Autoencoders and GANs

Materials for this section:

- https://www.deeplearningbook.org/contents/autoencoders.html
- https://www.youtube.com/watch?v=BUNl0To1IVw
- https://arxiv.org/pdf/1703.10593.pdf
- https://arxiv.org/pdf/1611.07004.pdf

Autoencoders

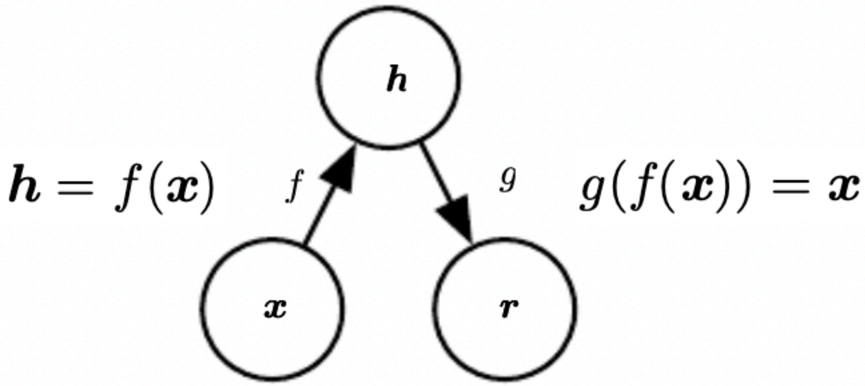
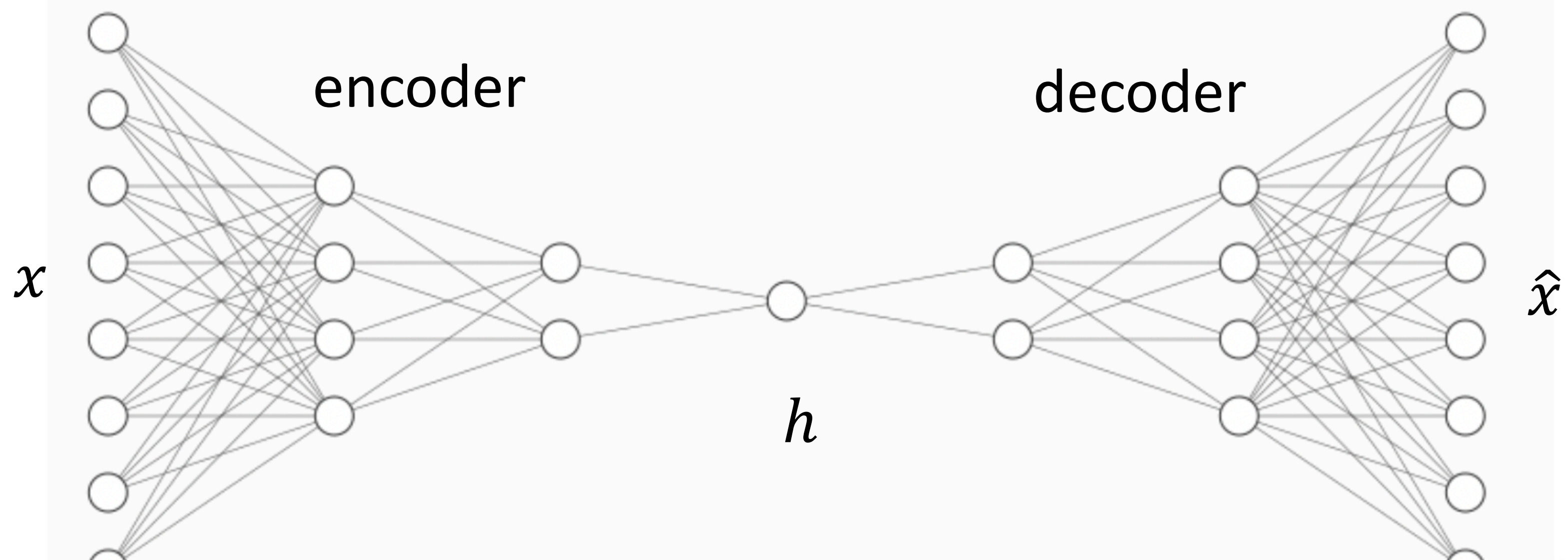


Figure 14.1: The general structure of an autoencoder, mapping an input $\boldsymbol{x}$ to an output (called reconstruction) $\boldsymbol{r}$ through an internal representation or code $\boldsymbol{h}$. The autoencoder has two components: the encoder f (mapping $\boldsymbol{x}$ to $\boldsymbol{h}$) and the decoder g (mapping $\boldsymbol{h}$ to $\boldsymbol{r}$).

- autoencoders are unsupervised learning models

www.deeplearningbook.org/contents/autoencoders.html

Autoencoders



embedded space = *latent* space = *bottleneck*

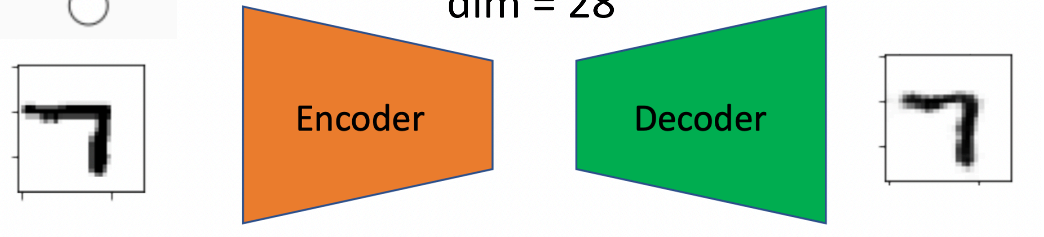
Typically, *latent space* h have a smaller dimension than input x

