Is it REAL?
Challenges in Image Synthesis

Lihi Zelnik-Manor

EE, Technion

INTERACTING WITH VIRTUAL WORLDS

Does it Look REAL?

[Gatys, et al. CVPR 2016]


Does it Look REAL?
[Mechrez, et al. BMVC 2017]


Does it Feel REAL?

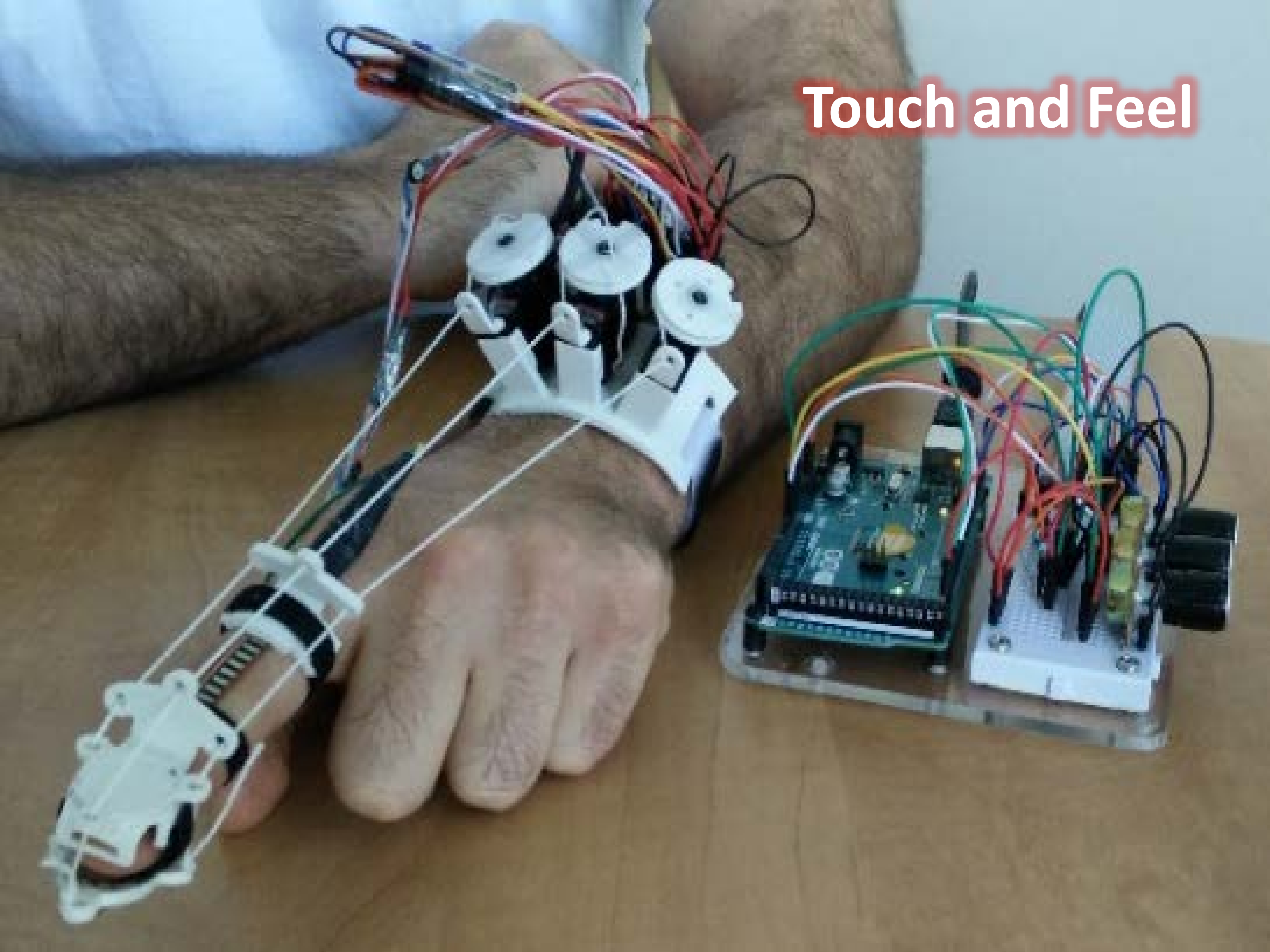

Touch and Feel

Image to Surface

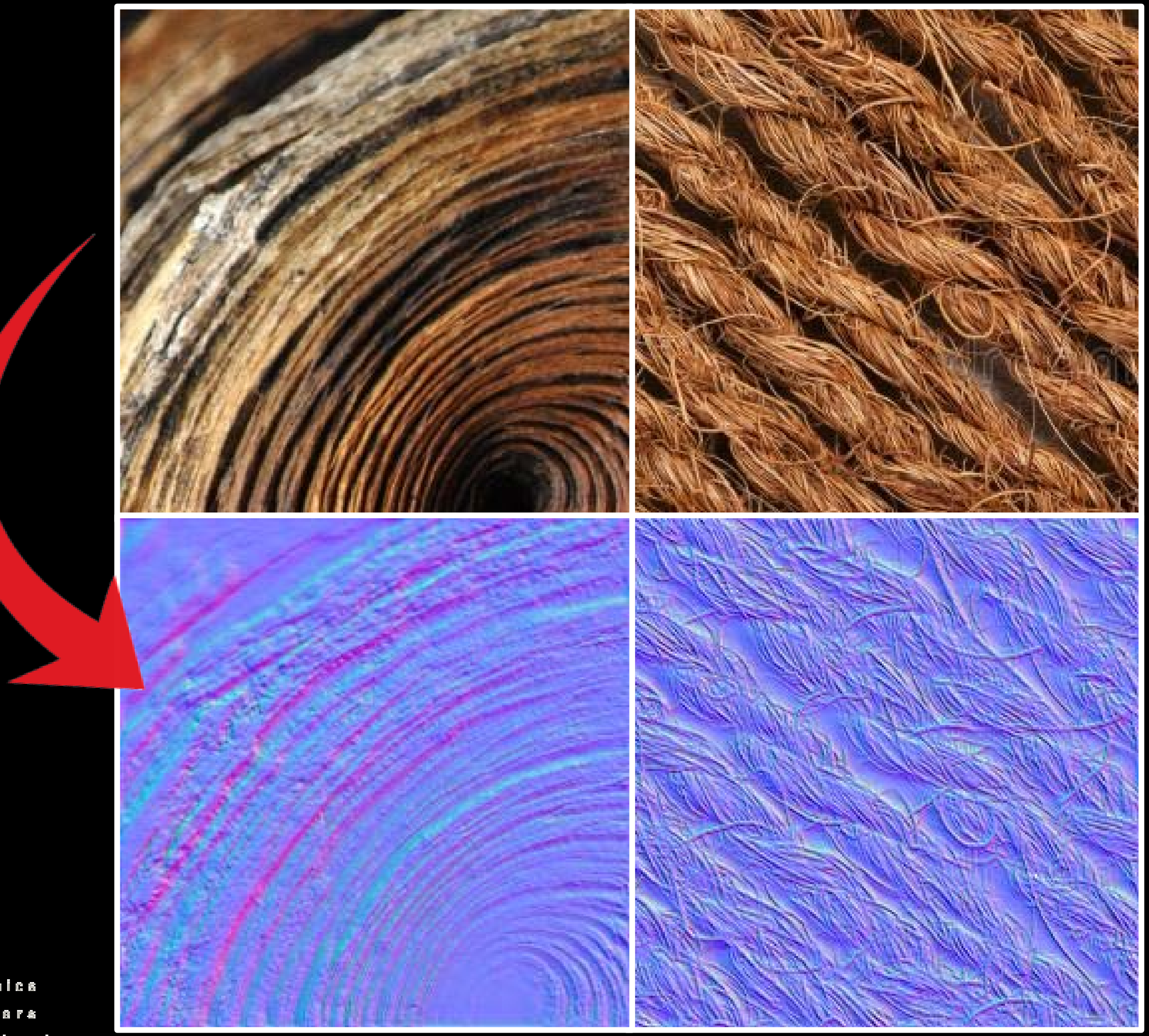
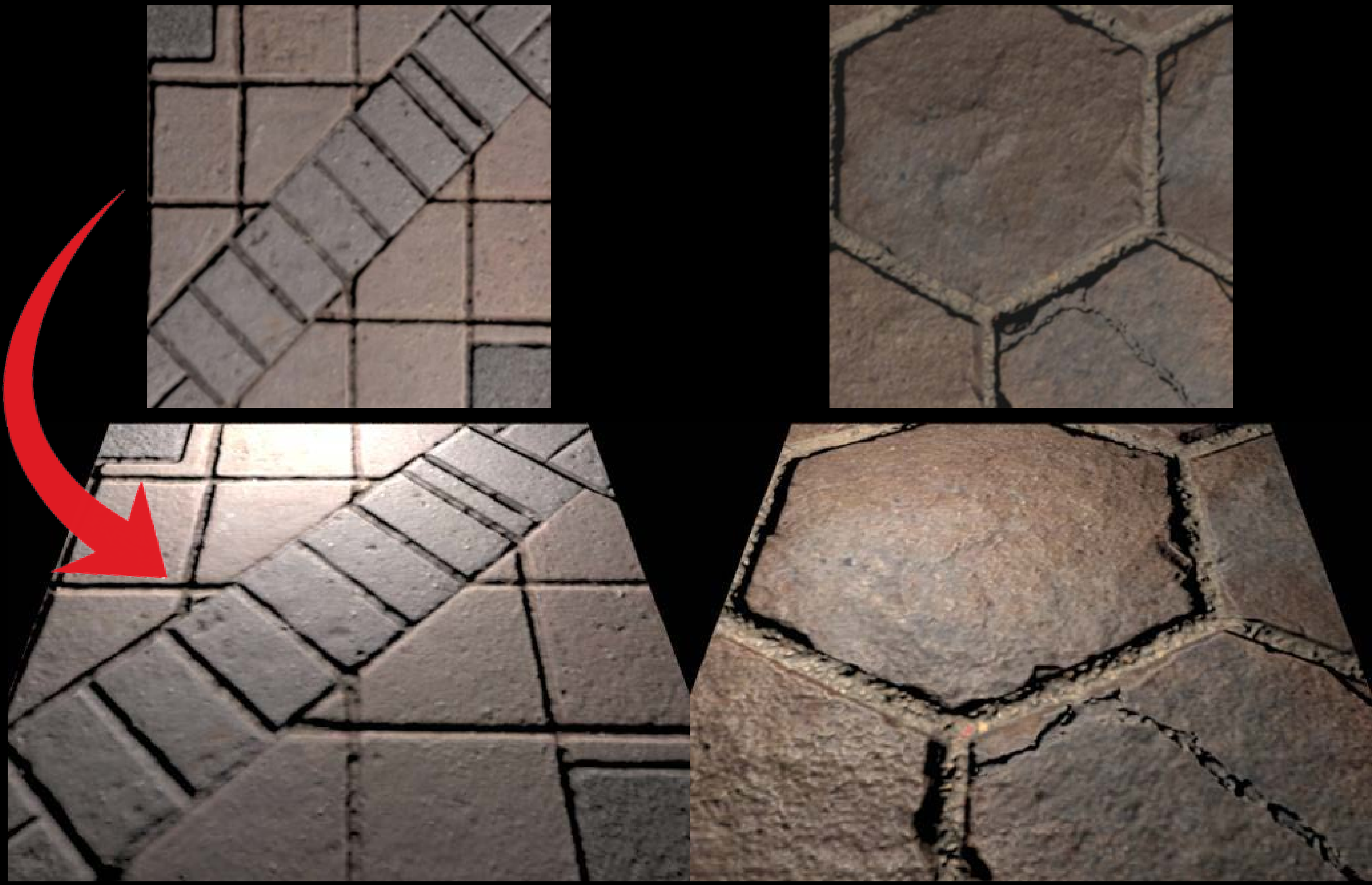


