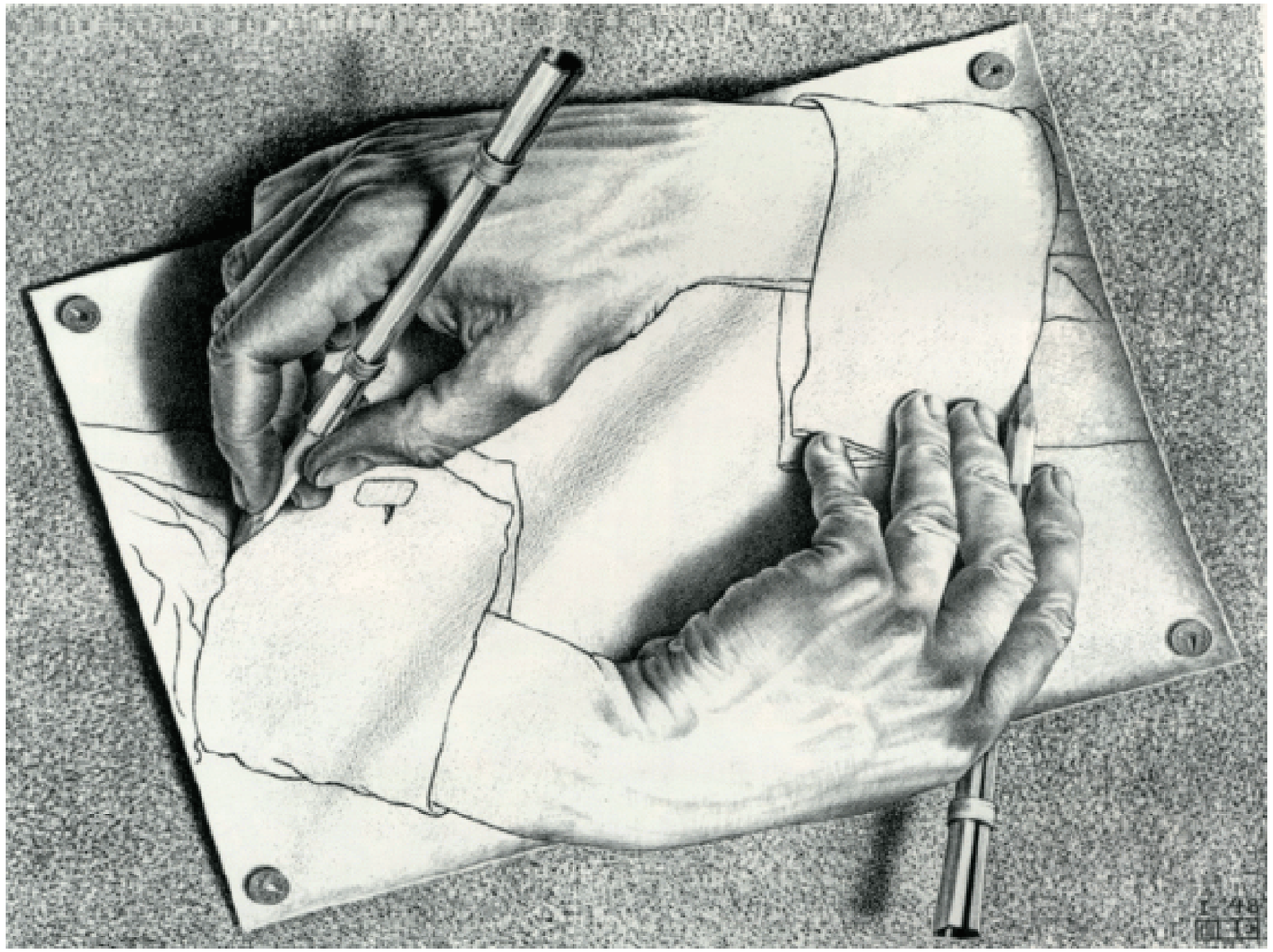
Reasons to Love GANs



Alexei (Alyosha) Efros
UC Berkeley

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Reasons to hate GANs

- They are a devil to train!
  - The discriminator nearly always wins
  - Sometimes, training longer makes it worse
  - Sometimes, more data doesn’t make it better
  - You need to boil the ocean to train the best ones
- For any given problem, application-tailored solutions might work better
  - e.g. “perceptual losses” (Gatys, Johnson, etc)
- Do they really generate a distribution?
- Do they span a subspace?



Karras et al, “*Progressive growing of GANs for improved quality, stability, and variation”, ICLR’18*



Karras et al, “*Progressive growing of GANs for improved quality, stability, and variation”, ICLR’18*

BigGAN: GANs grow up!



- Results looks very impressive!
- Discriminator is at chance!
  - At least on validation set
- Funny thing: technically no big changes. Just larger batch-size.
  - Data finally allowed to speak?

But are GANs better than non-parametric sampling???

- Facebook anecdote:
  - 40 Billion images
  - Use doubt L2 on pixels

My reasons to love GANs

1. GANs set up an arms race
2. GANs can be used as a “learned loss function”
3. GANs are “meta-supervisors”
4. GANs are great data memorizers
5. GANs generalize texture synthesis
6. GANs are democratizing computer art

1. Arms Races


ПАРИКМАХЕРСКАЯ





The Five-Year Plan



The Five-Year Plan



HP-80 (1973)



Elektronika DD (1968-1973)

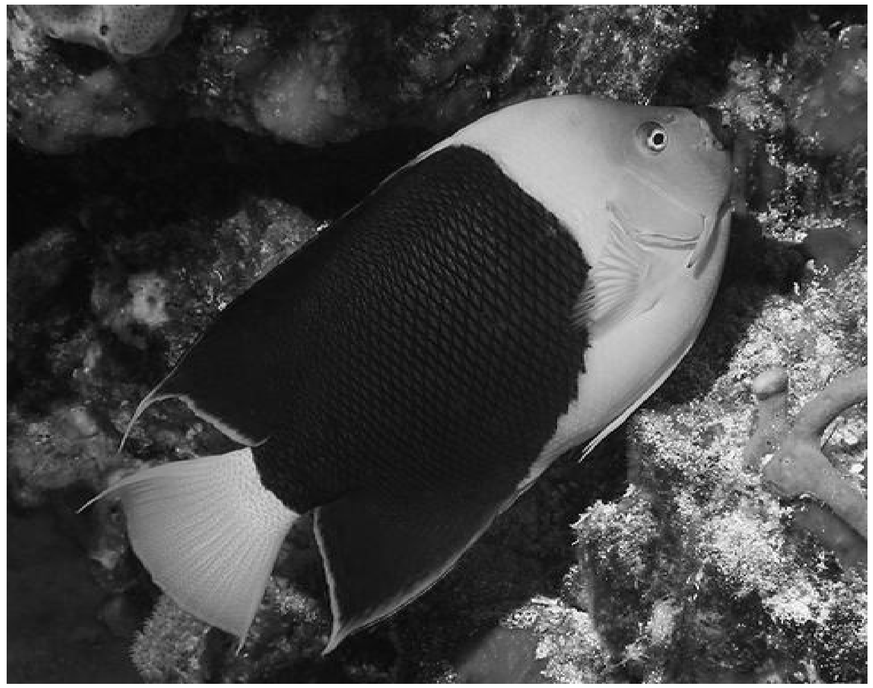
Arms races cause emergent properties

2. GANs as “learned loss functions”

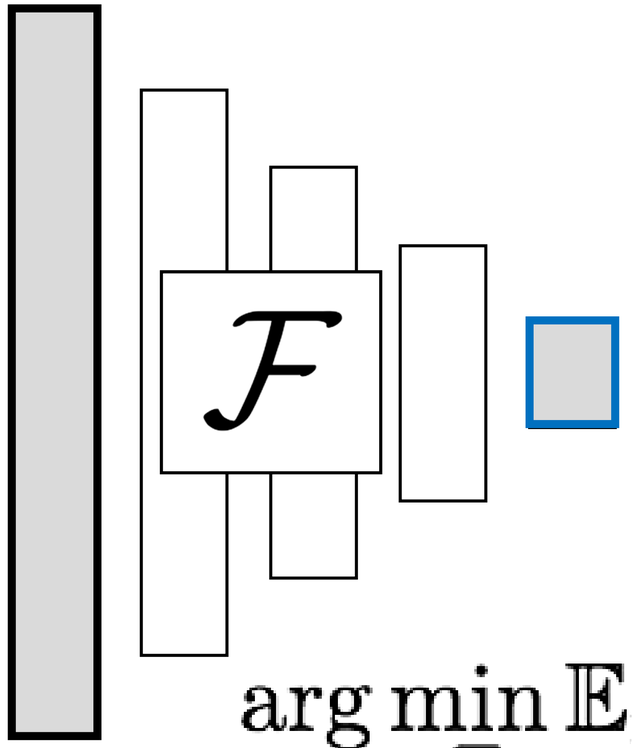


Discriminative Deep Networks

x



y



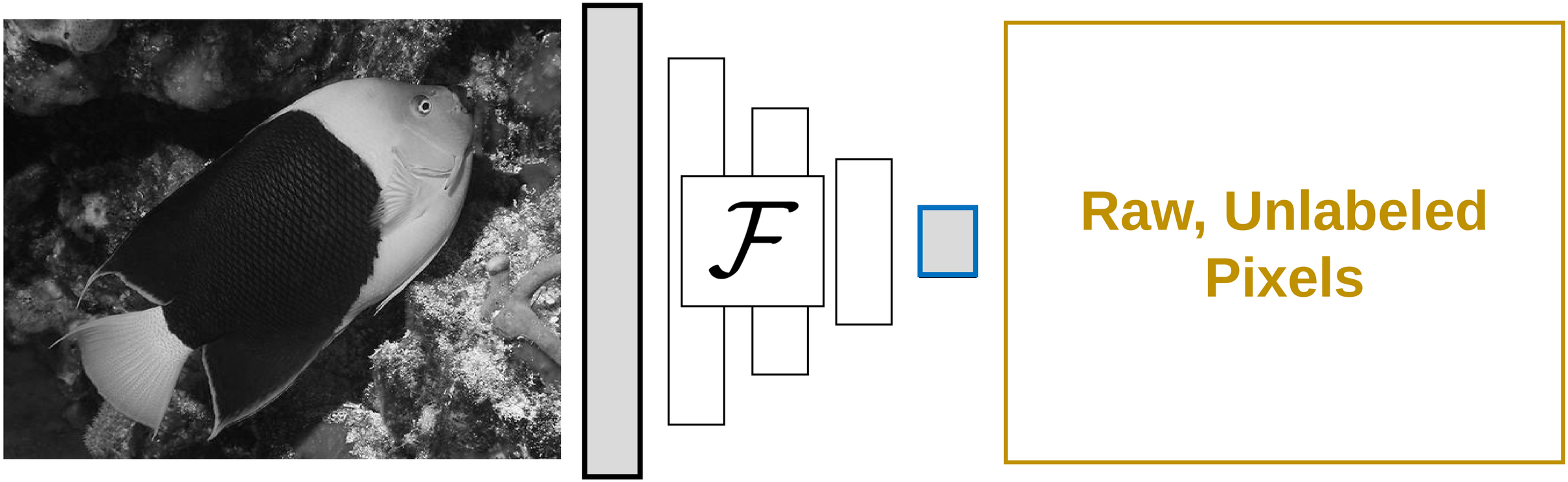
“Rockfish”

$$\arg\min_{\mathcal{F}} \mathbb{E}_{\mathbf{x},\mathbf{y}}[L(\mathcal{F}(\mathbf{x}),\mathbf{y})]$$

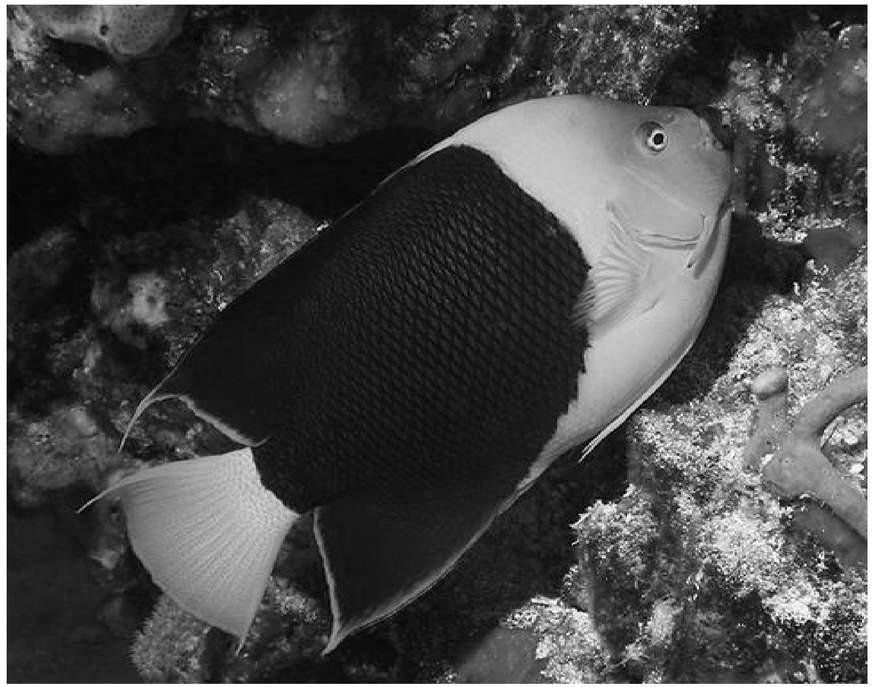
Objective function (loss)

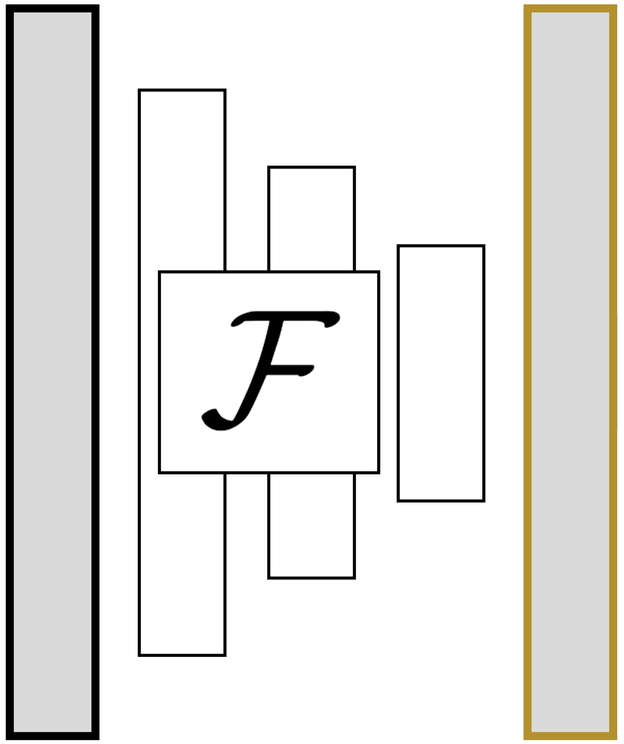
Neural Network

Discriminative Deep Networks



Generative Deep Networks





Raw, Unlabeled Pixels



$\mathcal{F}$

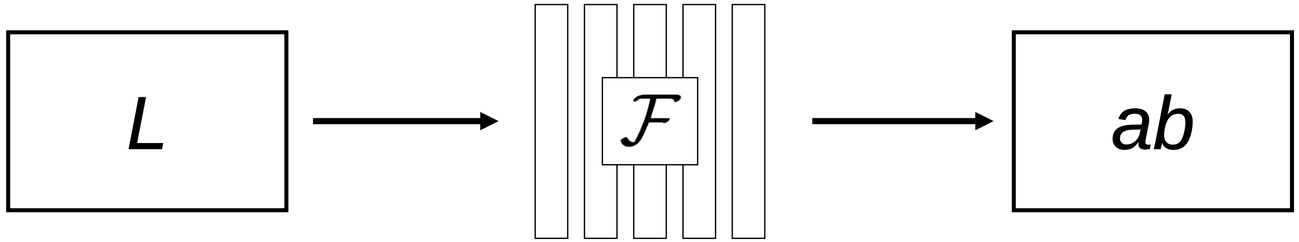


Grayscale image: *L* channel

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

Color information: *ab* channels

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



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



$\mathcal{F}$

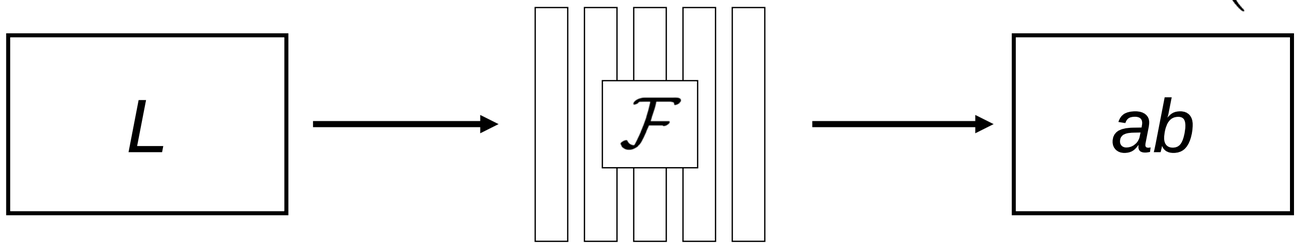


Grayscale image: *L* channel

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

Concatenate (*L*,*ab*) channels

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



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

Simple L2 regression doesn’t work ☹

Input



Output

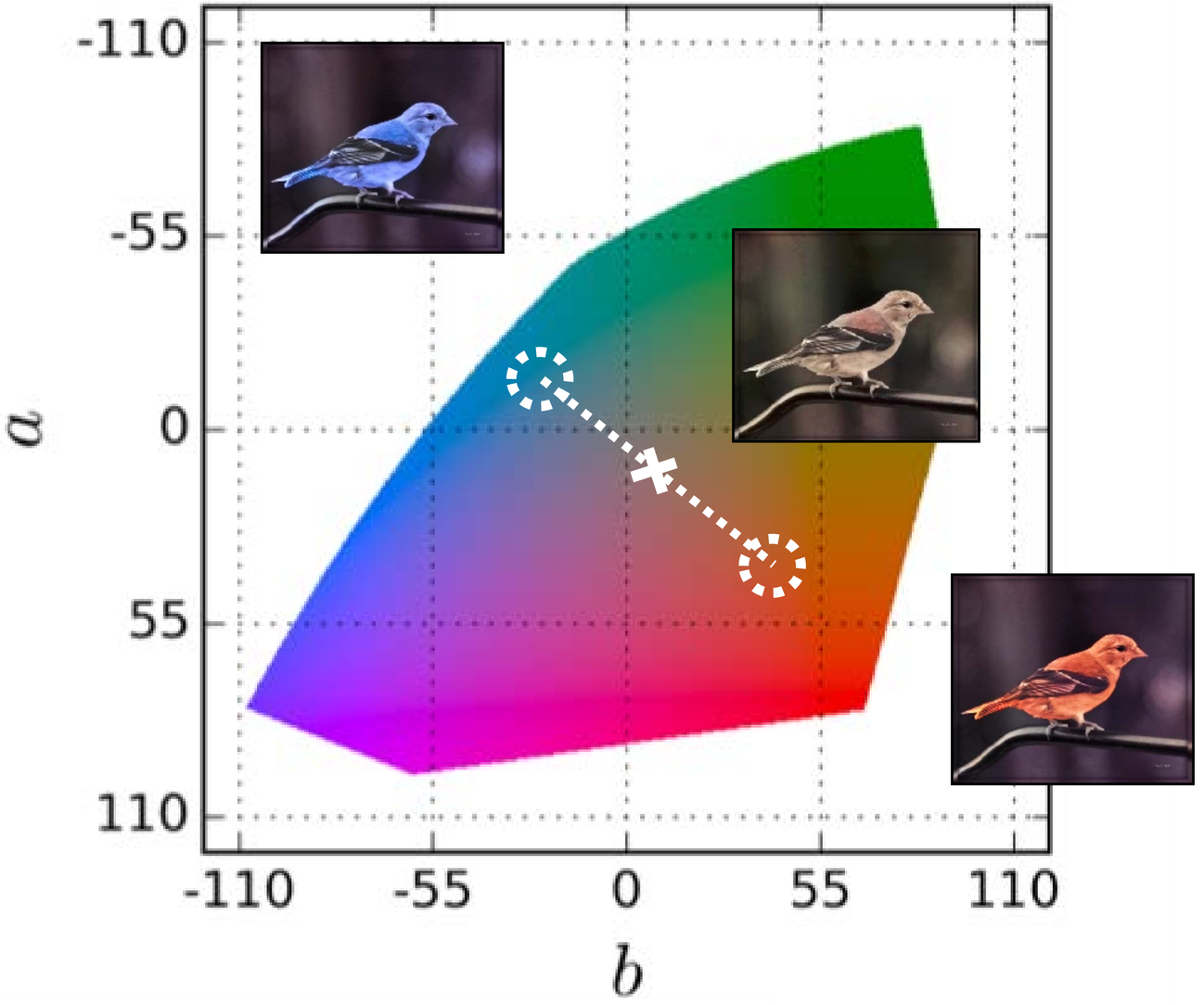


Ground truth



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

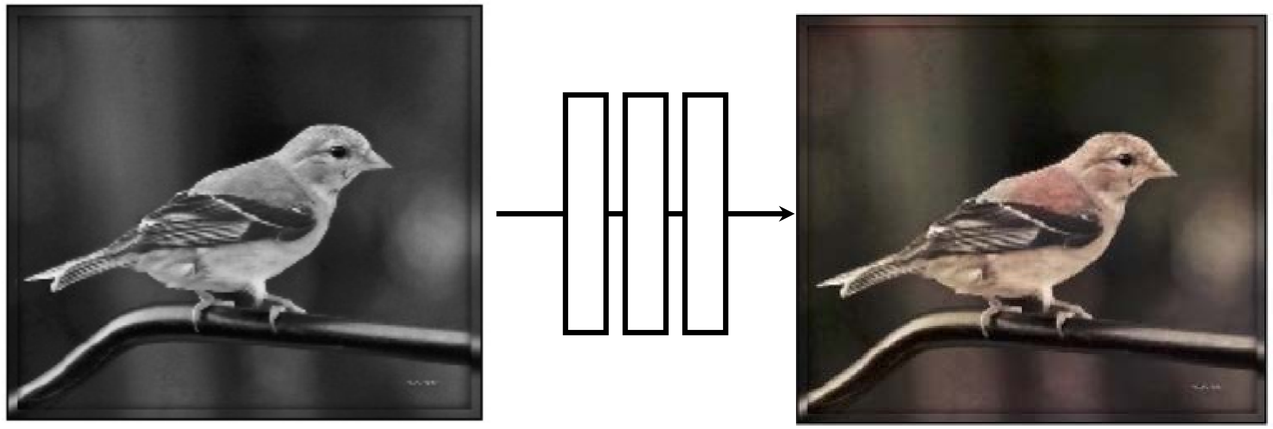




$$\mathrm{L}_2(\widehat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \|\mathbf{Y}_{h,w} - \widehat{\mathbf{Y}}_{h,w}\|_2^2$$