- An autoencoder whose code dimension is less than the input dimension is called **undercomplete**

$$\mathcal{L}(x, \hat{x}) = \frac{1}{N} \sum_{i=1}^{N} ||x_i - \hat{x}_i||^2$$

latent space dim = 28

input: 28x28

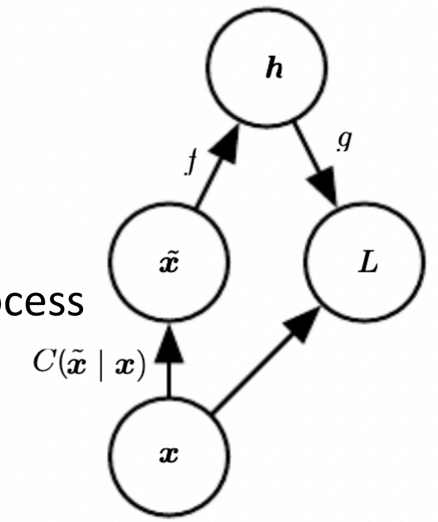


sebastianraschka.com/pdf/lecture-notes/stat479ss19/L15_autoencoder_slides.pdf

Autoencoders

Denoising autoencoders

corruption process

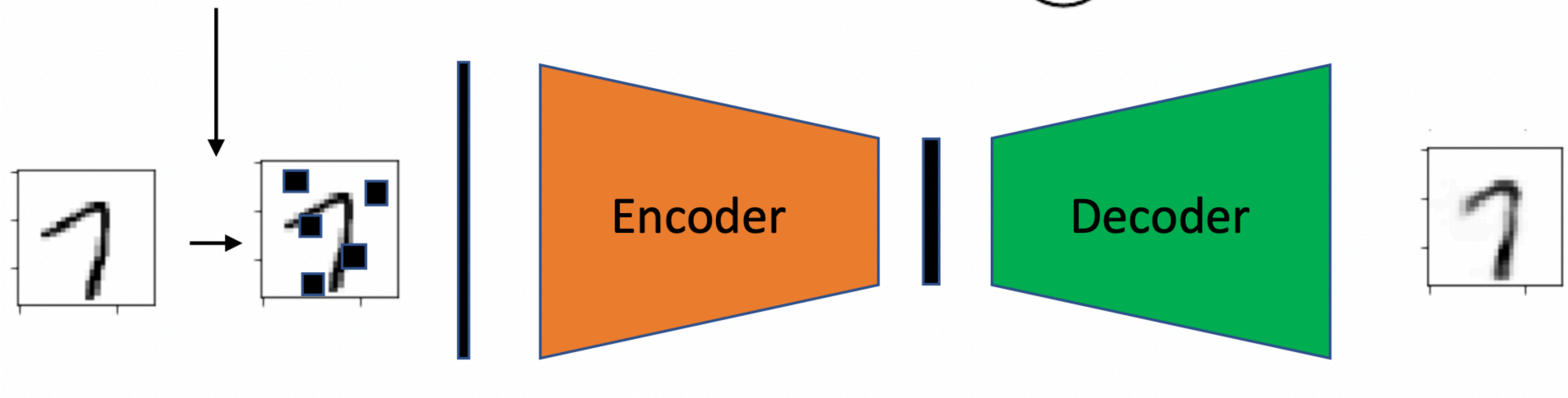


DAE minimizes:

$$L(\boldsymbol{x}, g(f(\tilde{\boldsymbol{x}})))$$

corrupted copy of input

Add noise to the input to learn how to denoise images

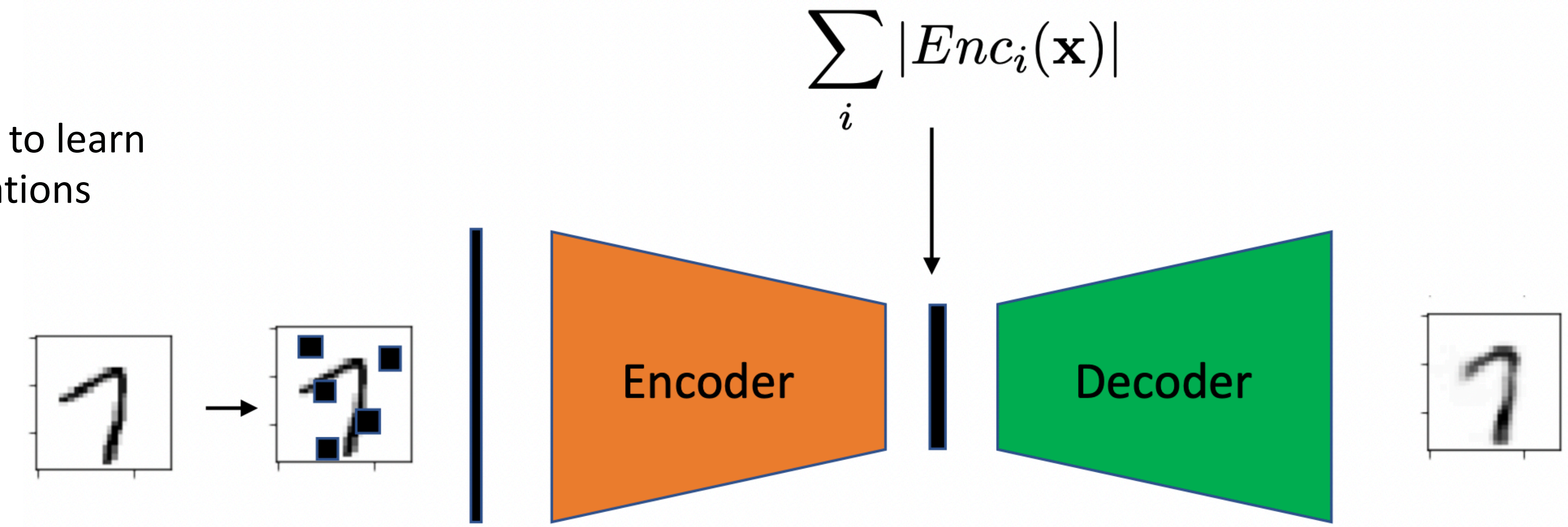


https://sebastianraschka.com/pdf/lecture-notes/stat479ss19/L15_autoencoder_slides.pdf

Autoencoders

Sparse autoencoders

Add L_1 penalty to the loss to learn **sparse** feature representations



- sparsity of a latent space enforces learning only useful features of an input
- after learning we take trained encoder and use as feature extractor e.g. in classification task

$$\mathcal{L} = ||\mathbf{x} - Dec(Enc(\mathbf{x}))||_2^2 + \sum_i |Enc_i(\mathbf{x})|$$

- other application: dimensionality reduction

https://sebastianraschka.com/pdf/lecture-notes/stat479ss19/L15_autoencoder_slides.pdf

Generative models

Which one is unreal?

