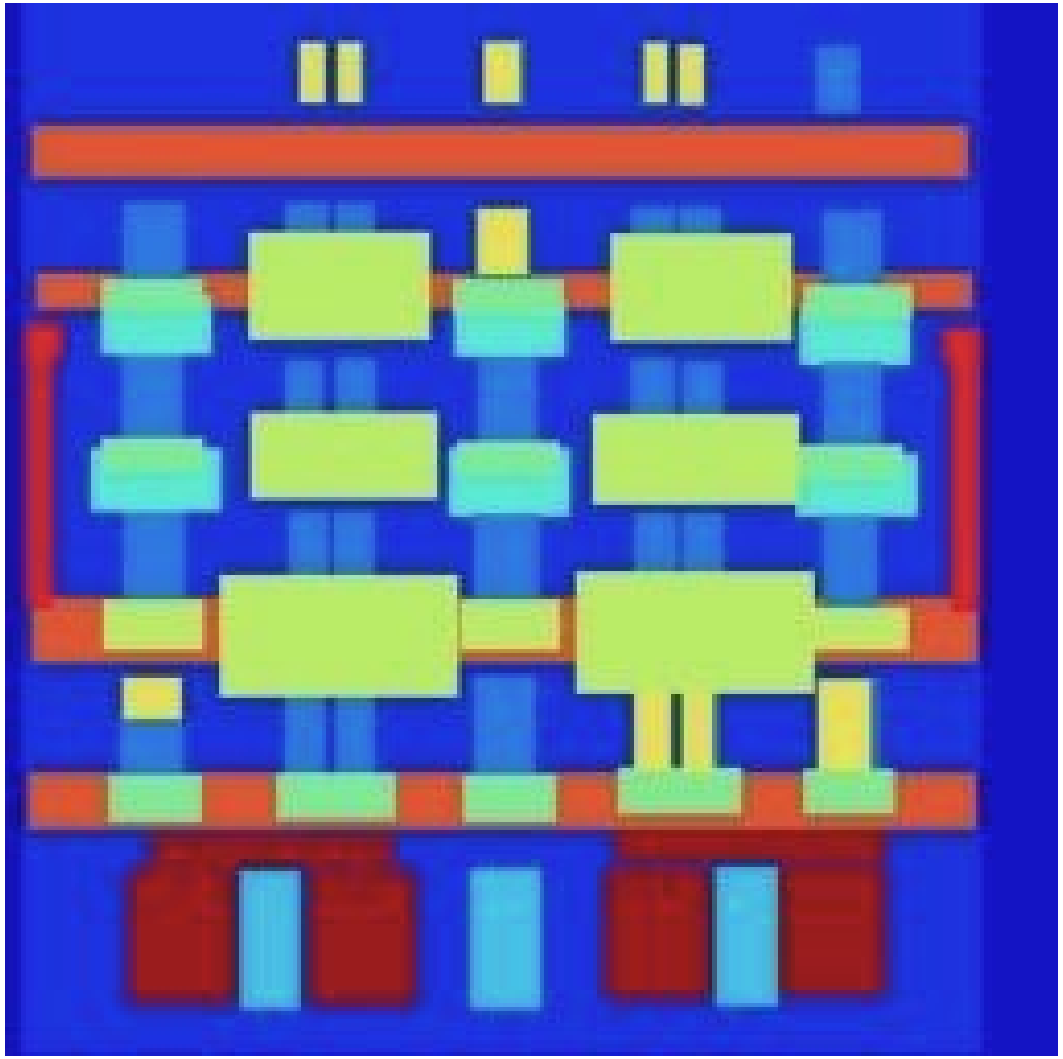
Output

Groundtruth

Data from [maps.google

Labels → Facades

Input

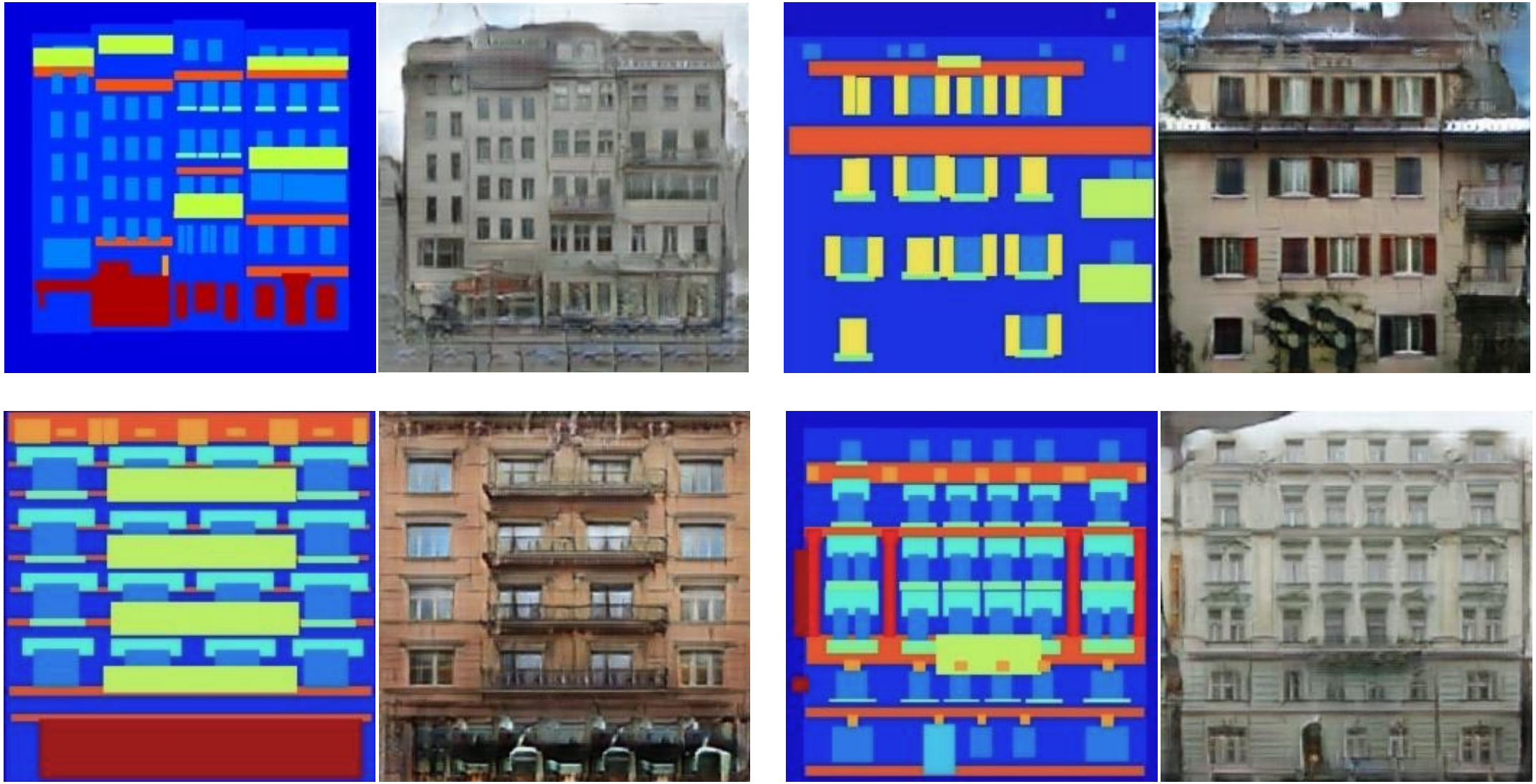