Designing loss functions

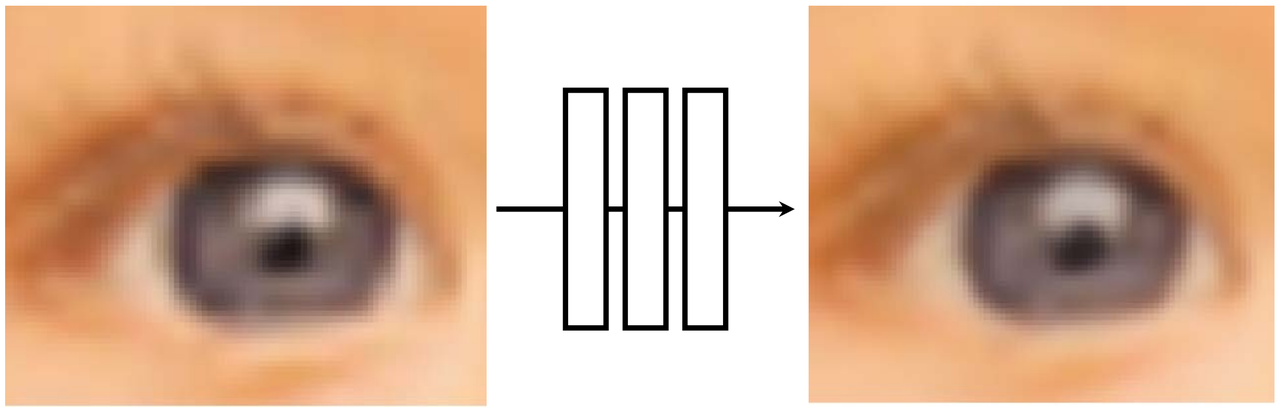
Image colorization



[Zhang et al. 2016]

L2 regression

Super-resolution



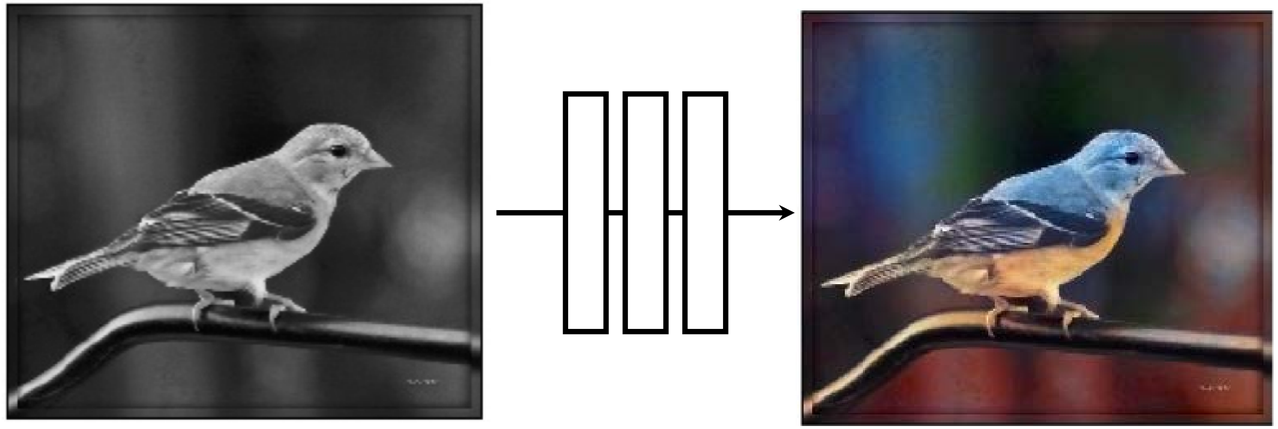
[Johnson et al. 2016]

L2 regression

slides credit: Phillip Isola

Designing loss functions

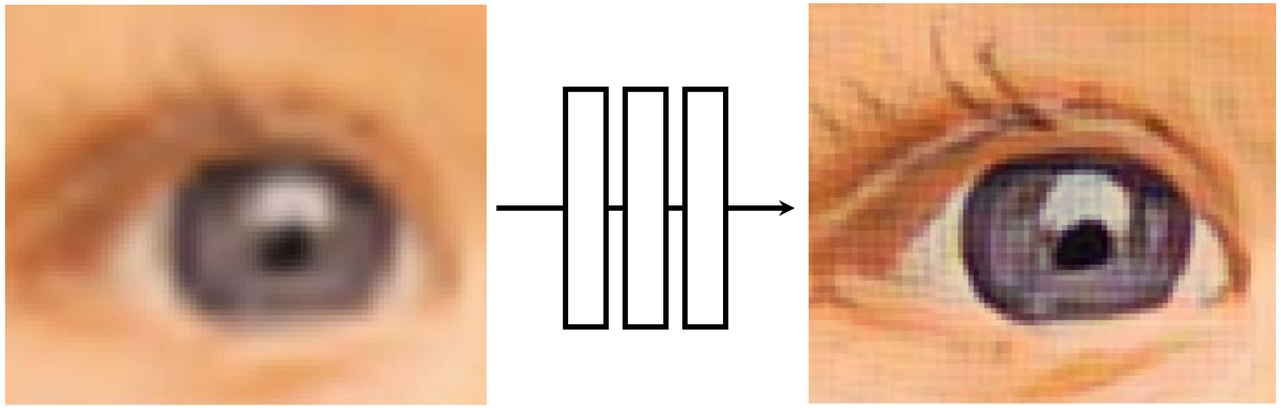
Image colorization



[Zhang et al. 2016]

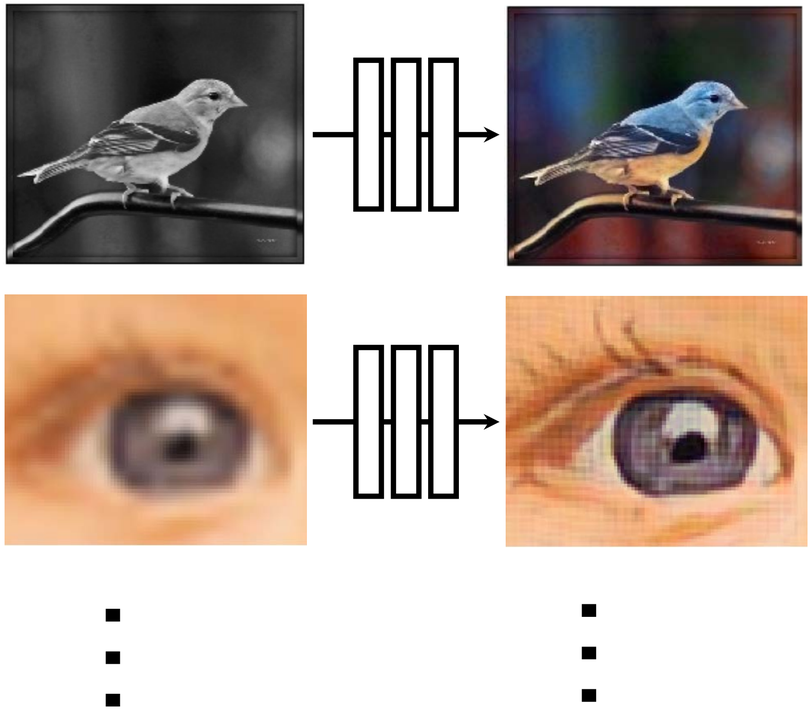
Cross entropy loss, with colorfulness term

Super-resolution



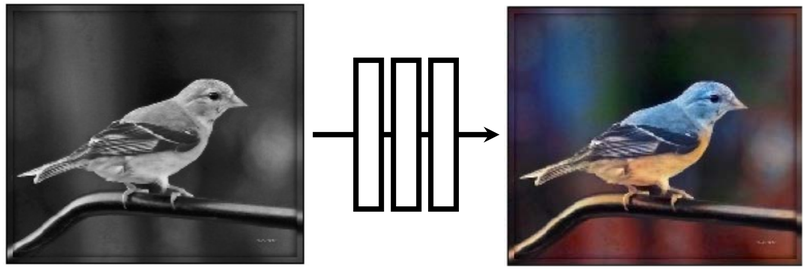
[Johnson et al. 2016]

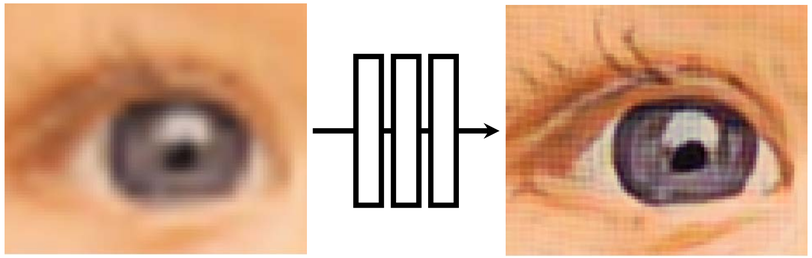
“semantic feature loss” (VGG feature covariance matching objective)



Universal loss?

Generated images





Generative Adversarial Network

Generated vs Real (classifier)

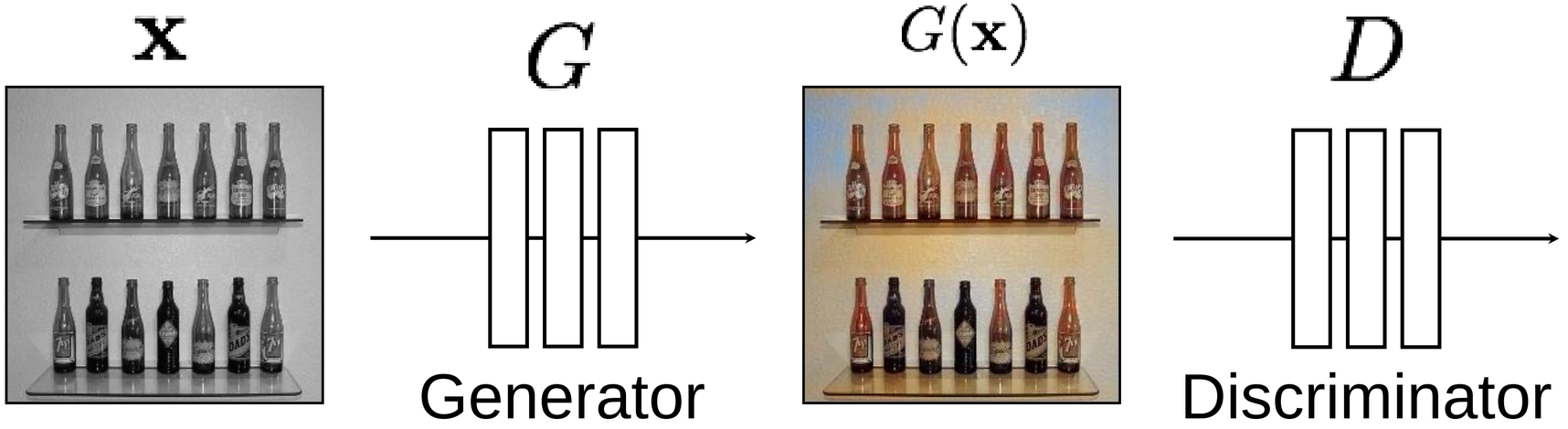
Real photos



…



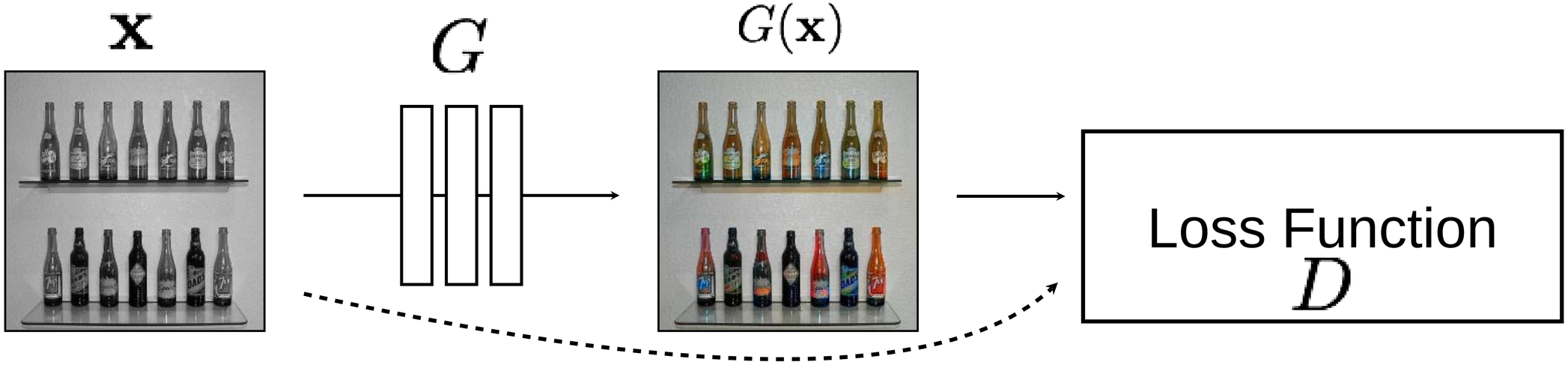
[Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, Bengio 2014]



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

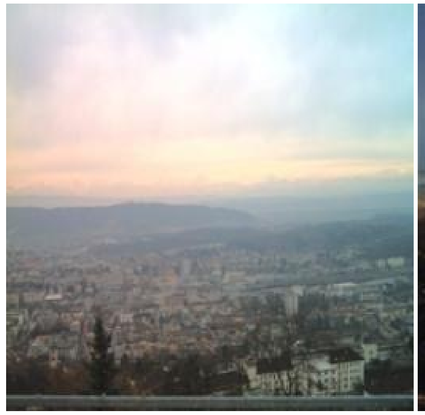
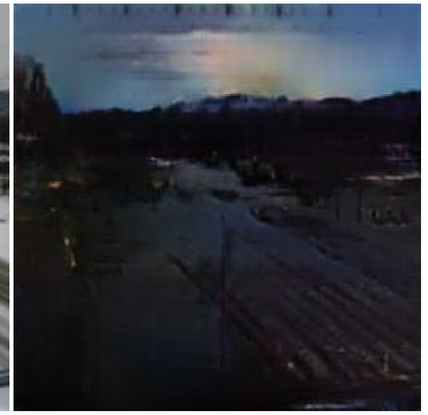
Output

Groundtruth



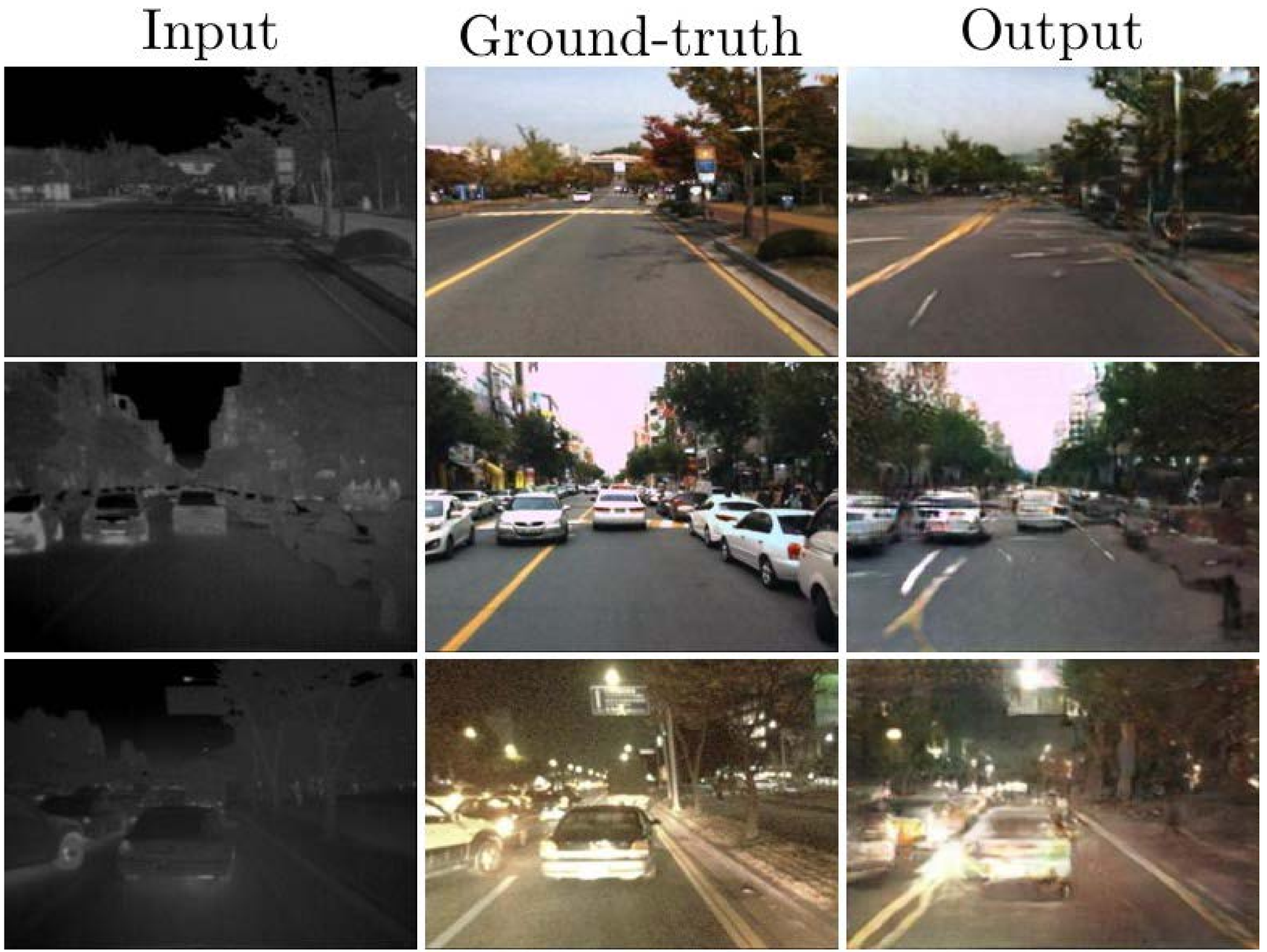
Data from [maps.google

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

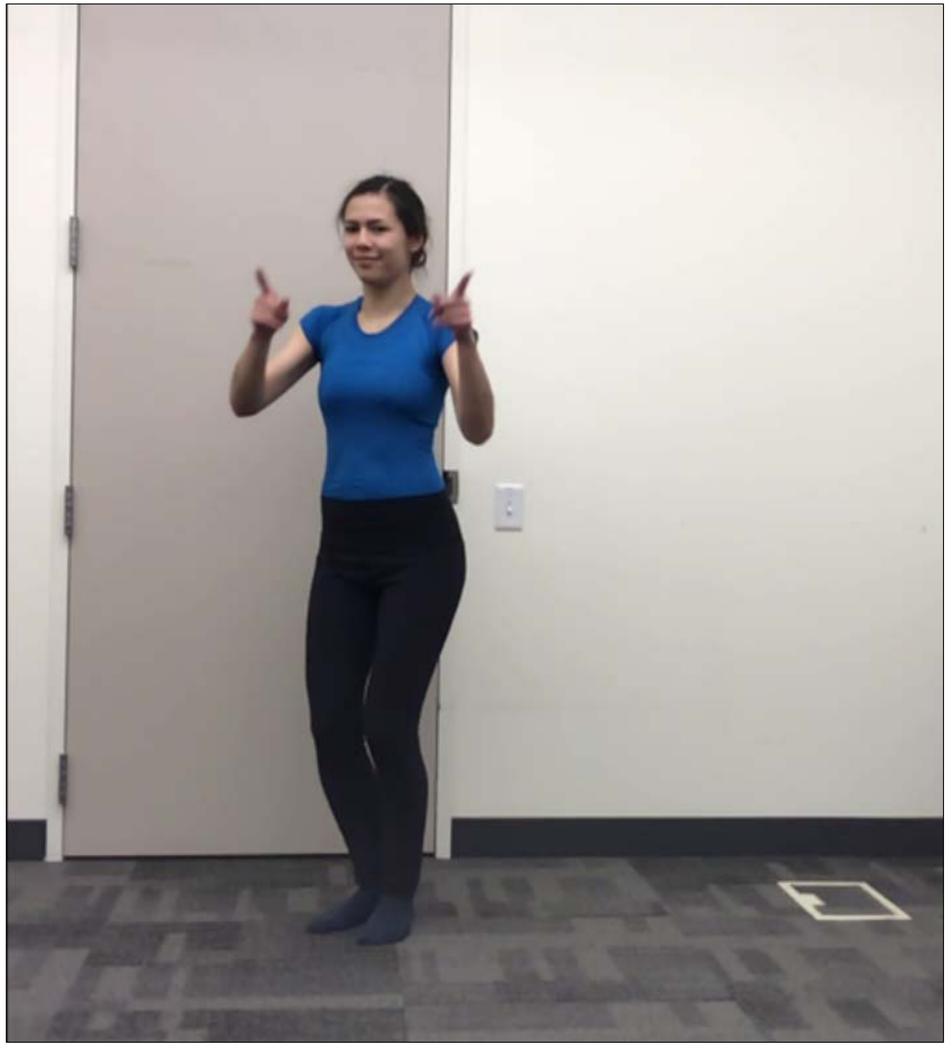
Input | Output | Input | Output | Input | Output



