EE-559
Deep Learning

What’s on today?

- Generative Adversarial Networks: on synthetizing data from noise
- Image-to-image GANs: on translating images across domains
- StyleGAN: on creating hyperrealistic images with control of details
- Exercises: implementing and training a GAN

Generative Adversarial Networks

$$\{x_i\}_{i=1}^{N}$$

$$\{x_j^*\}_{j=1}^{M}$$

Transforming random noise z_j
into data $\{x_j^*\}_{j=1}^{M}$ that are indistinguishable
from a training set $\{x_i\}_{i=1}^{N}$

generator

captures the training data distribution

discriminator

estimates the probability that a sample came from the training data rather than from the generator

game-theoretic formulation for training a data synthesis model (the generator)

Challenge

distance between two probability distributions

loss

balance between quality of discriminator and generator

if discriminator becomes too good, training updates of generator are attenuated

if discriminator perfectly separates generated and real samples, no change to generated data will change classification score

Wasserstein distance

quantity of work needed to equate probability mass across 2 distributions

work:
mass x distance moved

solution to a constrained minimization problem

mapping the mass of one distribution to the other

Concepts: Wasserstein distance is (a) well-defined even for disjoint distributions and (b) decreases smoothly as they become closer to one another.

Loss

$d(.)$ discriminator
$g(.)$ generator

$$L(\phi) = \sum_j d\left(x_j^*; \phi\right) - \sum_i d(x_i; \phi)$$

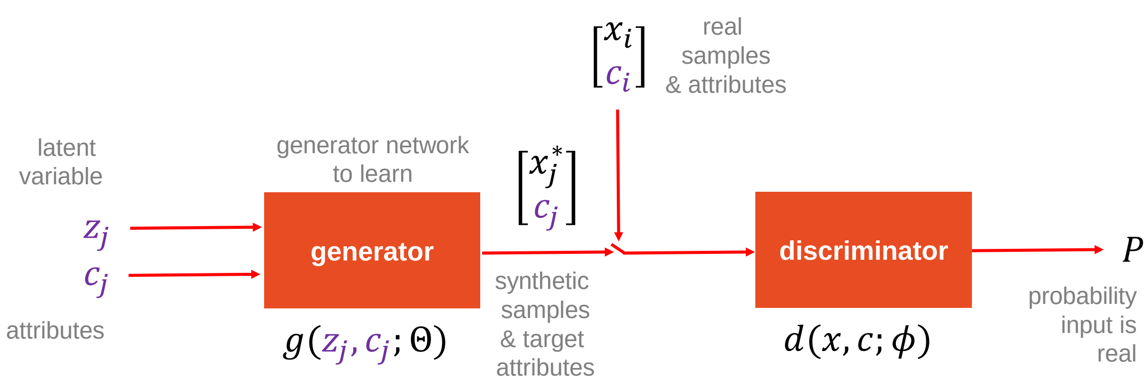
$$= \sum_j d\left(g(z_j; \Theta); \phi\right) - \sum_i d(x_i; \phi)$$

Constraint for the discriminator $\left|\frac{\partial d(x;\phi)}{\partial x}\right| < 1$