Generative adversarial networks

- artifacts

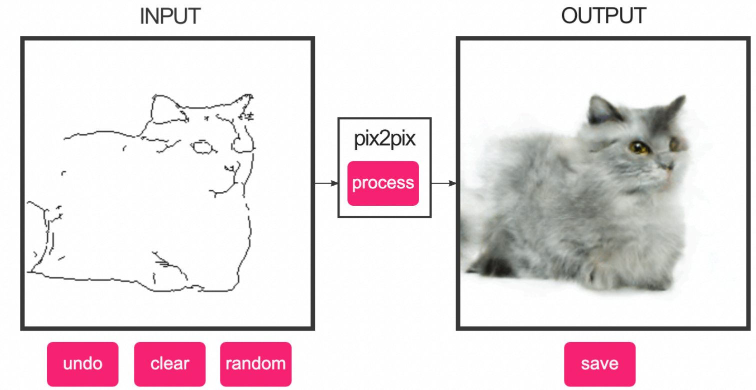
https://thispersondoesnotexist.com

Generative adversarial networks

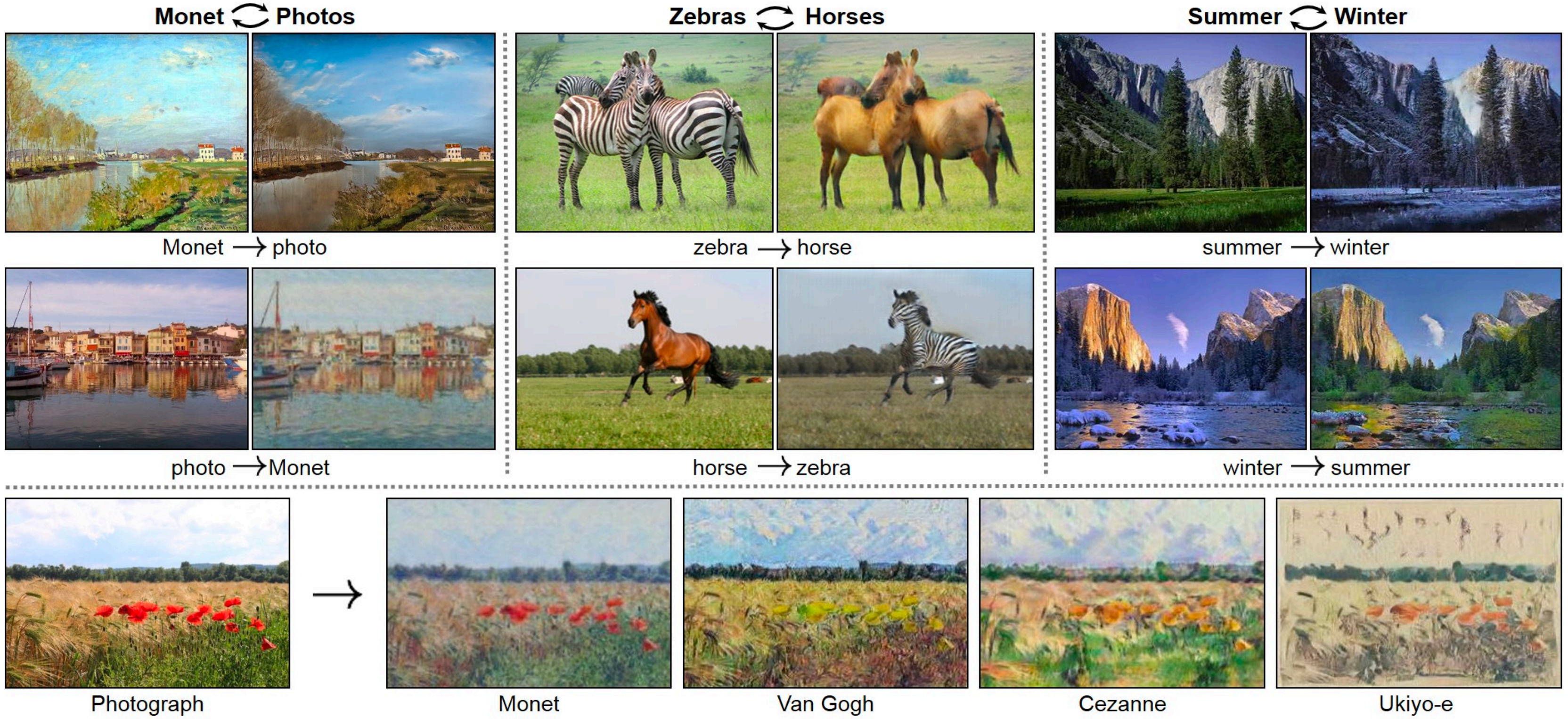
Another famous examples:

- cycle GANs
- pix2pix

github.com/junyanz/pytorch-CycleGAN-and-pix2pix



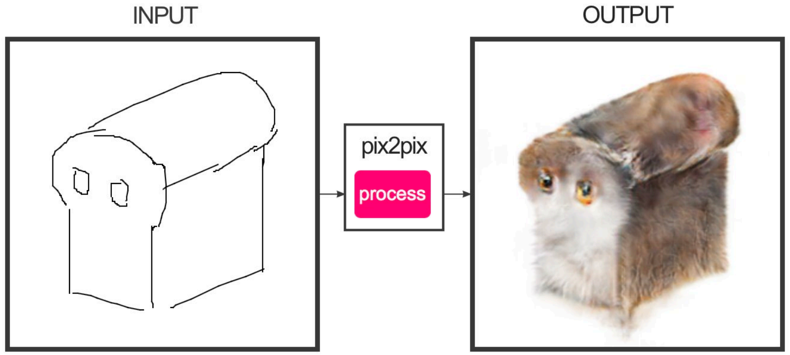
https://affinelayer.com/pixsrv/



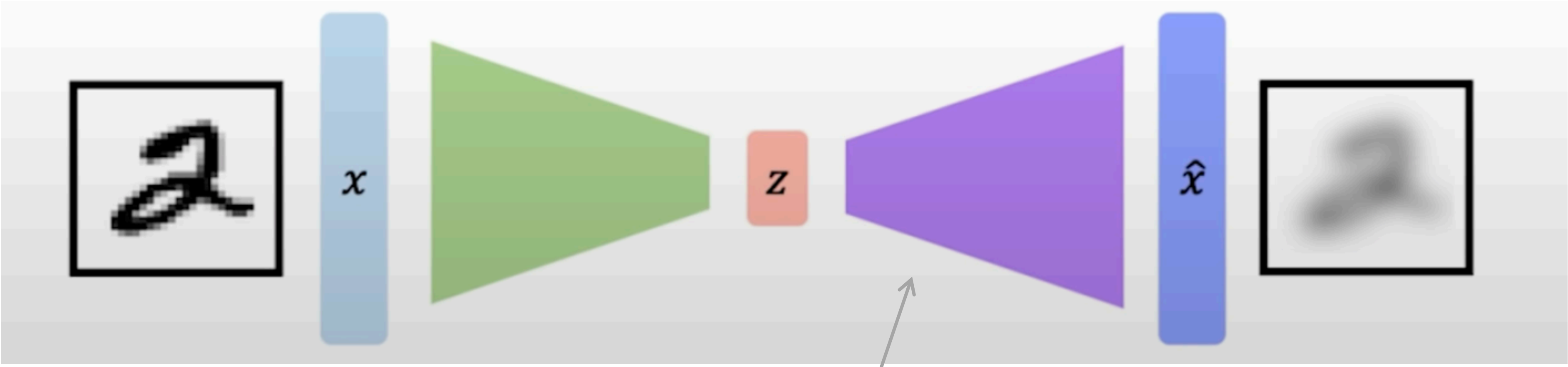
Generative adversarial networks

GAN = letting two neural networks **compete** with each other to **generate** new data

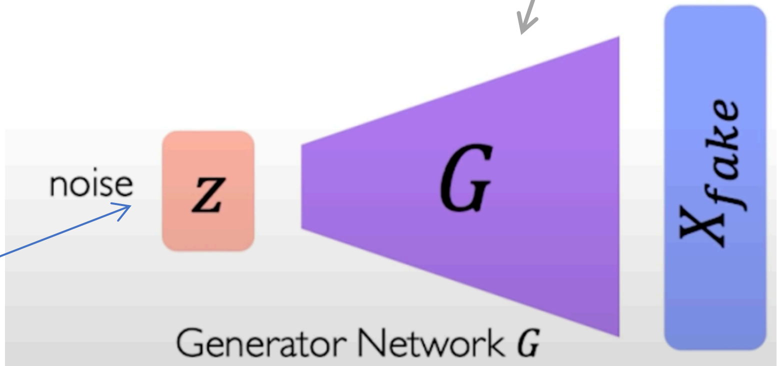
- The original purpose is to generate new data
- Classically for generating new images, but applicable to wide range of domains
- Learns the training set distribution and can generate new images that have never been seen before
- In contrast to e.g., autoregressive models or RNNs (generating one word at a time), GANs generate the whole output all at once



Generative adversarial networks



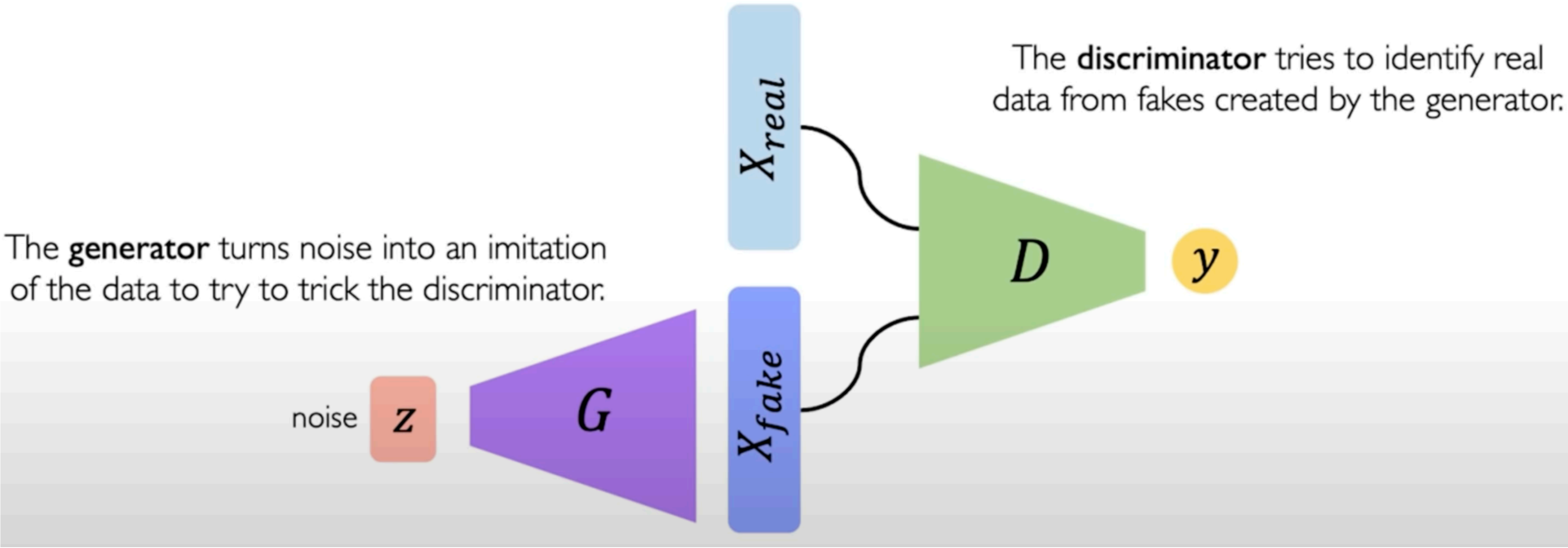
well organized latent space from which we can sample -> Variational Autoencoders



