Autoencoders

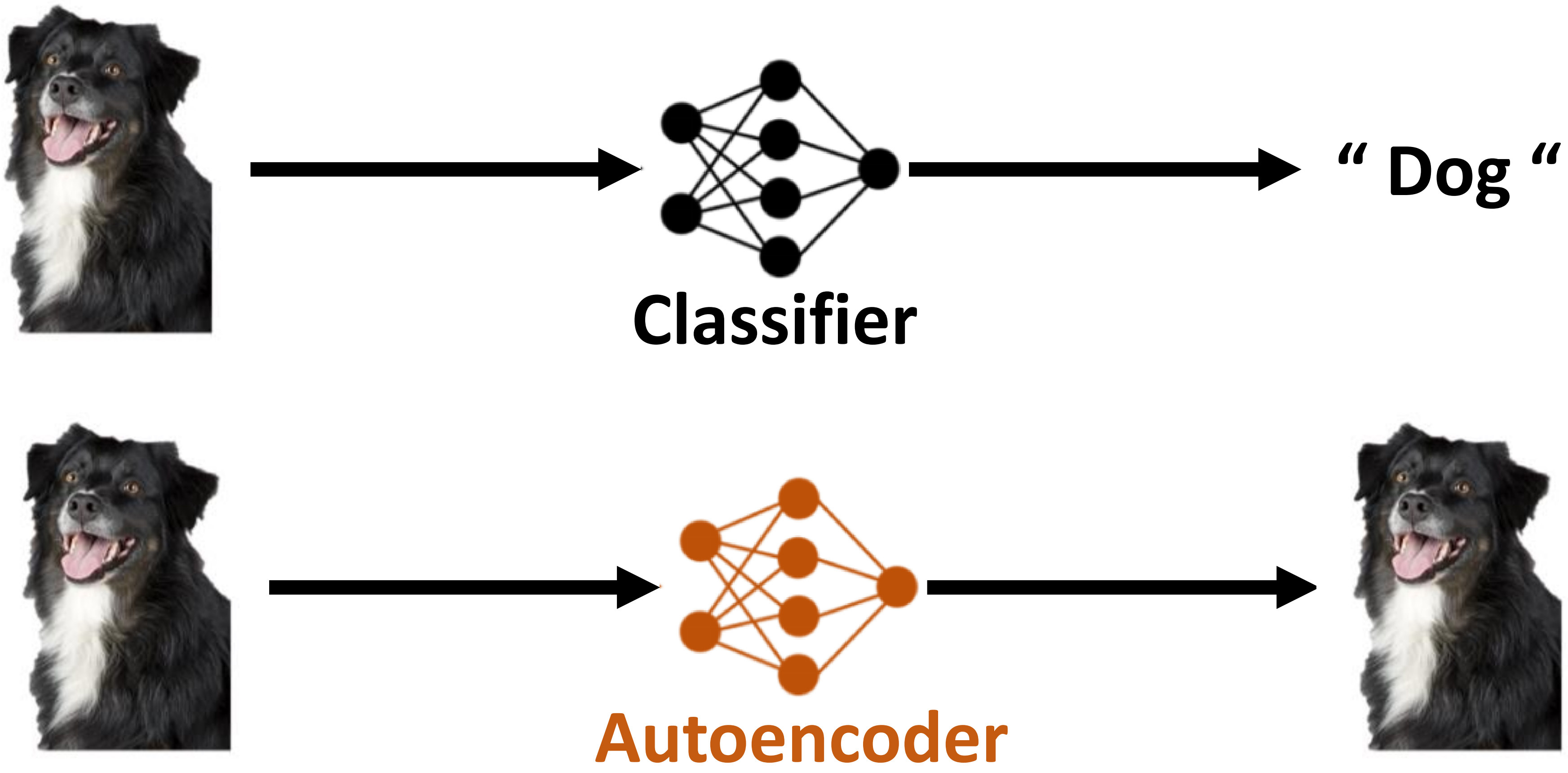
A brief introduction

Overview

- What are autoencoders?
- Toy Examples
- Neural Network Autoencoder
- PCA, and K-Means as an Autoencoder
- Variational Autoencoders
- Applications

What are autoencoders?

- Autoencoders are a type of neural networks that try to **reconstruct the provided input**



Where can I use autoencoders?