Trained on Edges → Images

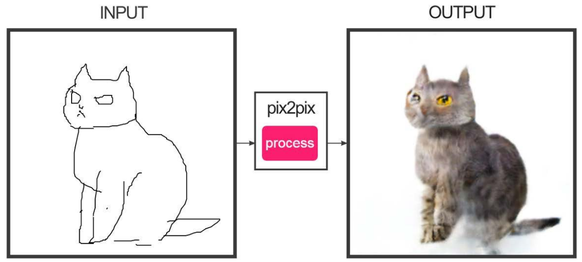
Data from [Eitz, Hays, Alexa, 2012]

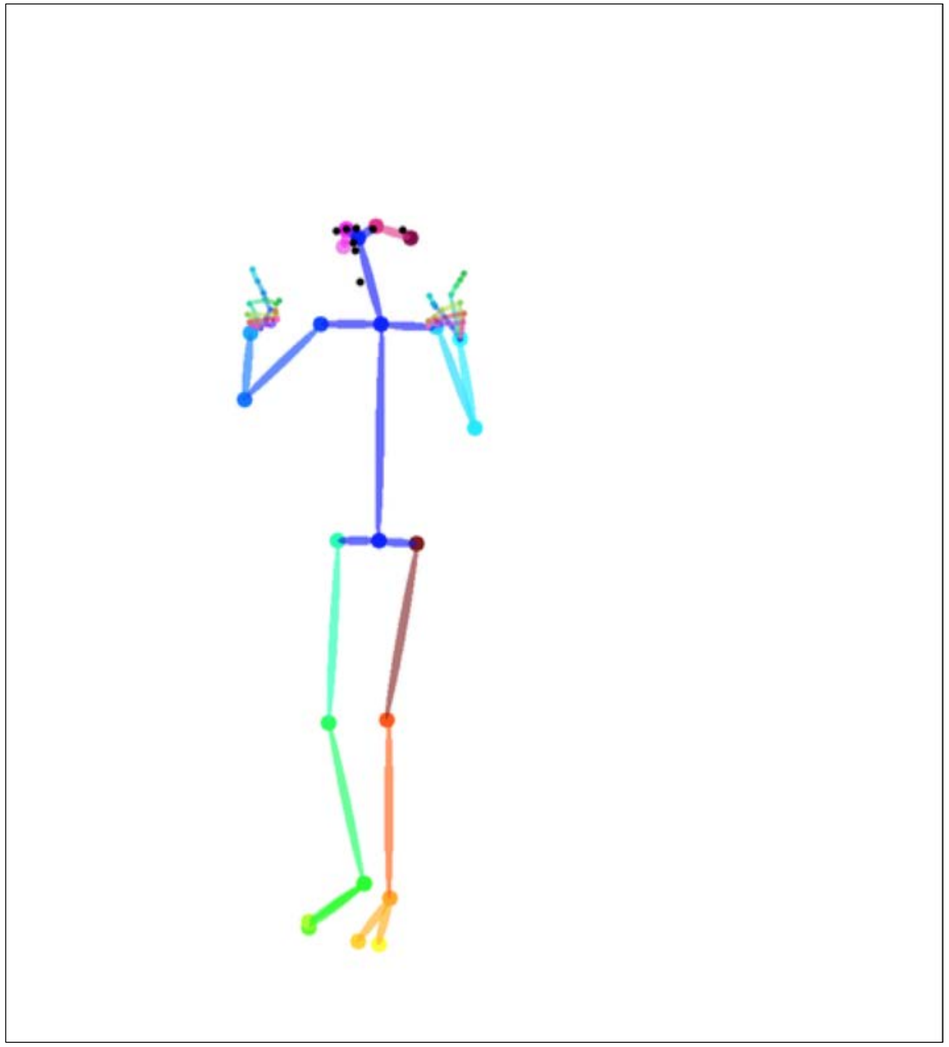
“Do as I Do”



OpenPose

pix2pix





Everybody Dance Now

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



Source Subject



Target Subject

Results



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

3. GANs as “meta-supervision”

Direct vs. Meta Supervision

$F(x) = y$

- direct supervision

$F(x) \in Y$

- GANs

$G(F(x)) = x$

- cycle-consistency

Cycle-consistency as Supervision

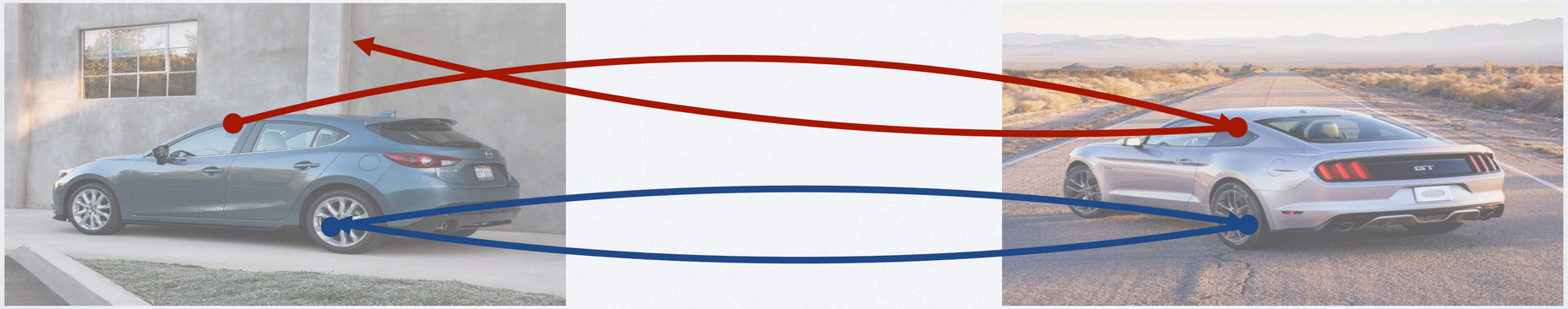
- Composite flows along a cycle should be zero





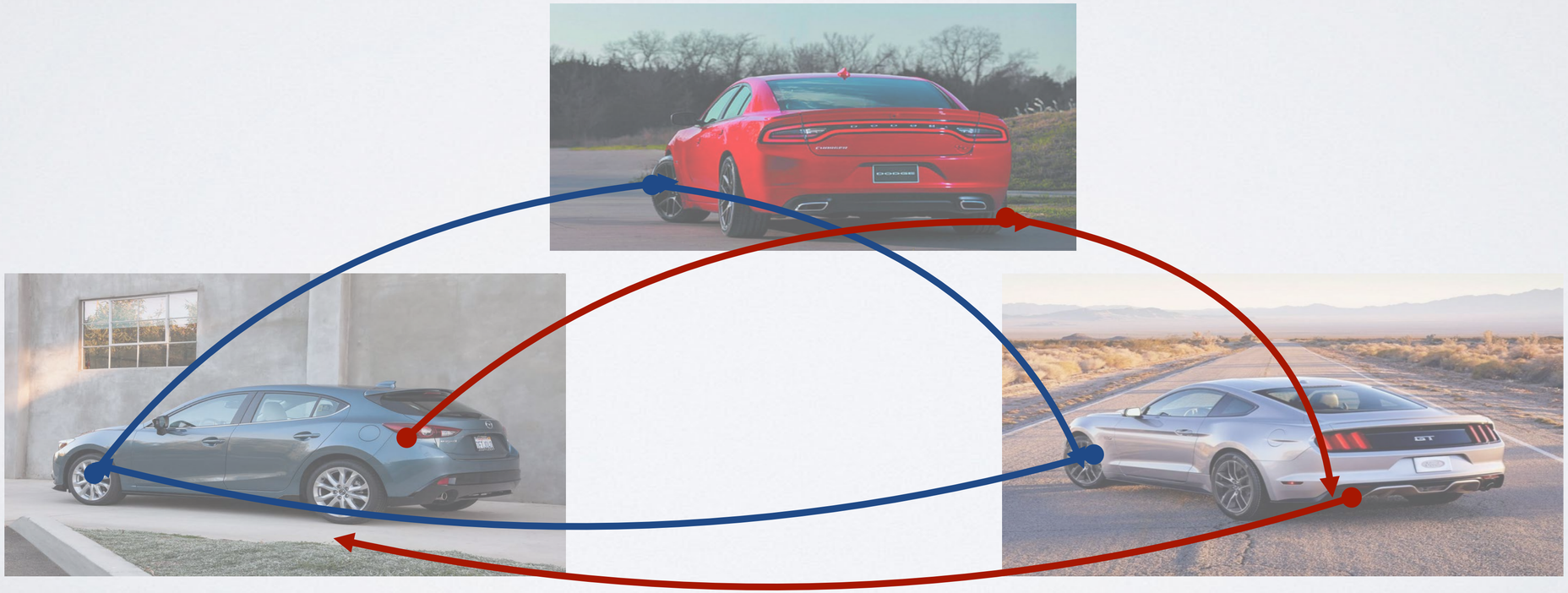
Cycle-consistency as Supervision

- Composite flows along a cycle should be zero
- 2-cycle consistency: $F_{i,j} \circ F_{j,i} = 0$



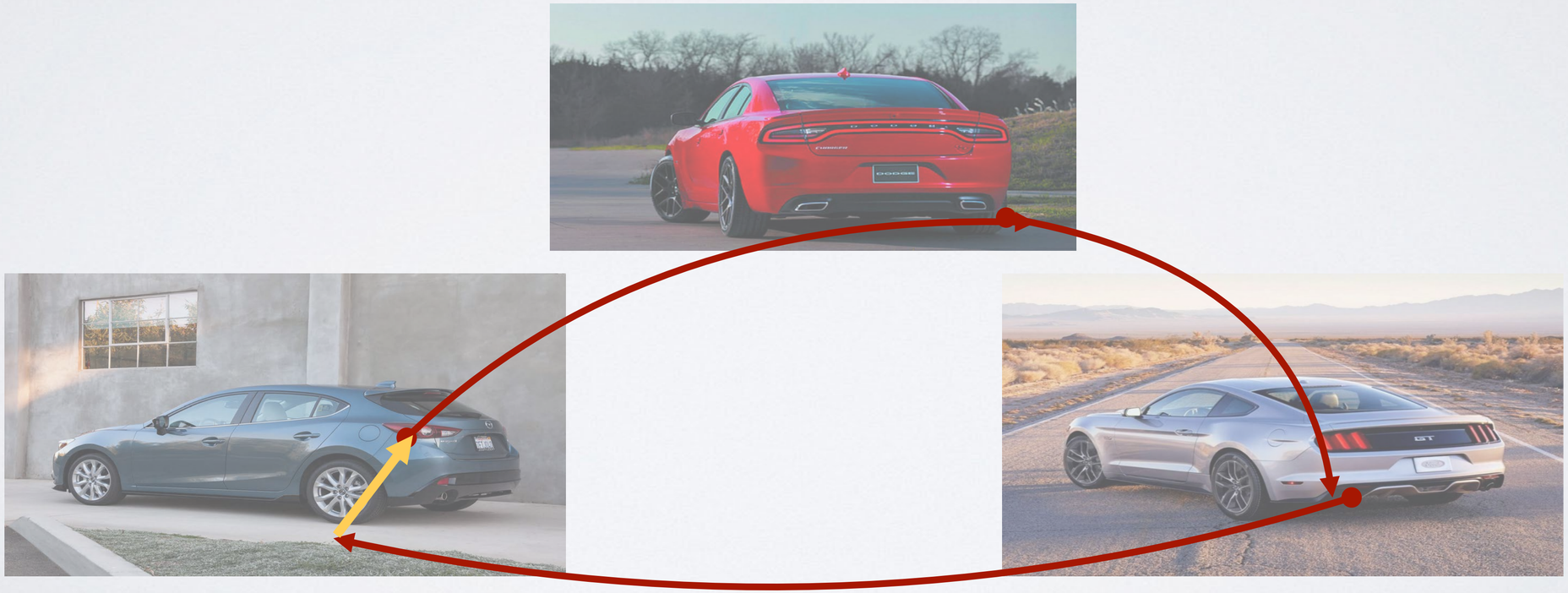
Cycle-consistency as Supervision

- Composite flows along a cycle should be zero
- 2-cycle consistency: $F_{i,j} \circ F_{j,i} = 0$
- 3-cycle consistency: $F_{i,k} \circ F_{k,j} \circ F_{j,i} = 0$



Cycle-consistency as Supervision

- Composite flows along a cycle should be zero
- 2-cycle consistency: $F_{i,j} \circ F_{j,i} = 0$
- 3-cycle consistency: $F_{i,k} \circ F_{k,j} \circ F_{j,i} = 0$



Cycle-consistency as Supervision

- Composite flows along a cycle should be zero
- 2-cycle consistency: $F_{i,j} \circ F_{j,i} = 0$
- 3-cycle consistency: $F_{i,k} \circ F_{k,j} \circ F_{j,i} = 0$

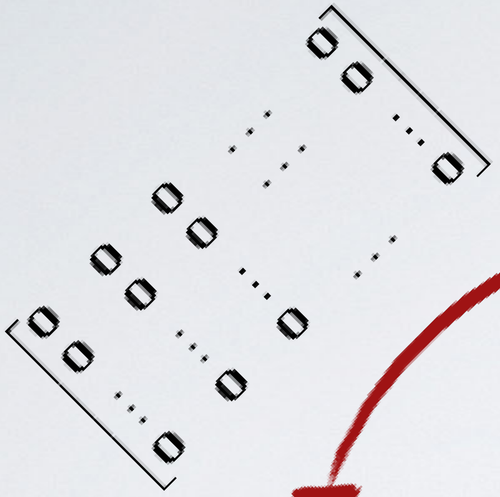




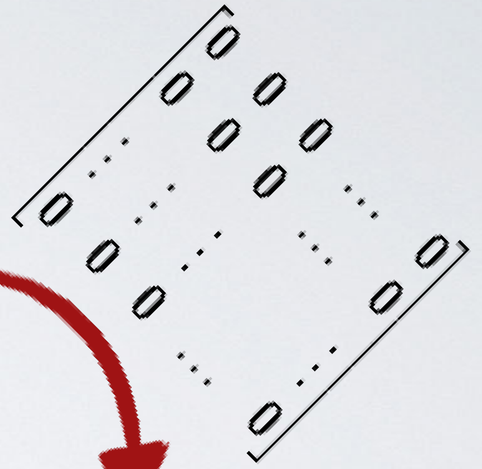
CNN

Amount of inconsistency

Could be consistent but wrong…







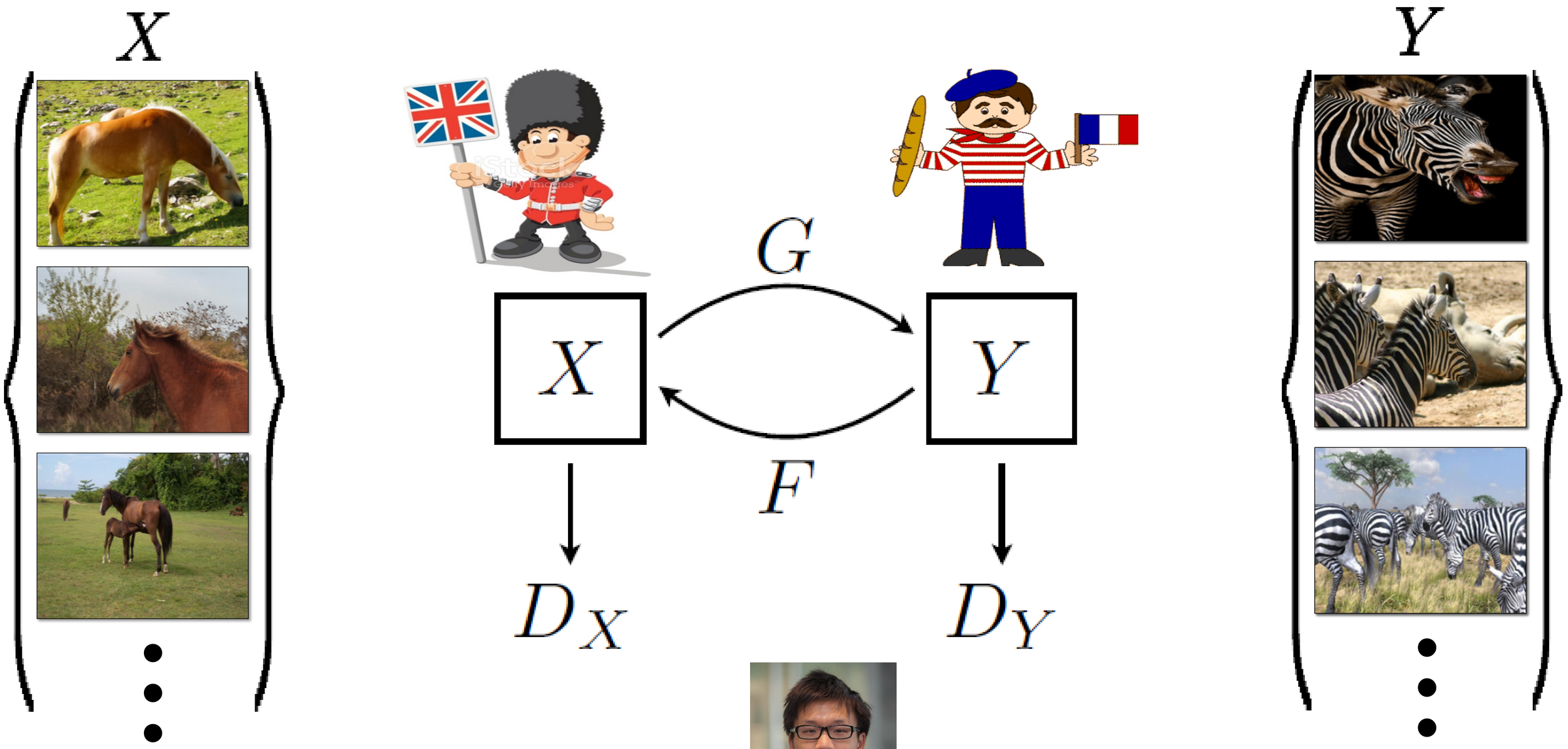
Need an anchor edge!