$$\{x_j^*, c_j\}_{j=1}^{M}$$

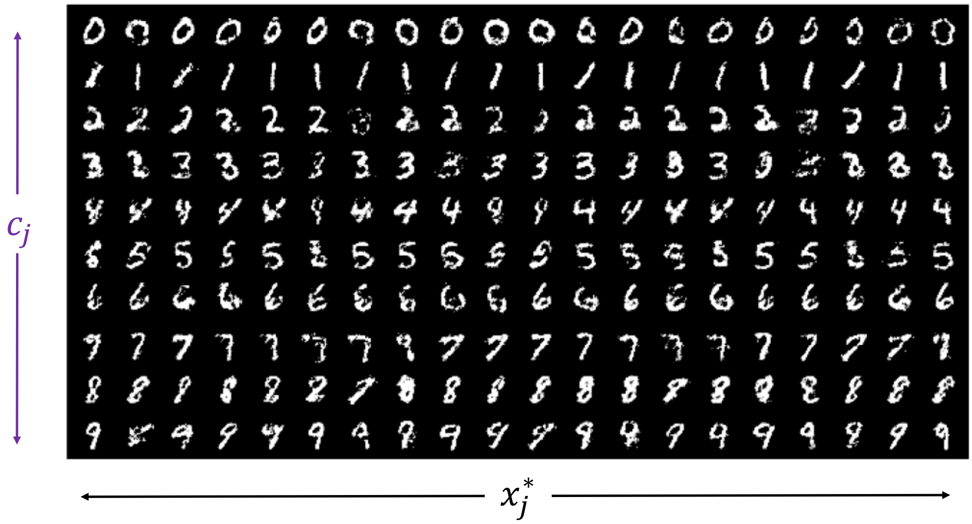
(controlled) attribute(s) of x_j^*
discrete or continuous

Conditional GAN



Concepts: Generator trained *using* a discriminator network whose task is to distinguish real vs generated samples, GAN augmented using side information (*recall*: multi-task training).

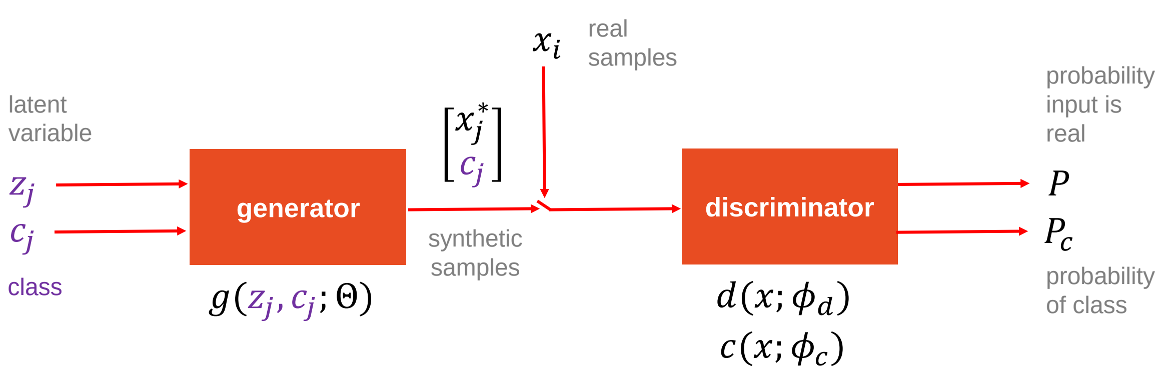
Conditional GAN: examples



MNIST dataset training (10 classes), 2014

arXiv:1411.1784

Auxiliary classifier GAN



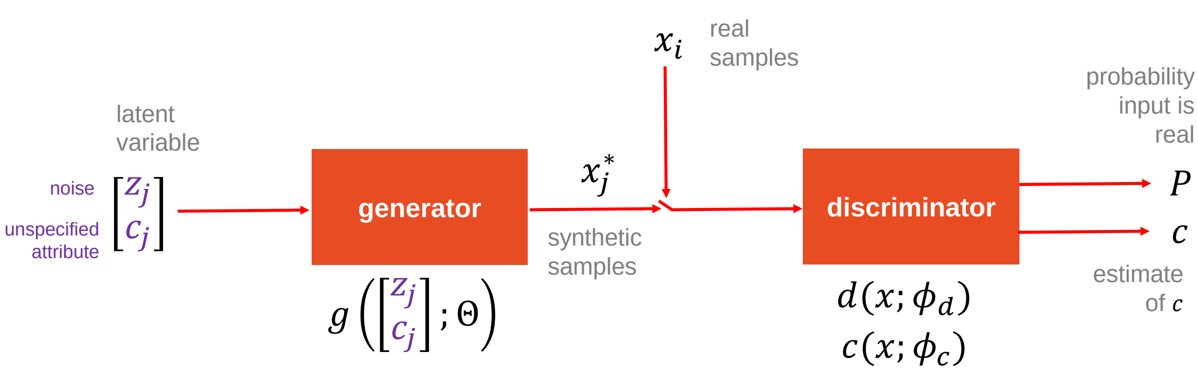
Concept:
Label conditioning (adding more structure to the latent space for higher quality samples)