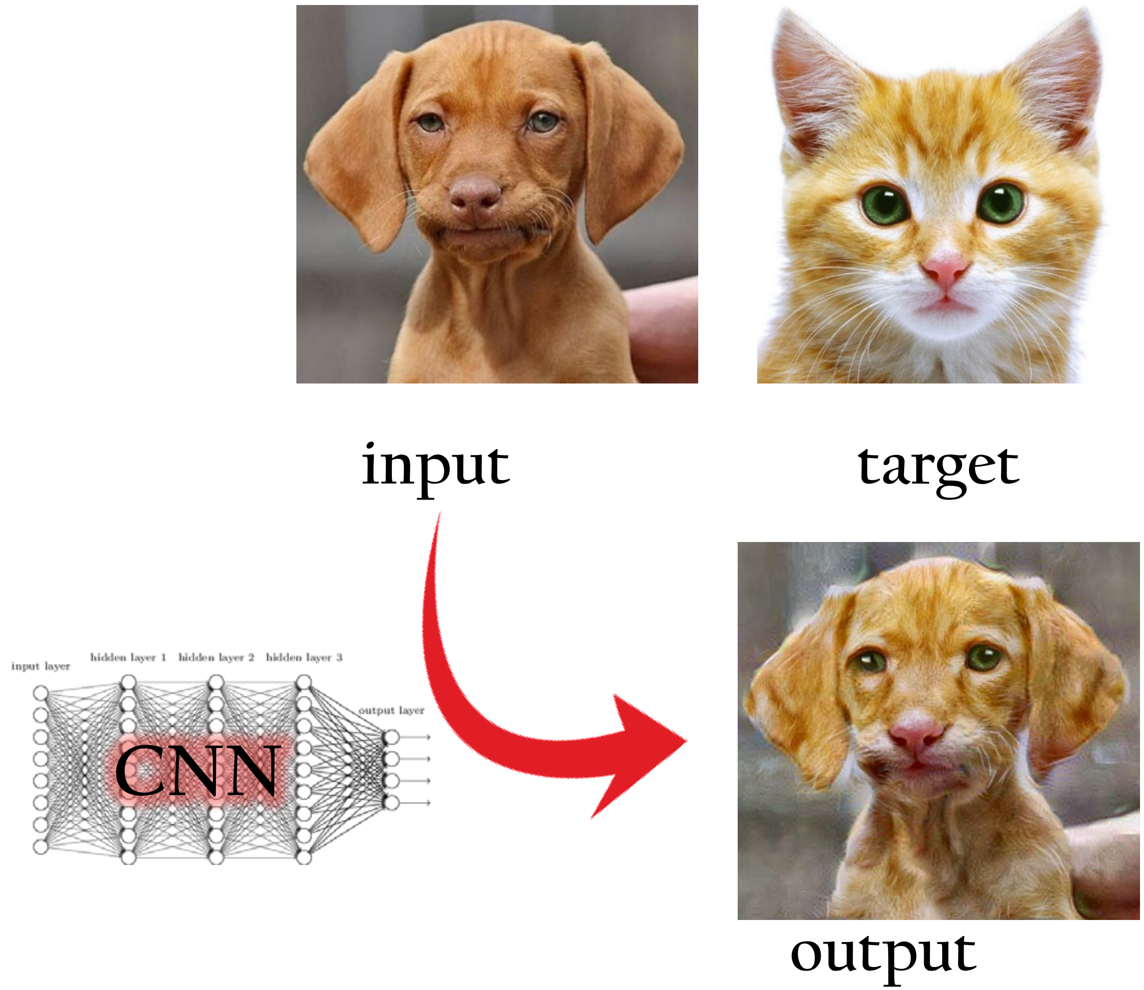
Image to Surface



Pushing the boundaries of CNNs

GENERATING IMAGES THAT LOOK REAL

Style Transfer as a Case Study

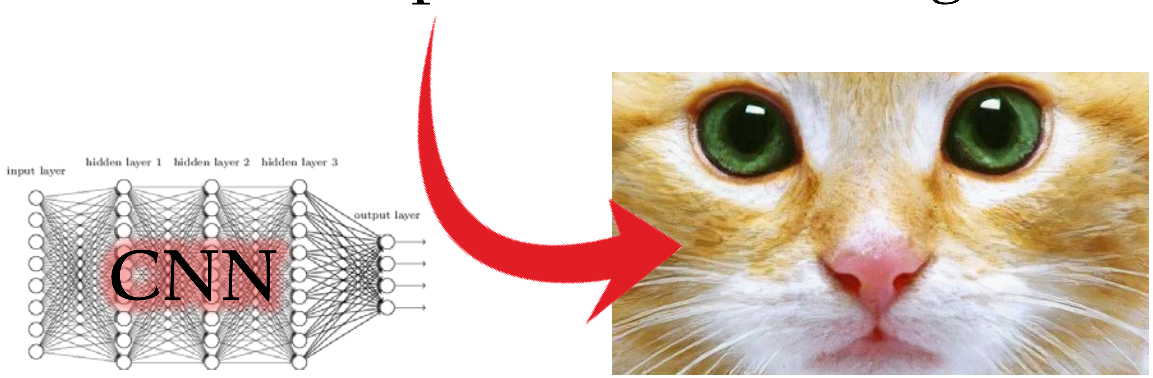


Electronics
Computers
Communications

Single-Image Super-Resolution

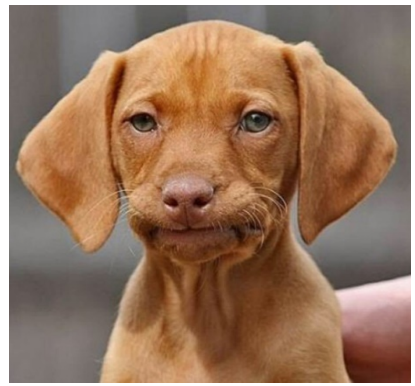
input

target



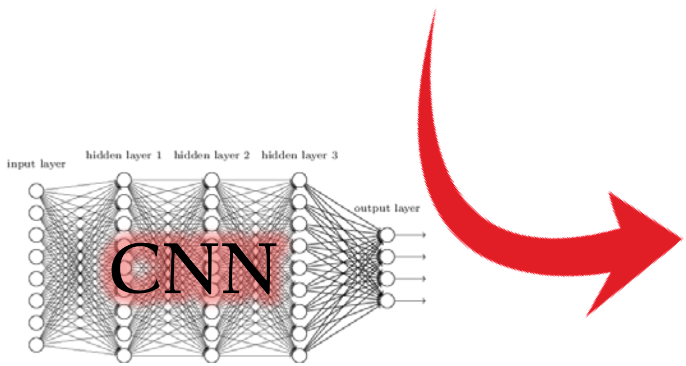
output

Style Transfer as a Case Study



input

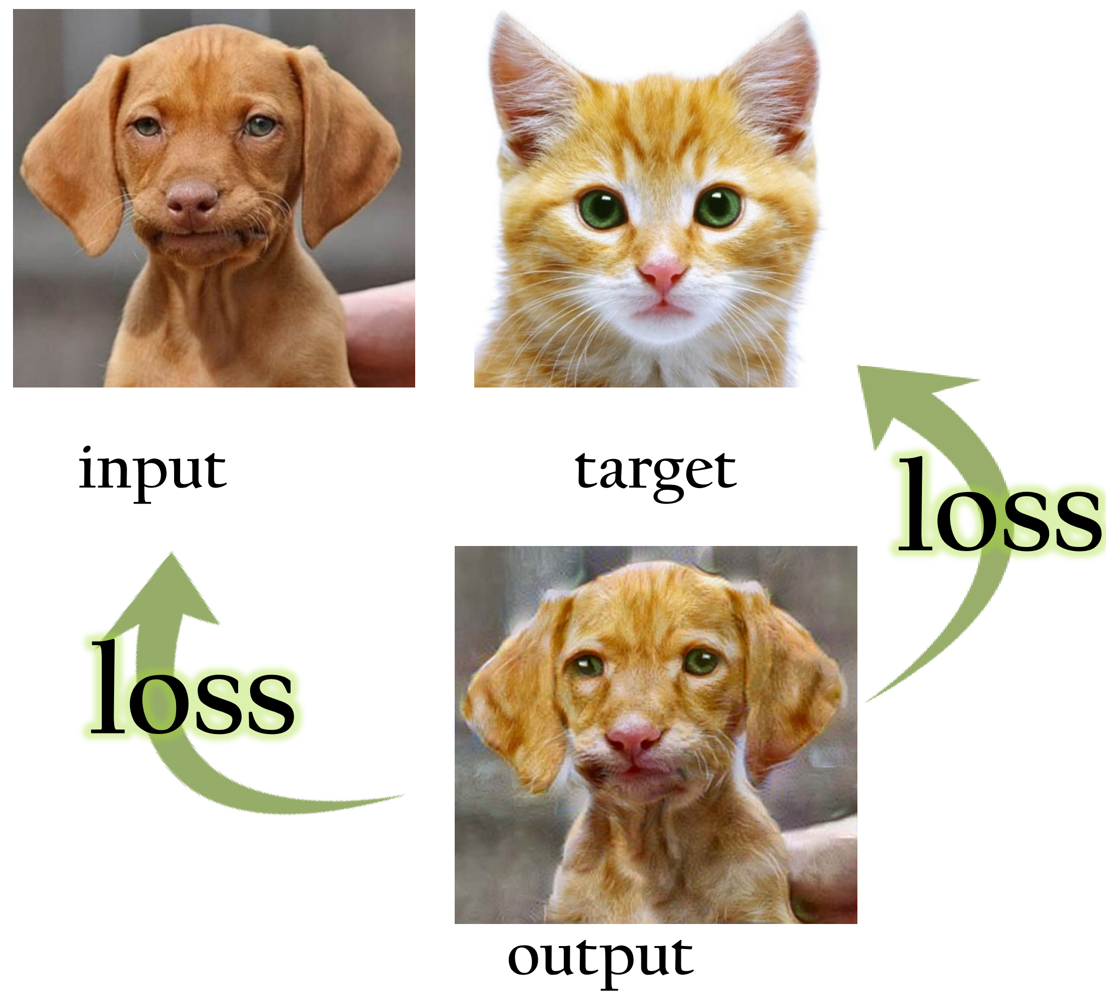
target



output

Solution with CNNs

It's all in the loss!!