Output

Data from [Tylecek, 2013]

Labels → Facades

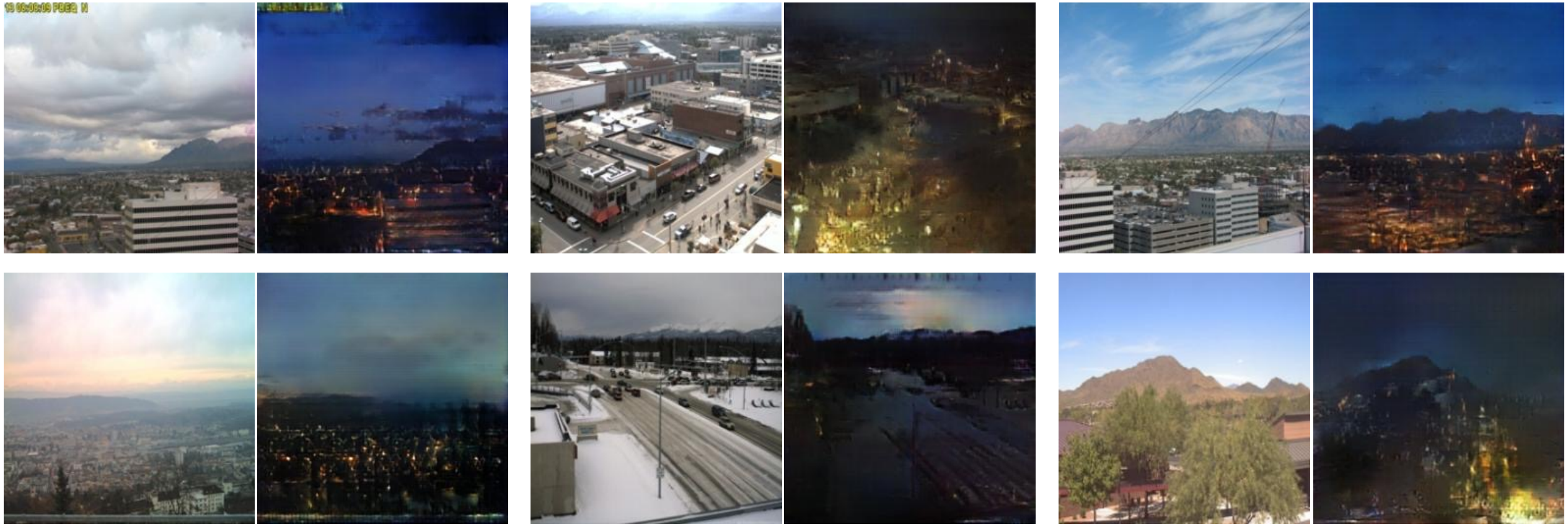
Input Output Input Output



Data from [Tylecek, 2013]