Auxiliary classifier GAN: examples

monarch butterfly | goldfinch | daisy | redshank

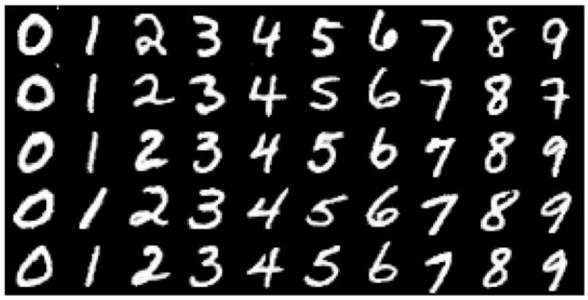
128x128 samples, ImageNet dataset training (1K classes), 2016/2017 — arXiv:1610.09585

InfoGAN



Concepts: Unsupervised learning of disentangled (*interpretable*) representations, information-theoretic extension to GAN, maximizing the mutual information between latent attribute and the generated data.

InfoGAN: examples



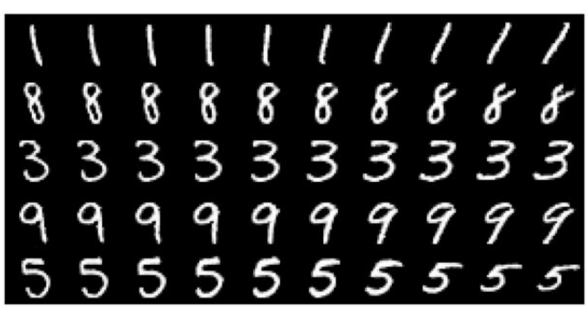
c_1

digit type
(discrete)

other latent codes & noise are fixed

arXiv:1606.03657

InfoGAN: examples



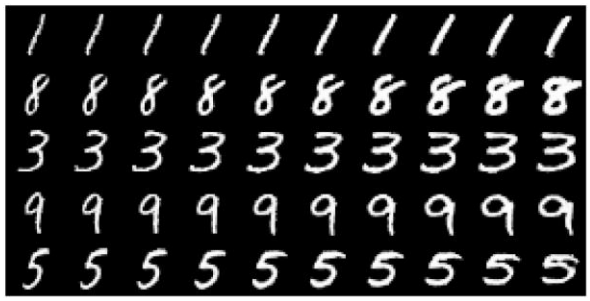
c_2

rotation
(continuous)

other latent codes & noise are fixed

arXiv:1606.03657

InfoGAN: examples



c_3

width
(continuous)

other latent codes & noise are fixed

arXiv:1606.03657

Desirable properties of synthetic data

(images)