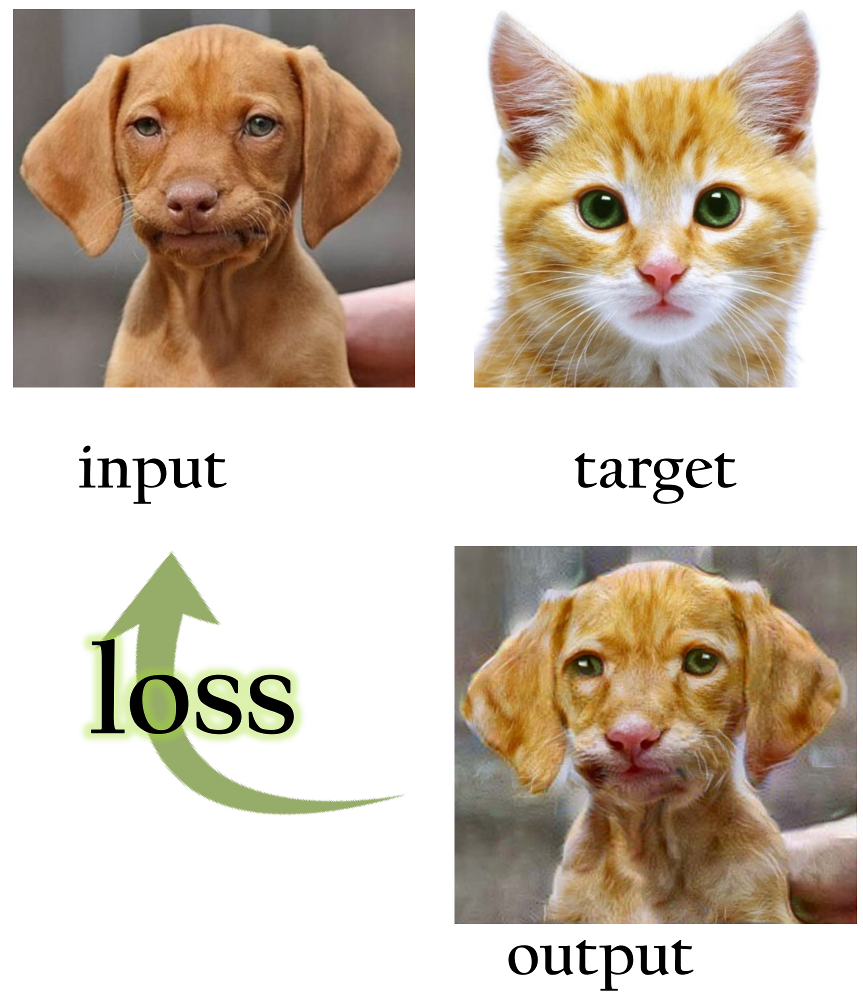


The current loss functions

<u>Pixel-to-pixel</u>

L1, L2, Perceptual, ...

Require alignment !!



The current loss functions

<u>Global style</u>

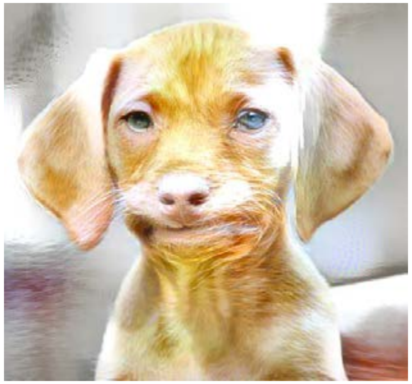
Gram

Global statistics !!

input

target

loss



Gram loss output

output

Wrong Semantics

input

target

loss

Global style

Gram

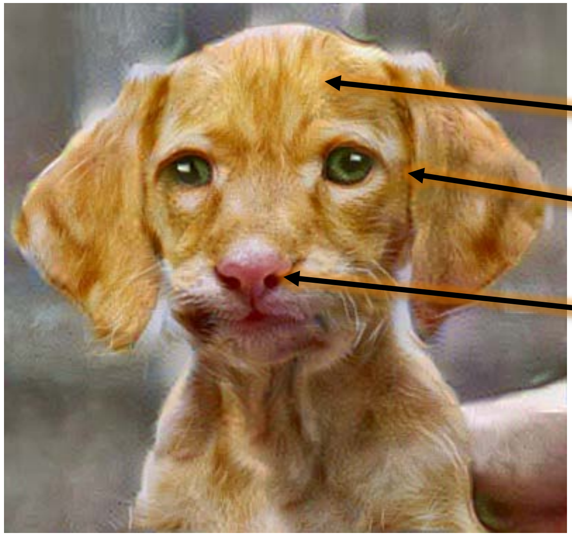
Global statistics !!

Gram loss
output

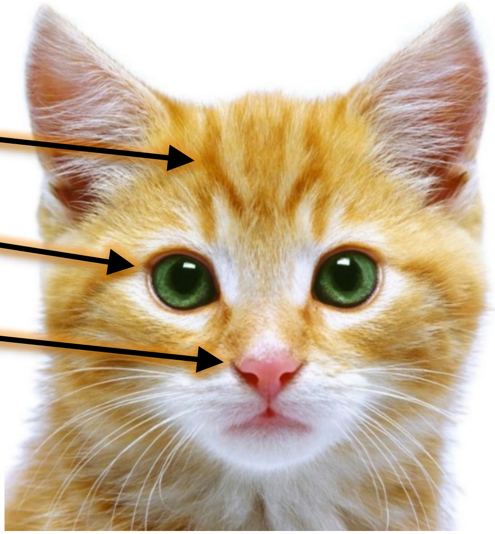
The goal

Compare regions based on semantics

(and NOT spatial position)