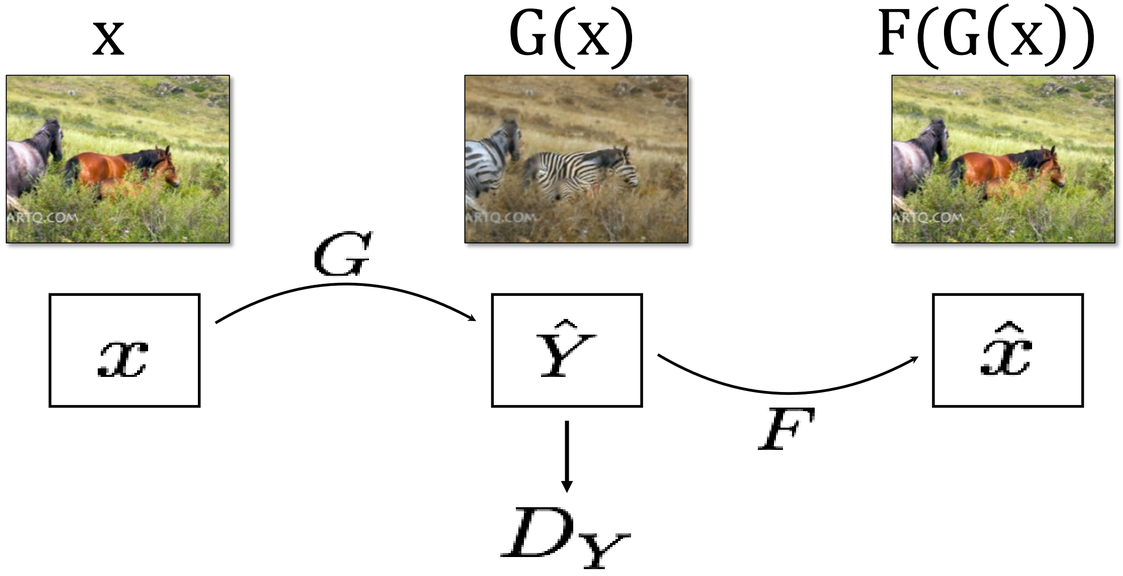
$$\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

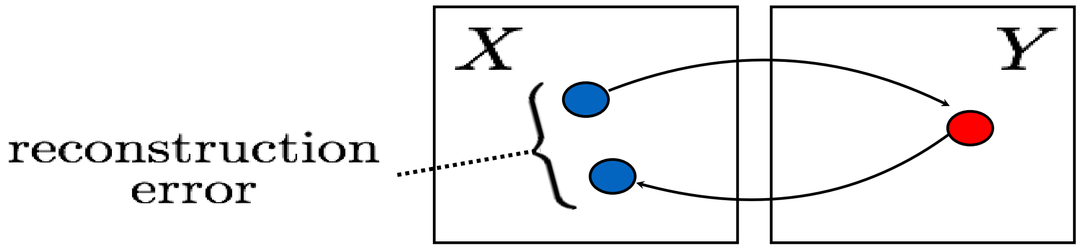
CycleGAN, or “there and back aGAN”



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

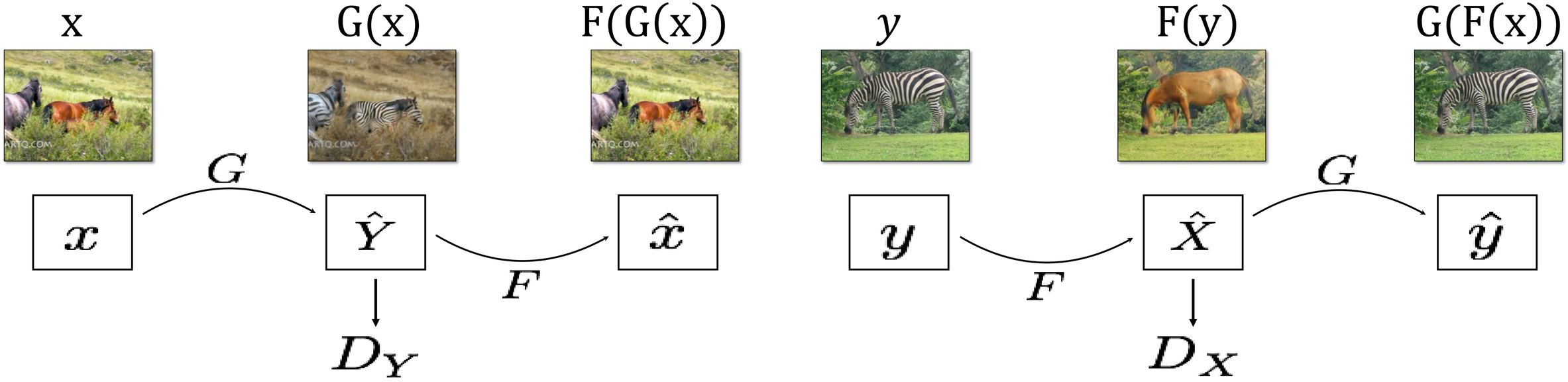
Cycle-Consistency Loss

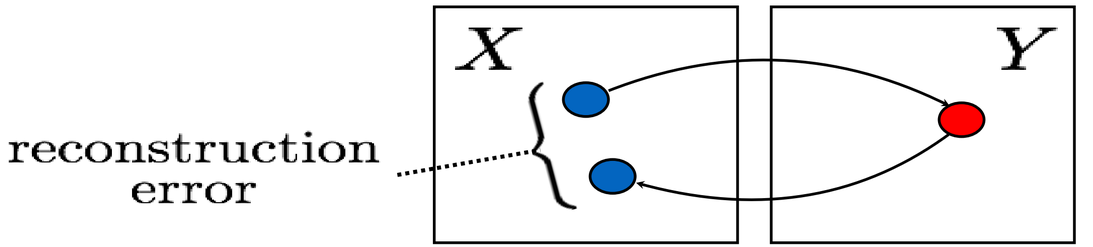




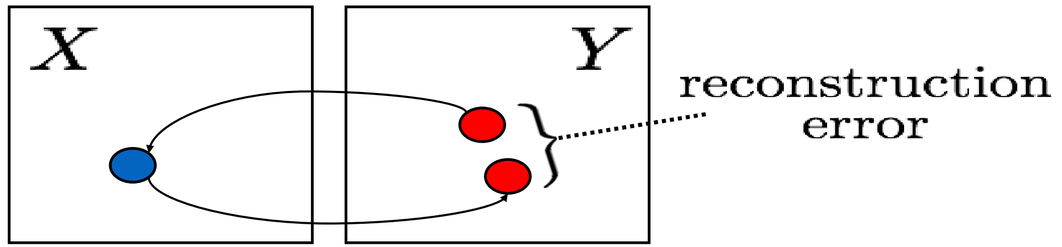
$$\|F(G(x)) - x\|_1$$

Cycle-Consistency Loss





$$\|\mathrm{F}(\mathrm{G}(\mathrm{x})) - \mathrm{x}\|_1$$



$$\|\mathrm{G}(\mathrm{F}(\mathrm{y})) - y\|_1$$





Video























Collection Style Transfer



Photograph

© Alexei Efros



Monet



Van Gogh



Cezanne



Ukiyo-e









CG to Real

Grand Theft Auto



→





→



Real to CG



→





→

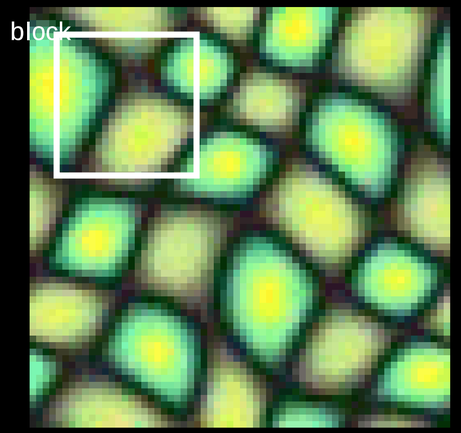


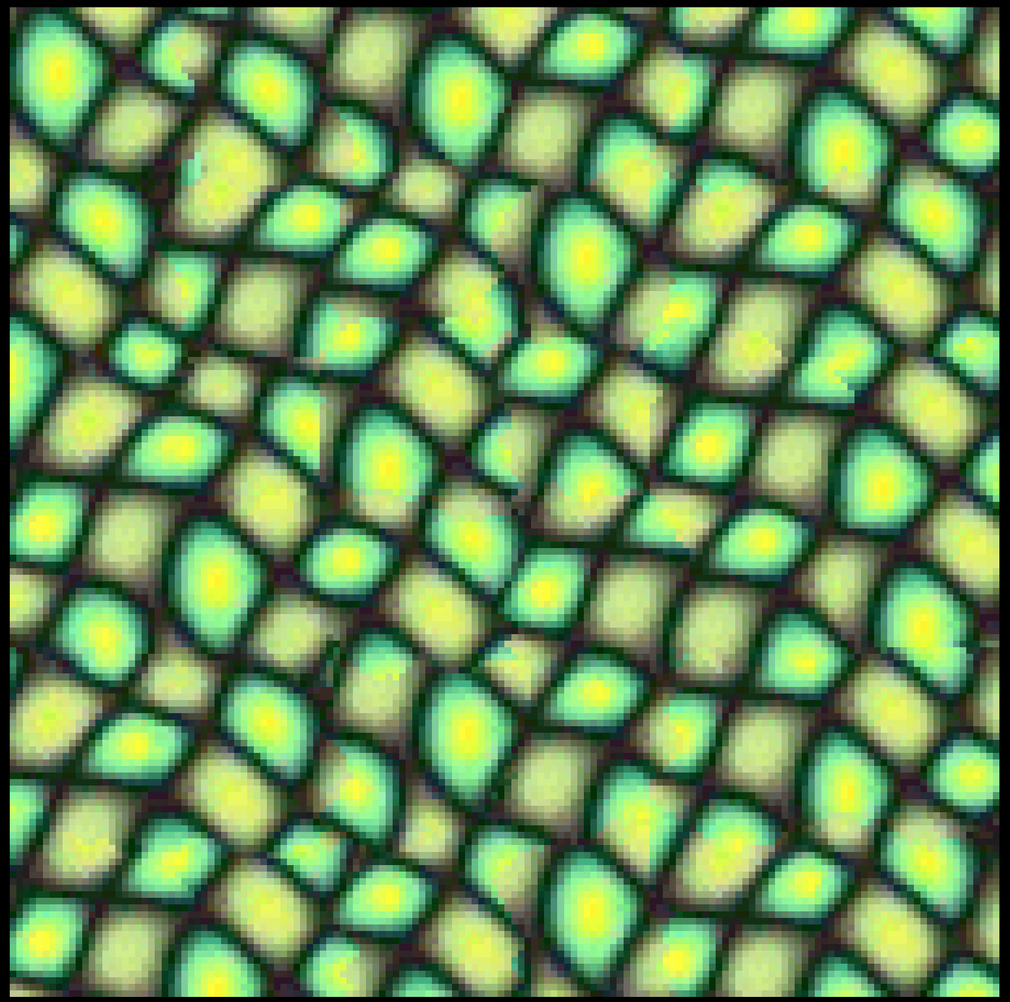
Failure case



4. GANs as compositional parametric data memorizers

Non-parametric image synthesis





[Efros & Freeman, 2001]

Image Analogies (Hertzmann, et al, 2001]



A



A’



B



B’



Wang et al, “High-Resolution Image Synthesis and Semantic Manipulatio with Conditional GANs”, CVPR’18



Chen et al, “Semi-parametric Image Synthesis”, CVPR’18

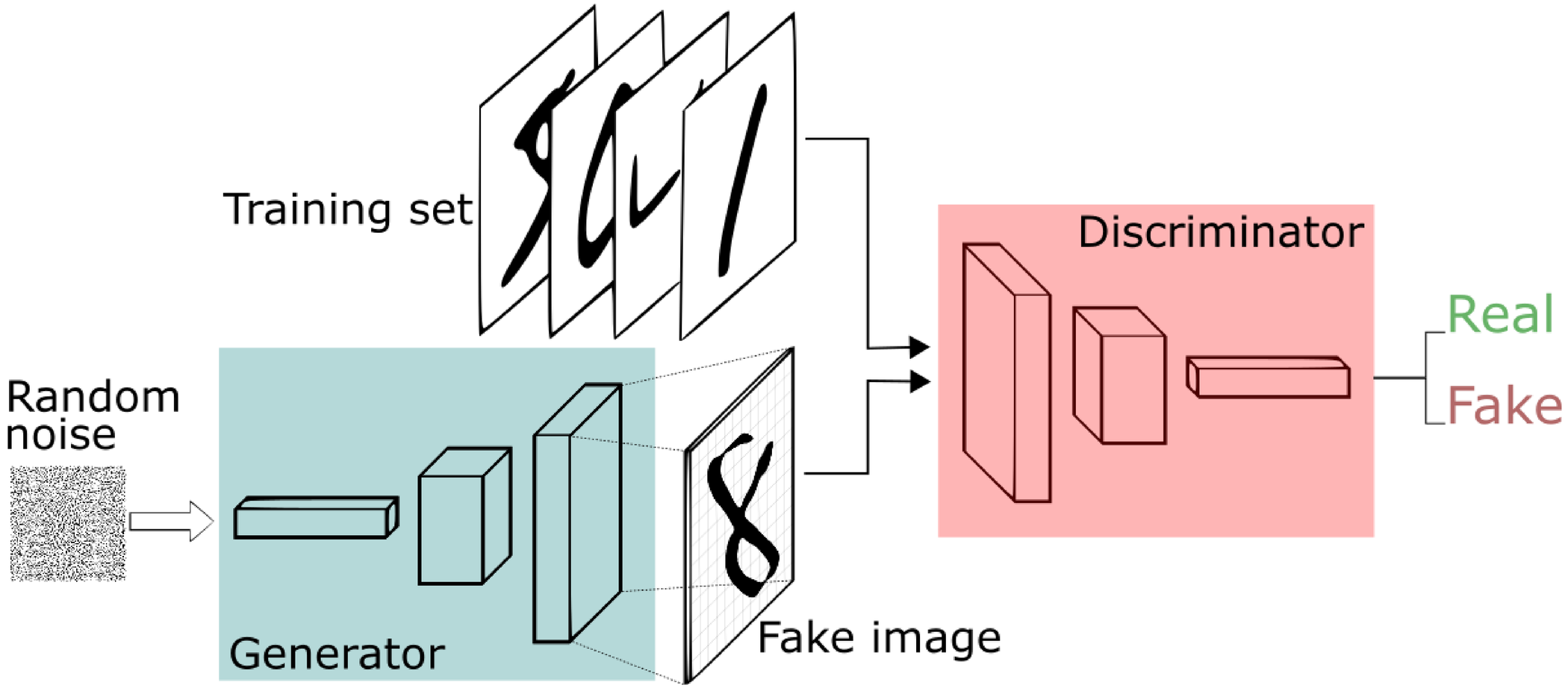
GANs are great data memorizers!

- Way faster than non-parametric nearest neighbors

- GANs seem to be able to choose “how much to steal”
  - “compositional nearest neighbors”

5. GANs as generalized texture synthesis

Funny thing: G never sees real images



- D instructs G on what makes an image real

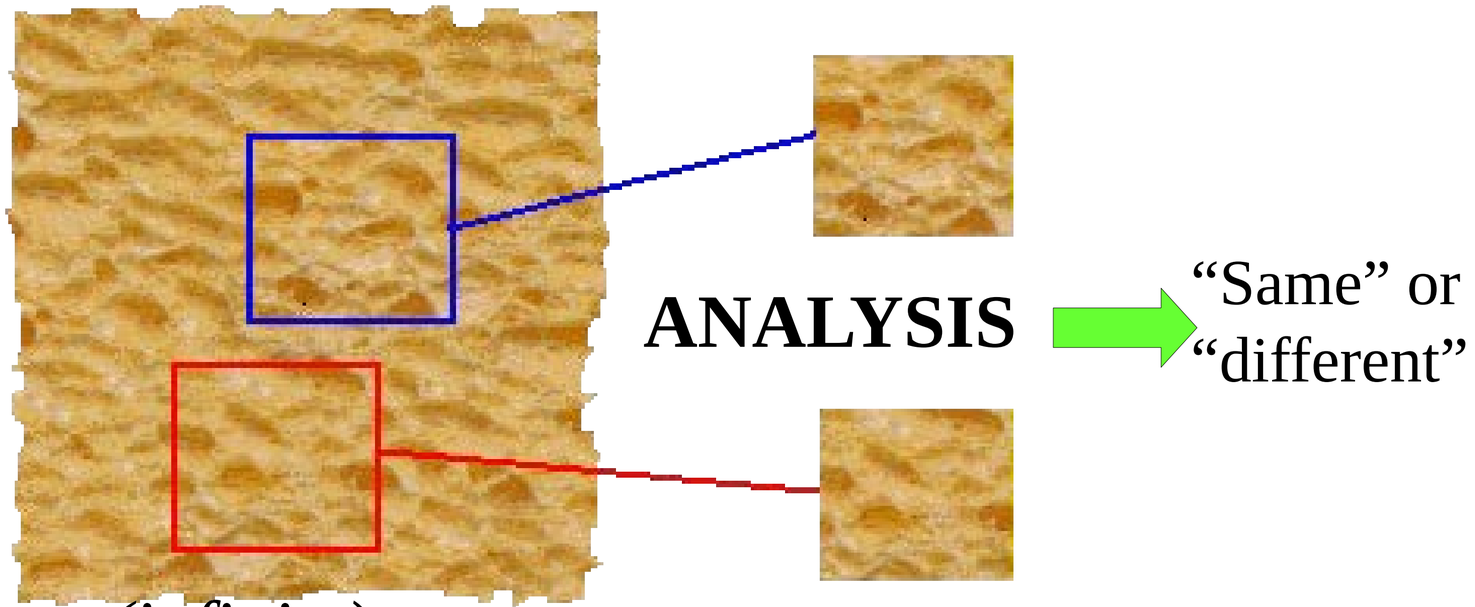
Visual texture







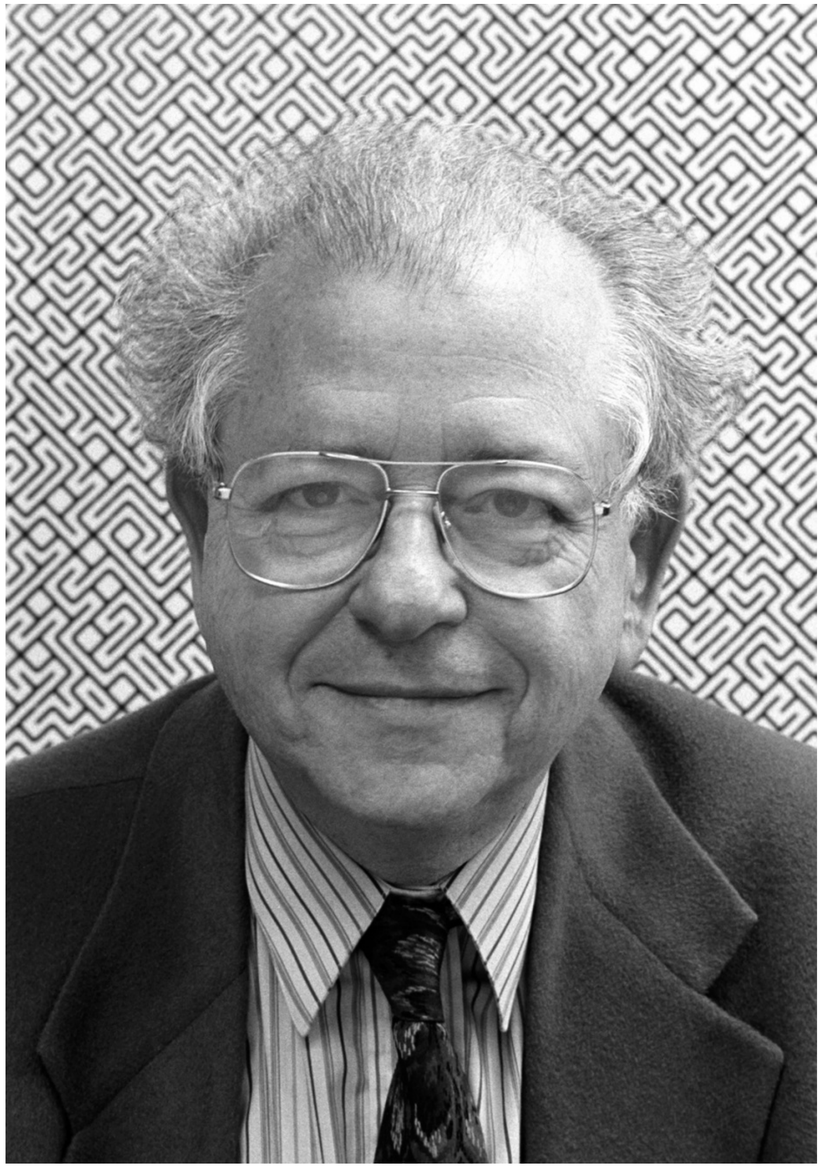
Texture Analysis



True (infinite) texture

Compare textures and decide if they’re made of the same “stuff”.