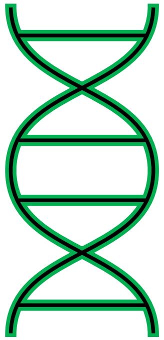
- Everywhere!

Image / Video

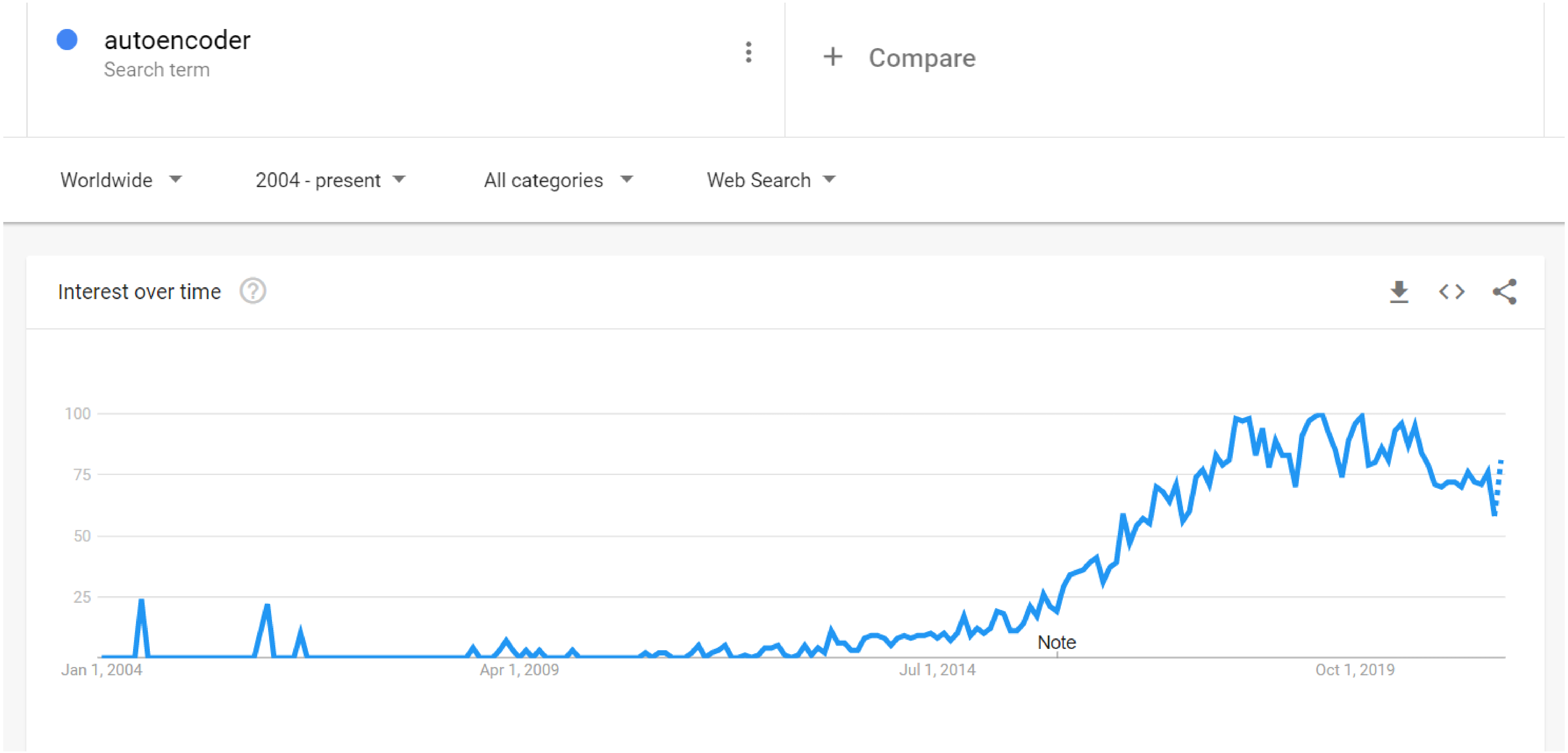
Lorem ipsum dolor sit am
eiusmod tempor incididun
enim ad minim veniam, qu
nisi ut aliquip ex ea co
reprehenderit in volupta
nulla pariatur. Excepteu
sunt in culpa qui offici
Sed ut perspiciatis unde
accusantium doloremque l