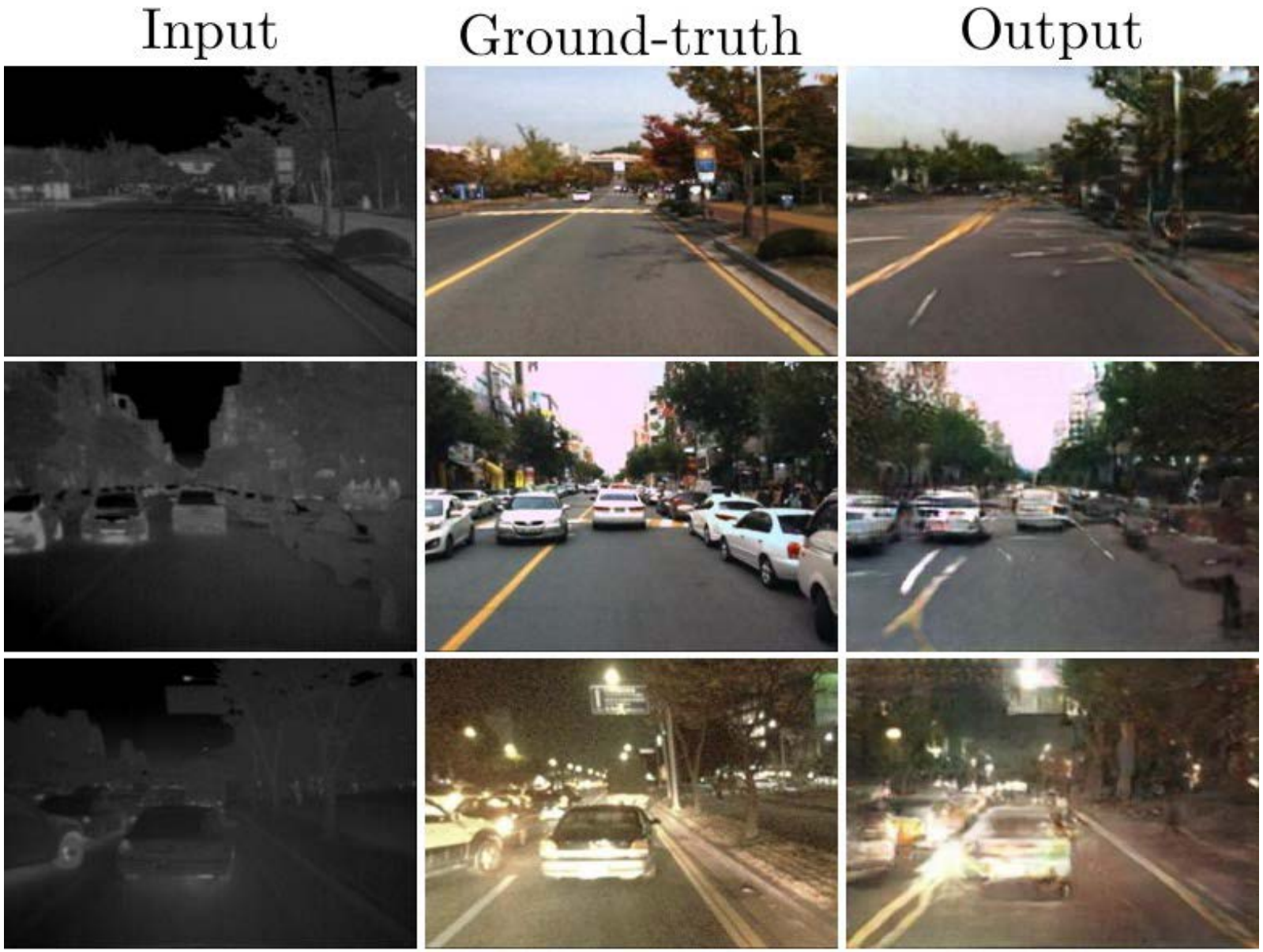
Day → Night

Input Output Input Output Input Output



Data from [Laffont et al., 2014]

Thermal → RGB



Edges → Images

Input Output Input Output Input Output

Edges from [Xie & Tu, 2015]

Sketches → Images

Input Output Input Output Input Output

Trained on Edges → Images

Data from [Eitz, Hays, Alexa, 2012]

junyanz / **pytorch-CycleGAN-and-pix2pix**

Watch 176 | Star 4,351 | Fork 982

Code | Issues 28 | Pull requests 3 | Projects 0 | Insights

Image-to-image translation in PyTorch (e.g., horse2zebra, edges2cats, and more)

pytorch gan cyclegan pix2pix deep-learning computer-vision computer-graphics image-manipulation image-generation generative-adversarial-network gans

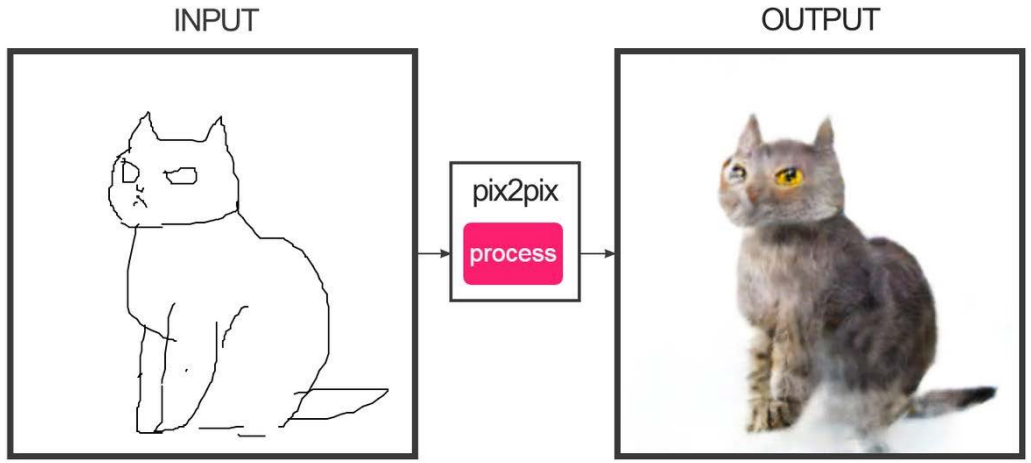
223 commits | 3 branches | 0 releases | 26 contributors

Branch: master | New pull request | Find file | Clone or download

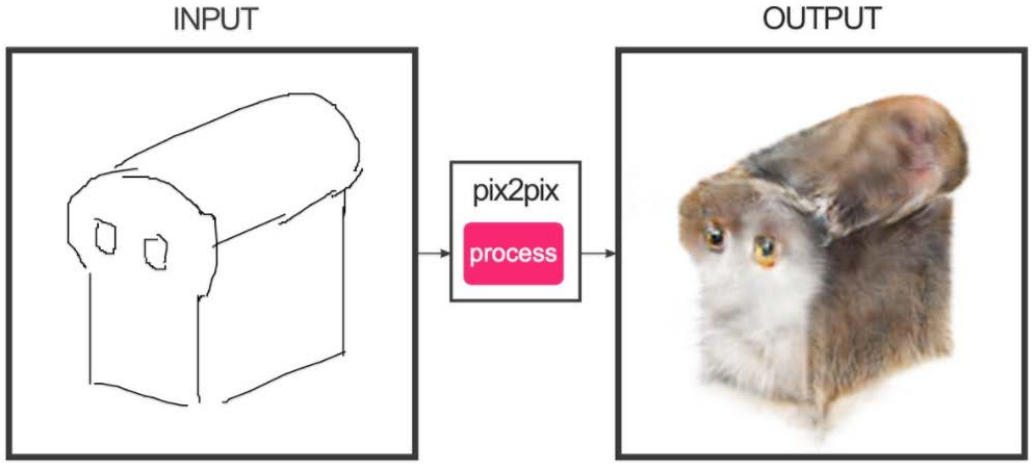
taesung89 Update README.md — Latest commit 6d9e173 10 Jun 13, 2018, 12:04 AM PDT

data	1. datasets are now configured automatically based on dataset_mode op...	10 days ago
datasets	Multiple changes regarding option management. See below.	15 days ago
imgs	add edges2cats demo	a year ago
models	TestModel now supports model_suffix option that can change the name o...	10 days ago

#edges2cats [Christopher Hesse]