Béla Julesz, father of texture

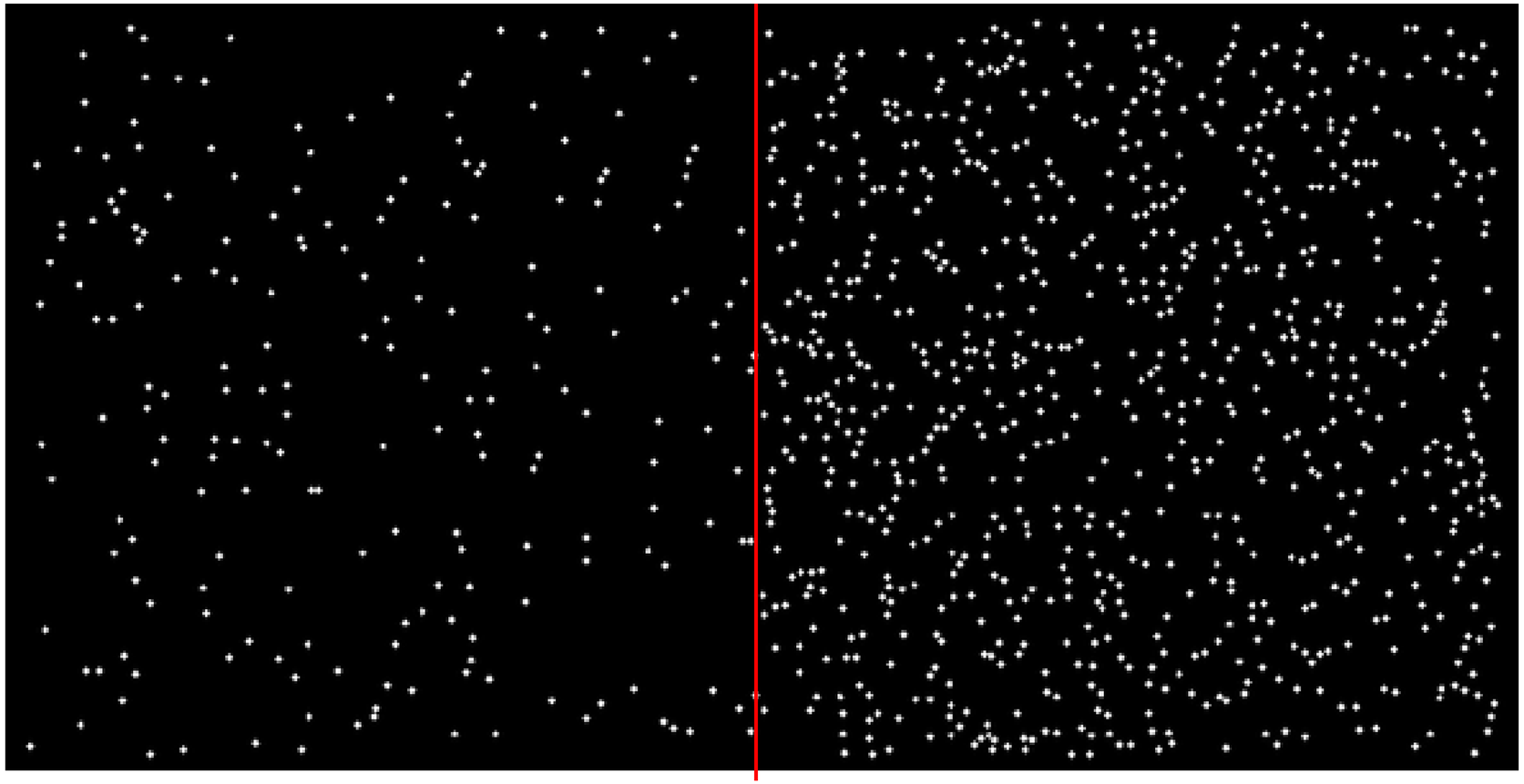


Julesz Conjecture

Textures cannot be spontaneously discriminated if they have the same first-order and second-order statistics and differ only in their third-order or higher-order statistics.

(later proved wrong)

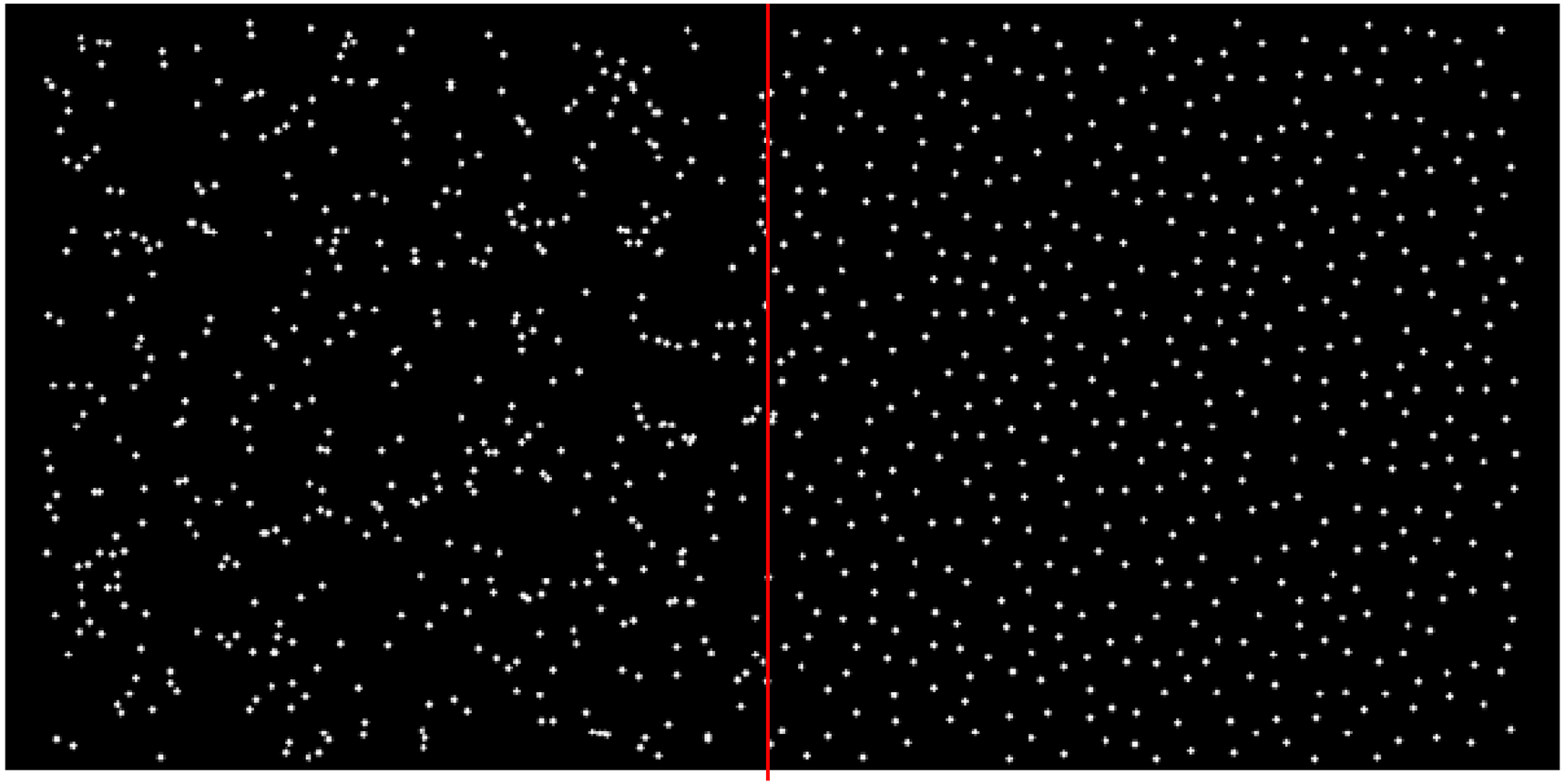
1st Order Statistics



5% white

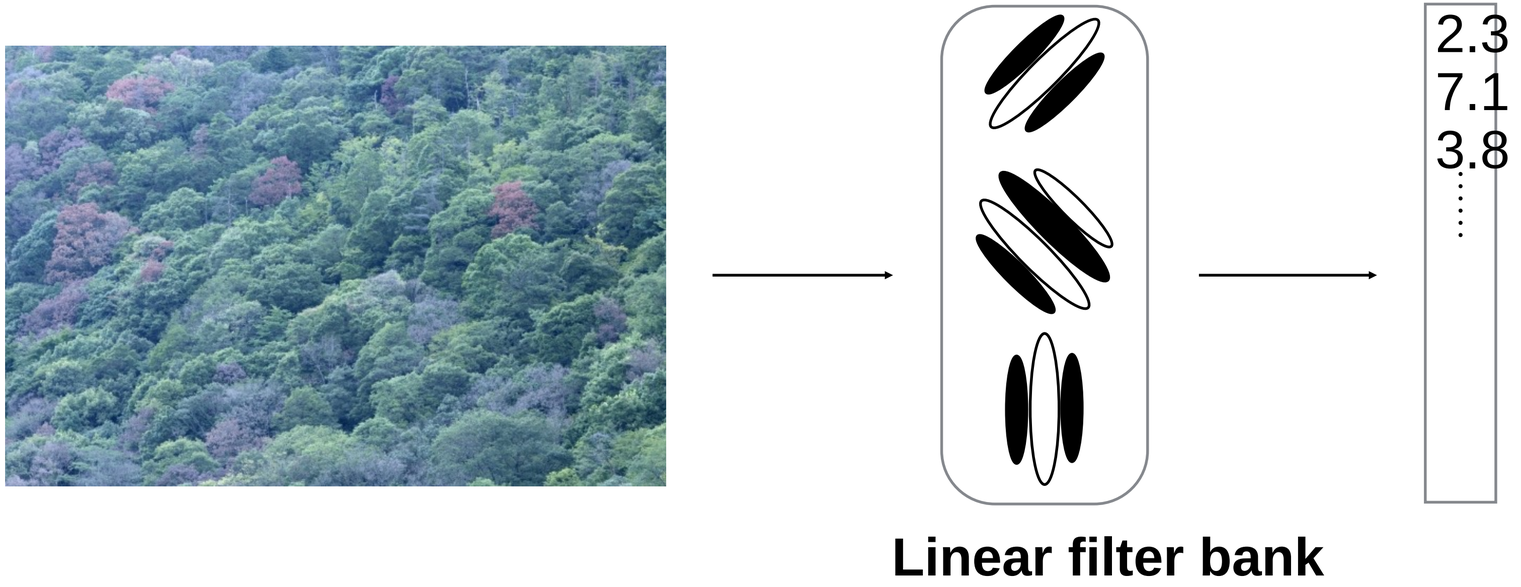
20% white

2nd Order Statistics



10% white

But what are the features?



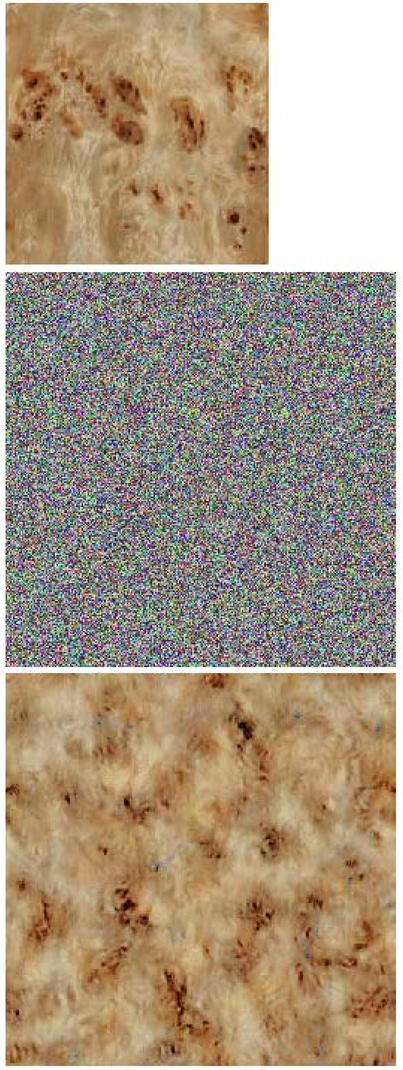
Heeger & Bergen (1995)
Portilla & Simoncelli (2000)

Heeger & Bergen, SIGGRAPH‘95

Start with a noise image as output

Main loop:

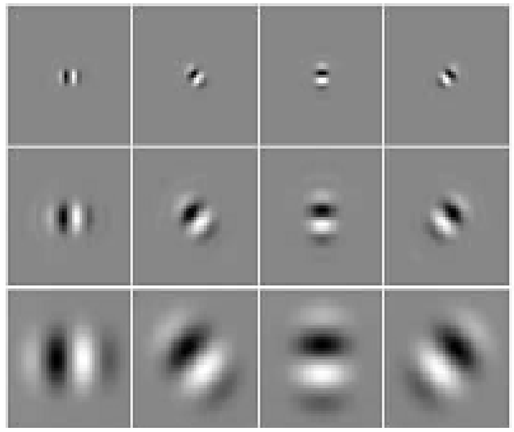
- Match pixel histogram of output image to input
- Decompose input and output images using multi-scale filter bank (Steerable Pyramid)
- Match subband histograms of input and output pyramids
- Reconstruct input and output images (collapse the pyramids)