add noise to
latent space

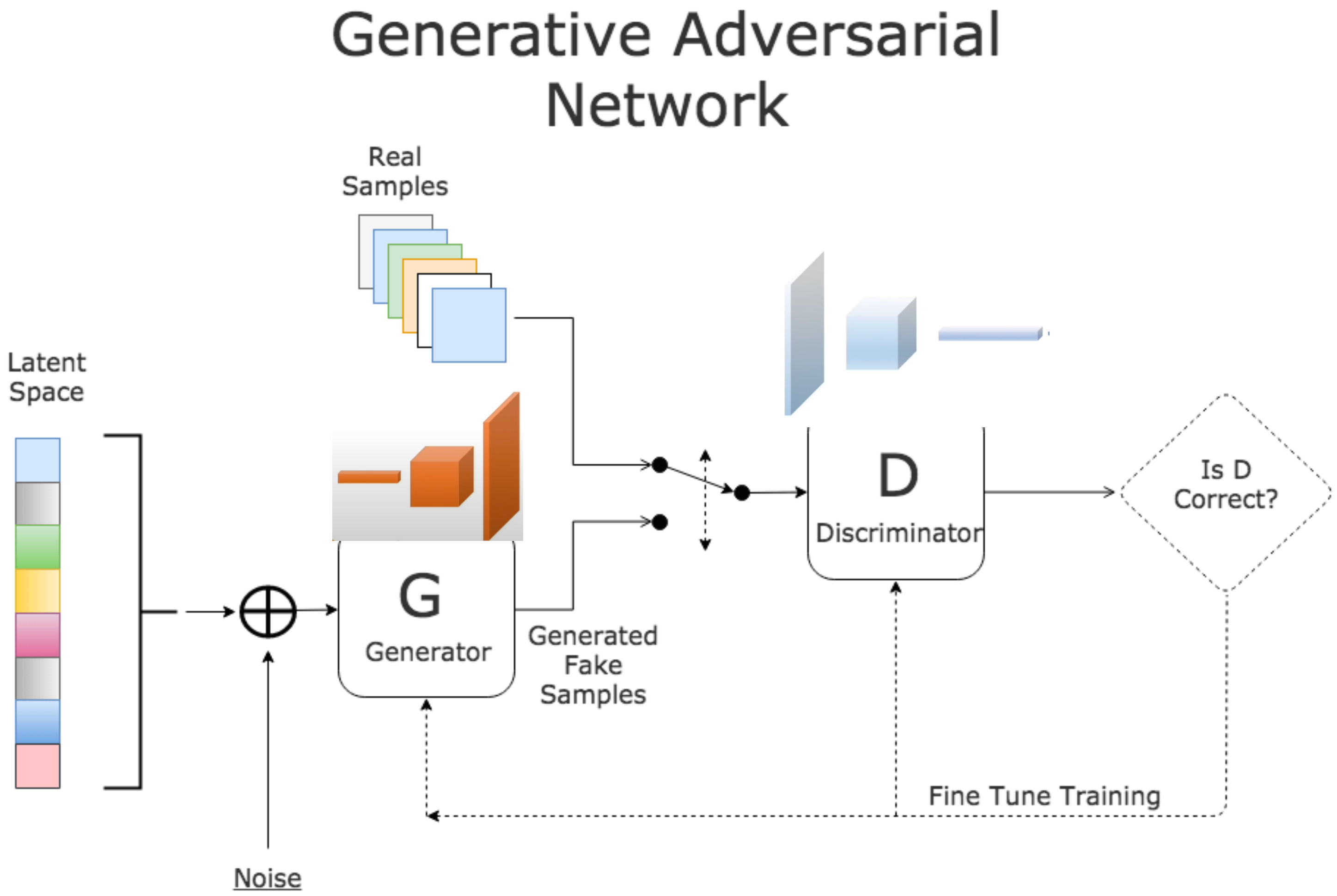
youtube.com/watch?v=BUNl0To1IVw

Generative adversarial networks



youtube.com/watch?v=BUNl0To1IVw

Generative adversarial networks



www.kdnuggets.com/2017/01/generative-adversarial-networks-hot-topic-machine-learning.html

Generative adversarial networks

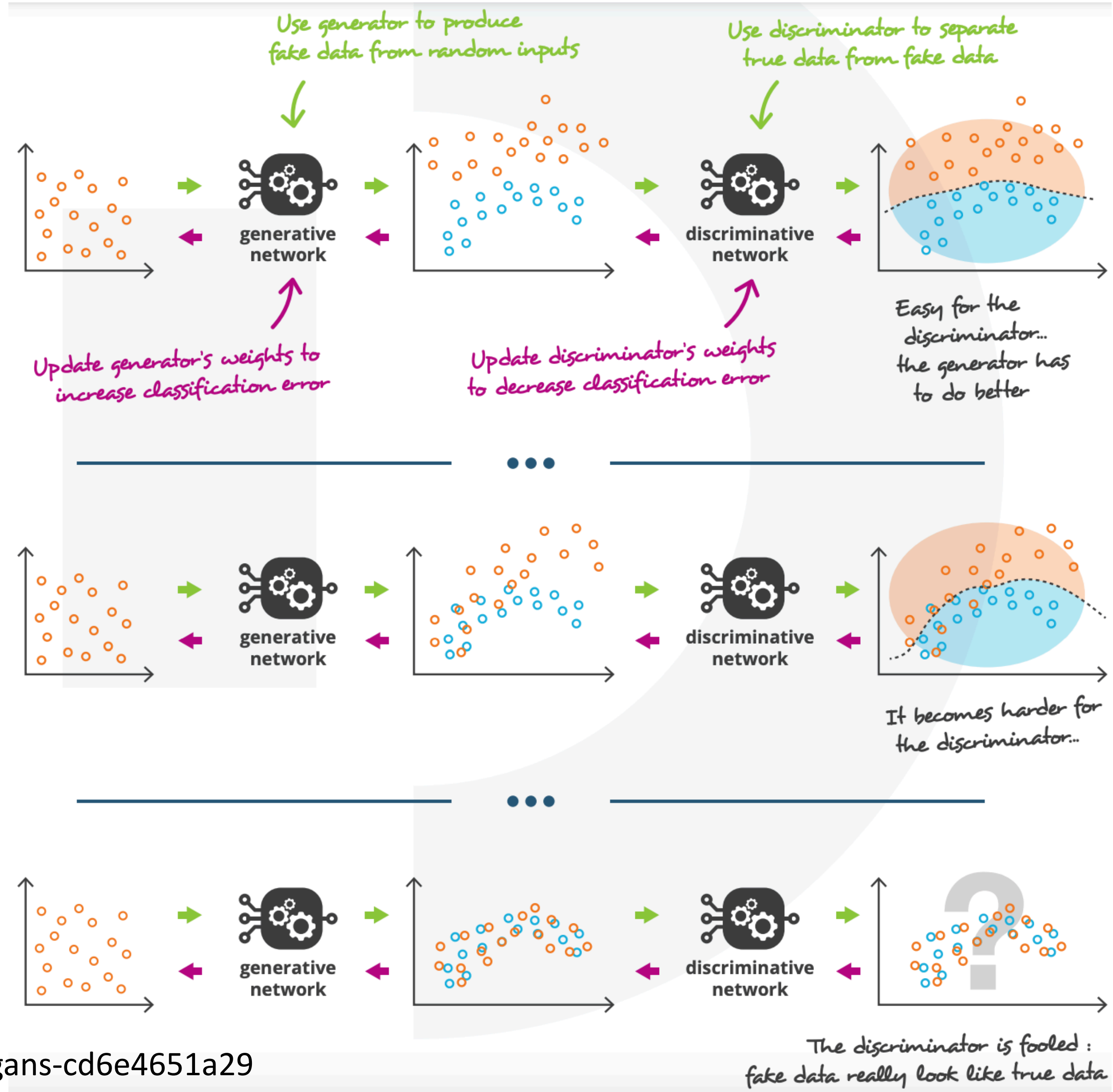
Training GANs

probability that D-fake is actually fake

$$\arg\min_{G}\max_{D}\ \mathbb{E}_{\mathbf{z},\mathbf{x}}\left[\ \log D(G(\mathbf{z})) + \log\left(1 - D(\mathbf{x})\right)\ \right]$$