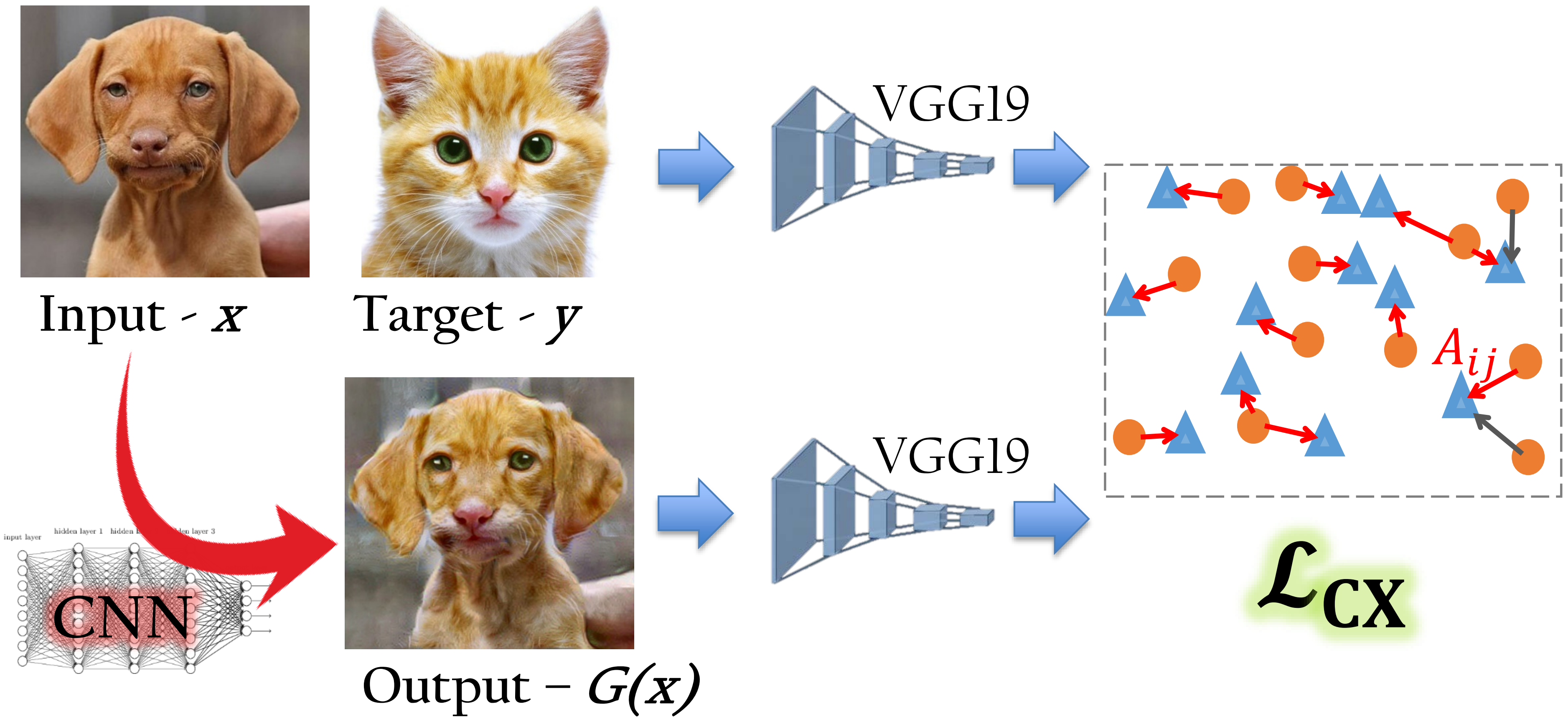


Generated image

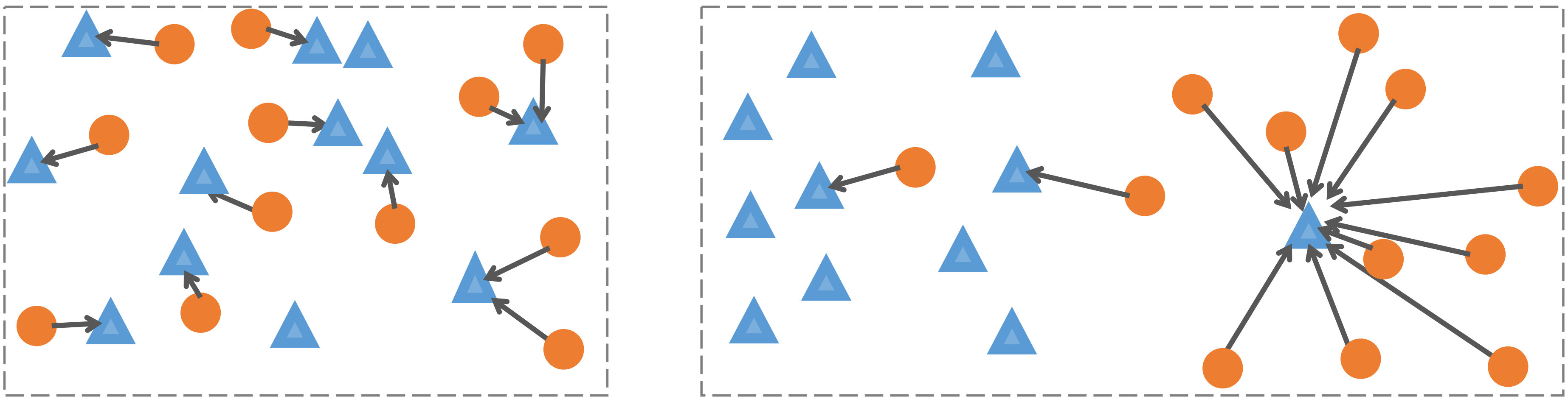


target

The Contextual loss



A Statistical Measure



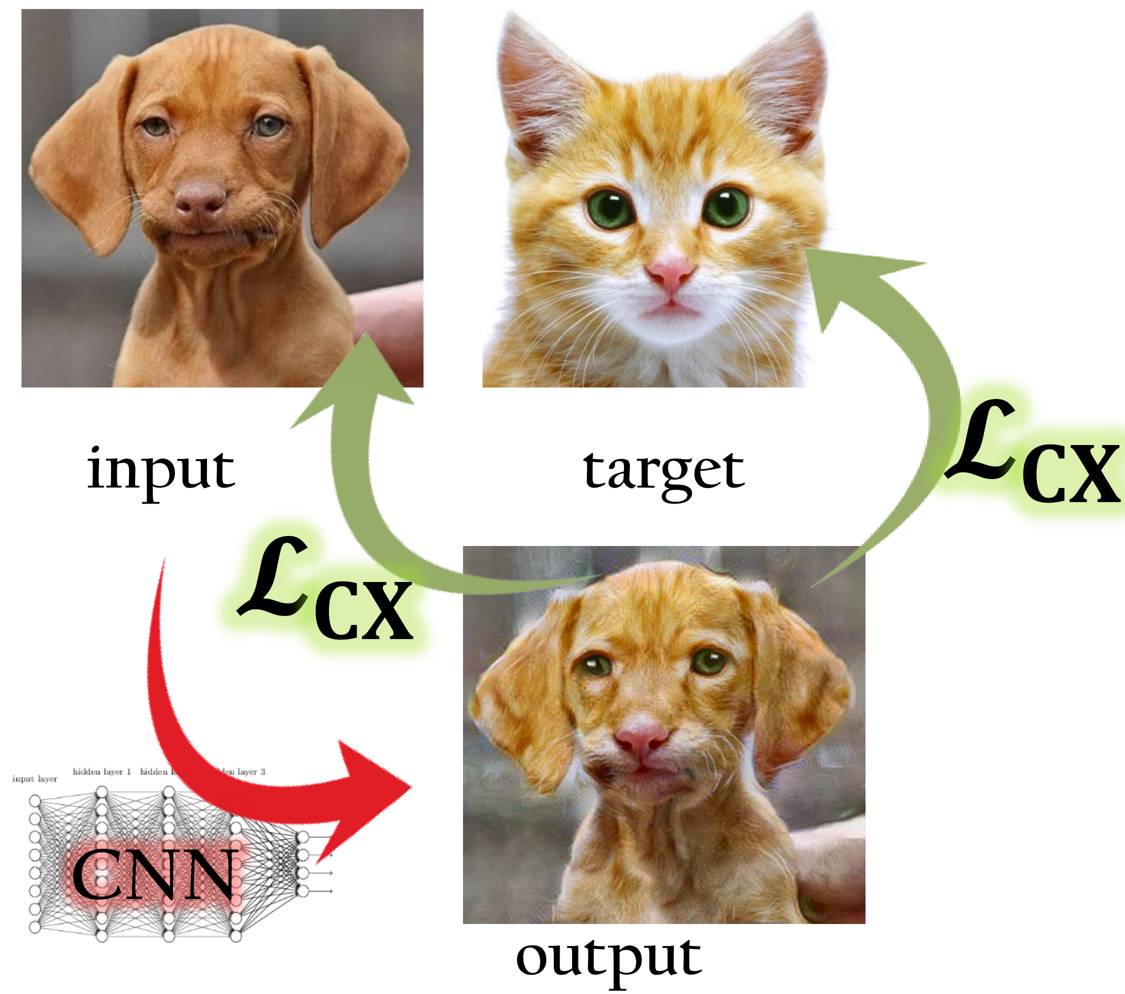
similar distributions

different distributions

Generated “patch”

Target “patch”

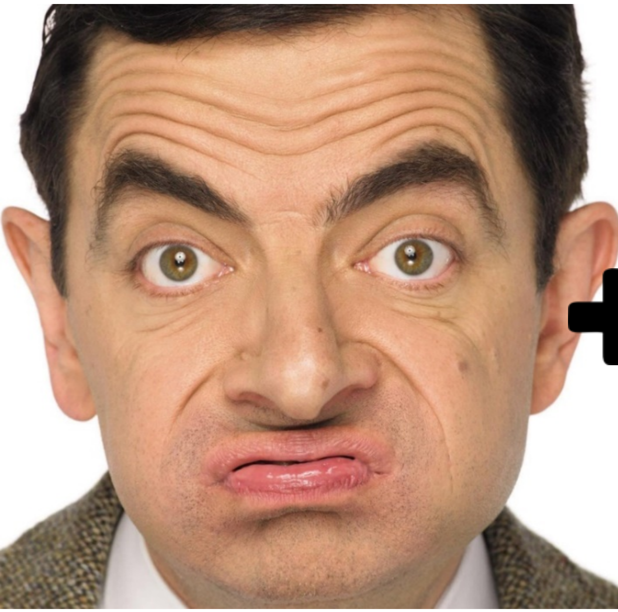
Semantic Style Transfer



Architecture:
Gatys (NS)

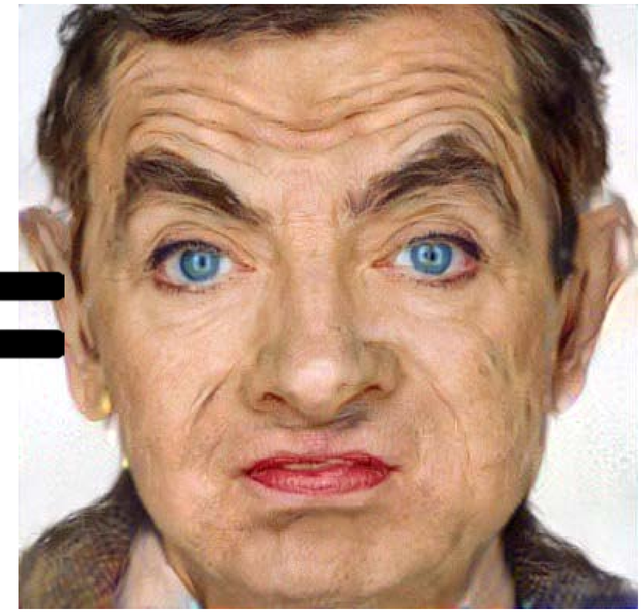
Electronics
Computers
Communications

Semantic Style Transfer



Input

Style



Mr. Beanary

Semantic Style Transfer

Input

Style

Obamary

Electronics
Computers
Communications

Semantic Style Transfer

Input

+

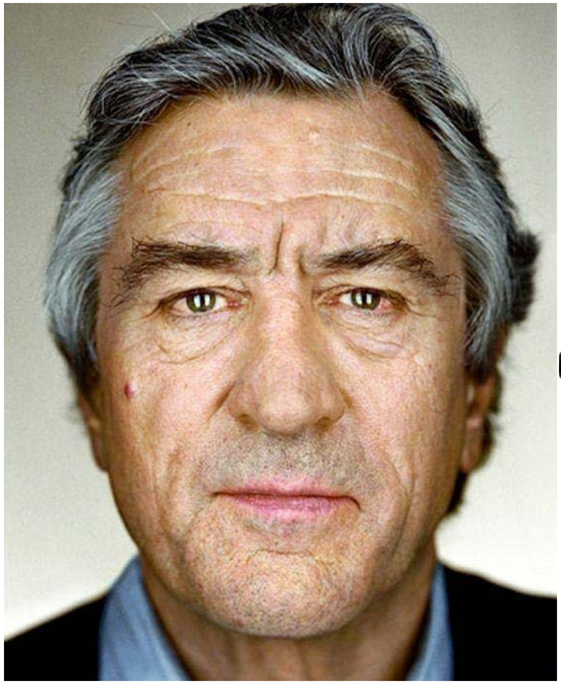
Style

=

Mr. Beanama

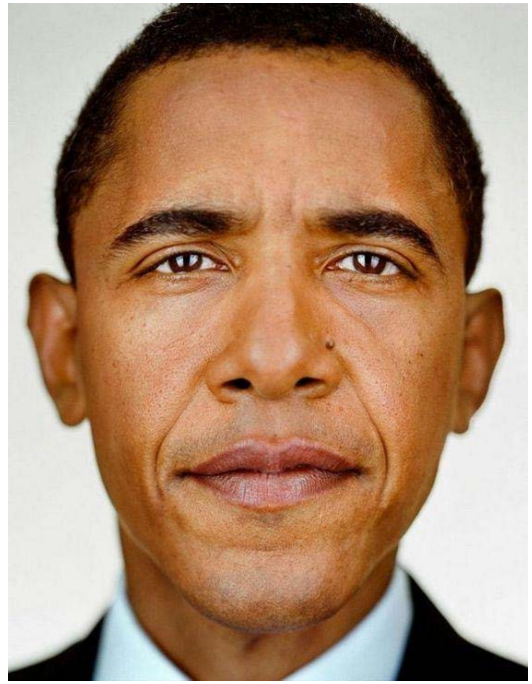
Electronics
Computers
Communications

Semantic Style Transfer



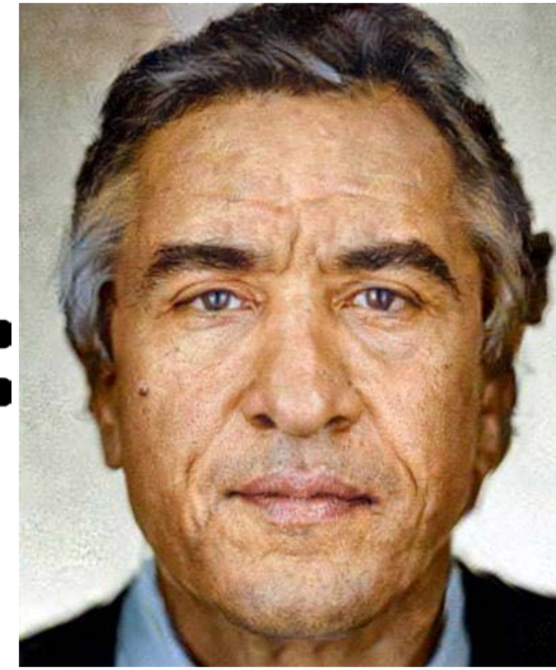
Input

+



Style

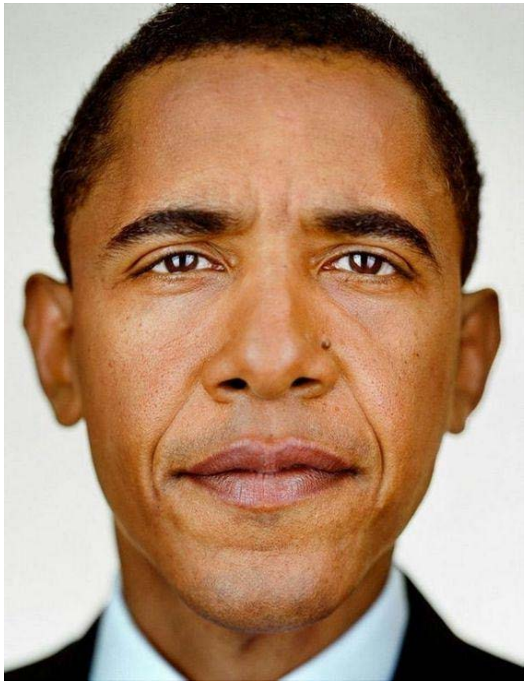
=



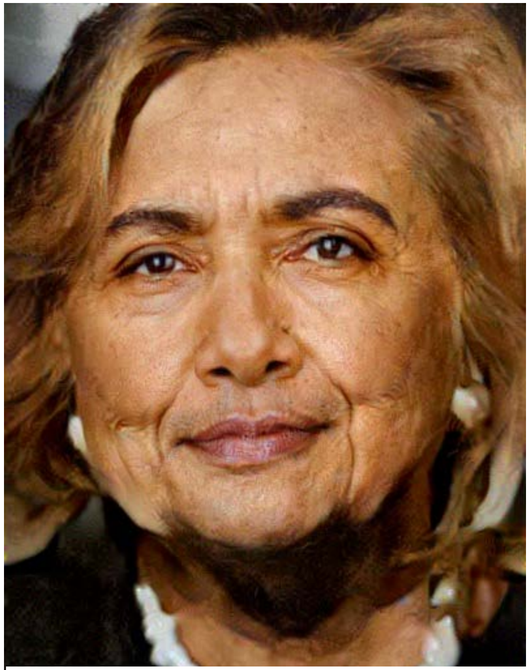
Barack De Niro

Semantic Style Transfer

Input

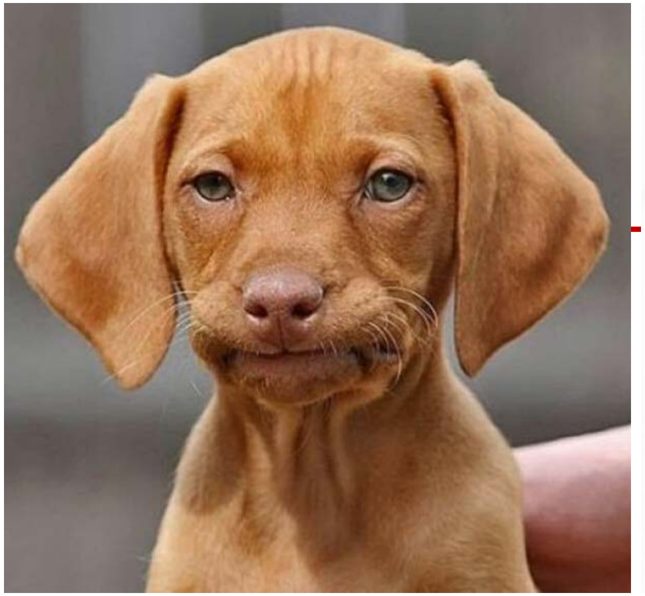


Style



Hillabama

Does it Look REAL?