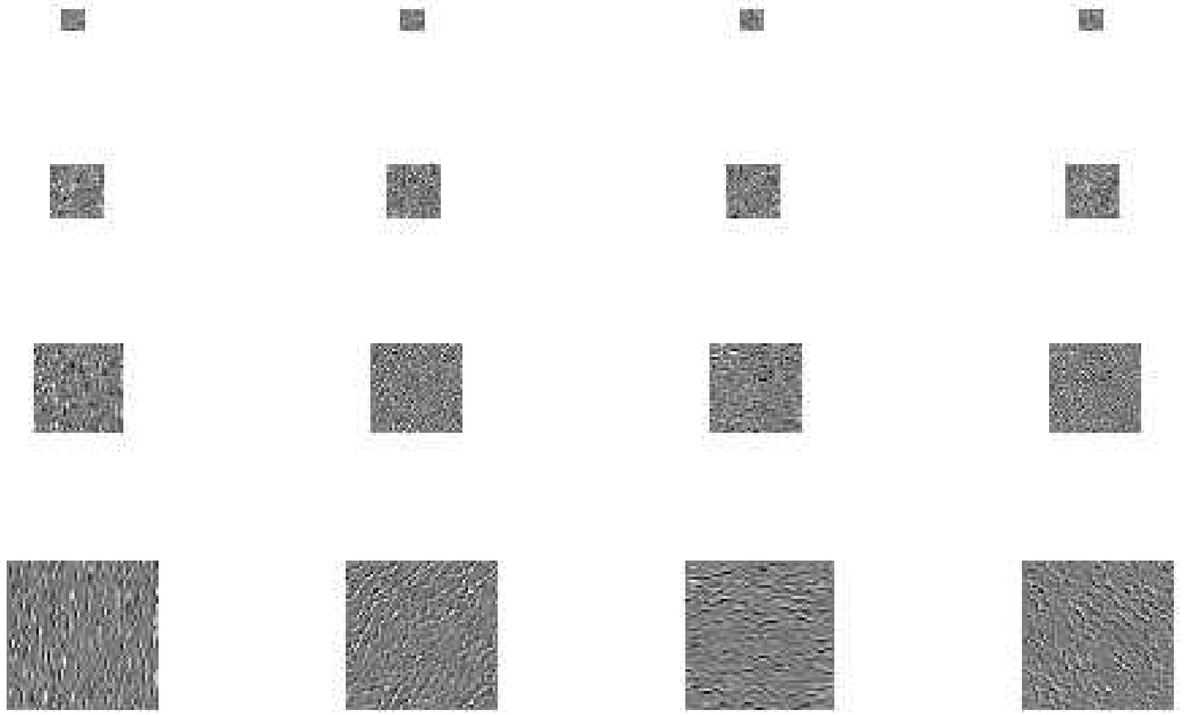


Heeger, Bergen, Pyramid-based texture analysis/synthesis, SIGGRAPH 1995

Multi-scale filter decomposition

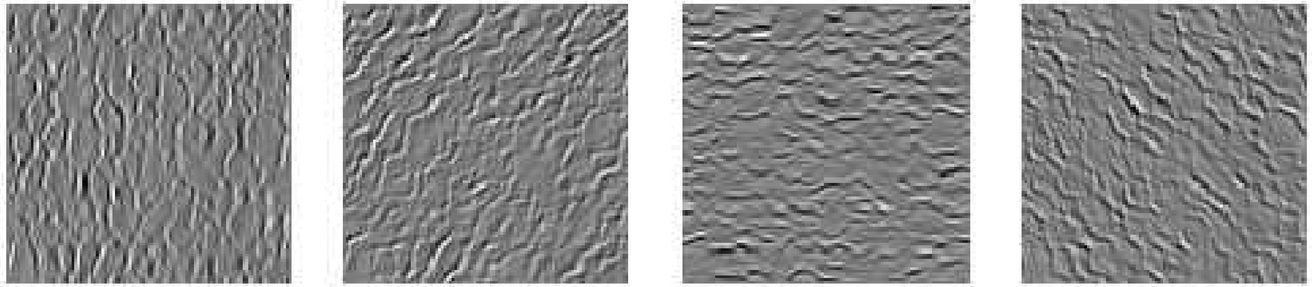
Filter bank



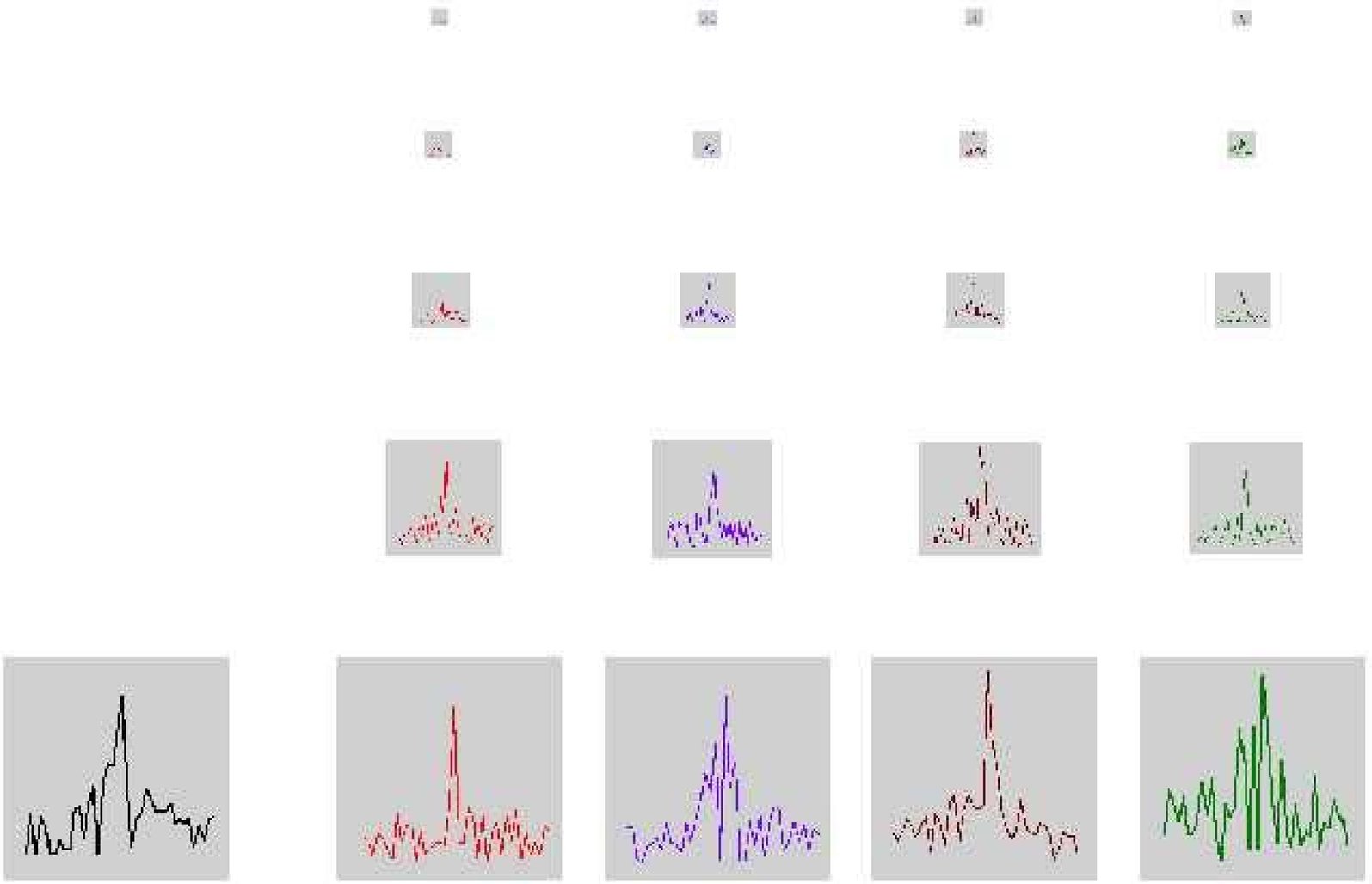


Input image

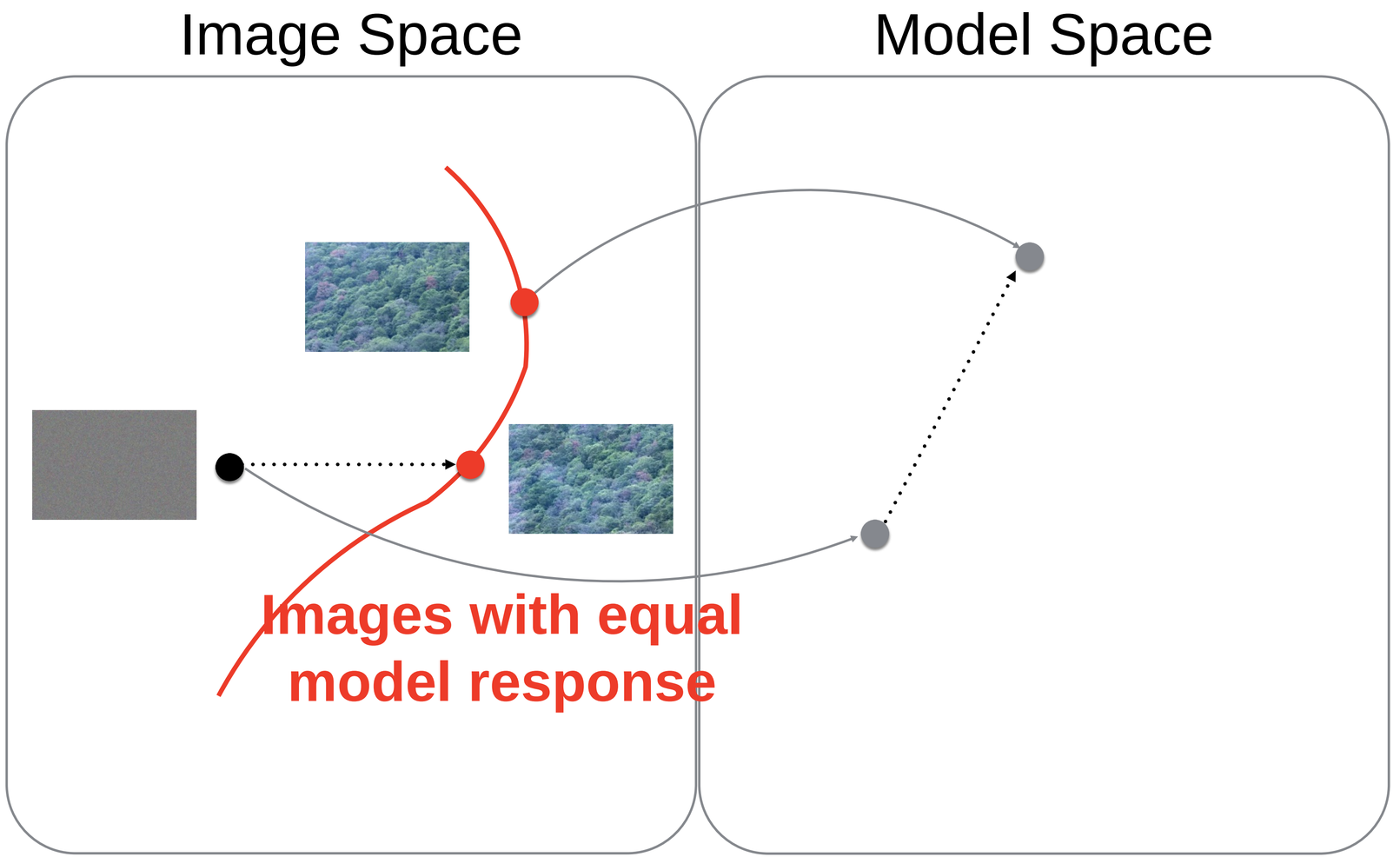




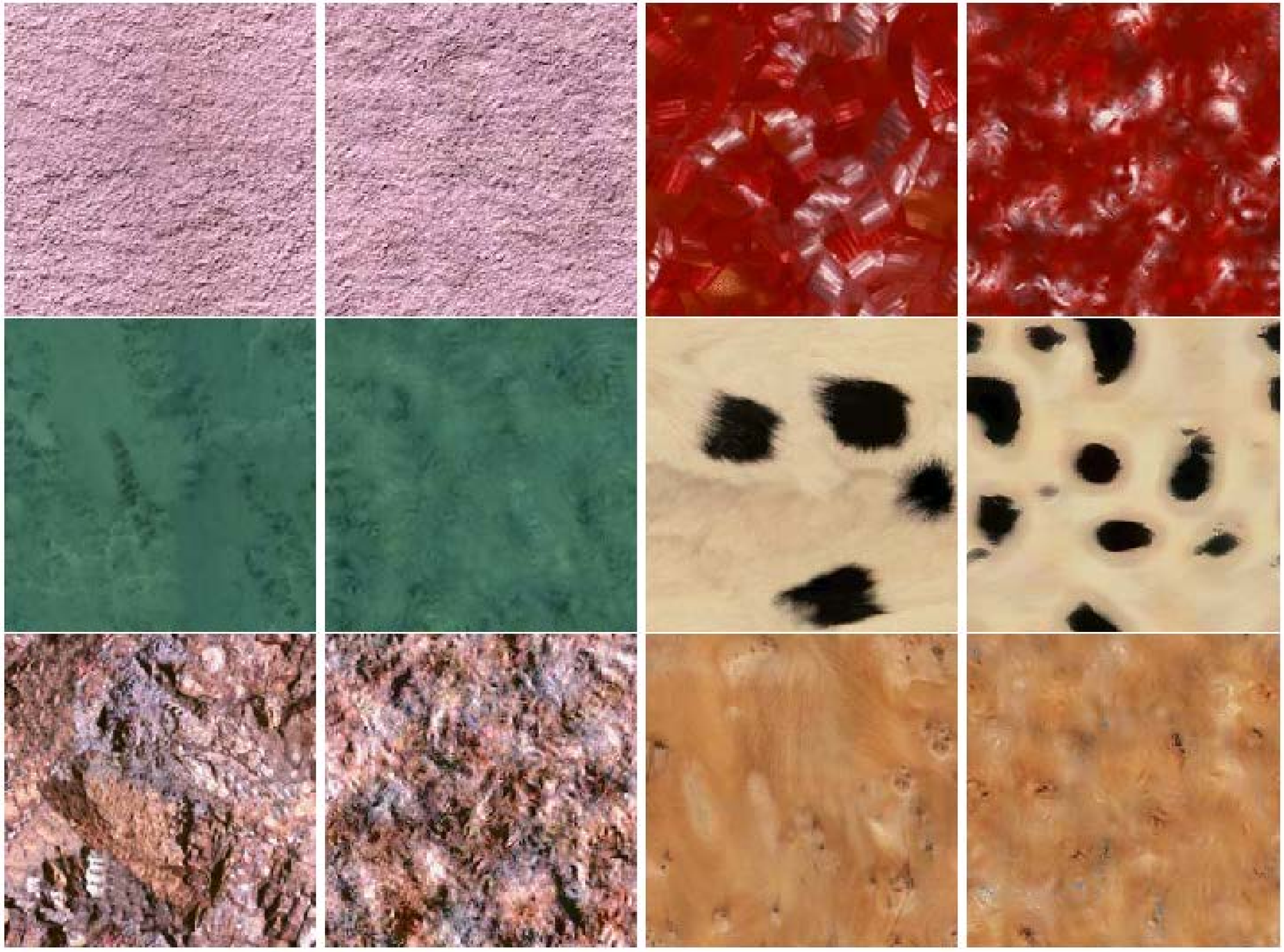
Filter response histograms

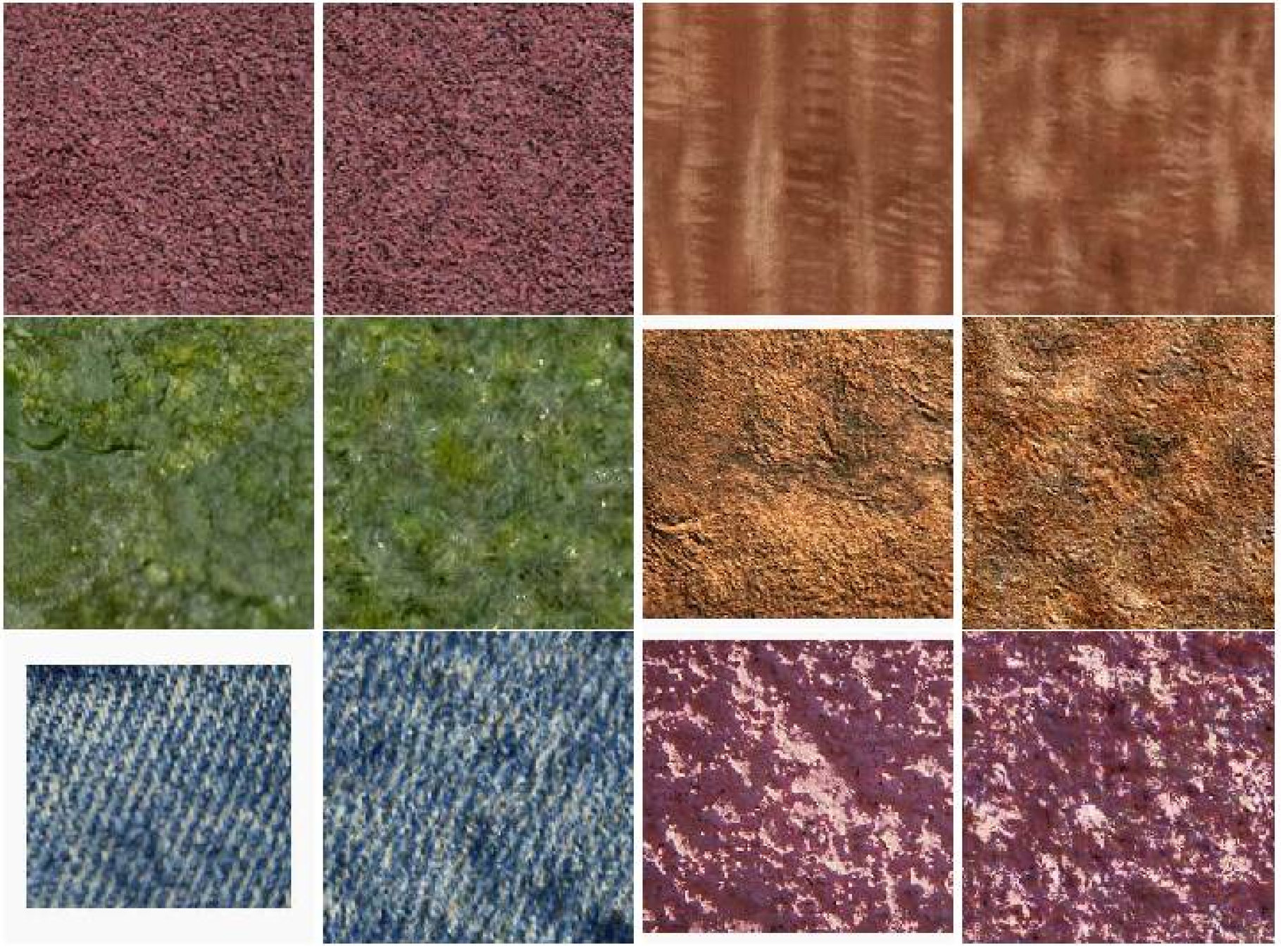


Texture Synthesis



Slide by Portilla & Simoncelli (2000)





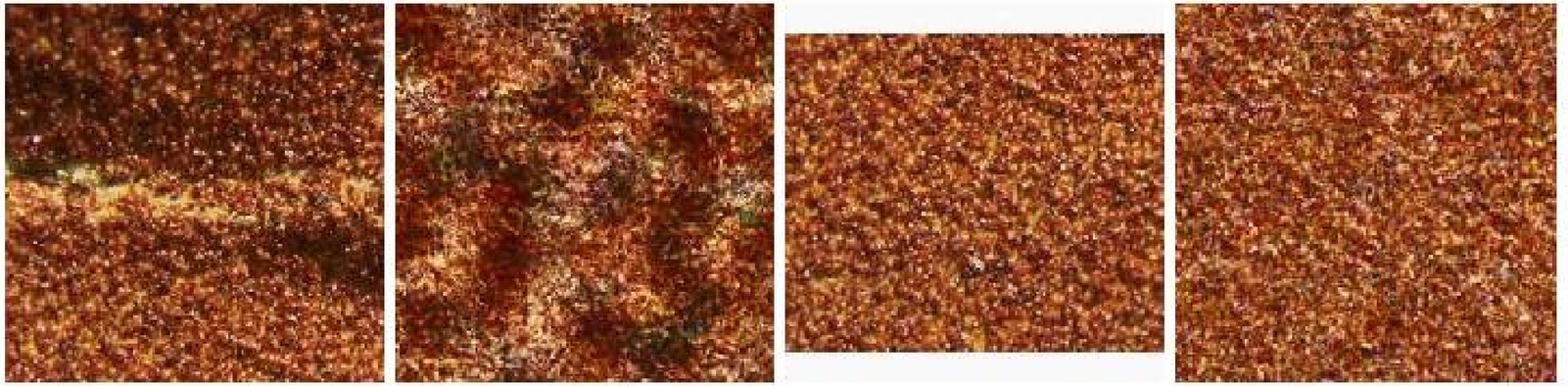


Figure 7: (Left pair) Inhomogeneous input texture produces blotchy synthetic texture. (Right pair) Homogeneous input.

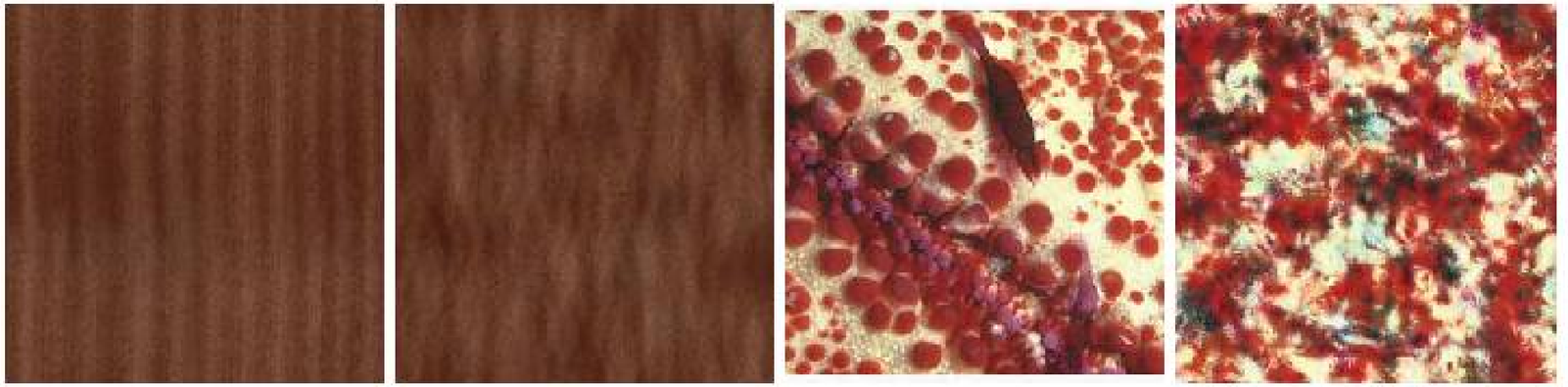


Figure 8: Examples of failures: wood grain and red coral.



Figure 9: More failures: hay and marble.

Simoncelli & Portilla ’98+

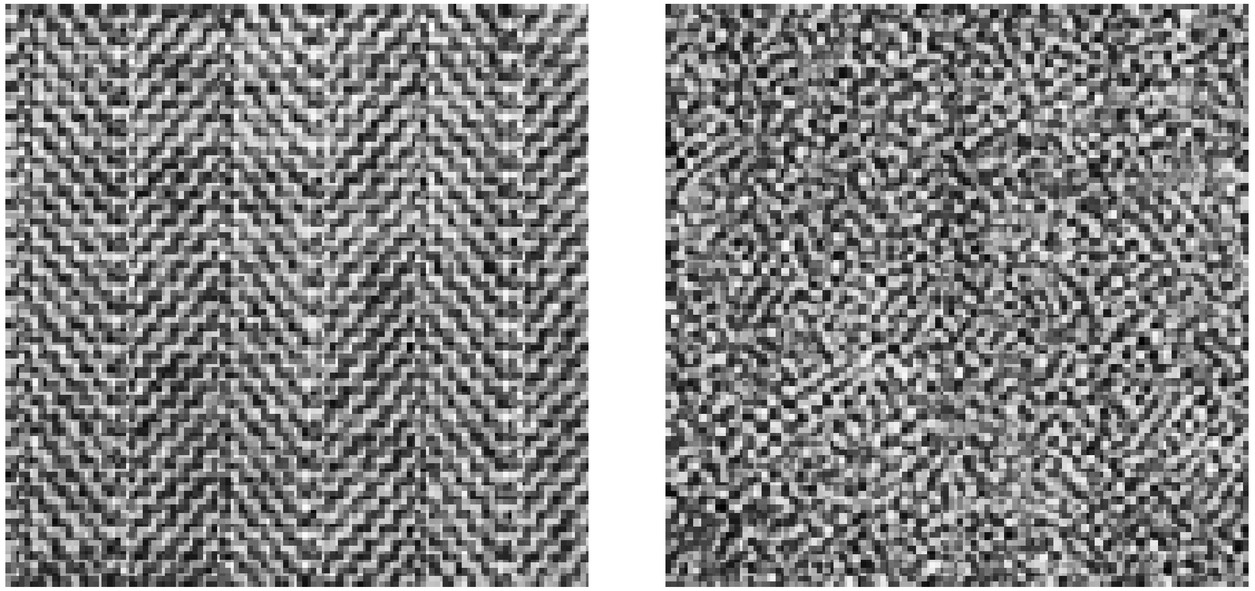


Figure 1. Textures with matching marginal statistics.

Marginal statistics are not enough

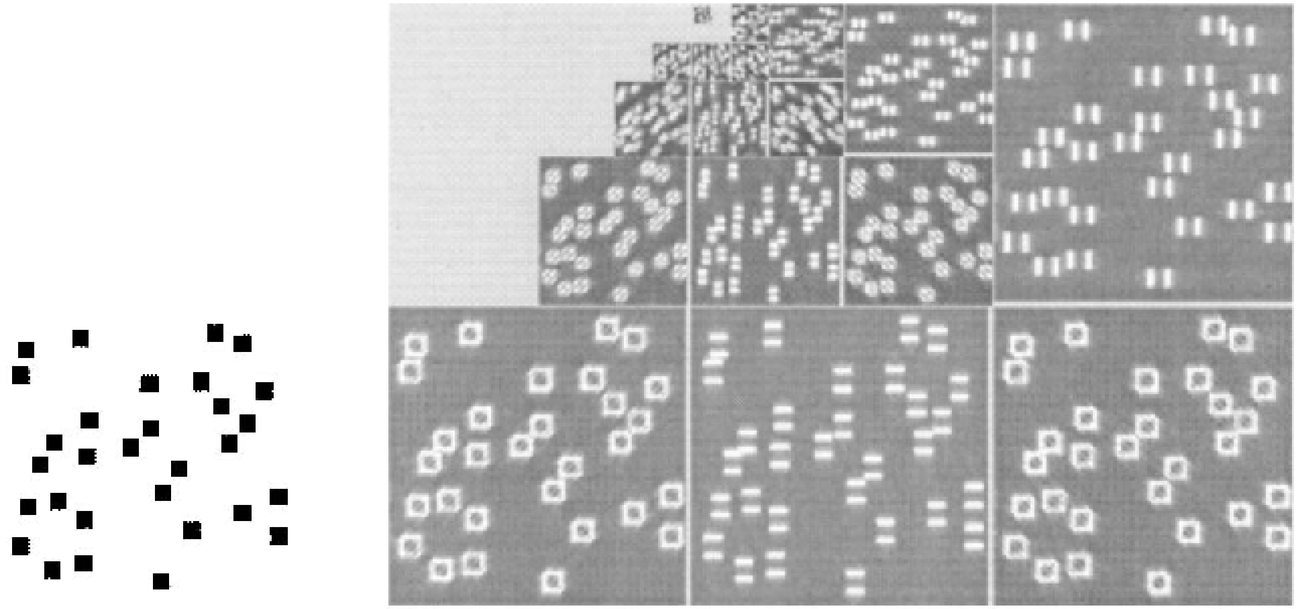
Neighboring filter responses are highly correlated

- an edge at low-res will cause an edge at high-res

Let’s match 2nd order statistics too!

