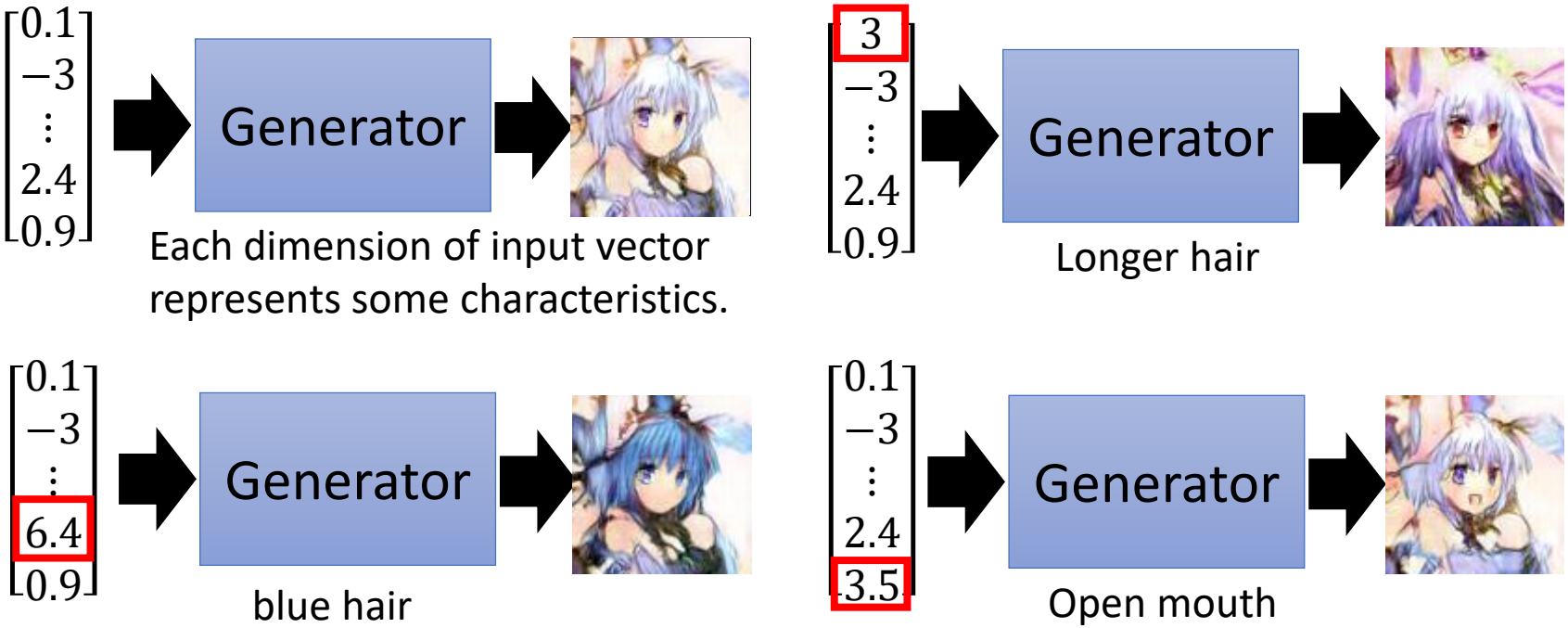
Intelligent Photo Editing

Modifying Input Code



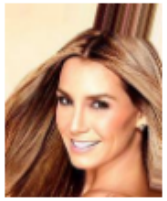
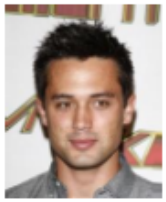
- The input code determines the generator output.
- Understand the meaning of each dimension to control the output.