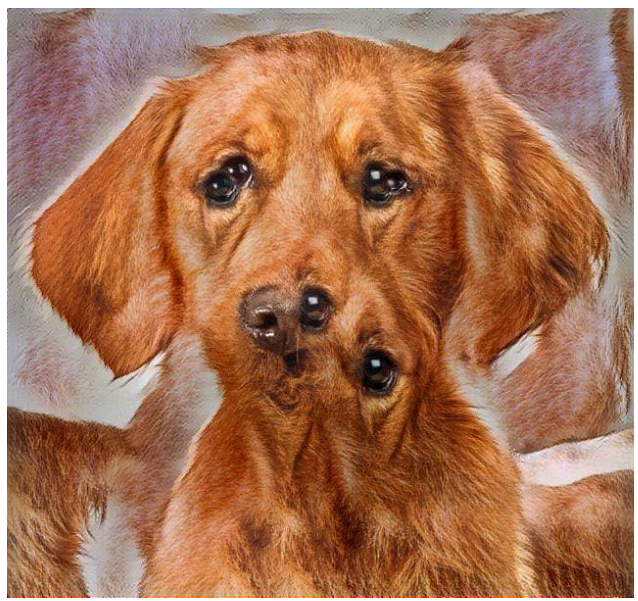


Input

Style

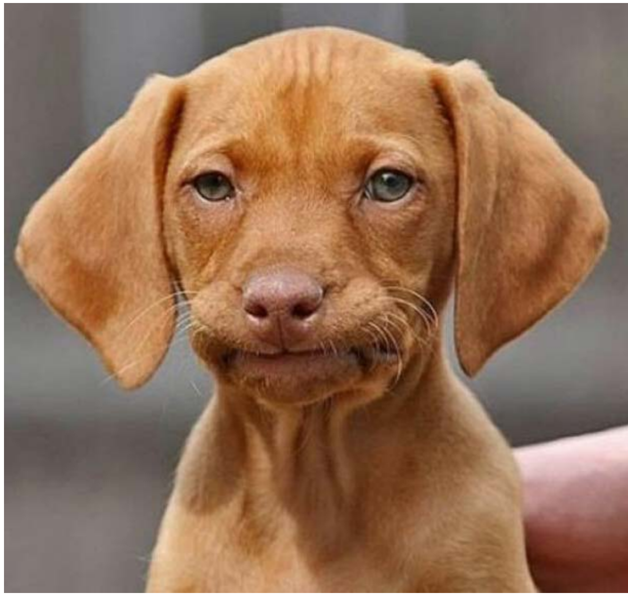
Output

Does it Look REAL?



CNNMRF

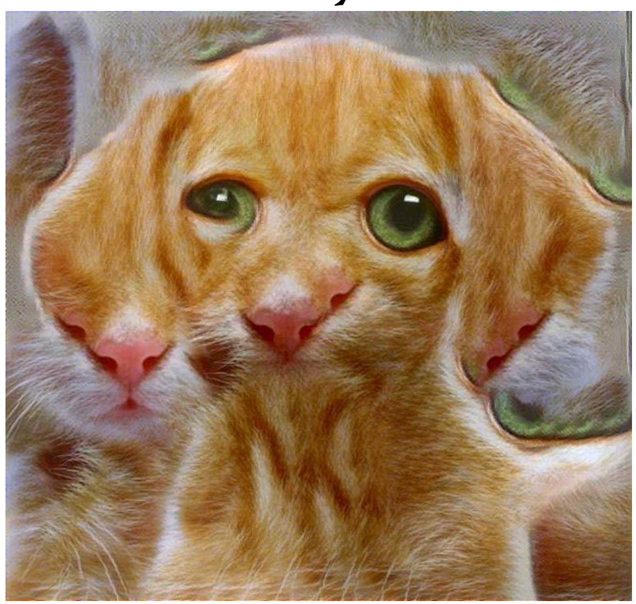
Gatys

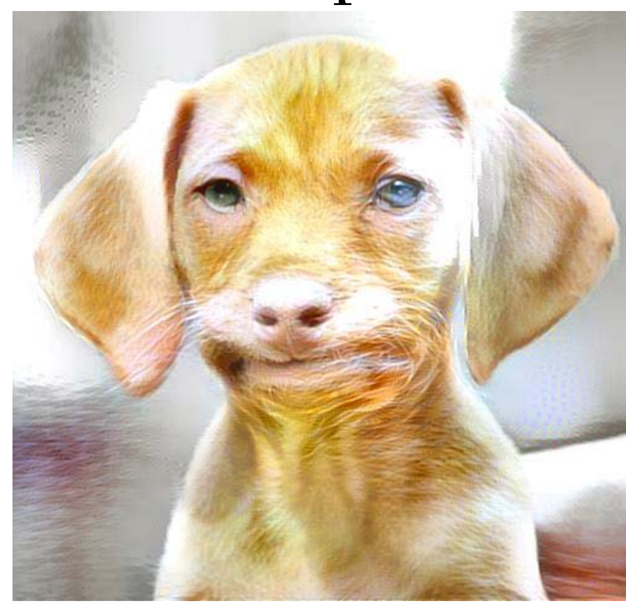

Input

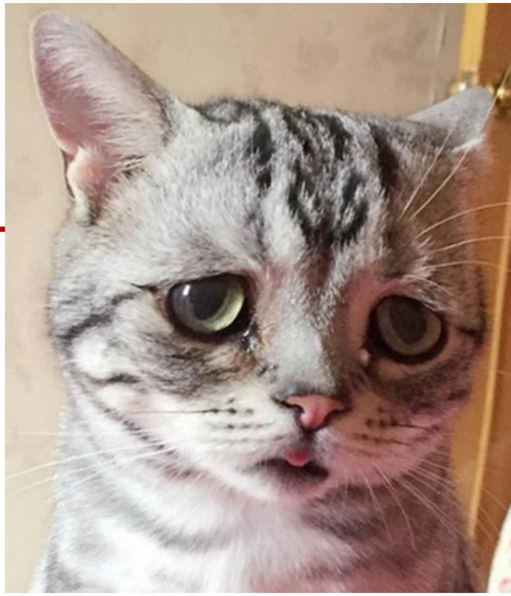

Style


Output

Does it Look REAL?

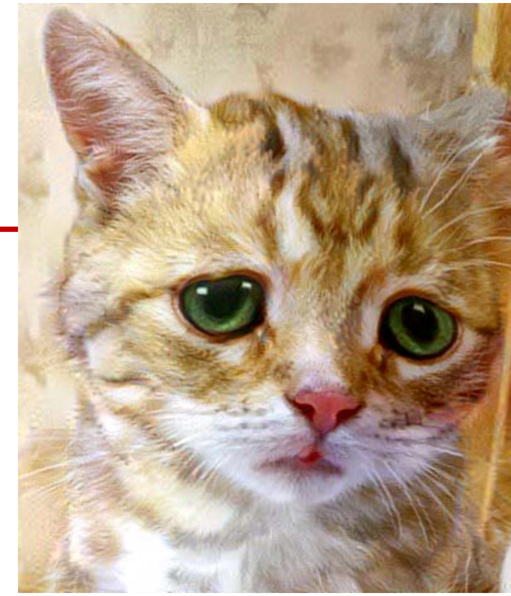

CNNMRF


Gatys



Input

Style



Output

Does it Look REAL?

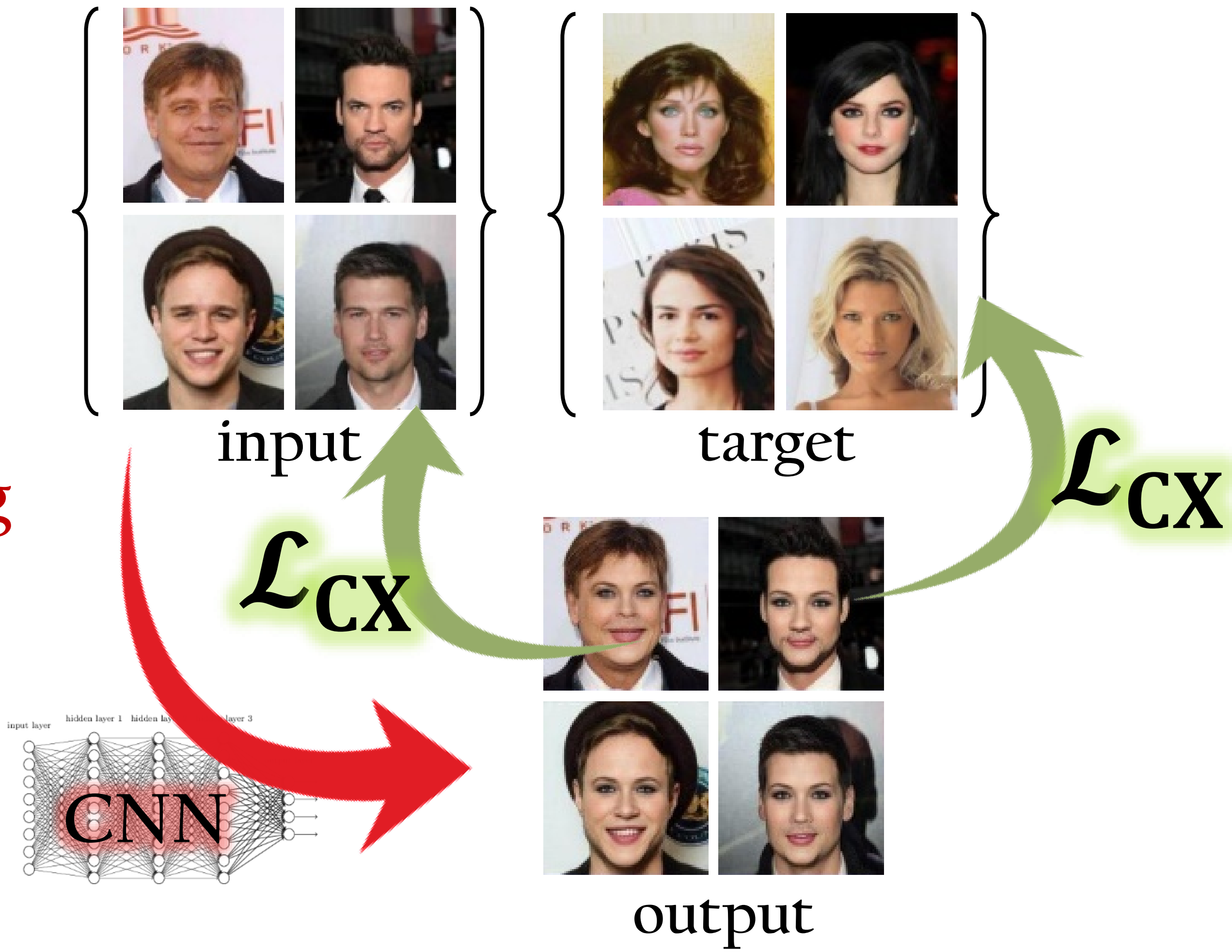
CNNMRF

Gatys

Gender Change

Use random
male-female
pairs for training

Architecture:
CRN



Male-to-Female

Does it Look REAL?