Text

Audio

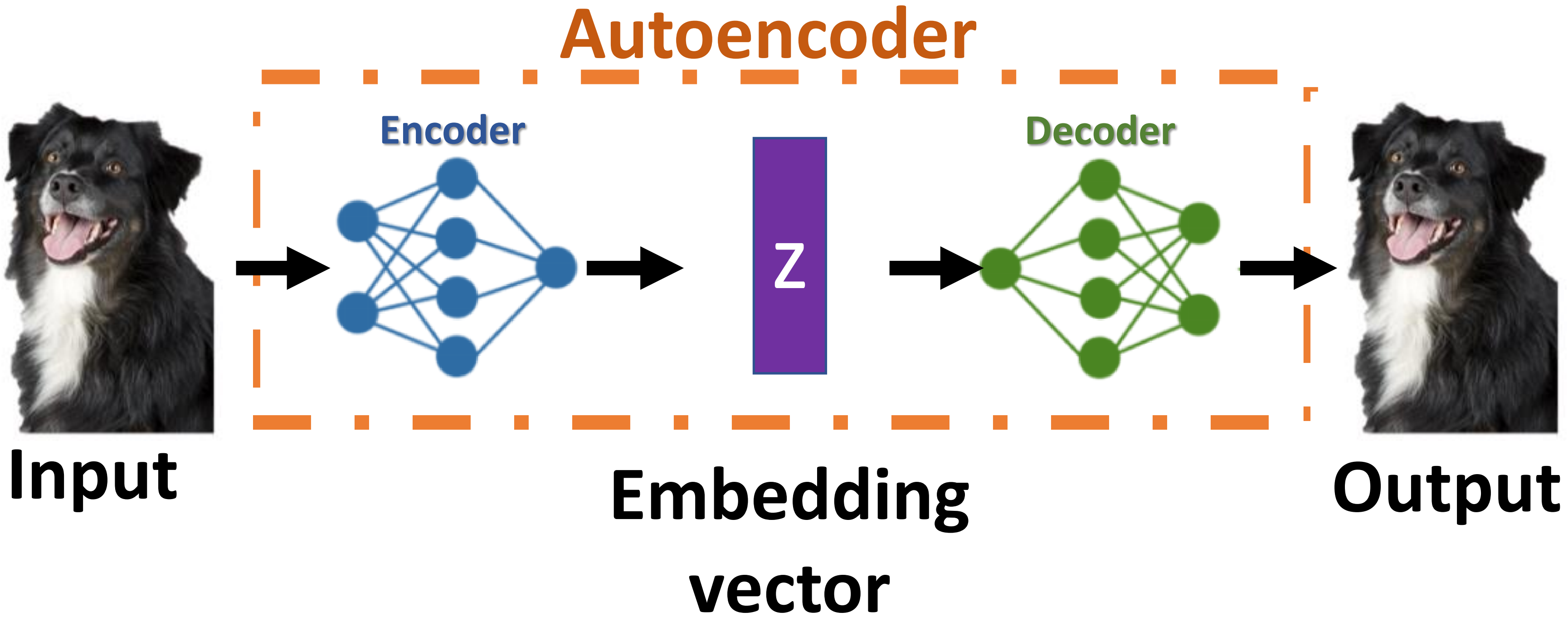


DNA


autoencoder
Search term
+ Compare
Worldwide
2004 - present
All categories
Web Search
Interest over time
100
75
50
25
Note
Jan 1, 2004
Apr 1, 2009
Jul 1, 2014
Oct 1, 2019