Connecting Code and Attribute

(c) Hair style

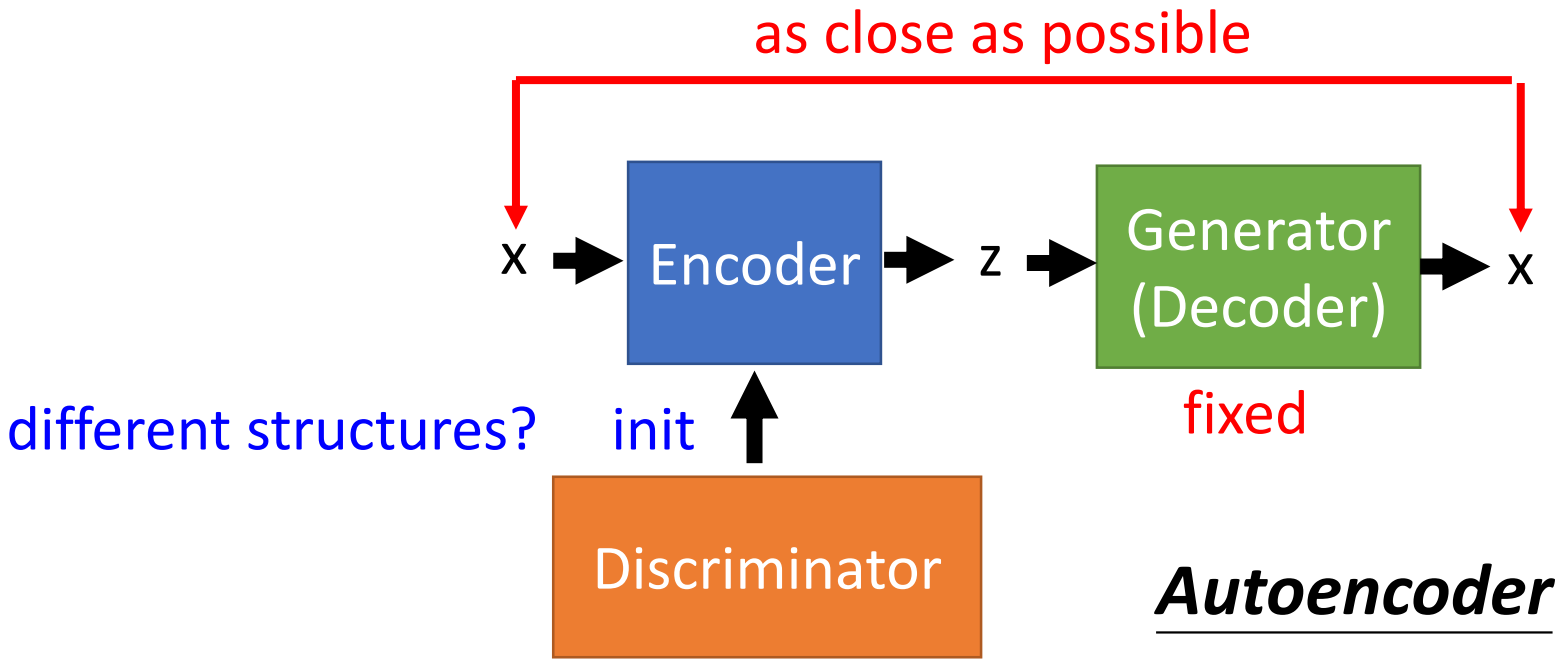
(d) Emotion

CelebA

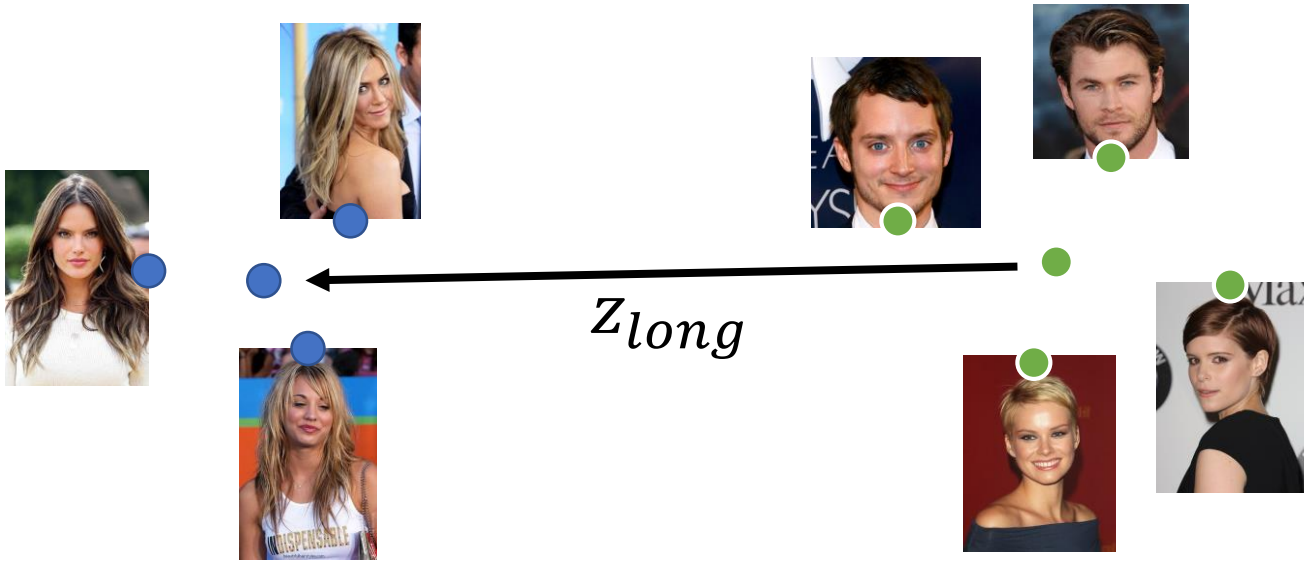
Image	Attributes
	Arched eyebrows, attractive, brown hair, heavy makeup, high cheekbones, mouth slightly open, no beard, pointy nose, smiling, straight hair, wearing earrings, wearing lipstick, young.
	5 o'clock shadows, attractive, bags under eyes, big lips, big nose, black hair, bushy eyebrows, male, no beard, pointy nose, straight hair, young.

GAN+Autoencoder

- We have a generator (input z, output x)
- However, given x, how can we find z?
  - Learn an encoder (input x, output z)