discriminability

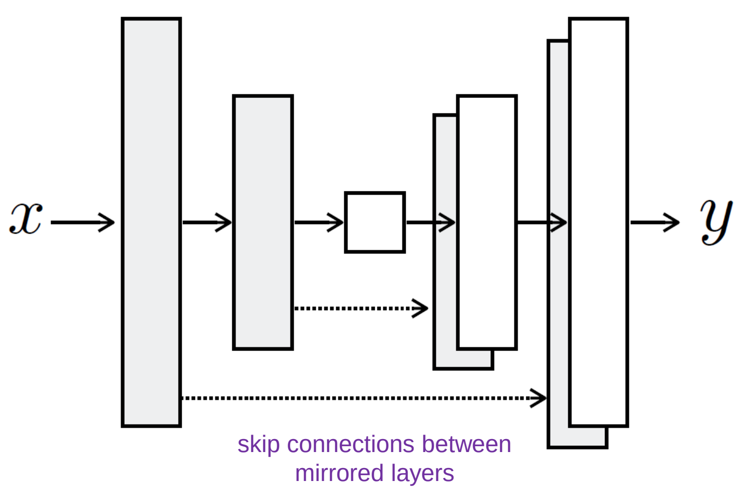
diversity

coherency

high resolution

Image-to-image GANS

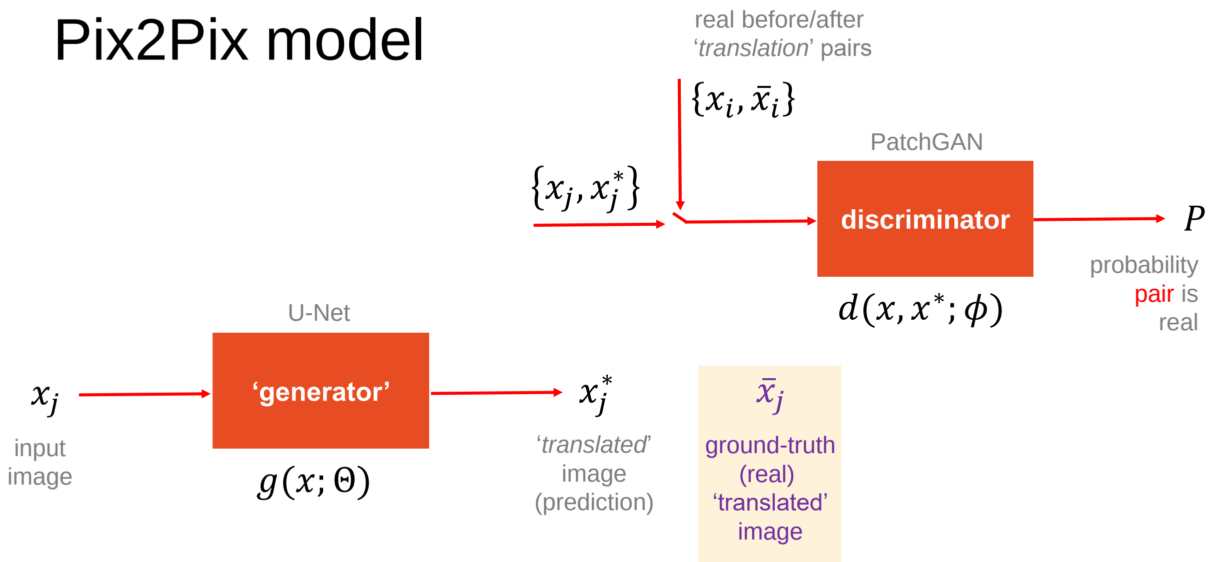
U-Net



concatenates all channels at layer i with those at layer $N - i$
(N: number of layers)