Female-to-Male

Does it Look REAL?

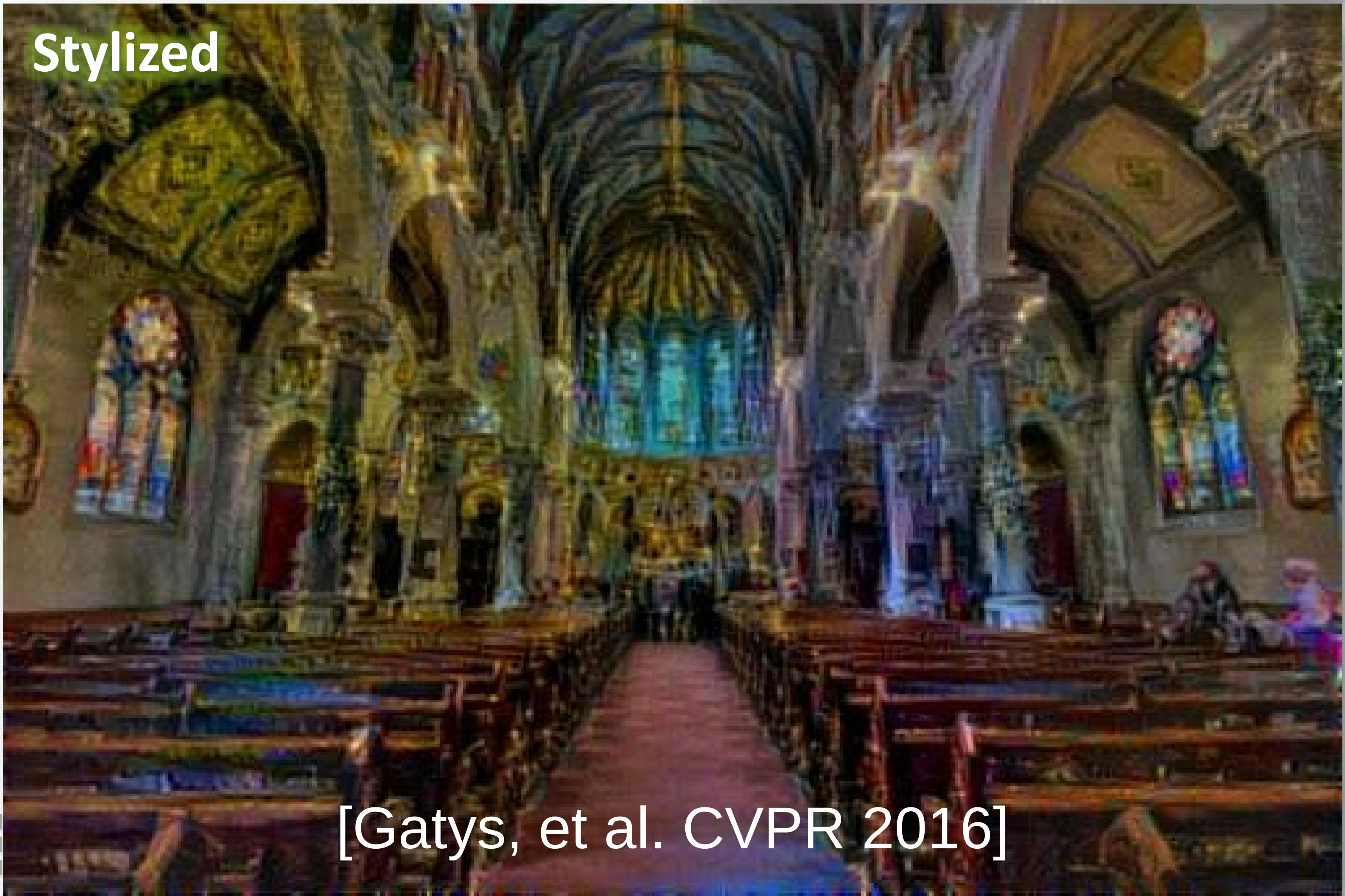
Male-to-Female

Female-to-Male

Pushing the boundaries of CNNs

GENERATING IMAGES THAT LOOK REAL

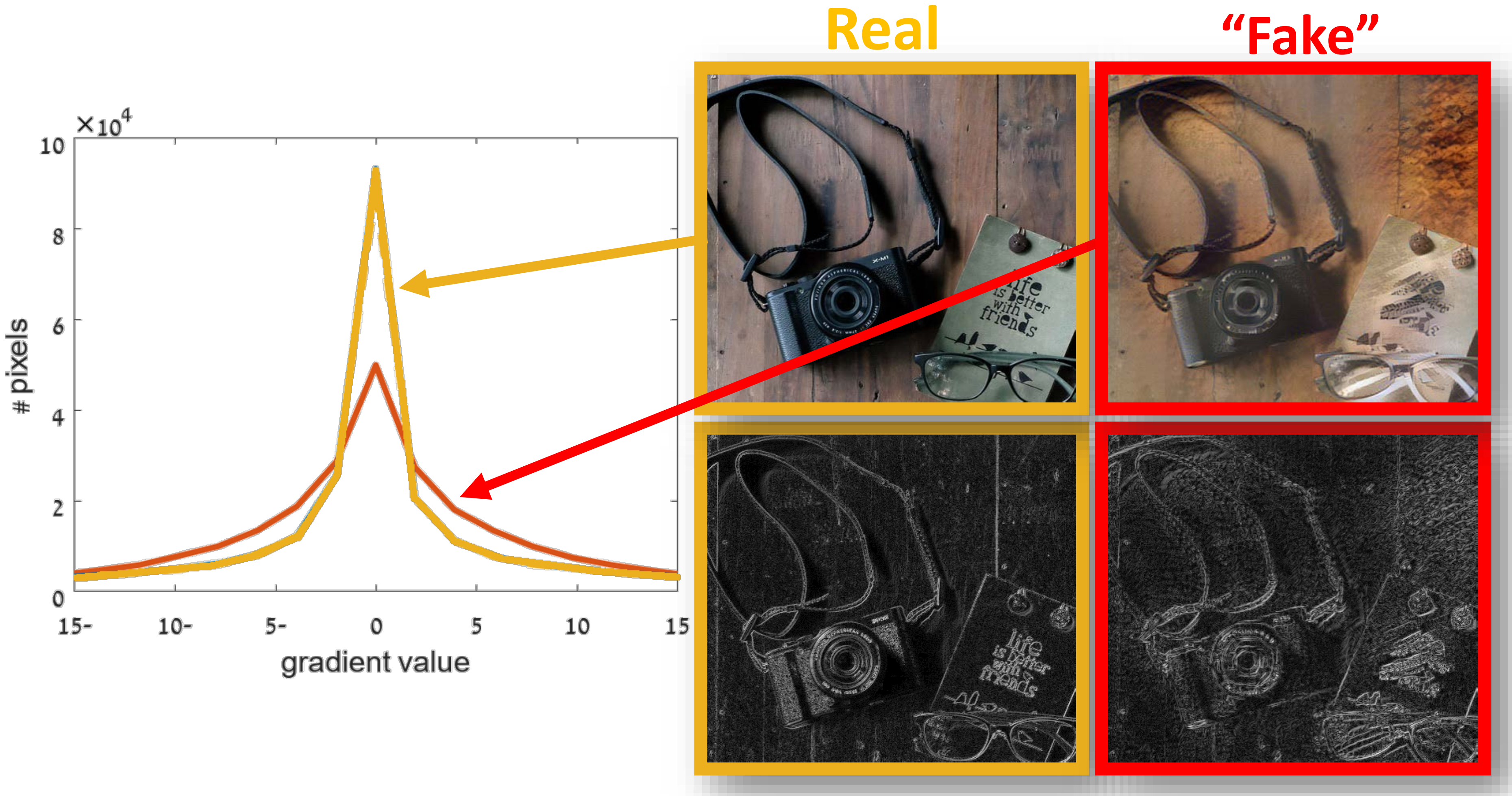
Artifacts



[Gatys, et al. CVPR 2016]

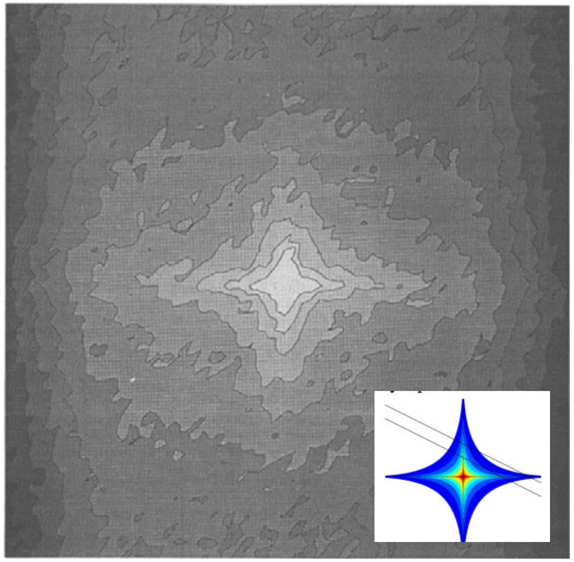
More Realistic

Statistics of Natural Images



Statistics of Natural Images

Power Spectrum of a Natural Image

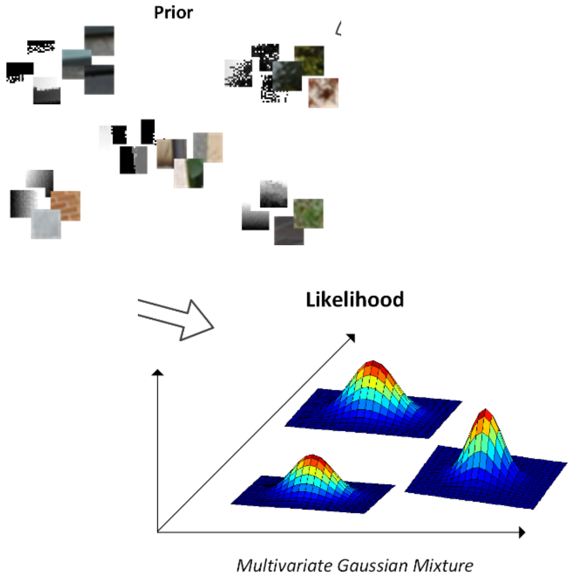


[Ruderman 1994]

Patch recurrence in a Natural Image

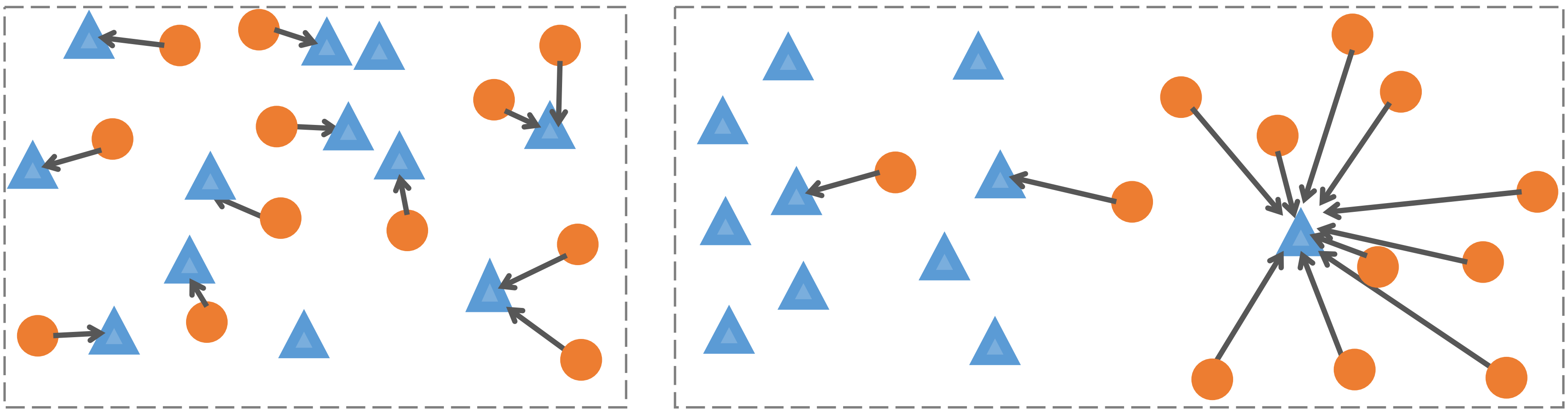
[Irani et al.]