Attribute Representation



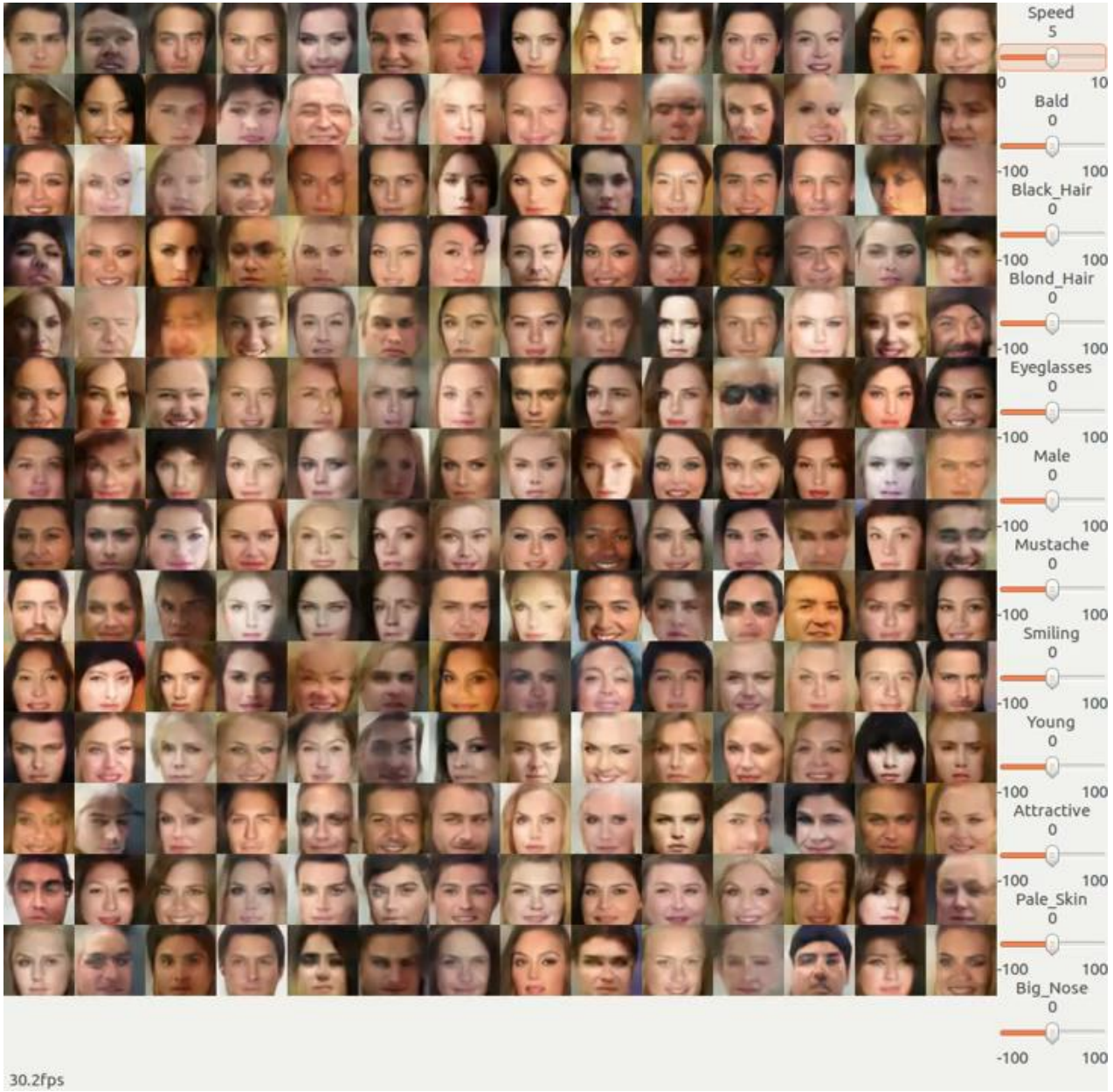
$$z_{long} = \frac{1}{N_1}\sum_{x \in long} En(x) - \frac{1}{N_2}\sum_{x' \notin long} En(x')$$