D-real is real

saddle point



Convergence means that the G produces realistic images and the D outputs random predictions (probabilities close to 0.5)

Traininng problems:

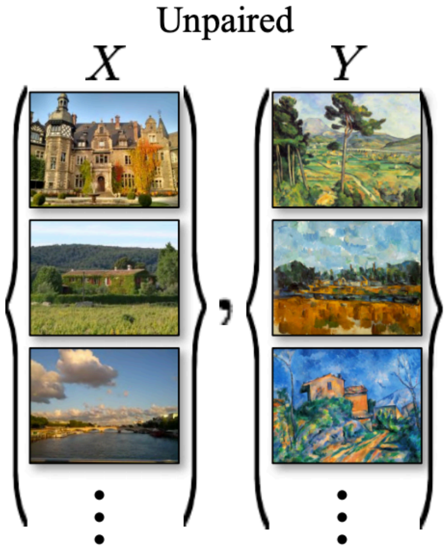
- oscillation between generator and discriminator loss
- mode collapse (generator produces examples of a particular kind only)

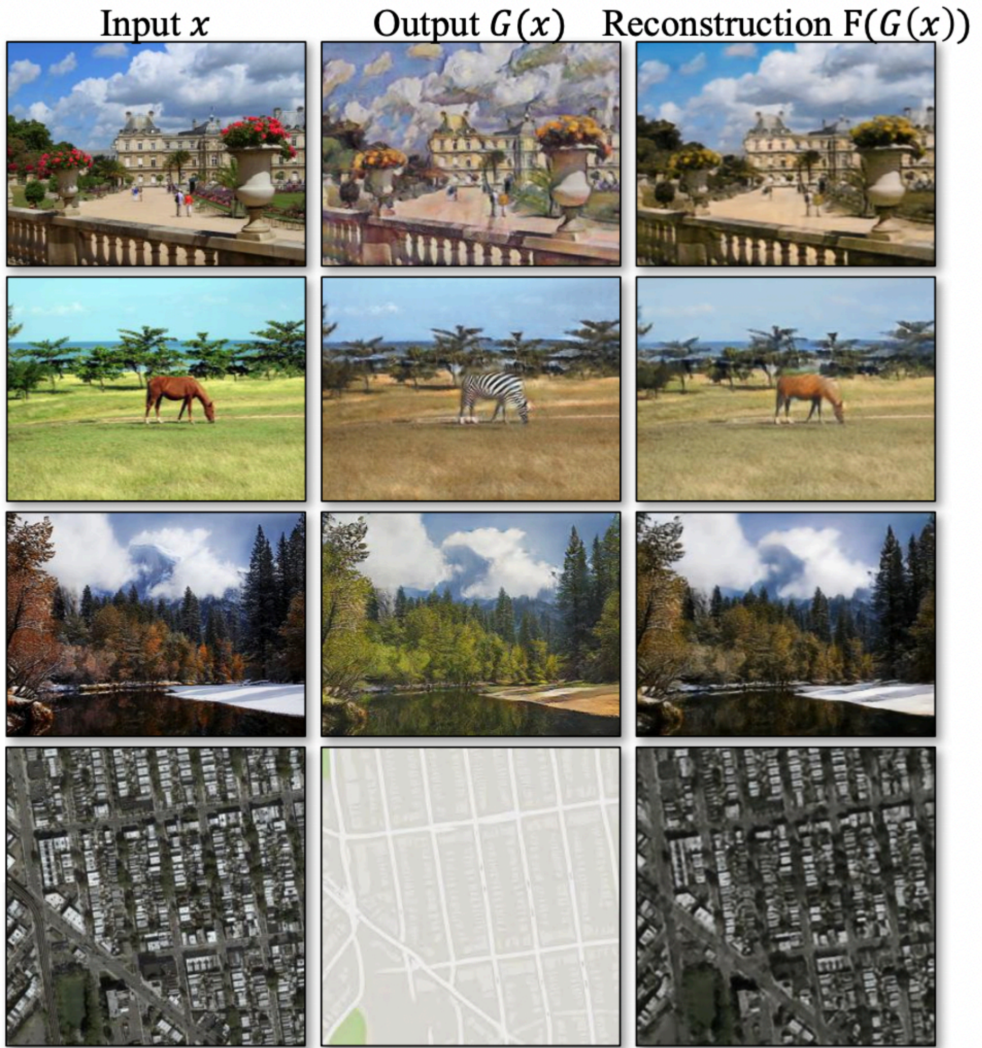
towardsdatascience.com/understanding-generative-adversarial-networks-gans-cd6e4651a29