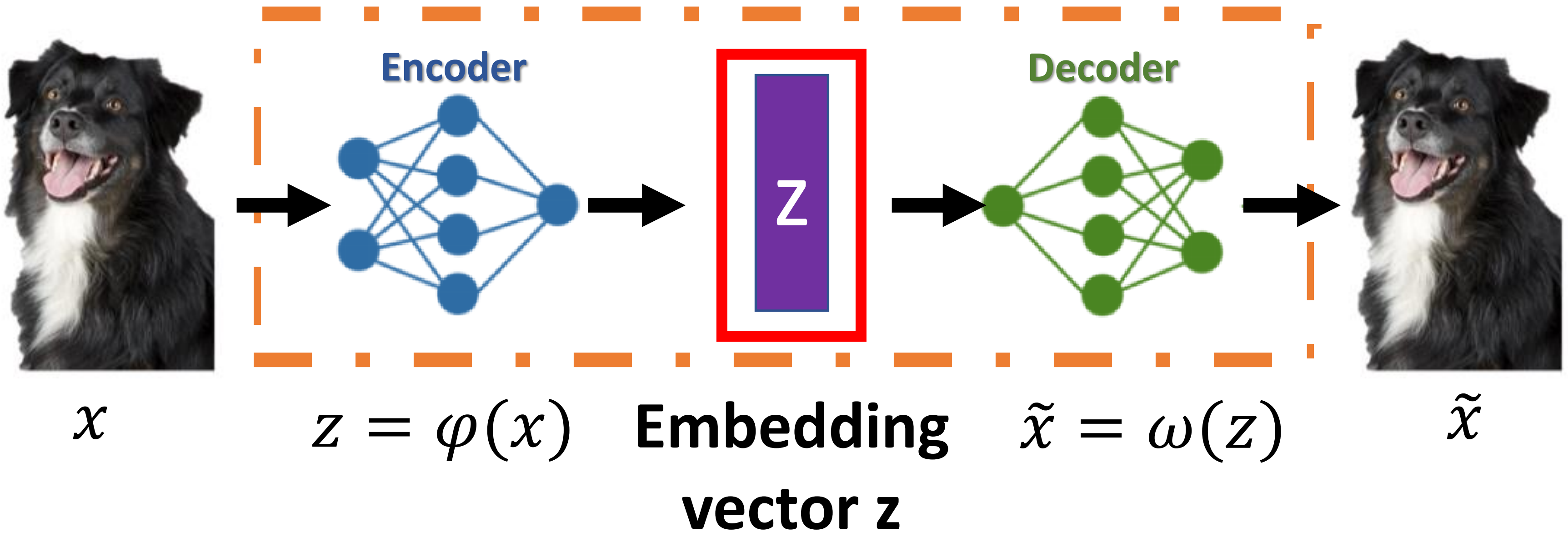
What are autoencoders?

- Autoencoders typically have two components:
  - **Encoder:** maps input (x) into an intermediate representation (z)
  - **Decoder:** maps the intermediate representation (z) into the input (x)



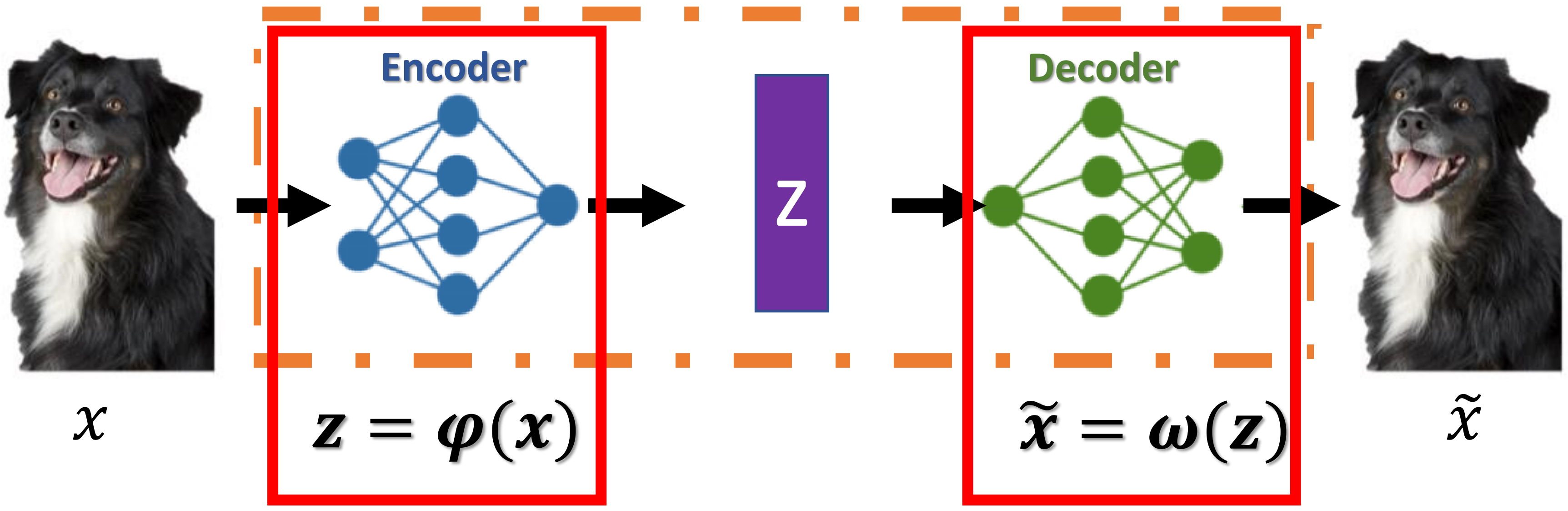
What are autoencoders?