Short Hair

$$x \Rightarrow En(x) + z_{long} = z' \Rightarrow Gen(z')$$

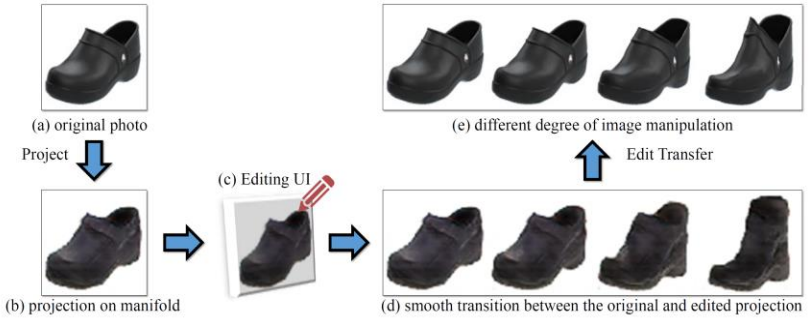
Long Hair

Photo Editing



https://www.youtube.com/watch?v=kPEIJJsQr7U

Basic Idea

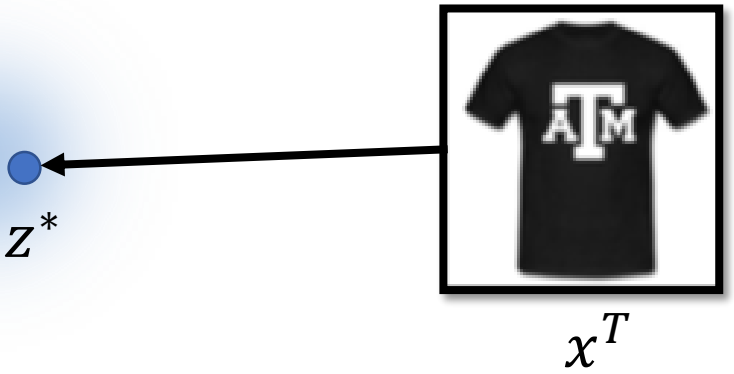


space of code z

Fulfill the constraint

Why move on the code space?