@gods_tail

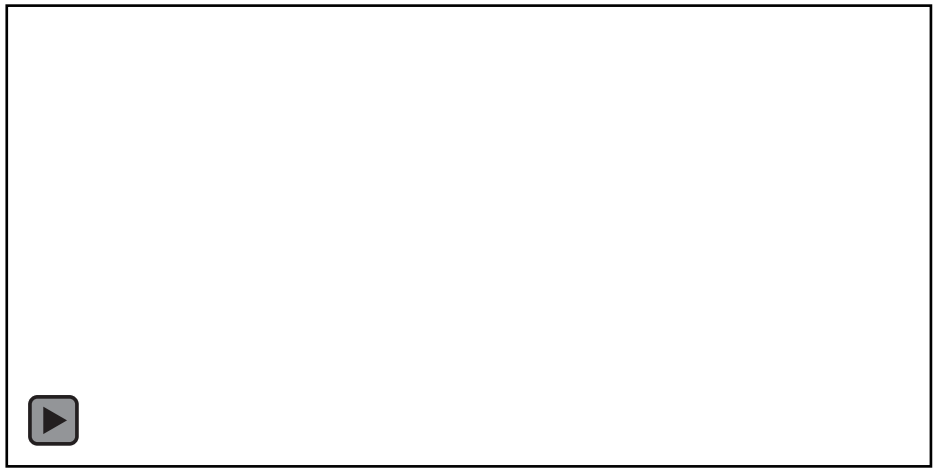


Ivy Tasi @ivymyt

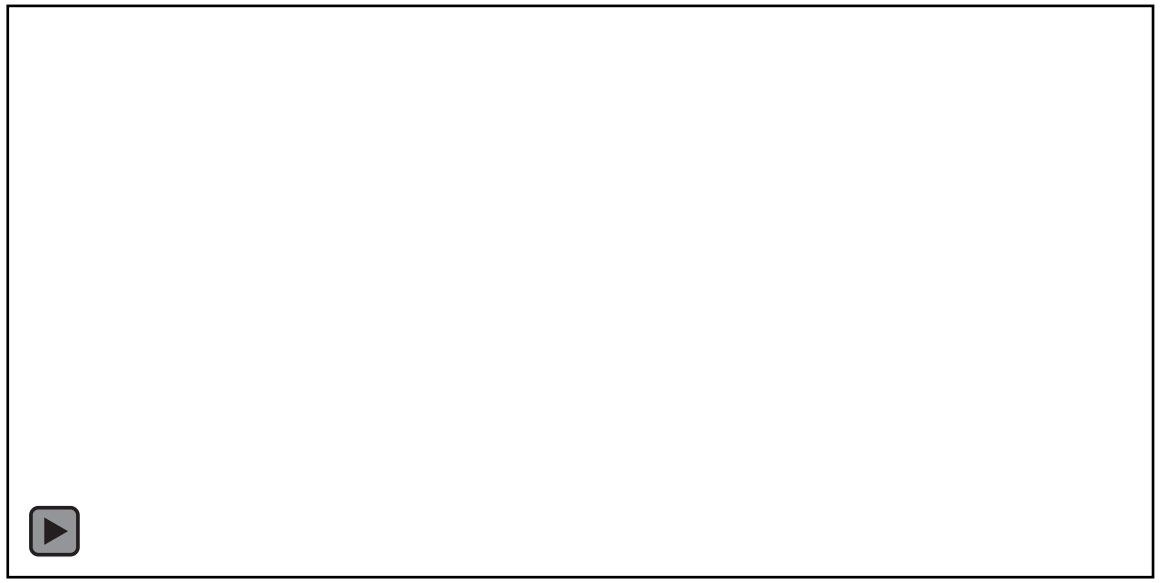
Vitaly Vidmirov @vvid

@ka92

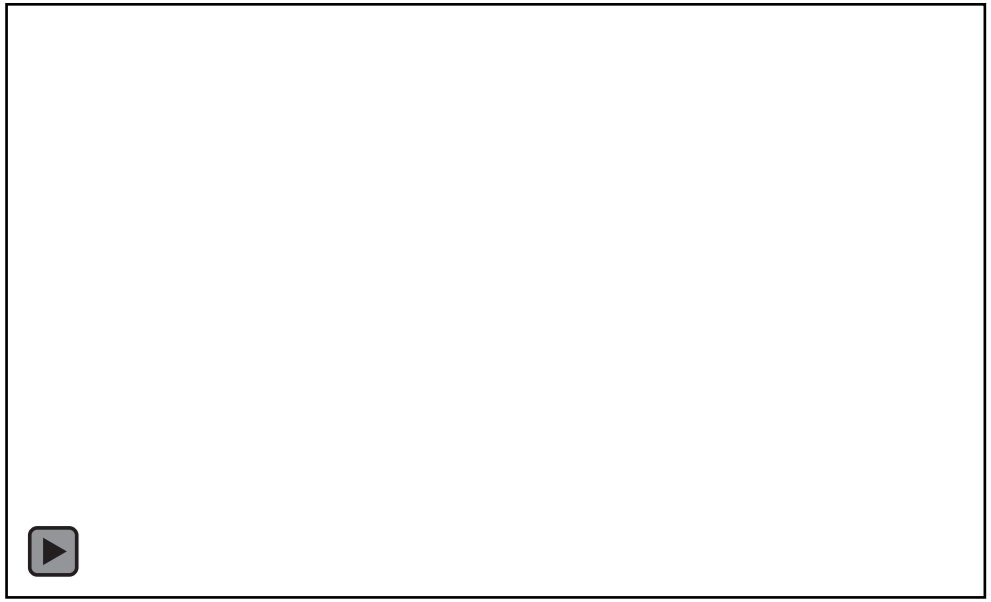
Twitter-driven research: #pix2pix



Mario Klingemann @quasimondo