- Intermediate representation (z) = embedding vector, hidden representation, bottleneck, latent space, code, …
- Ideally, z will capture the **essential information** of the data



What are autoencoders?

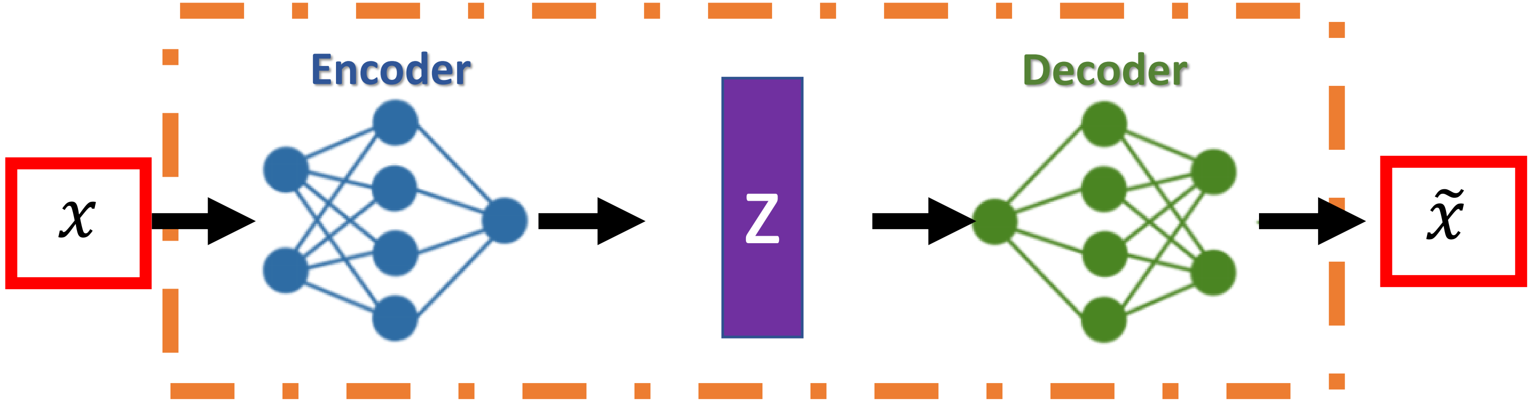
- The **encoder, $\boldsymbol{\varphi(x)}$,** and **decoder, $\boldsymbol{\omega(z)}$,** are typically neural networks: *Multi-layer perceptron (MLP), convolutional neural networks (CNN), recurrent neural networks (RNN), graph neural networks (GNN), transformers…*



What are autoencoders?

- The parameters of **encoder, $\boldsymbol{\varphi(x)}$,** and **decoder, $\boldsymbol{\omega(z)}$,** are trained by minimizing the reconstruction error

$$Error = \|x - \tilde{x}\|^2 = \|x - \boldsymbol{\omega(\varphi(x))}\|^2$$



$\boldsymbol{z = \varphi(x)}$

$\boldsymbol{\tilde{x} = \omega(z)}$

Toy examples

$\underline{\boldsymbol{z = \varphi(x)}}$ $\underline{\boldsymbol{\tilde{x} = \omega(z)}}$