Back to z



- ***Method 1***

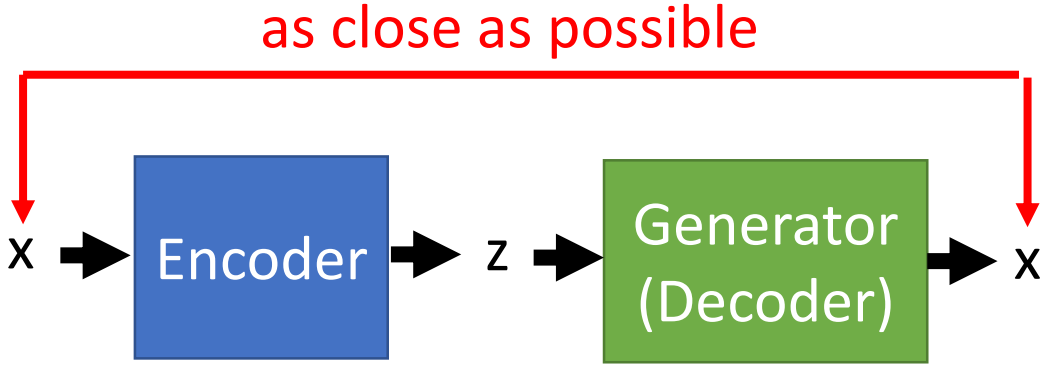
$$z^* = arg\min_z L(G(z), x^T)$$

Gradient Descent

Difference between G(z) and x^T

- Pixel-wise
- By another network

- ***Method 2***



- ***Method 3***

Using the results from ***method 2*** as the initialization of ***method 1***

Editing Photos

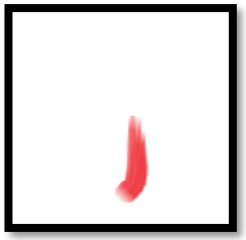
- z_0 is the code of the input image

image

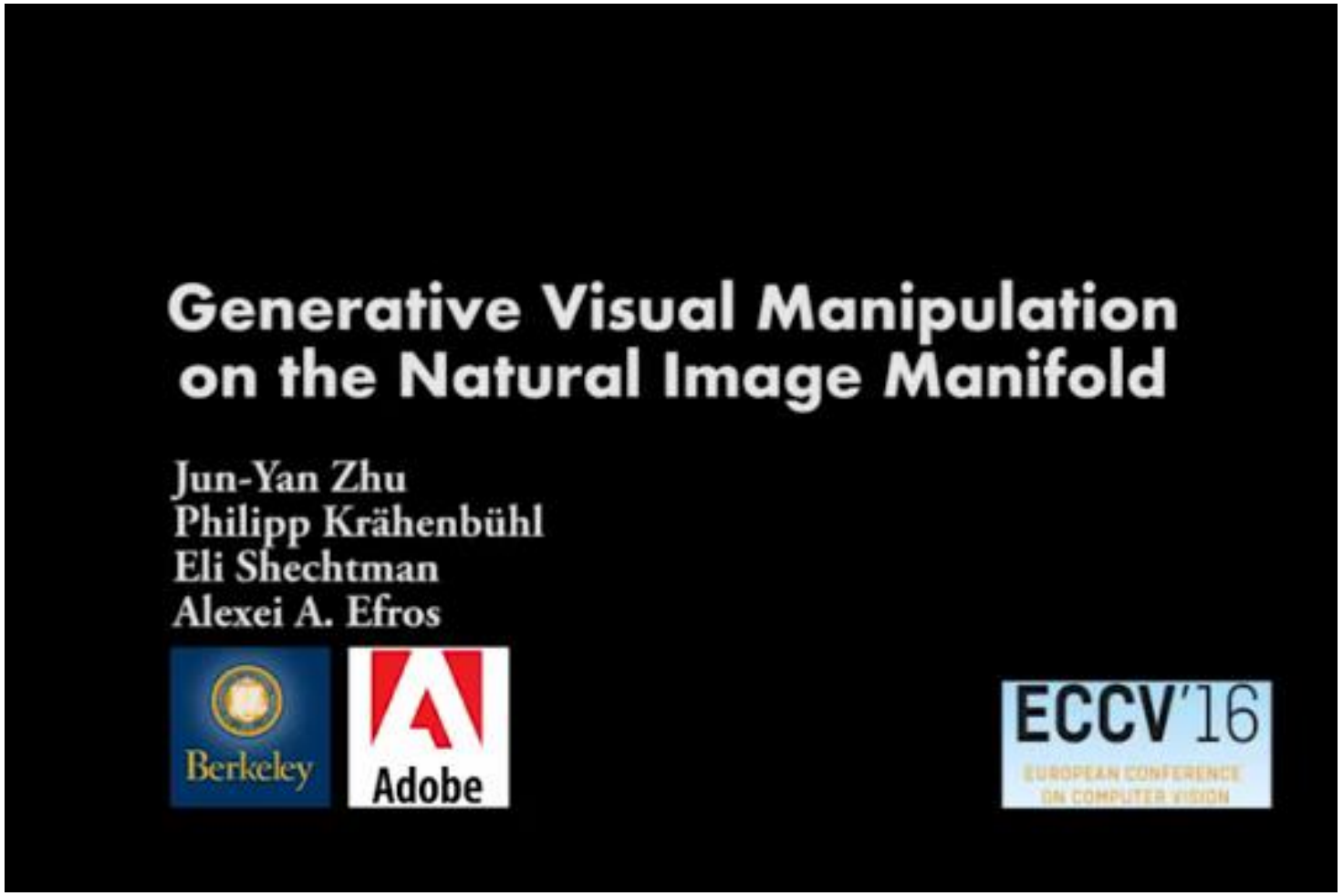
Using discriminator to check the image is realistic or not

$$z^* = arg\min_z U\big(G(z)\big) + \lambda_1 \|z - z_0\|^2 - \lambda_2 D\big(G(z)\big)$$