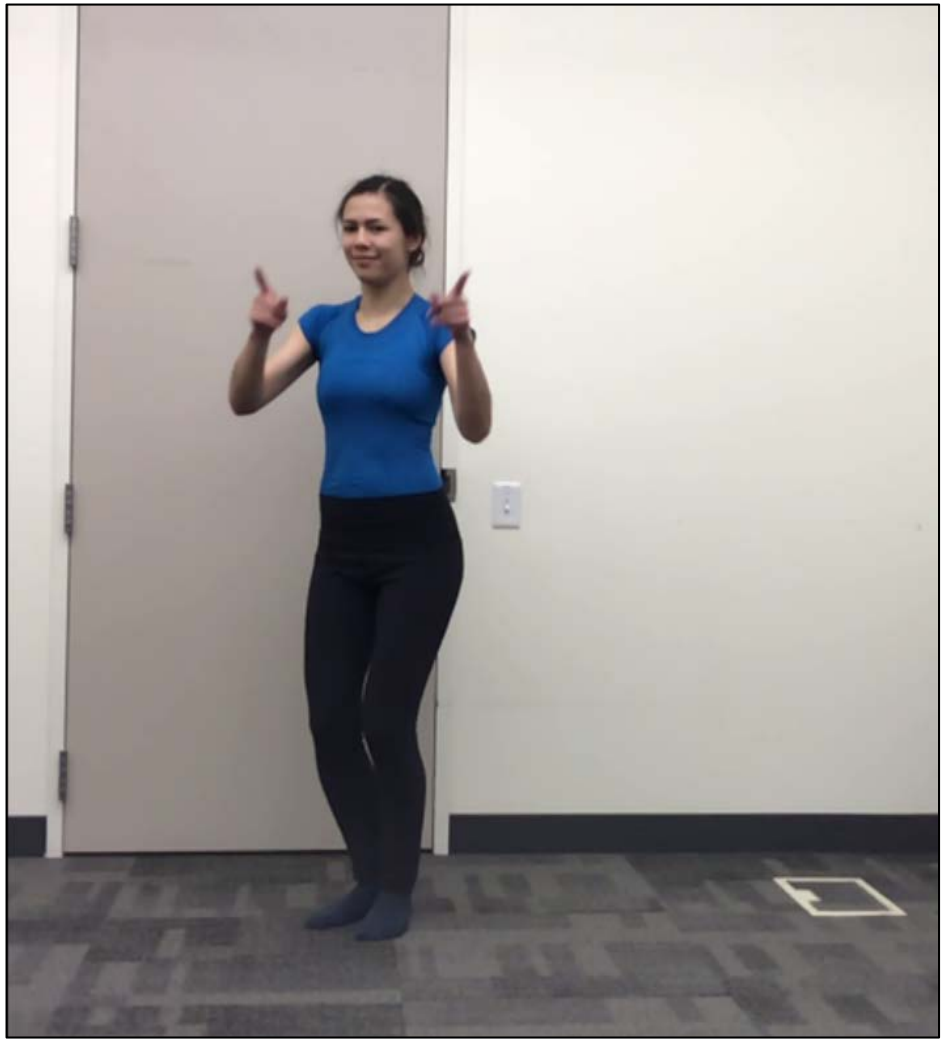
Brannon Dorsey @brannondorsey



Bertrand Gondouin @bgondouin

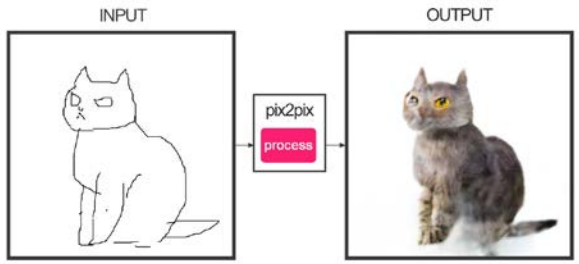
© **Memo Akten,** “Learning to See: Gloomy Sunday”

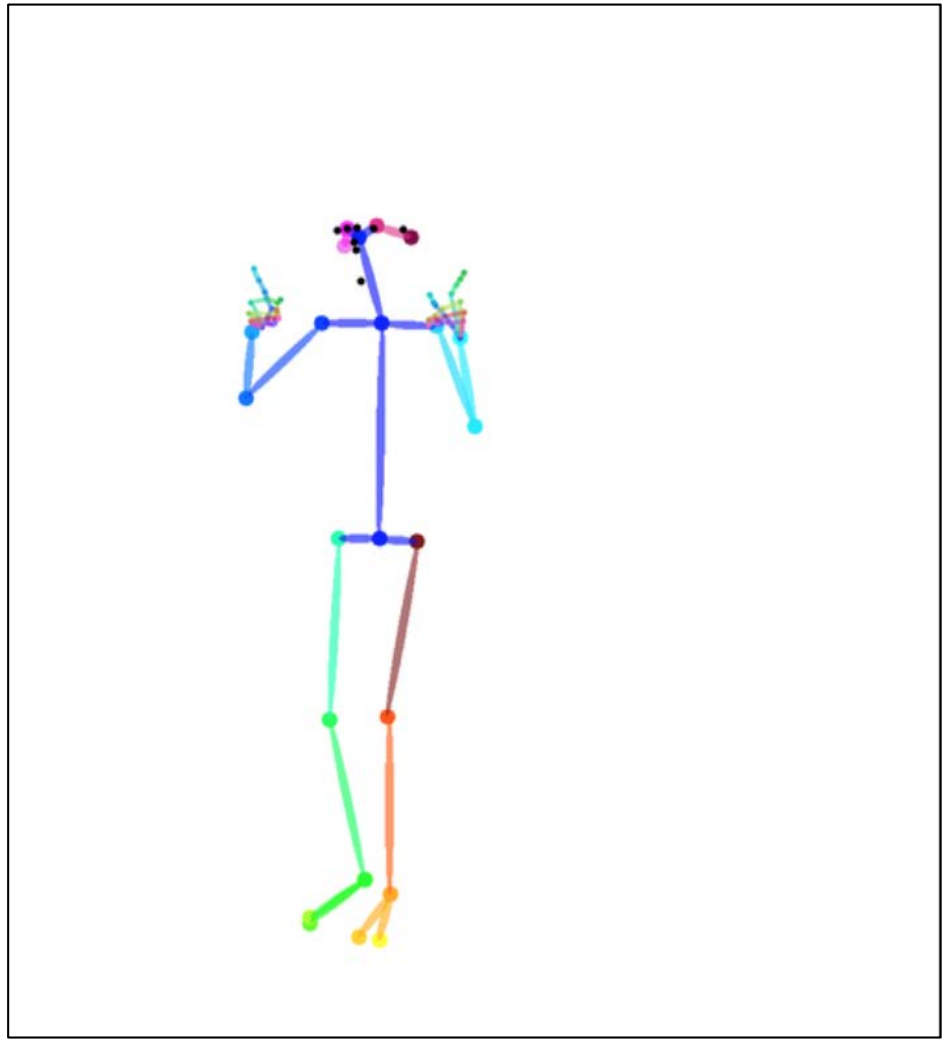
“Do as I Do”