Prior on Patch Likelihood



[Weiss et al.]

A Statistical Measure



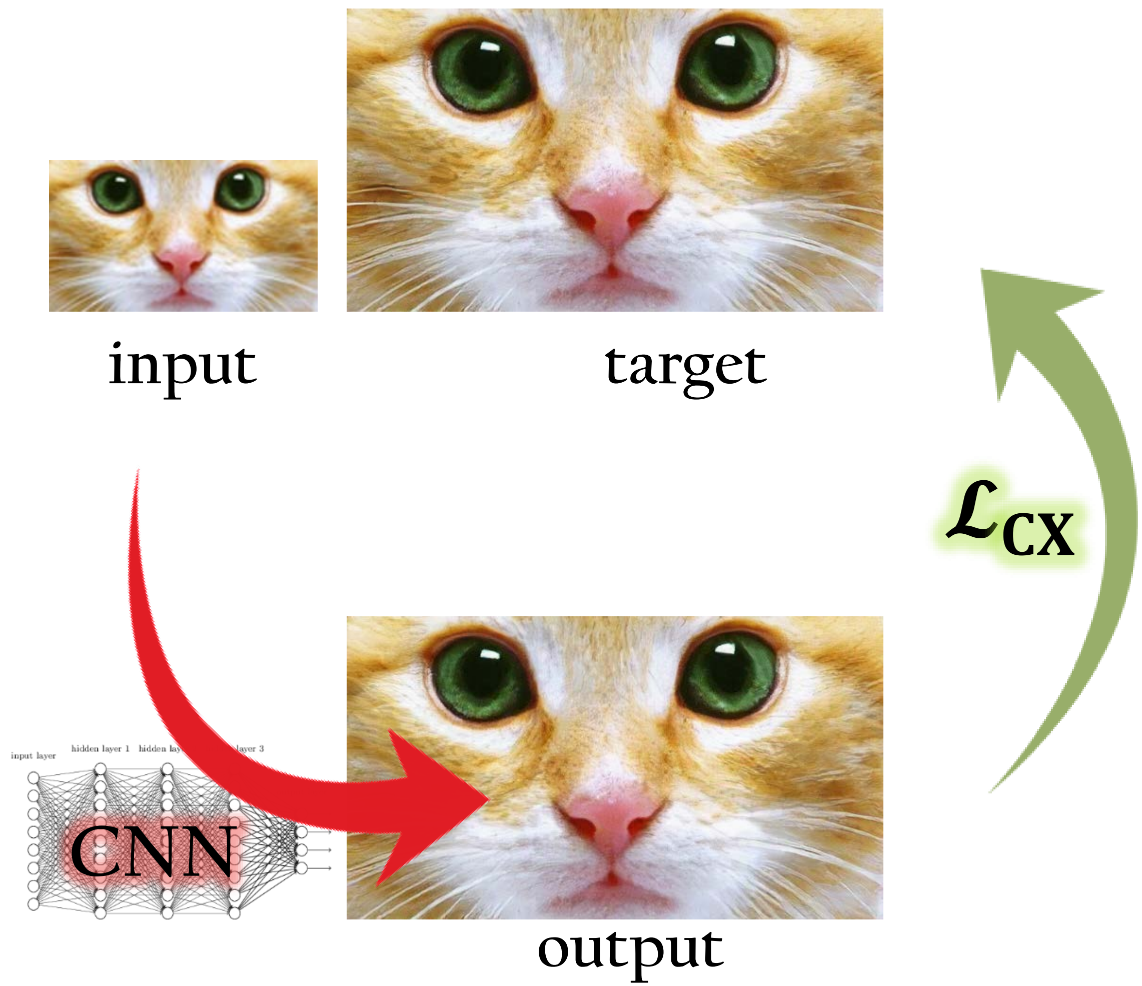
similar distributions | **different** distributions

Generated “patch”

Target “patch”

Single-Image Super-Resolution

Can quantify success



Architecture:
SRGAN
$\mathcal{L}_{GAN}$

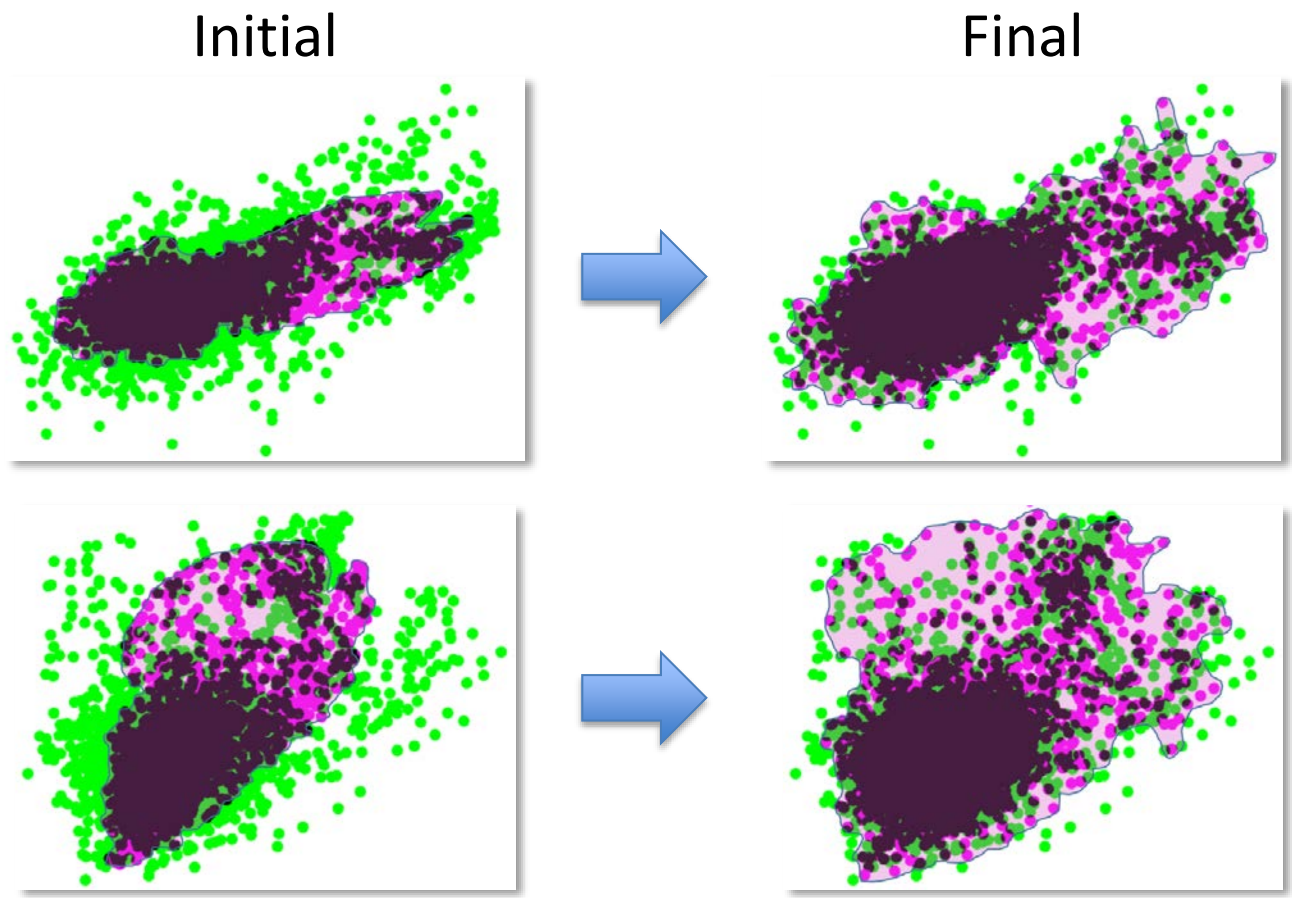
Typical Input

Typical Result

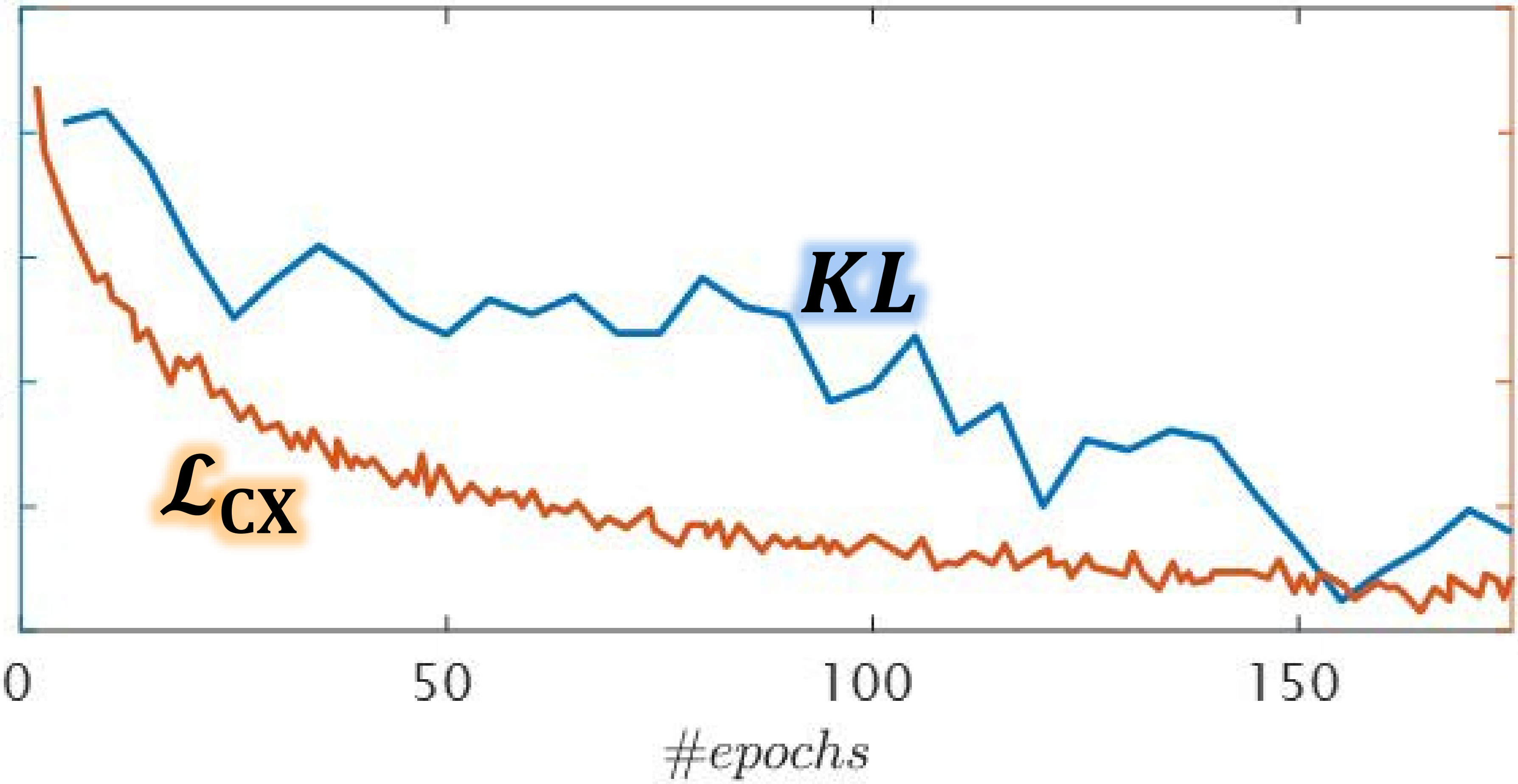
Internal Image Statistics

Random projections of image patches

- Target
- Generated Image
- Overlap



Minimizing the KL-divergence



Bicubic

Glasner et. al.