Not too far away from the original image

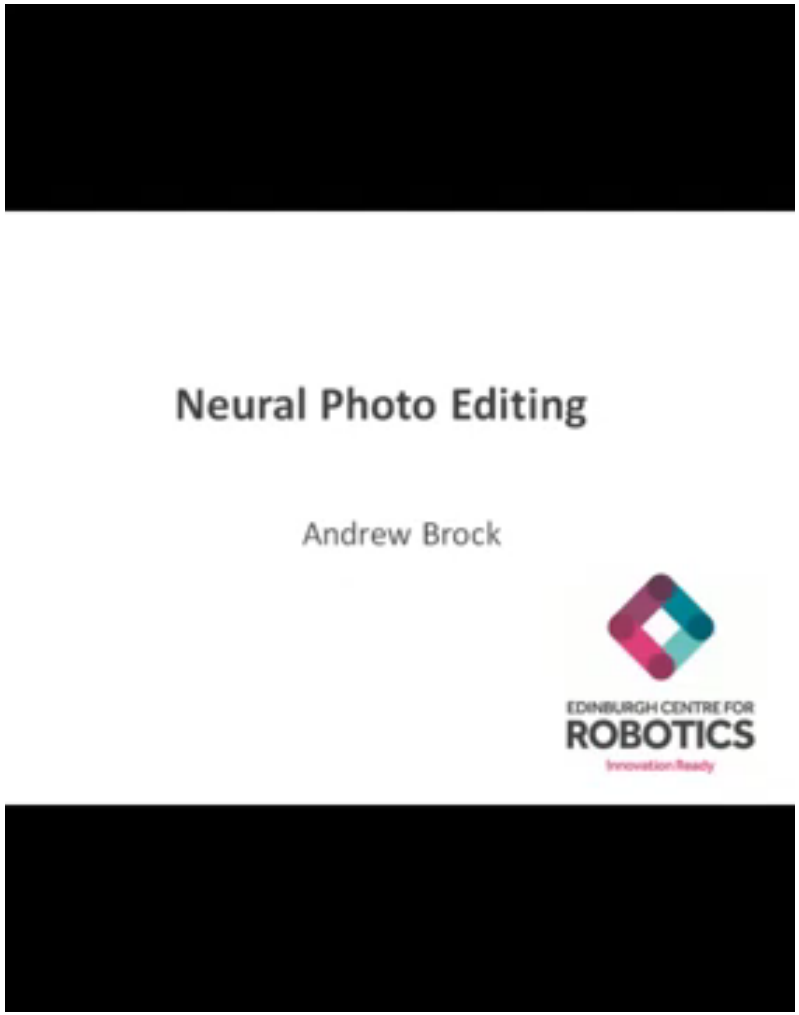


Does it fulfill the constraint of editing?



https://www.youtube.com/watch?v=9c4z6YsBGQ0

Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman and Alexei A. Efros. "Generative Visual Manipulation on the Natural Image Manifold", ECCV, 2016.



Andrew Brock, Theodore Lim, J.M. Ritchie, Nick Weston, **Neural Photo Editing with Introspective Adversarial Networks,** arXiv preprint, 2017

Image super resolution

- Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi, “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, CVPR, 2016