OpenPose

pix2pix





Everybody Dance Now

Caroline Chan, Shiry Ginosar, Tinghui Zhou, Alexei A. Efros
UC Berkeley

Source Subject

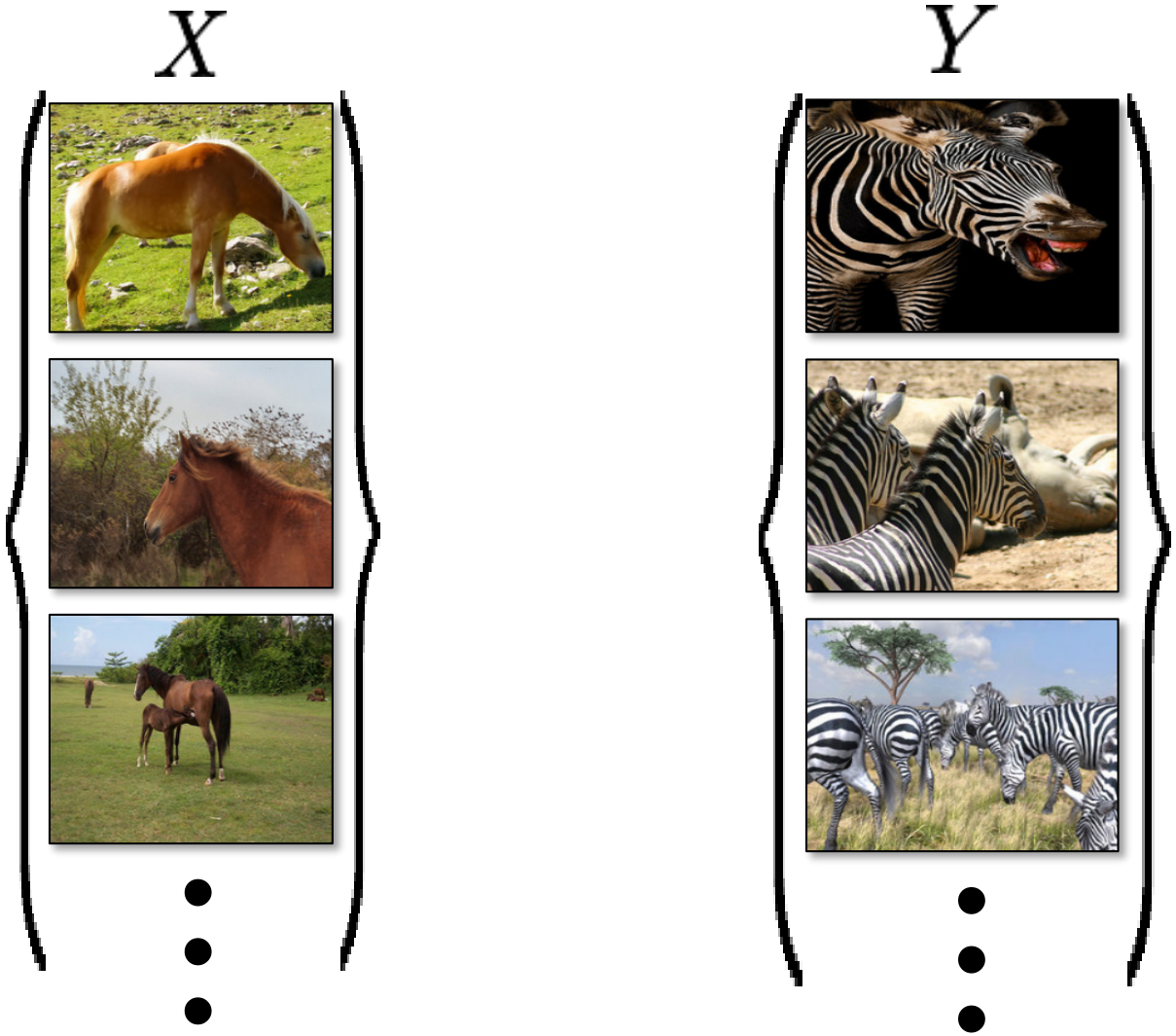
Target Subject

Results

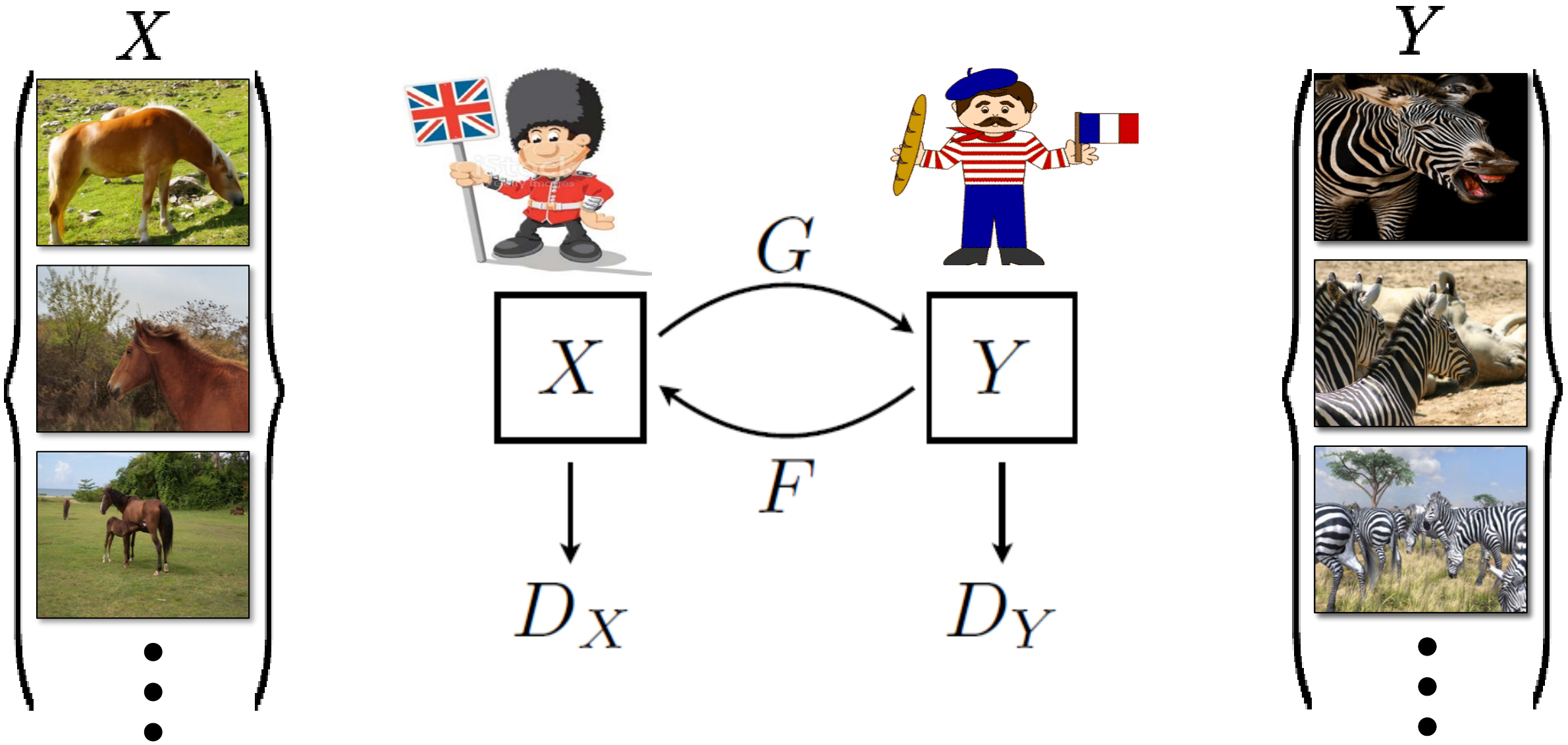