arXiv:1611.07004

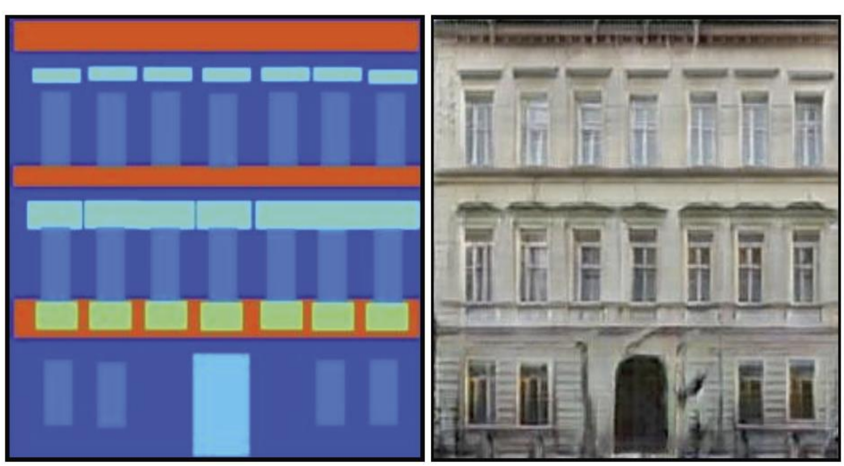
Pix2Pix model



Note: *input does not include noise*

arXiv:1611.07004

Pix2Pix model: example



x_j x_j^*

arXiv:1611.07004

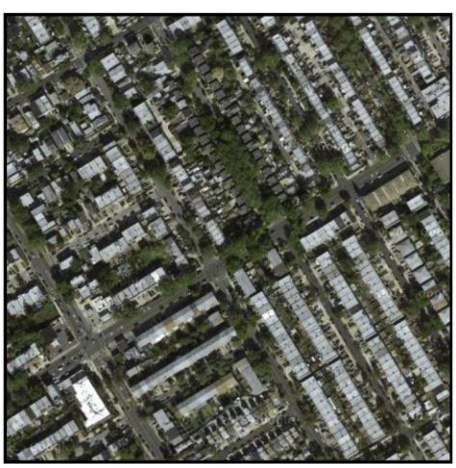
Pix2Pix model: example

x_j

x_j^*

arXiv:1611.07004

Pix2Pix model: example



x_j

x_j^*

arXiv:1611.07004

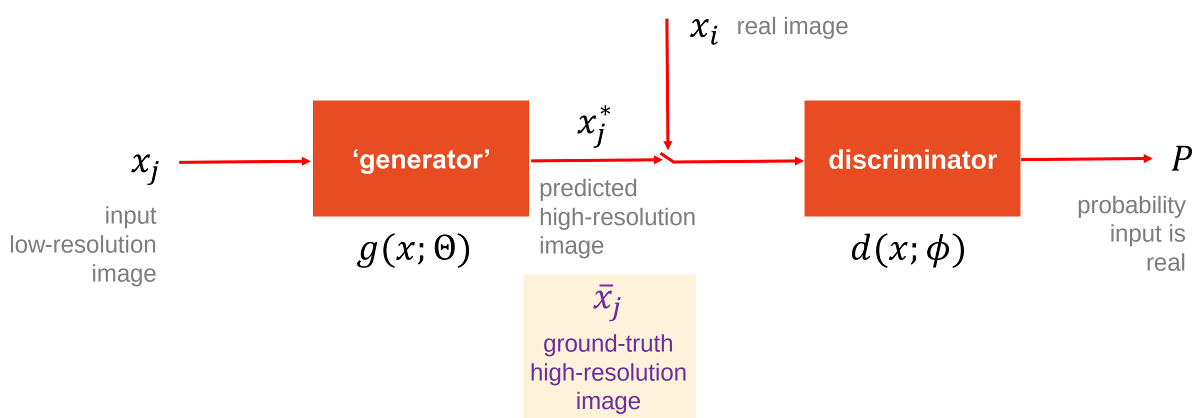
Pix2Pix model: example

x_j

x_j^*

arXiv:1611.07004

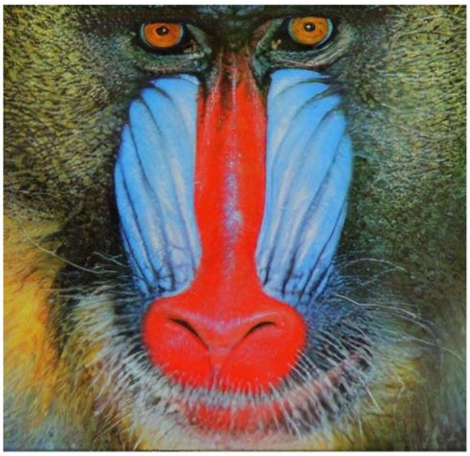
Super-resolution GAN



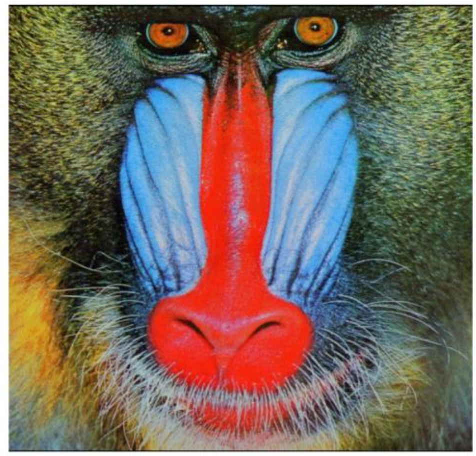