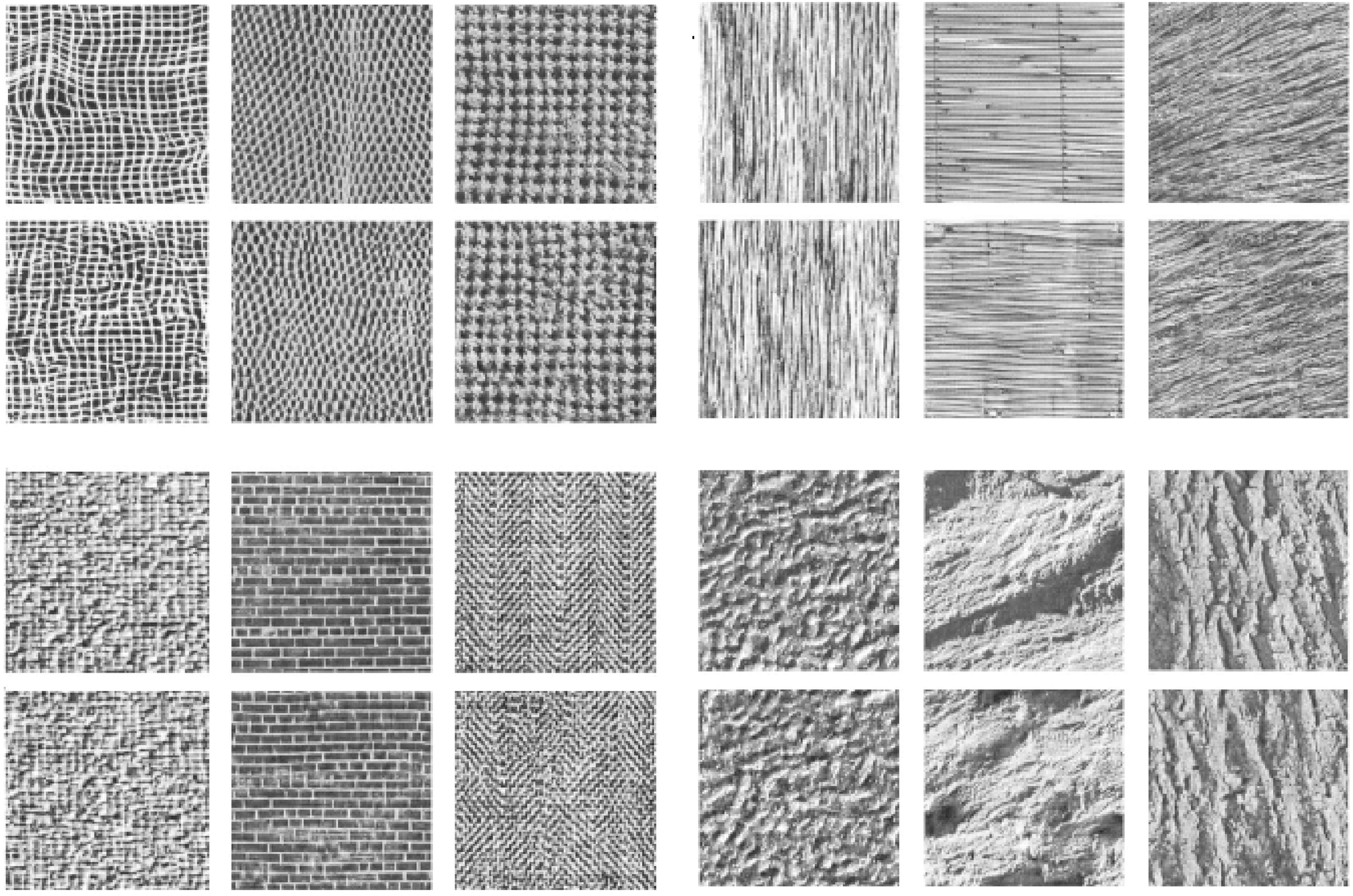
J Portilla and E P Simoncelli. *A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients.* Int'l Journal of Computer Vision. 40(1):49-71, October, 2000.

Simoncelli & Portilla ’98+

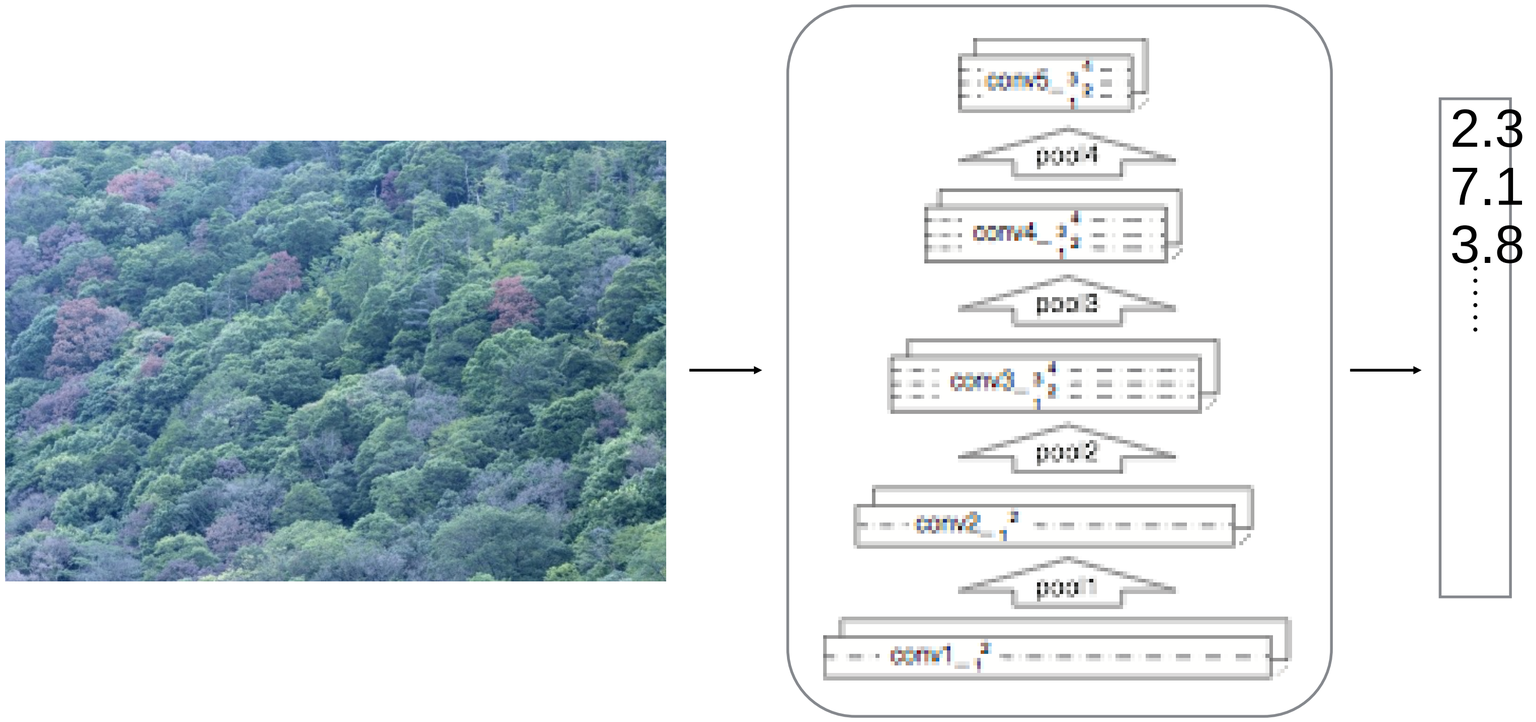


Match joint histograms of pairs of filter responses at adjacent spatial locations, orientations, and scales.

Optimize using repeated projections onto statistical constraint sufraces



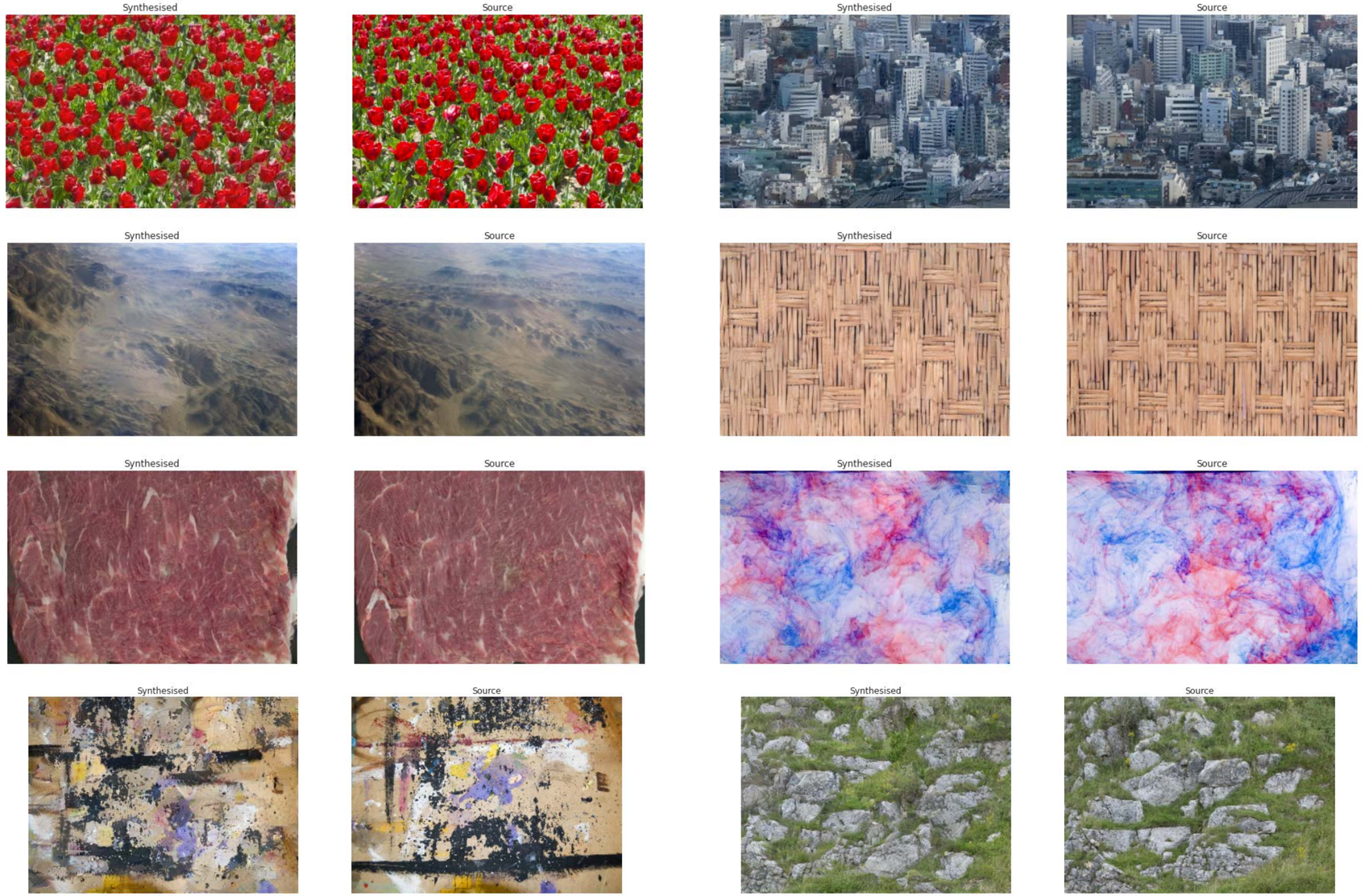
Convolutional Neural Network Texture Model



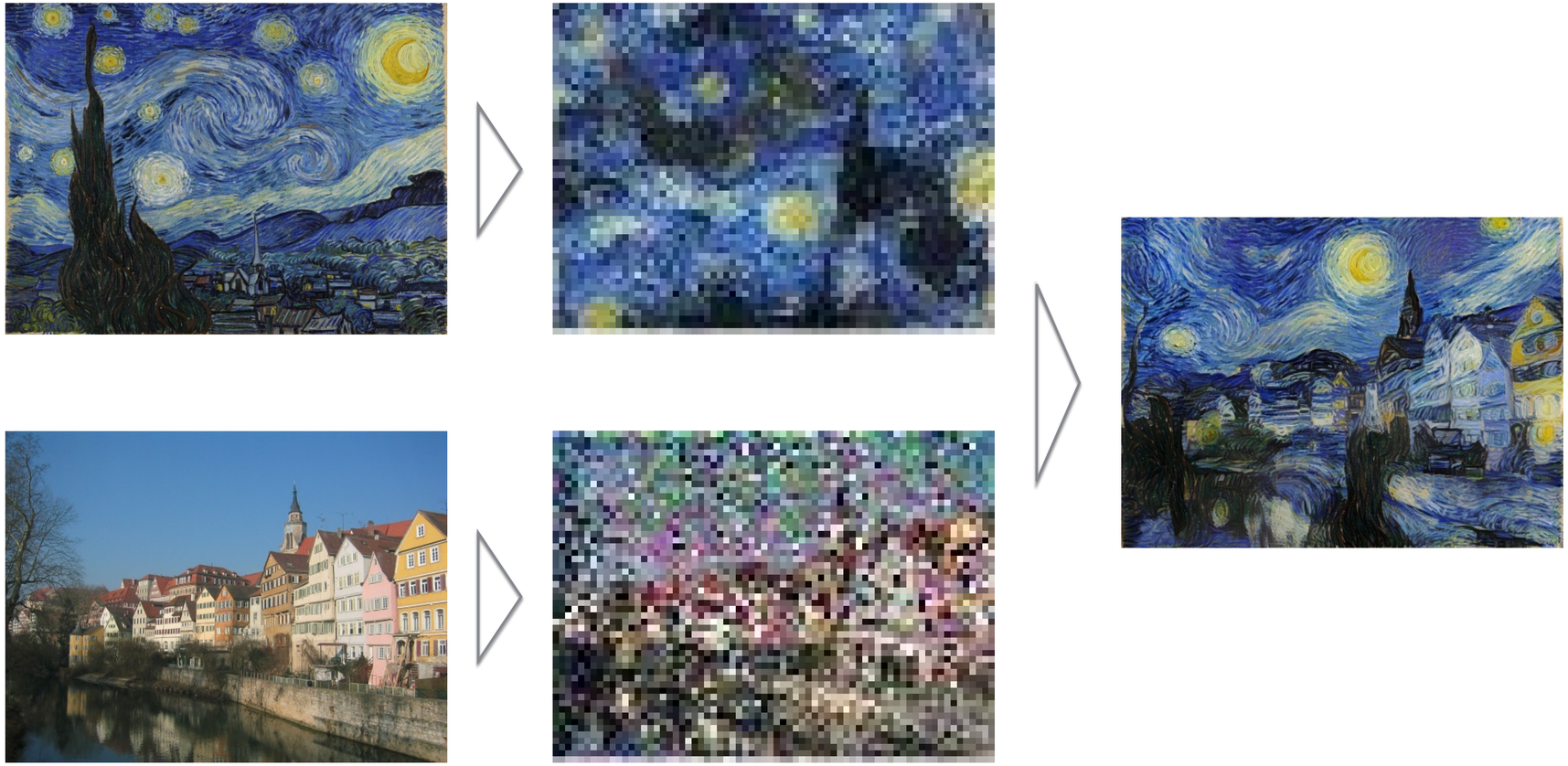
Convolutional Neural Network

Gatys et al. (NIPS 2015)

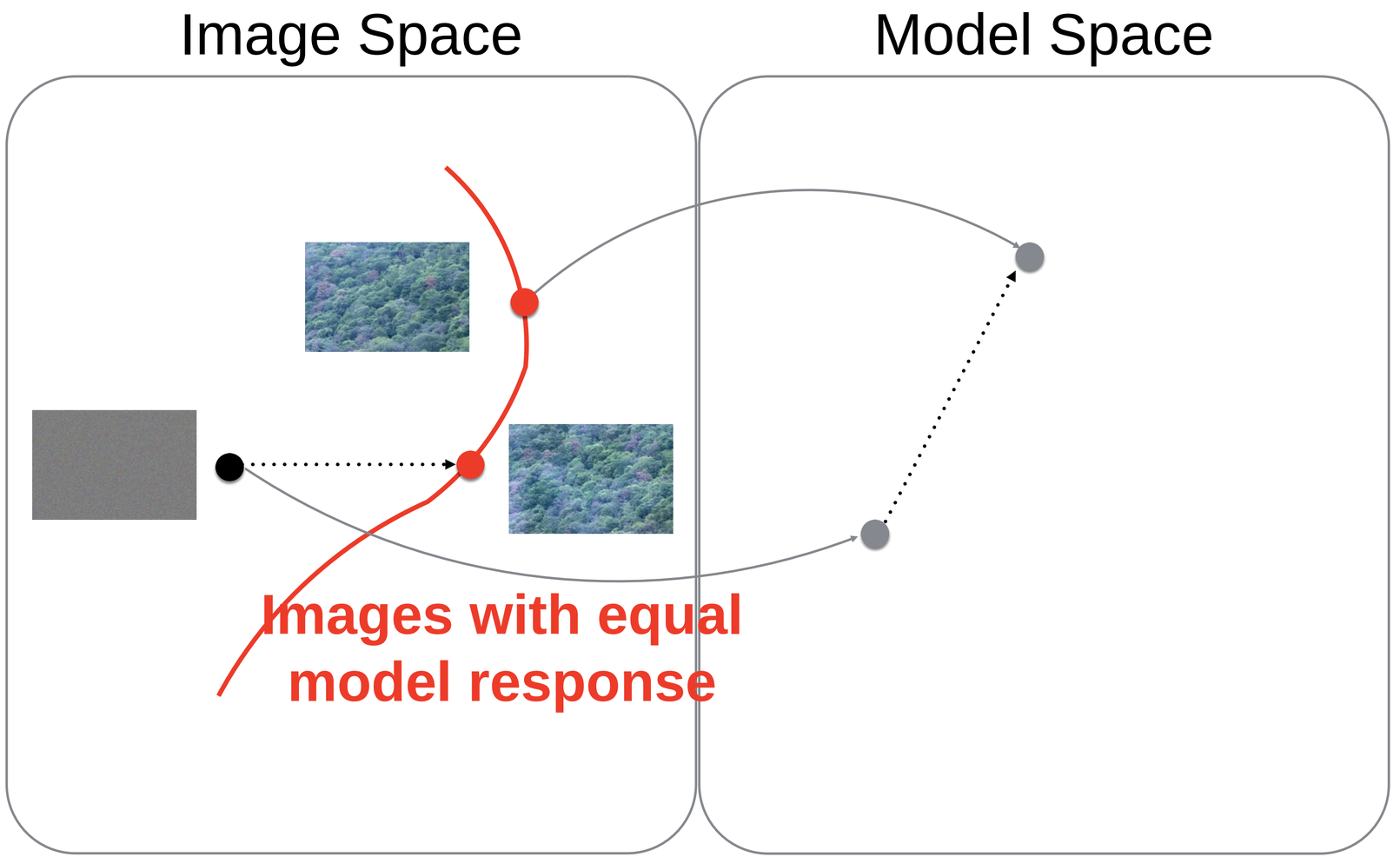
Gatys texture synthesis



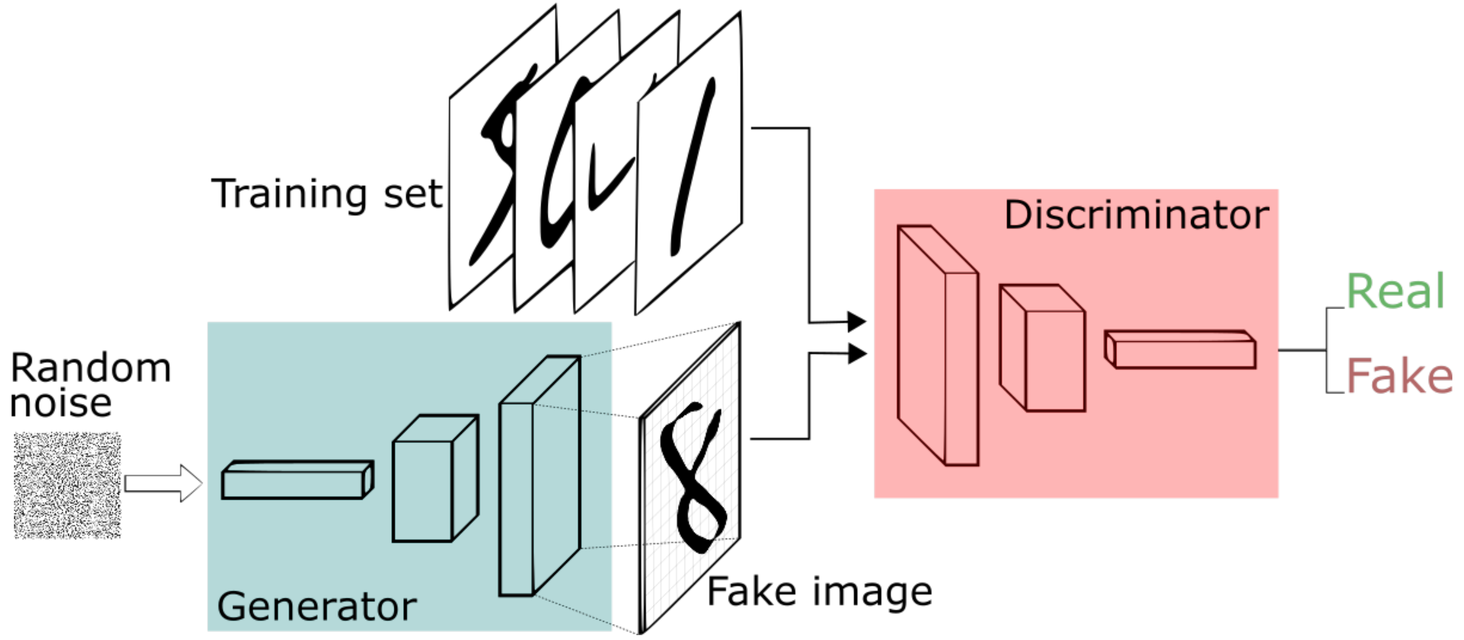
Gatys Texture Transfer



GANs?



Slide by Portilla & Simoncelli (2000)



- **Conjecture:** GANs might be learning the “right” features to match for natural images



6. GANs as a new artistic tool

GitHub - junyanz/pytorch

GitHub, Inc. [US] | https://github.com/junyanz/pytorch-CycleGAN-and...