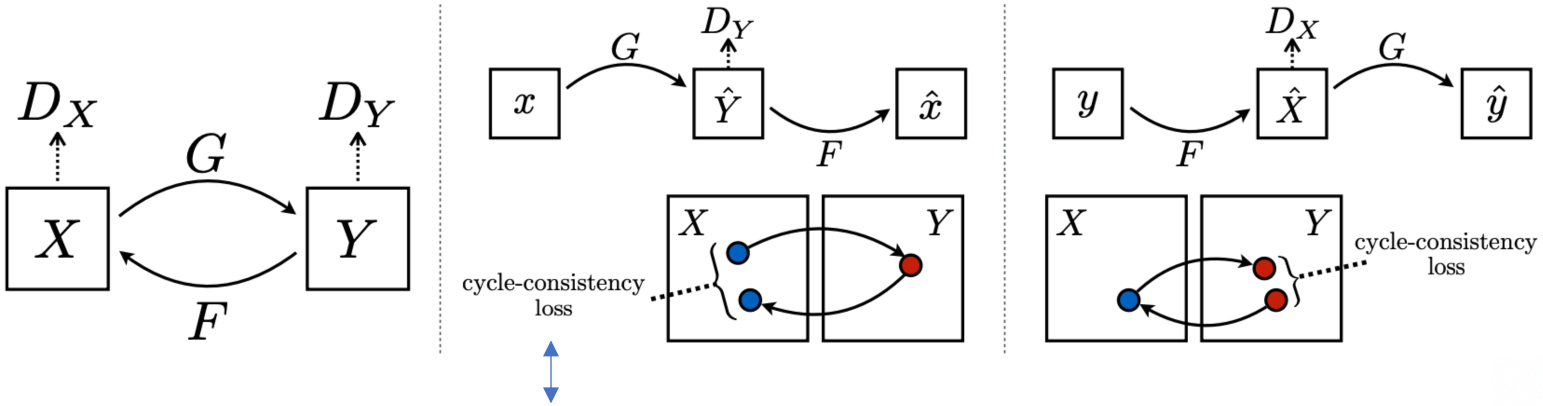
Cycle GANs

Cycle GAN
-> unpaired Image-to-Image translation

https://arxiv.org/pdf/1703.10593.pdf



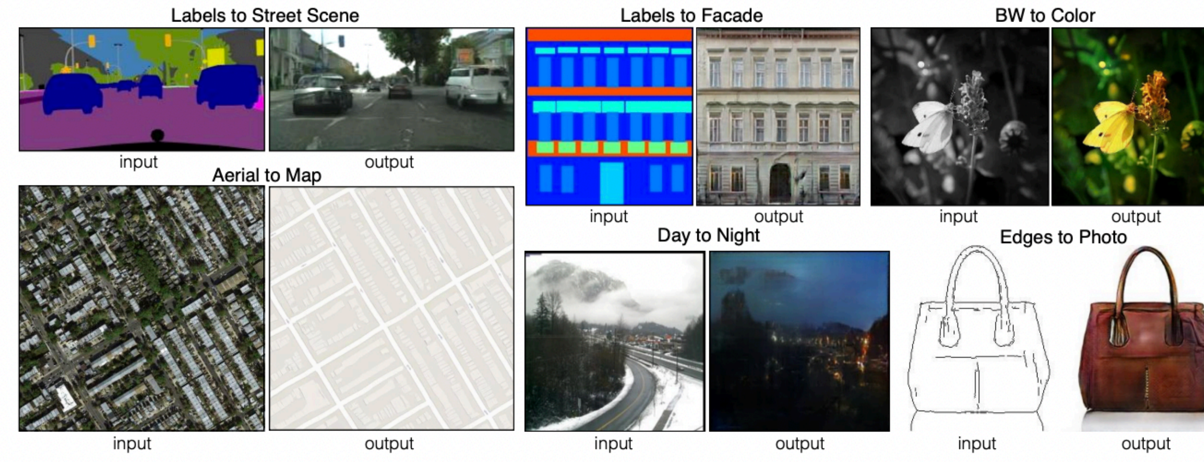




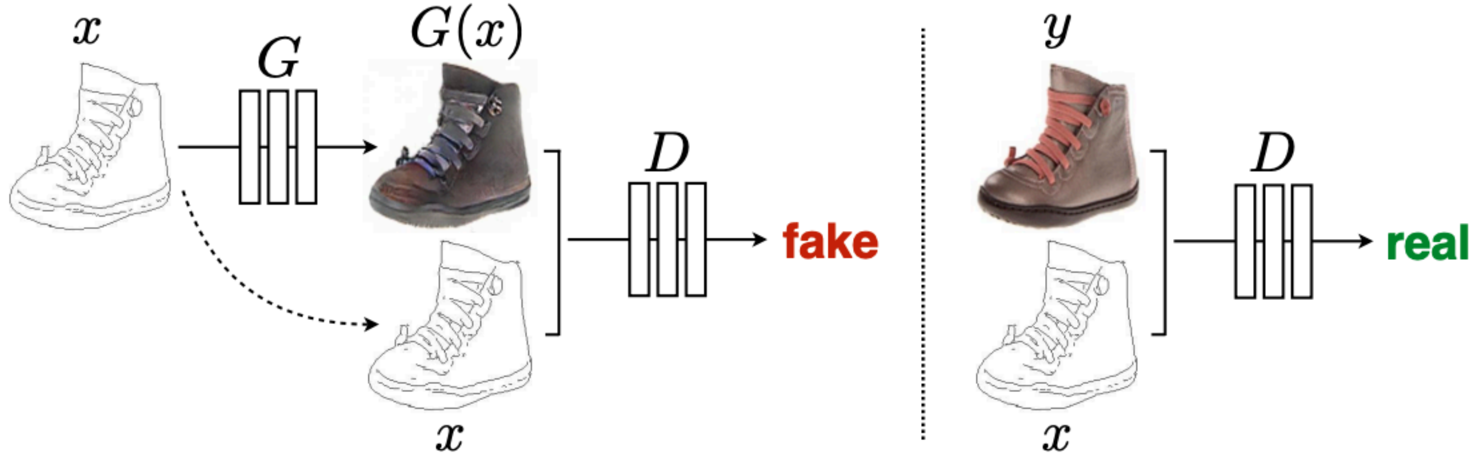
$$x \rightarrow G(x) \rightarrow F(G(x)) \approx x$$

Conditional GANs

Conditional GAN
-> e.g. pix2pix model



Unlike an unconditional GAN, in CGANs both the generator and discriminator observe the input edge map:

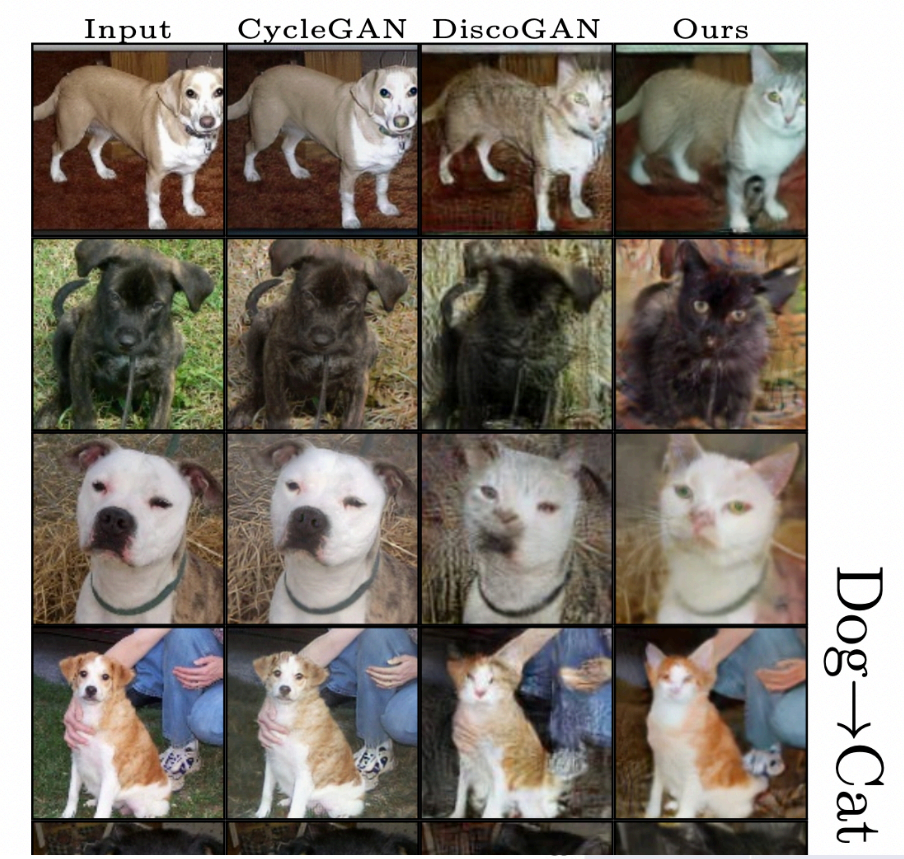


https://arxiv.org/pdf/1611.07004.pdf https://affinelayer.com/pixsrv/

Generative adversarial networks zoo

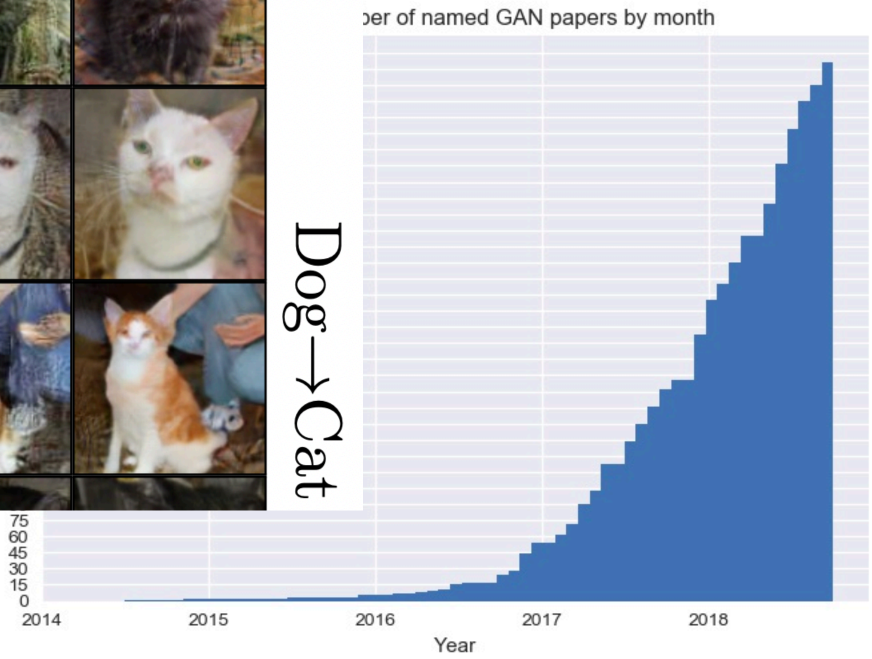
Many more flavors of GANs

github.com/hindupuravinash/the-gan-zoo



GANimorph

> 500 @ 2019



- 3D-ED-GAN - Shape Inpainting using 3D Generative Adversarial Network and Recurrent Convolutional Networks
- 3D-GAN - Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling (github)
- 3D-IWGAN - Improved Adversarial Systems for 3D Object Generation and Reconstruction (github)
- 3D-PhysNet - 3D-PhysNet: Learning the Intuitive Physics of Non-Rigid Object Deformations
- 3D-RecGAN - 3D Object Reconstruction from a Single Depth View with Adversarial Learning (github)
- ABC-GAN - ABC-GAN: Adaptive Blur and Control for improved training stability of Generative Adversarial Networks (github)
- ABC-GAN - GANs for LIFE: Generative Adversarial Networks for Likelihood Free Inference
- AC-GAN - Conditional Image Synthesis With Auxiliary Classifier GANs
- acGAN - Face Aging With Conditional Generative Adversarial Networks
- ACGAN - Coverless Information Hiding Based on Generative adversarial networks
- acGAN - On-line Adaptative Curriculum Learning for GANs
- ACtuAL - ACtuAL: Actor-Critic Under Adversarial Learning
- AdaGAN - AdaGAN: Boosting Generative Models
- Adaptive GAN - Customizing an Adversarial Example Generator with Class-Conditional GANs
- AdvEntuRe - AdvEntuRe: Adversarial Training for Textual Entailment with Knowledge-Guided Examples
- AdvGAN - Generating adversarial examples with adversarial networks
- AE-GAN - AE-GAN: adversarial eliminating with GAN
- AE-OT - Latent Space Optimal Transport for Generative Models
- AEGAN - Learning Inverse Mapping by Autoencoder based Generative Adversarial Nets
- AF-DCGAN - AF-DCGAN: Amplitude Feature Deep Convolutional GAN for Fingerprint Construction in Indoor Localization System
- AffGAN - Amortised MAP Inference for Image Super-resolution
- AIM - Generating Informative and Diverse Conversational Responses via Adversarial Information Maximization
- AL-CGAN - Learning to Generate Images of Outdoor Scenes from Attributes and Semantic Layouts
- ALI - Adversarially Learned Inference (github)
- AlignGAN - AlignGAN: Learning to Align Cross-Domain Images with Conditional Generative Adversarial Networks
- AlphaGAN - AlphaGAN: Generative adversarial networks for natural image matting
- AM-GAN - Activation Maximization Generative Adversarial Nets