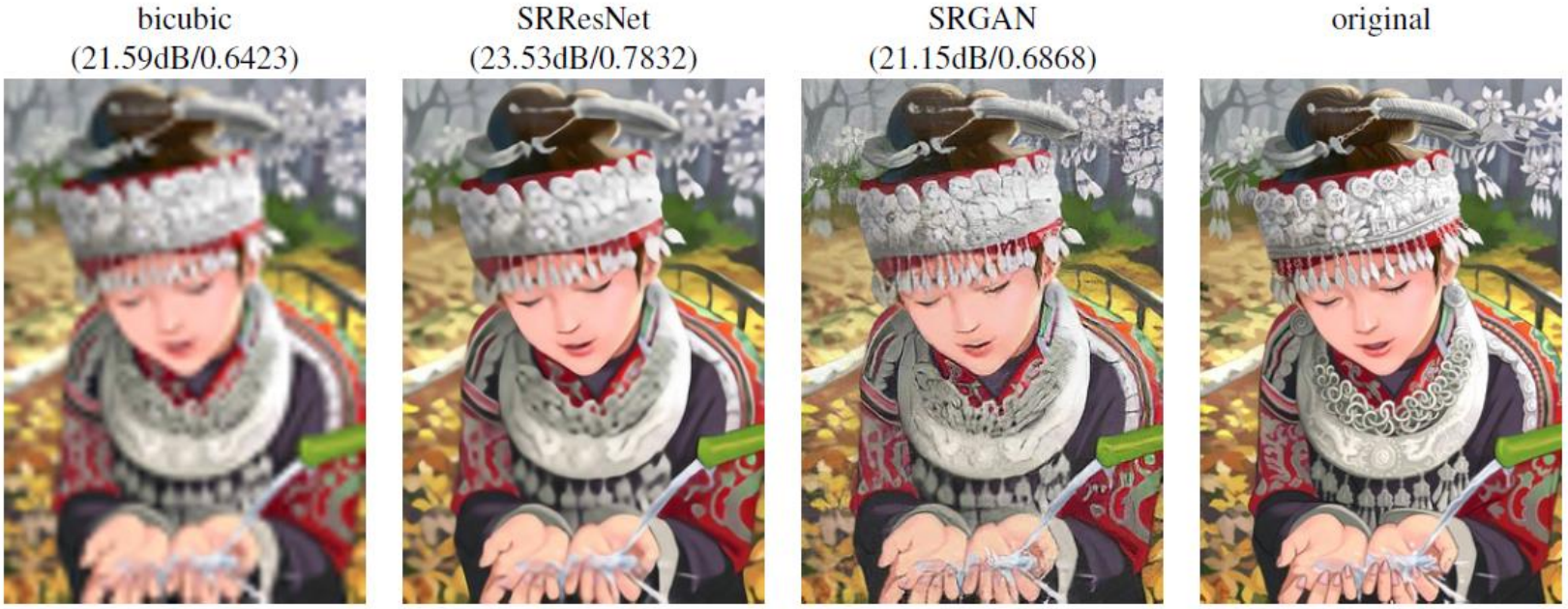
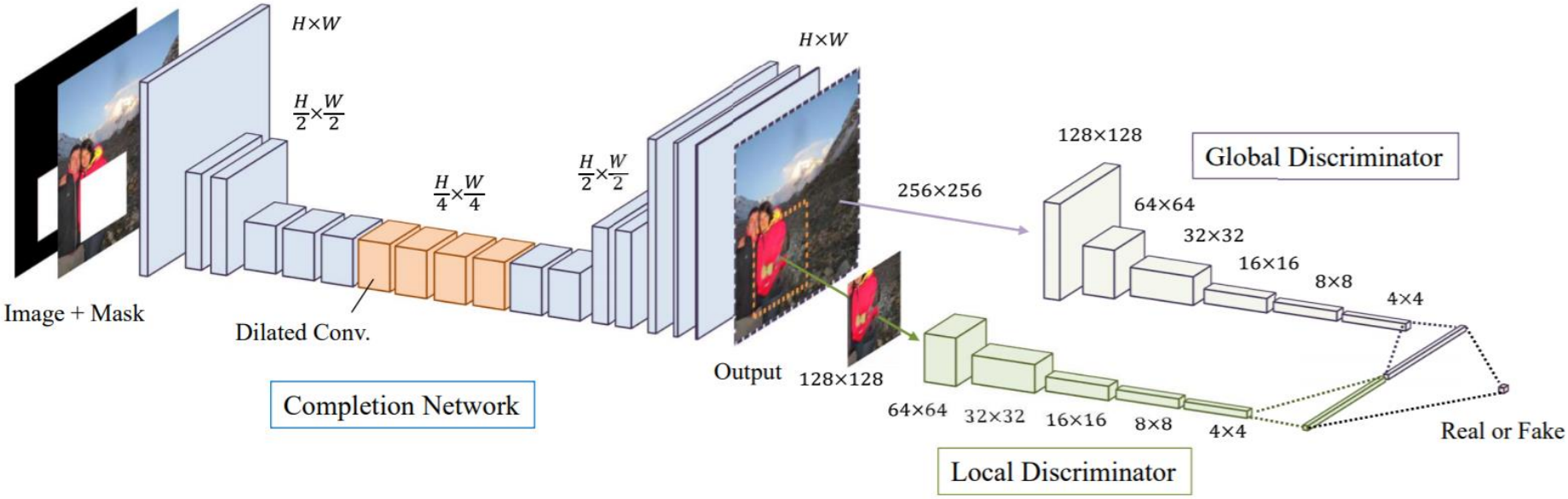


Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

Image Completion

http://hi.cs.waseda.ac.jp/~iizuka/projects/completion/en/



Demo

https://www.youtube.com/watch?v=5Ua4NUKowPU