Apps | Research | Paris Stuff | Кино онлайн беспл | teaching | Bay Area Stuff | Other bookm

Features Business Explore Marketplace Pricing Search / Sign in or Sign up

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

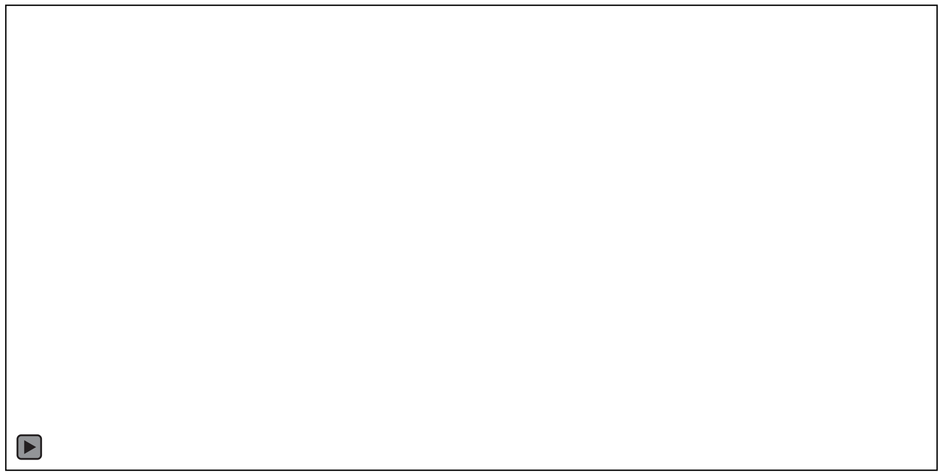
223 commits | 3 branches | 0 releases | 26 contributors

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

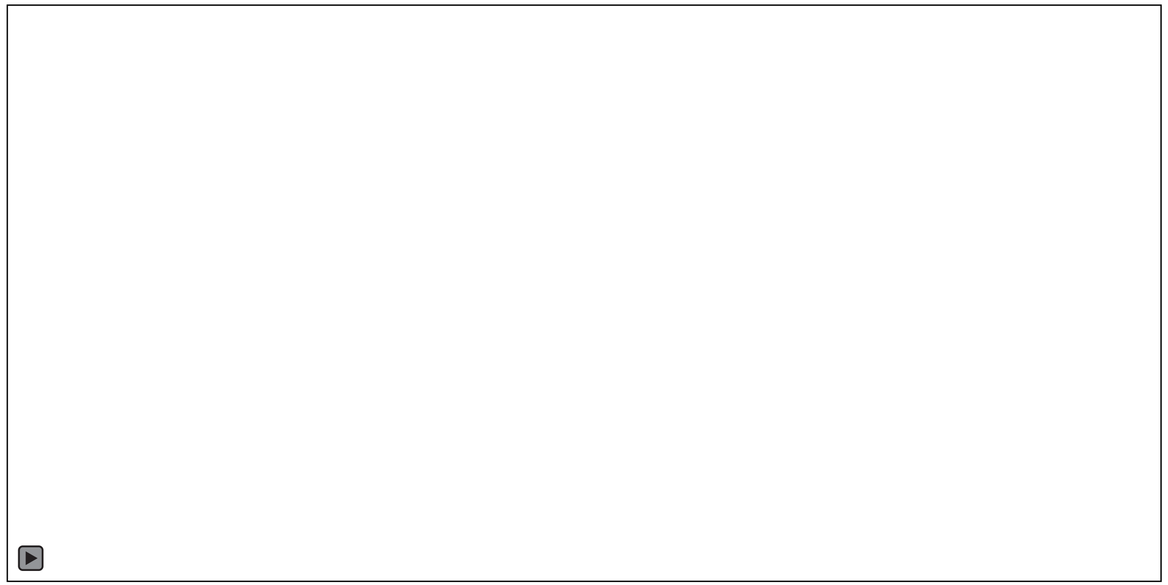
taesung89 Update README.md — Latest commit 6d9e173 10 Jun 13, 2018, 12:04 AM PDT

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

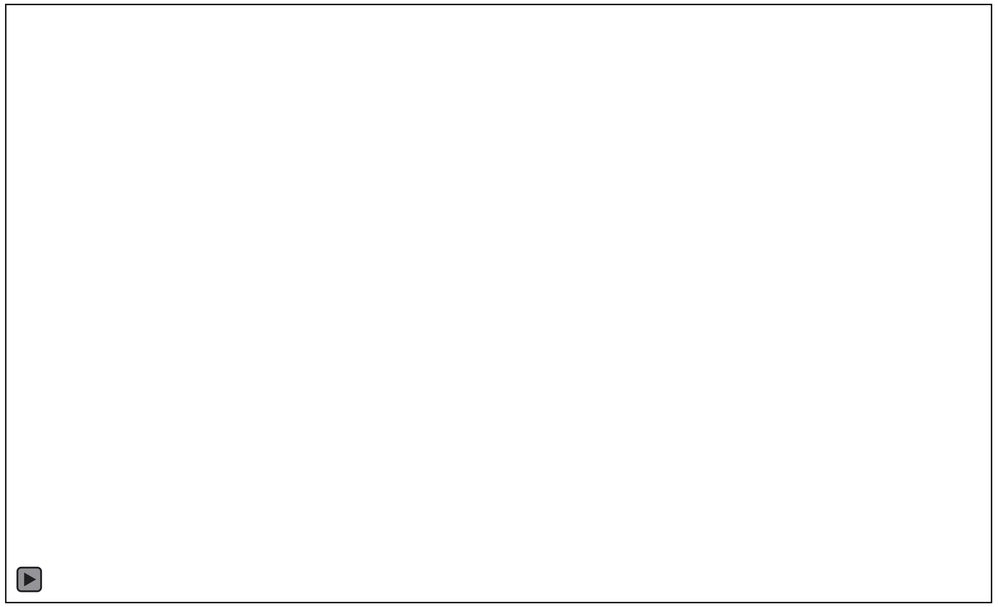
Twitter-driven research: #pix2pix



Mario Klingemann @quasimondo

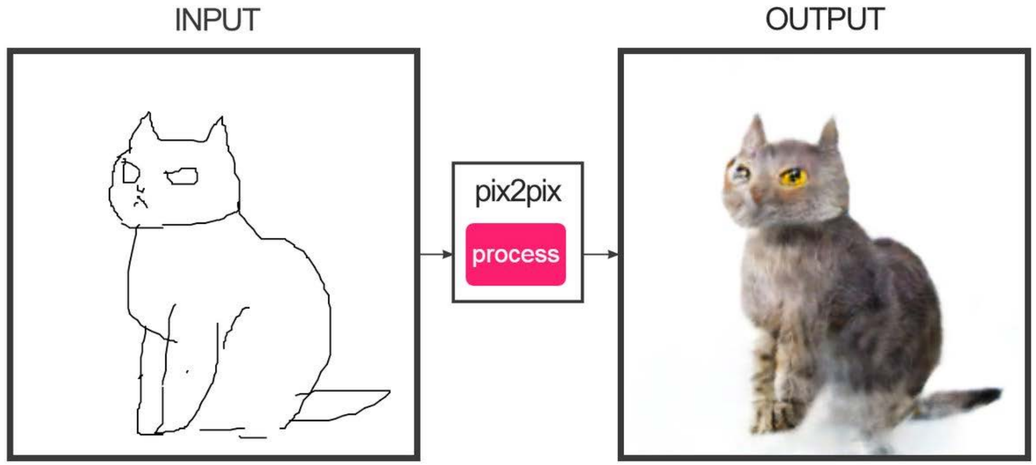


Brannon Dorsey @brannondorsey

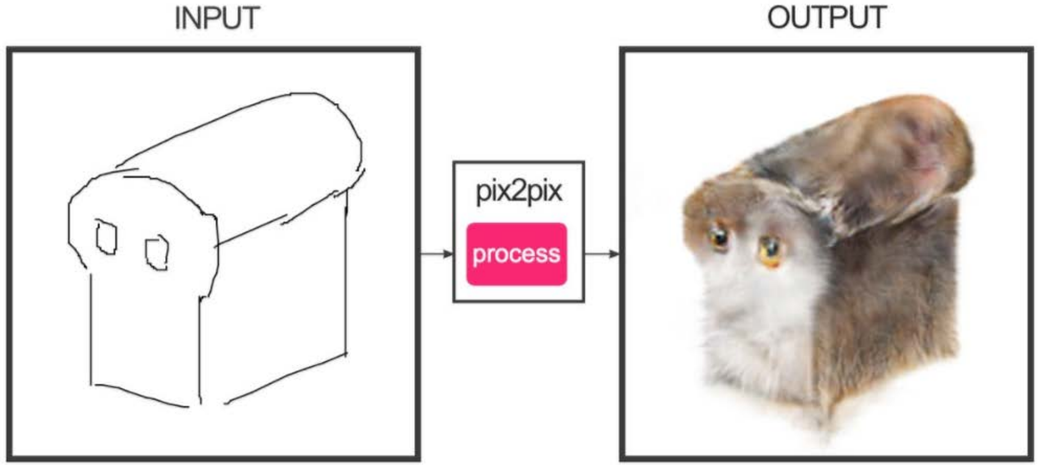


Bertrand Gondouin @bgondouin

#edges2cats [Christopher Hesse]



@gods_tail



Ivy Tasi @ivymyt

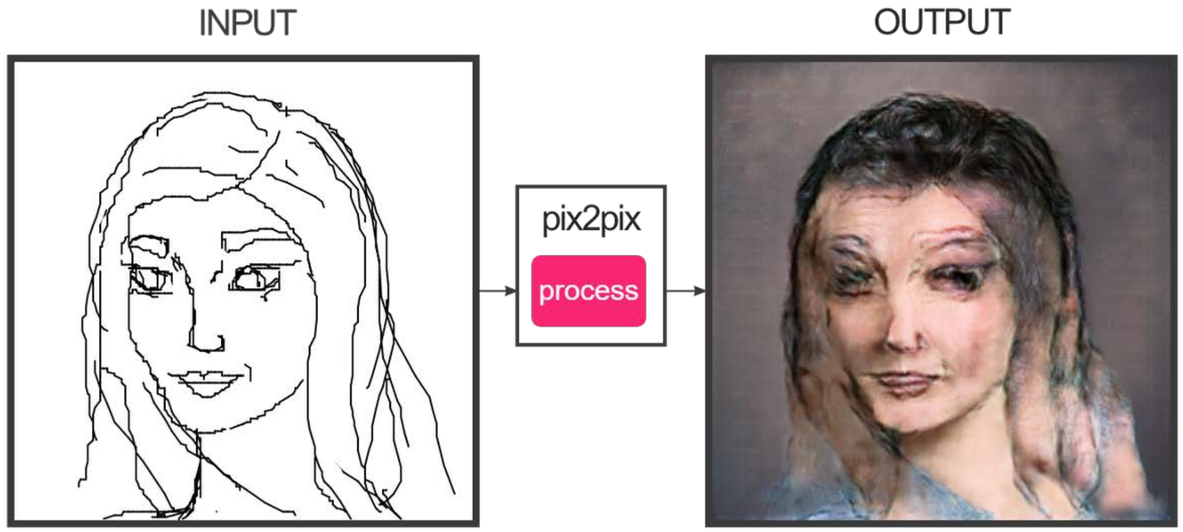


Vitaly Vidmirov @vvid

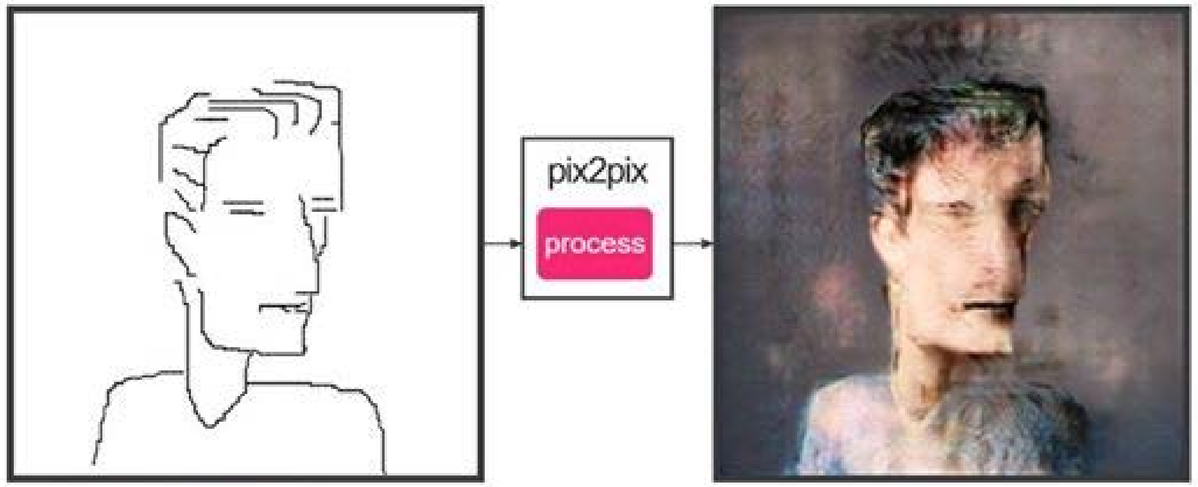


@ka92

sketch2face



Аня Кы @pure_memology



© Yann LeCun

- **Video game stylization** (cyclegan)
  - minecraft-->cezanne
  - fortnite-->pubg
- **Generating videos by predicting next frame**
  - https://twitter.com/matttrent/status/966584388883378177
  - https://twitter.com/quasimondo/status/948244469853360128
  - https://twitter.com/quasimondo/status/1000450809904222210t
- **Mario Klingemann**
  - https://twitter.com/quasimondo/status/985791855316078592
  - https://twitter.com/quasimondo/status/984704115262369792 (many more...)
- **Helen Sarin**
  - (trains cyclegan on her art, then uses it to make more art in her style)
  - https://twitter.com/glagolista/status/976876475810017280https://www.flickr.com/photos/tarelki/sets/72157694589553071

- **Memo Akten**
  http://www.memo.tv/category/work/by-type/ (uses pix2pix)
- **Jason Salavon**
  - https://twitter.com/Salavon/status/92536601762637824O
  - https://latentculture.com/genmo/
- **Martinez Alonso**
  - http://nono.ma/suggestive-drawing (uses pix2pix)
- **Gene Kogan**
  - http://genekogan.com/
  - https://twitter.com/genekogan/status/941674205329928192
  - https://twitter.com/genekogan/status/941663238713892865



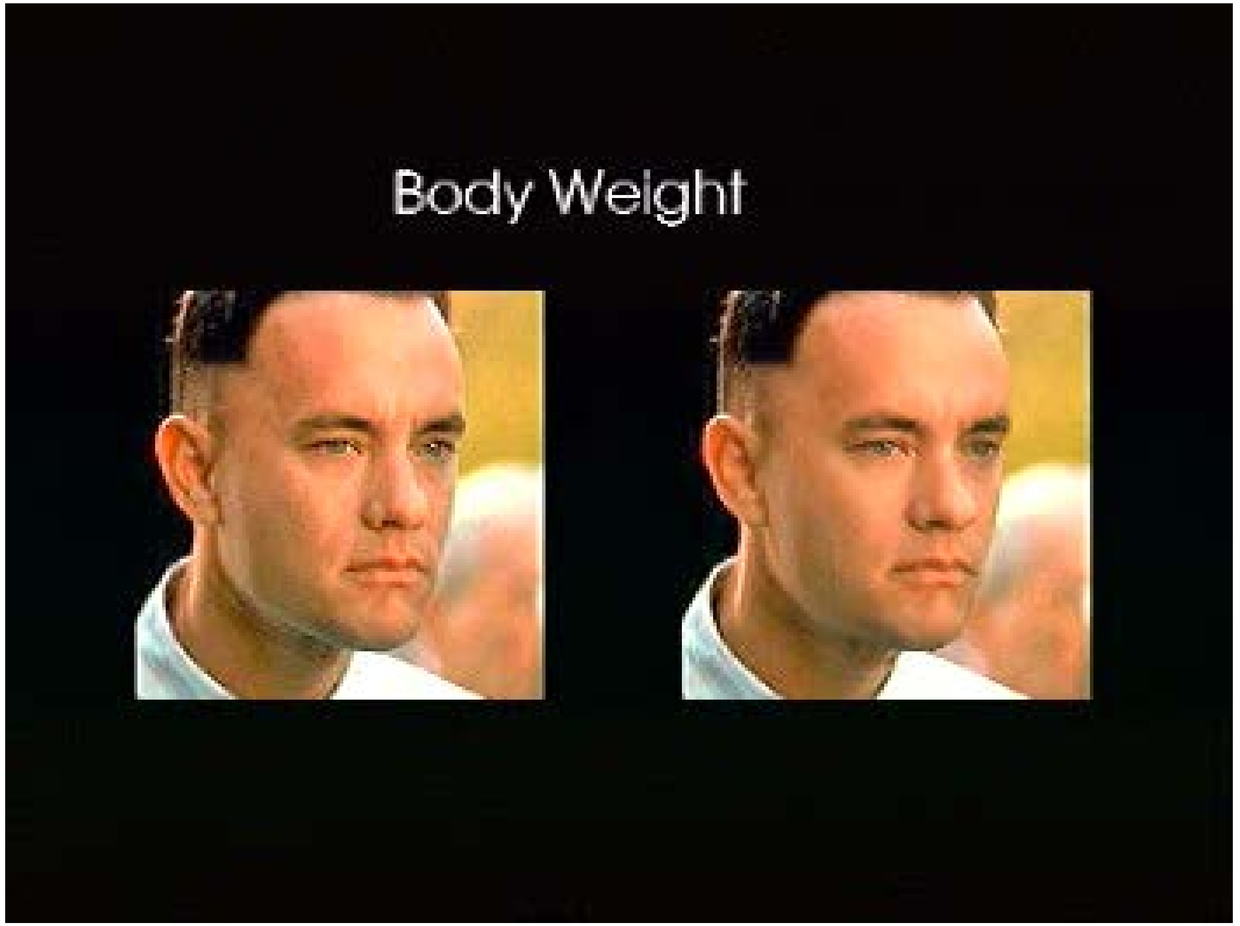
© Memo Akten
“Learning to See:
Gloomy Sunday”

Epiloge: GAN misconceptions

Nothing perceptual about GANs (yet)

- GANs just match distributions
- Happy to match JPEG artifacts before actual “realistic content”

GANs aren’t responsible for “DeepFakes” (yet)



GANs

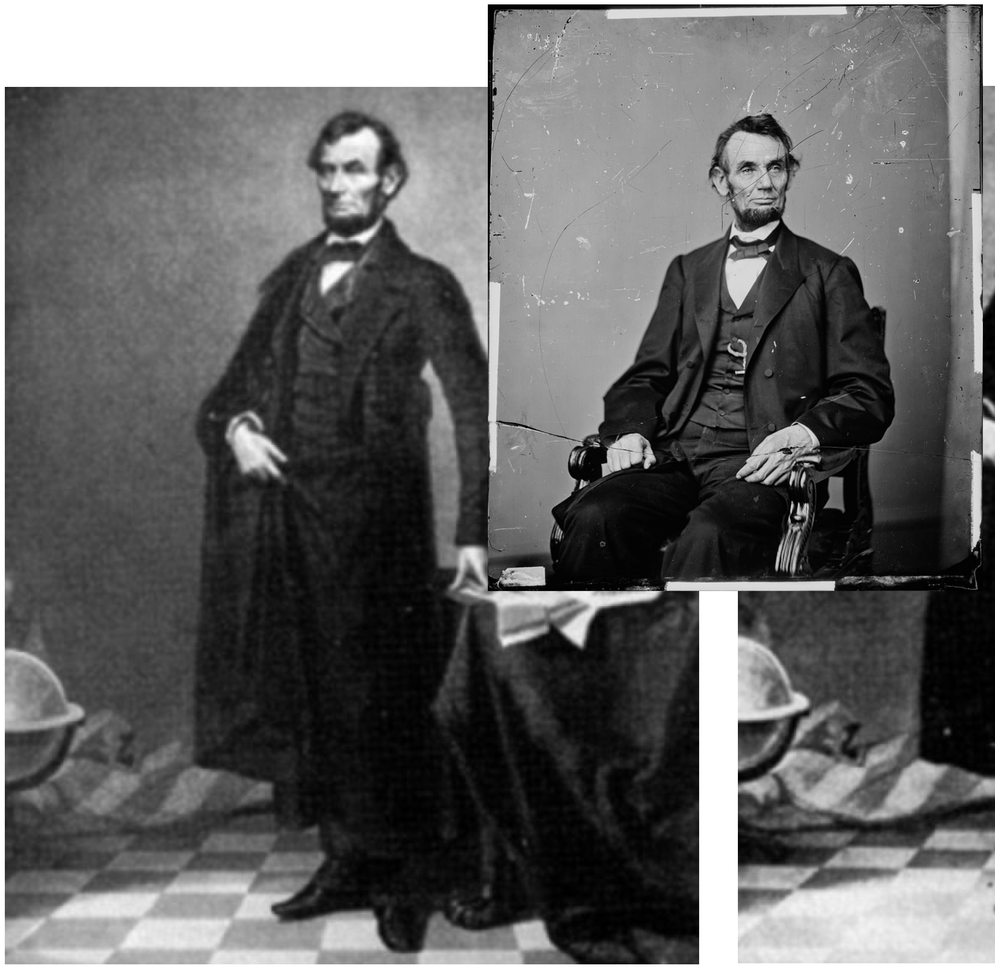


https://www.youtube.com/watch?v=PCBTZh41Ris&feature=youtu.be

photo manipulation is as old as photography



W.H. Mumler, “Mary Todd Lincoln with her dead husband”, circa 1869



classic Lincoln



John Calhoun



Thank You