https://www.youtube.com/watch?v=PCBTZh41Ris&feature=youtu.be

Paired training examples

Unpaired training examples

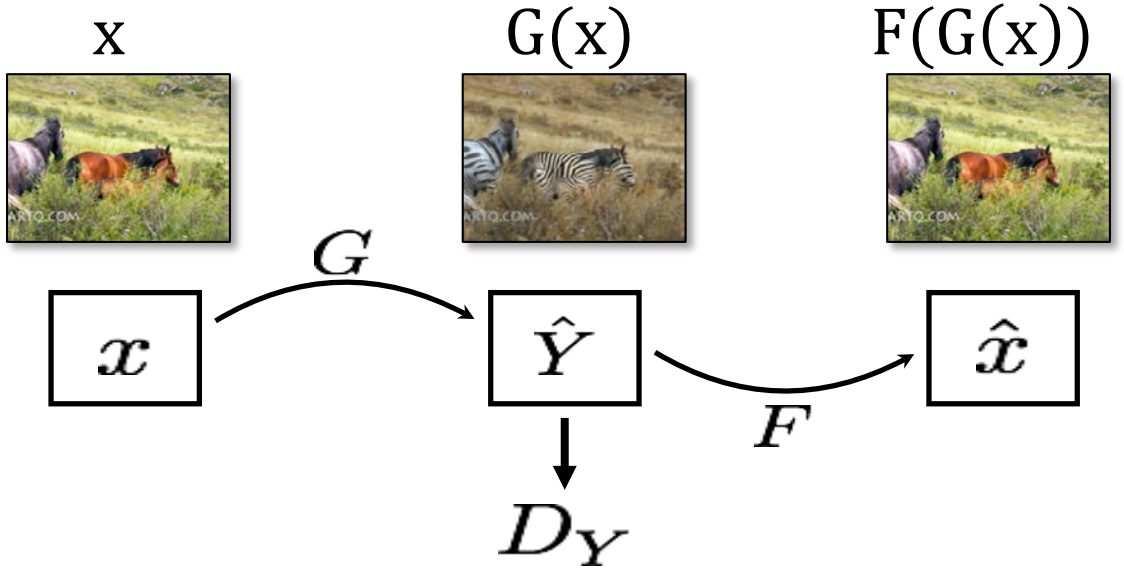


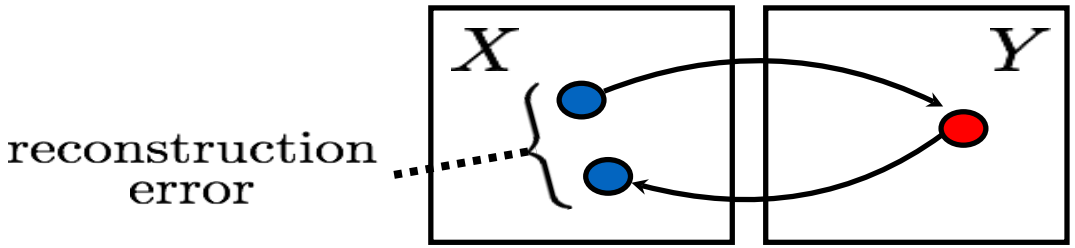
CycleGAN, or “there and back aGAN”