Note: *input does not include noise*

arXiv:1609.04802

Super-resolution GAN: example



x_j^*



$\bar{x}_j$

arXiv:1609.04802

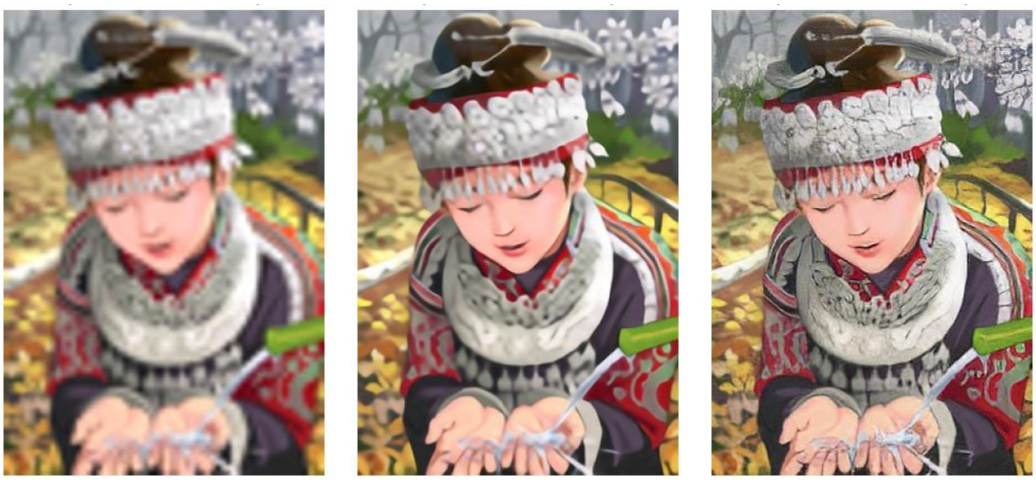
Super-resolution GAN: example

x_j^*

$\bar{x}_j$

arXiv:1609.04802

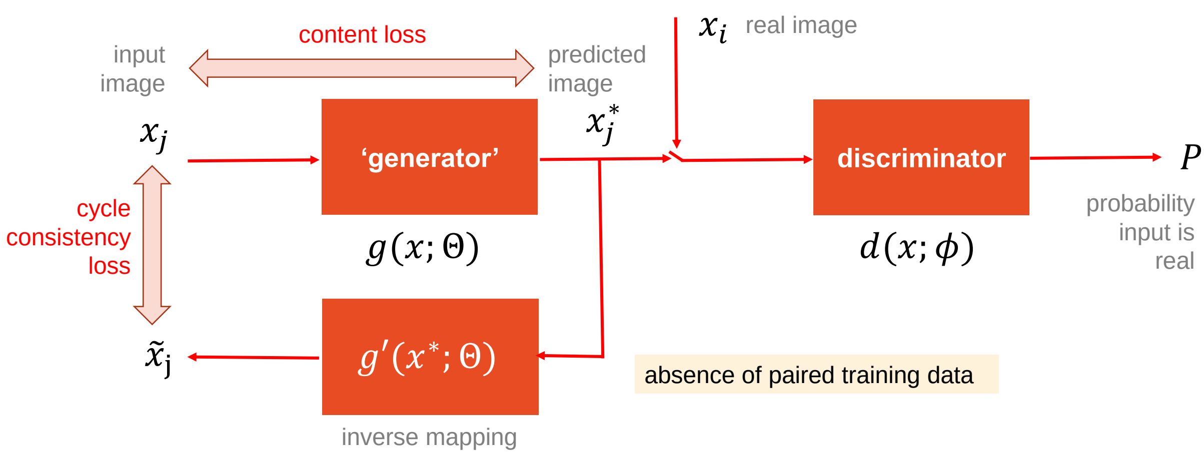
Super-resolution GAN: comparison



method 1 method 2 x_j^*

arXiv:1609.04802

CycleGAN

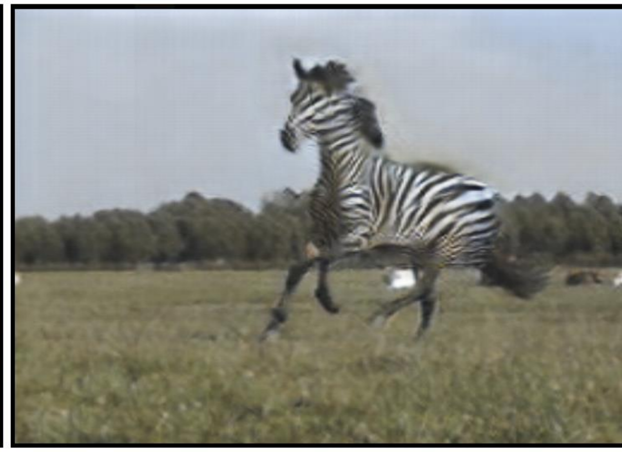


Note: *input does not include noise*

arXiv:1703.10593

CycleGAN: example

x_j

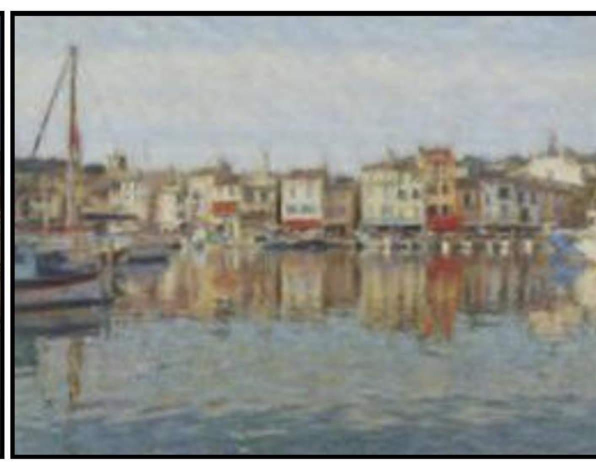


x_j^*

arXiv:1703.10593

CycleGAN: example