[Glasner , Bagon and Irani. Super-Resolution From a Single Image. ICCV 2009]

SRGAN

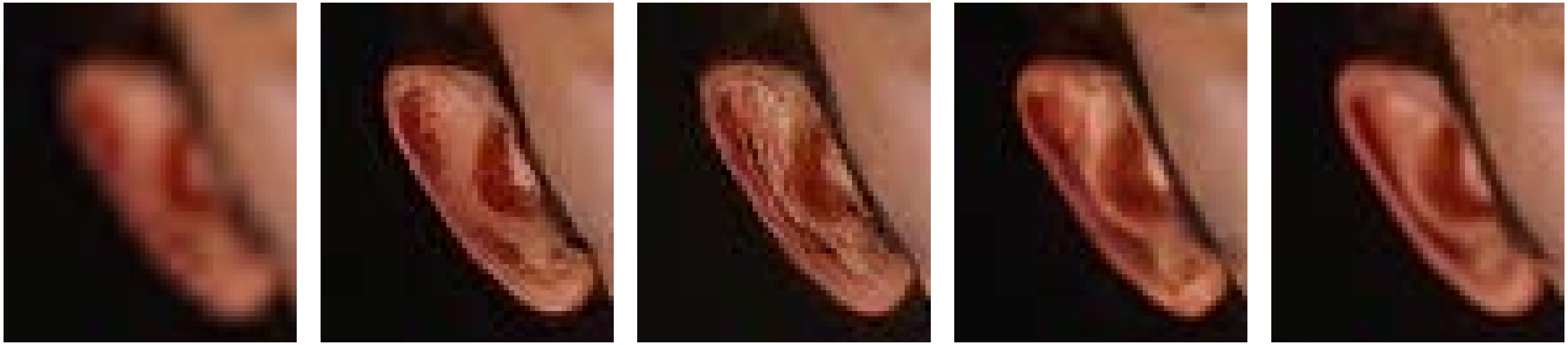
[Ledig et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. CVPR 2017]

Proposed

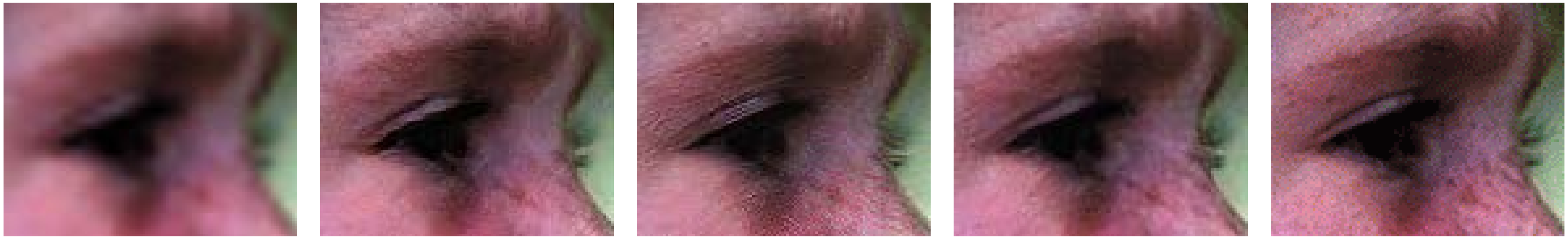
$\mathcal{L}_{CX}$

Ground Truth HD

Perceptual Super-Resolution

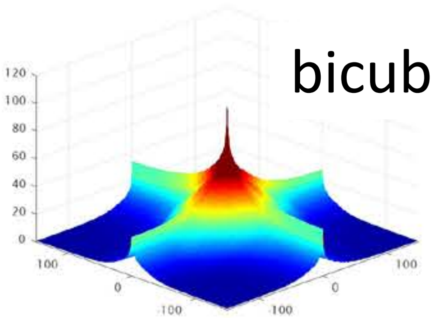


bicubic EnhanceNet SRGAN $\mathcal{L}_{CX}$ GT

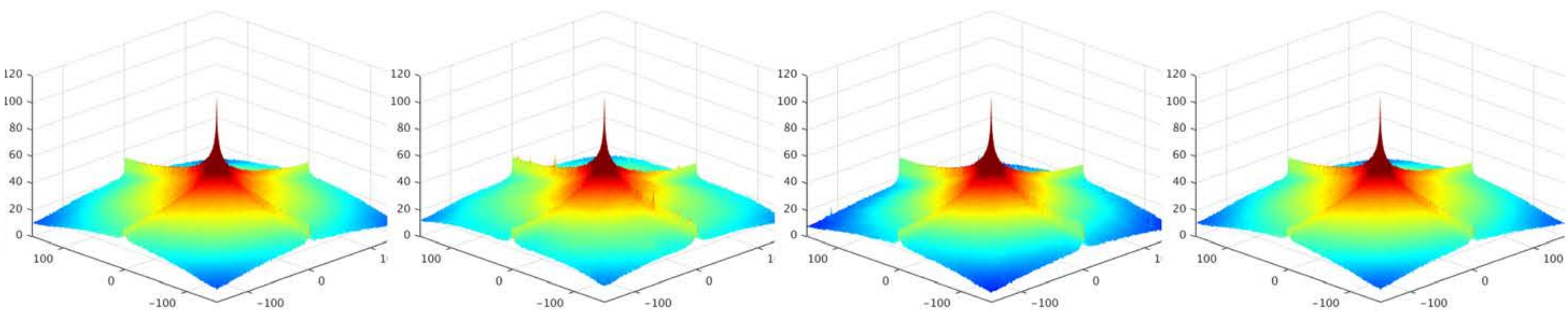


Electronics
Computers
Communications

Power Spectrum



bicubic



EnhanceNet

SRGAN

$\mathcal{L}_{\mathbf{CX}}$

GT

Electronics
Computers
Communications

Quantitative

Indice	Johnson et al. [16]	SRGAN [5]	EnhanceNet [20]	ours
Distortion: SSIM	0.631	0.640	0.624	**0.643**
Perceptual: Ma et al. [22]	7.800	8.705	8.719	**8.800**

Image to Surface

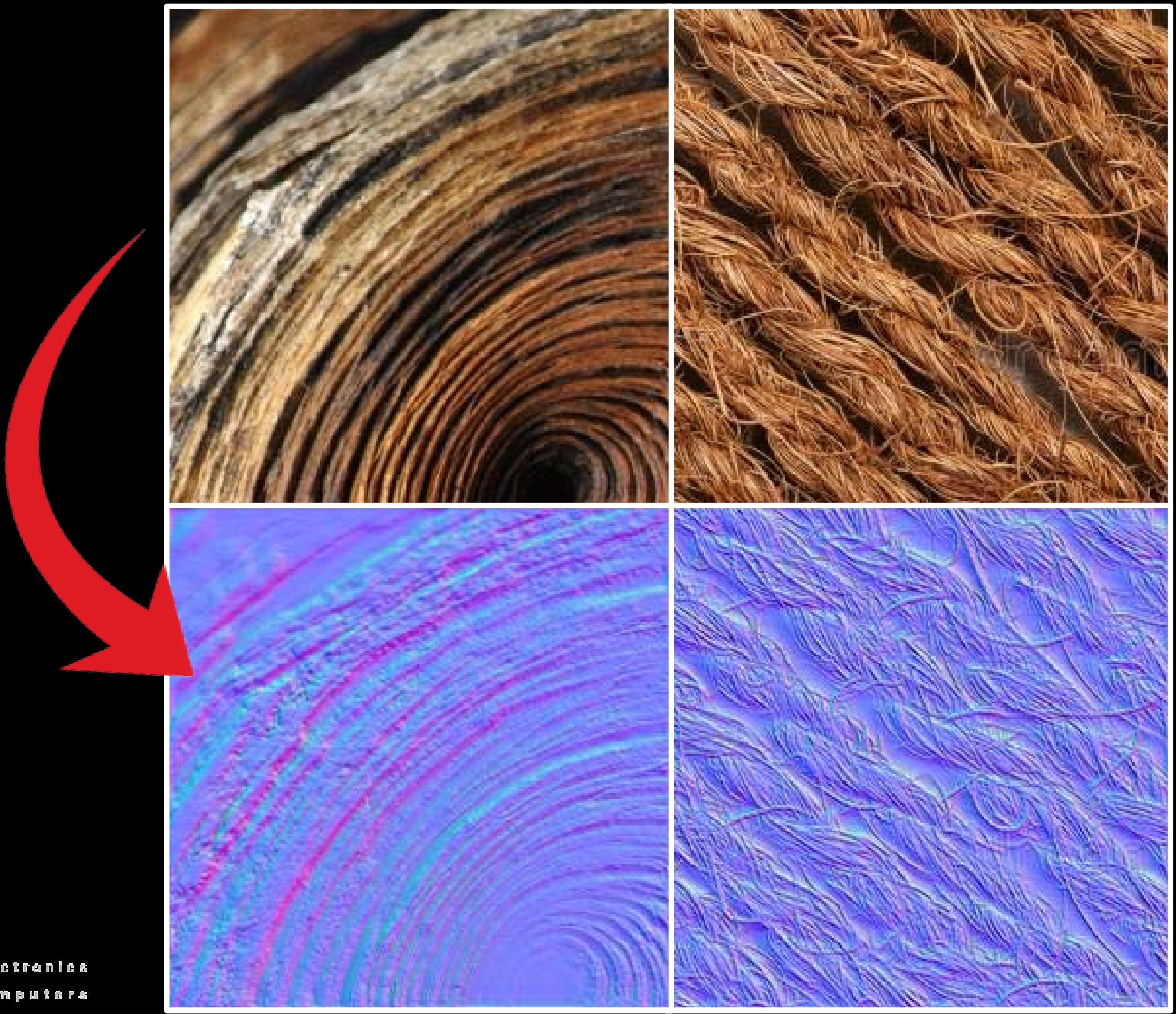


Image to Surface

L1 $\mathcal{L}_{CX}$ GT

Image to Surface

L1

$\mathcal{L}_{CX}$

GT

"Facts are stubborn things, but statistics are pliable."

Mark Twain

THANK YOU