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

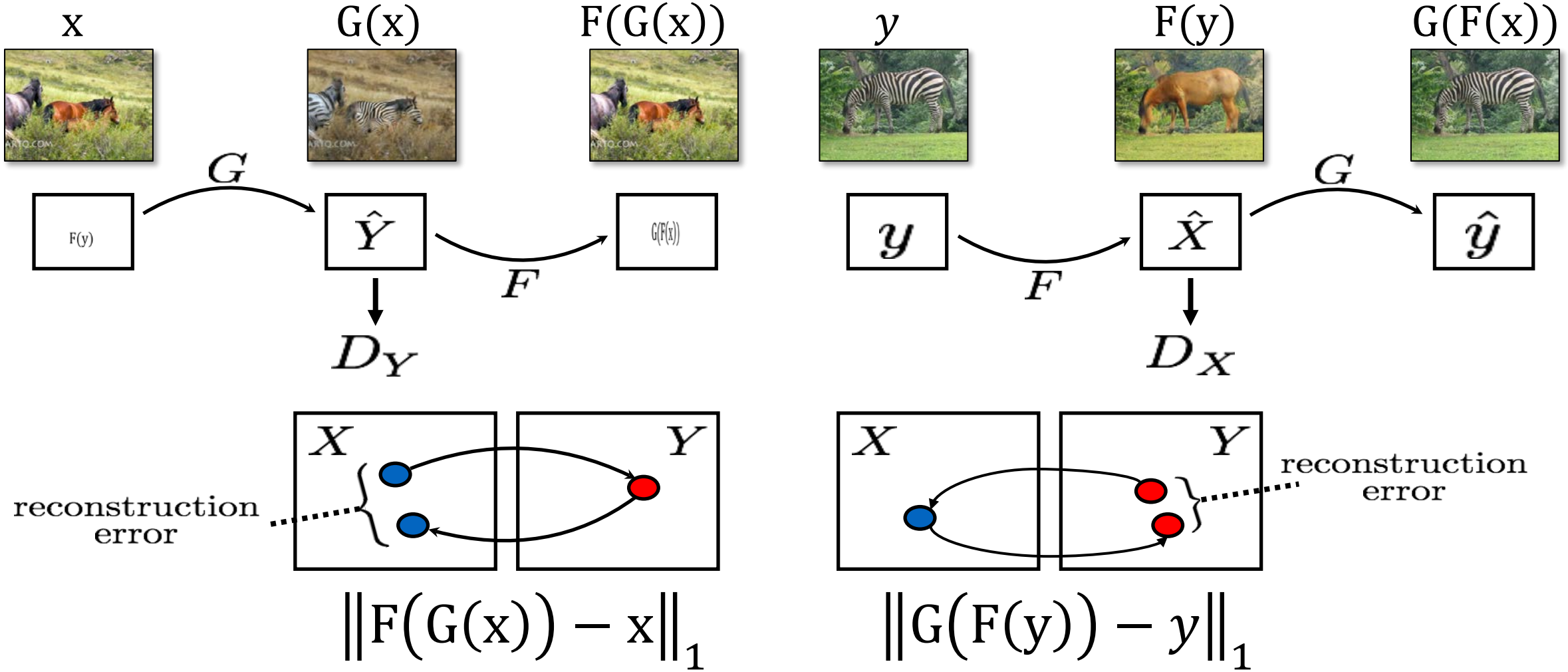
Cycle-Consistency Loss





$$\|F(G(x)) - x\|_1$$

Cycle-Consistency Loss



Video

Collection Style Transfer

Photograph

Monet

Van Gogh

Cezanne

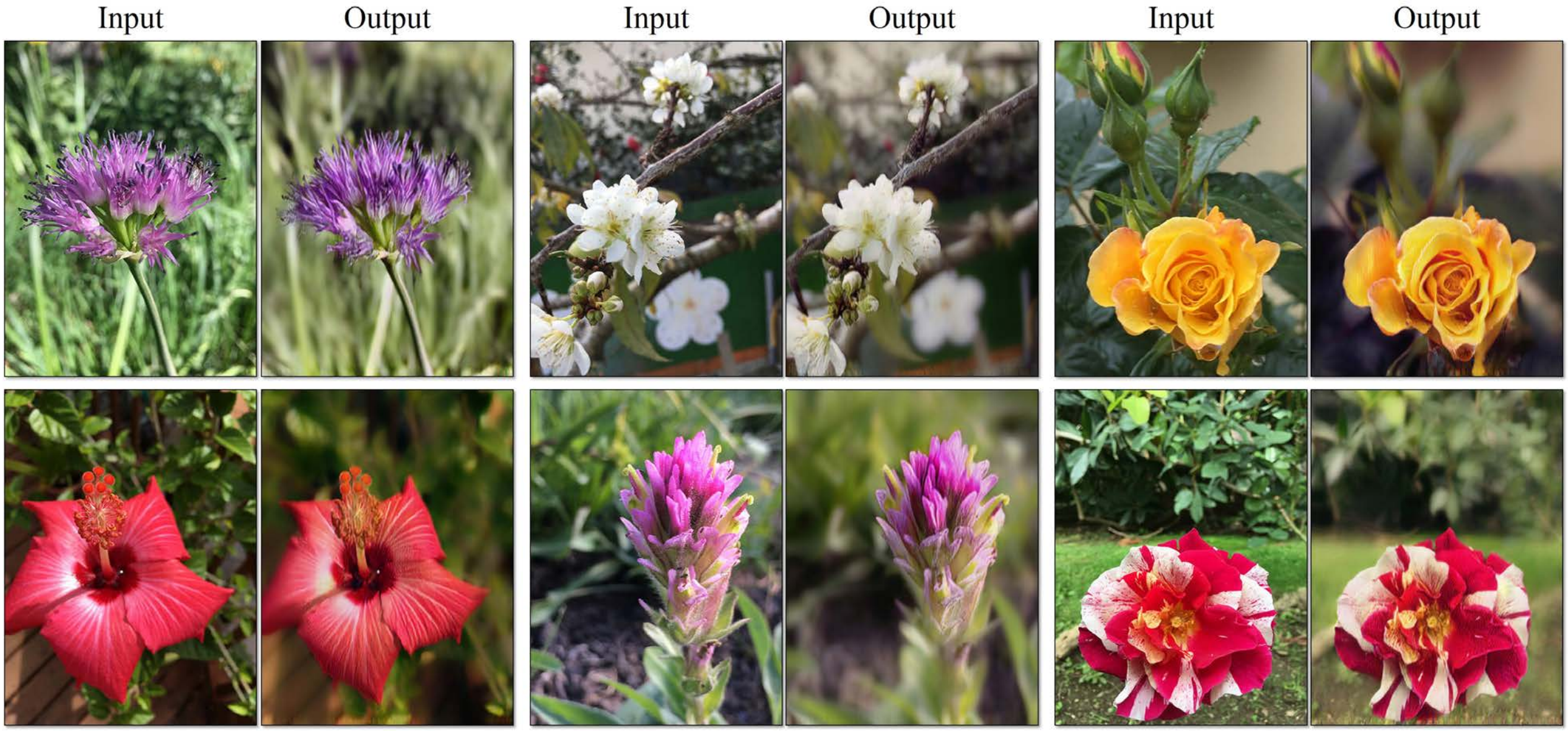
Ukiyo-e

CG to Real

Grand Theft Auto

Real to CG

Shallower depth of field



Failure case