x_j



x_j^*

arXiv:1703.10593

CycleGAN: example

x_j

x_j^*

arXiv:1703.10593

StyleGAN

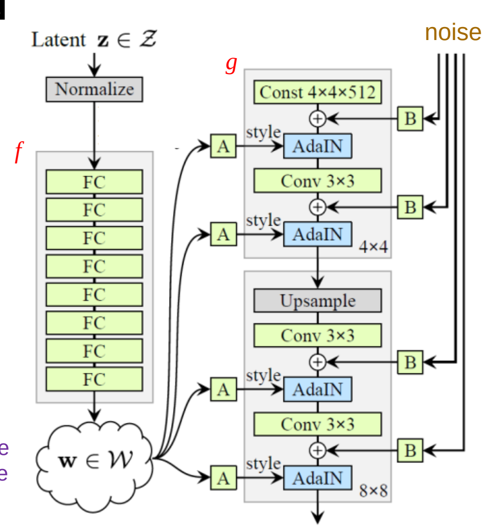
StyleGAN

scale-specific control of the synthesis

learned, unsupervised separation of
high-level attributes
(e.g. pose, identity)
& **stochastic variation**
(e.g. freckles, hair)

arXiv:1812.04948

StyleGAN



single-channel images consisting of uncorrelated Gaussian noise

A learned affine transform (specializes $\mathbf{w}$ to styles)
B learned per-channel scaling factors
f *mapping* network
g *synthesis* network

intermediate latent space

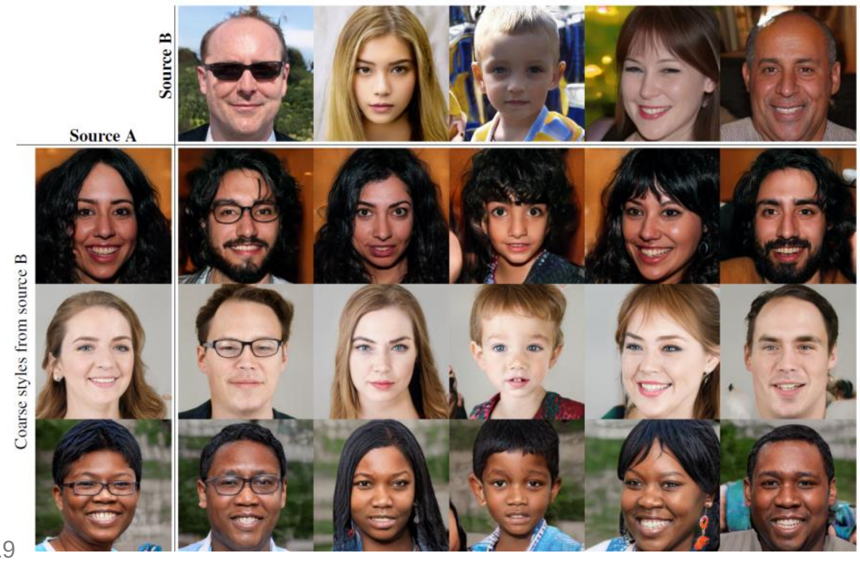
controlling the generator by adaptive instance normalization (AdaIN)

scale-specific modifications to the styles

draw samples for each style from a learned distribution

arXiv:1812.04948

StyleGAN: examples



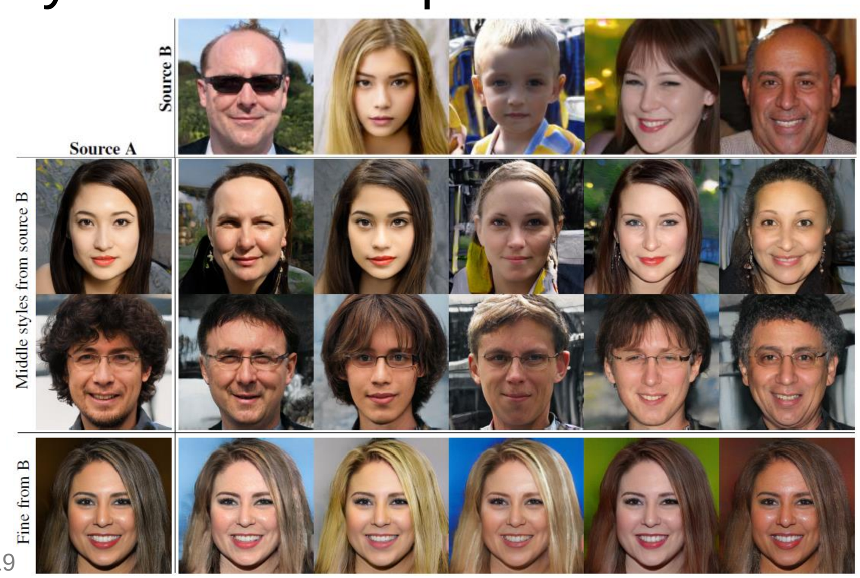
sources A and B: generated from their respective latent codes

other images: generated copying a specified subset of styles from source B and taking the rest from source A

2018/2019

arXiv:1812.04948

StyleGAN: examples



sources A and B: generated from their respective latent codes

other images: generated copying a specified subset of styles from source B and taking the rest from source A

2018/2019

arXiv:1812.04948

Marked exercise

Submission 2

- Implementation and training of a Wasserstein GAN (today's lab slot)

- Denoising Diffusion Probabilistic Models (from next week)
  - implementation of forward and reverse diffusion processes
  - implementation of training and sampling
  - answering two questions

What did we learn today?

- Generative Adversarial Networks
- Image-to-image GANs
- StyleGAN
- Exercises

EE-559

Deep Learning

andrea.cavallaro@epfl.ch