$z = 0.5\, x$ $\hat{x} = 2z = x$

$z = Ax$ $\hat{x} = Bz = BAx$

$z = Wx$ $\hat{x} = W^T z = W^T W x$

PCA is a linear autoencoder

- A linear autoencoder will learn a rotated Principal Component Analysis projection / a Singular Value Decomposition

$$\underline{\boldsymbol{z} = \boldsymbol{\varphi}(\boldsymbol{x})} \qquad \underline{\widetilde{\boldsymbol{x}} = \boldsymbol{\omega}(\boldsymbol{z})}$$

$$z = Wx \qquad \hat{x} = W^T z = W^T W x$$

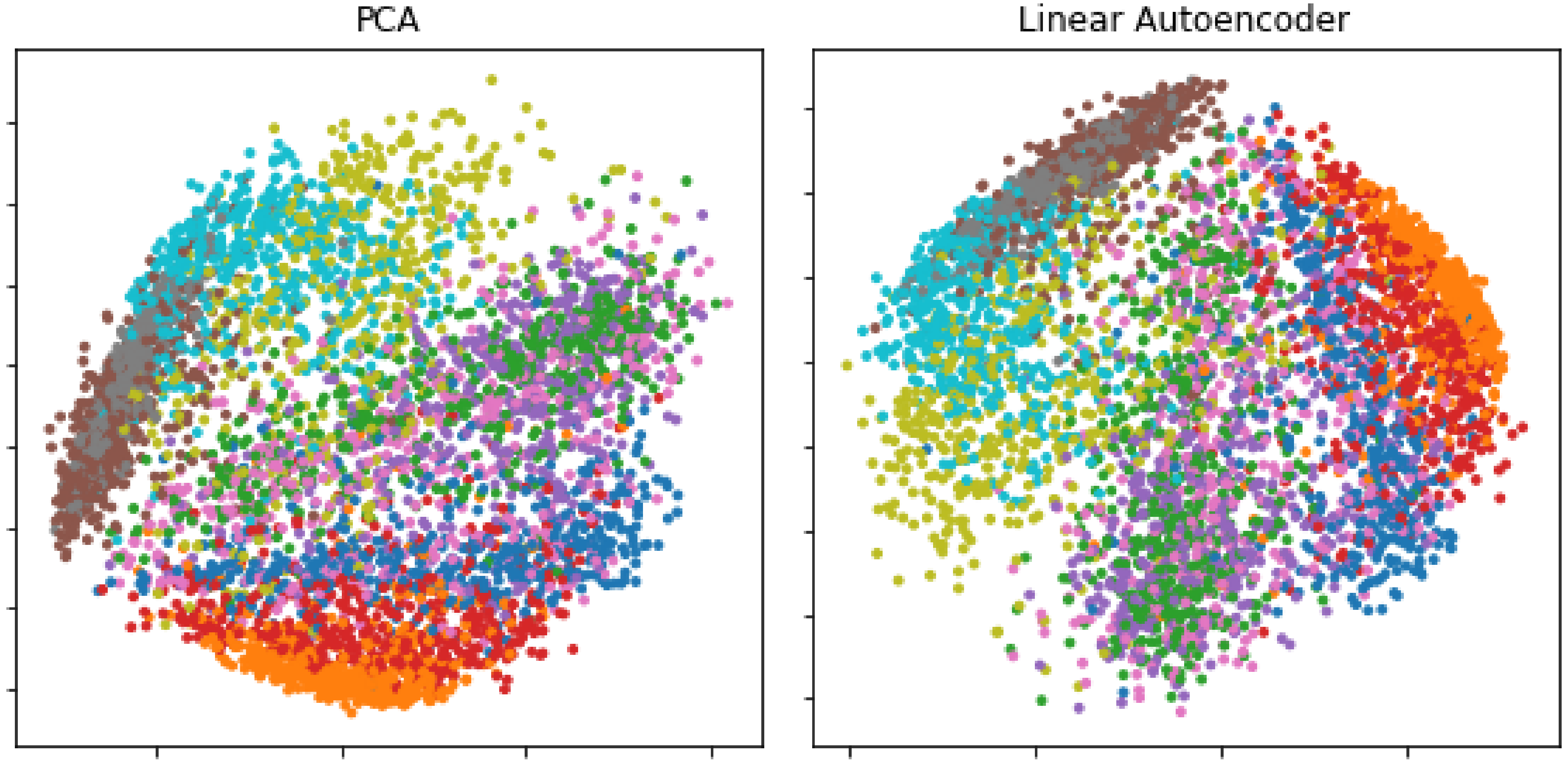
$$Z = XW$$

Principal Components

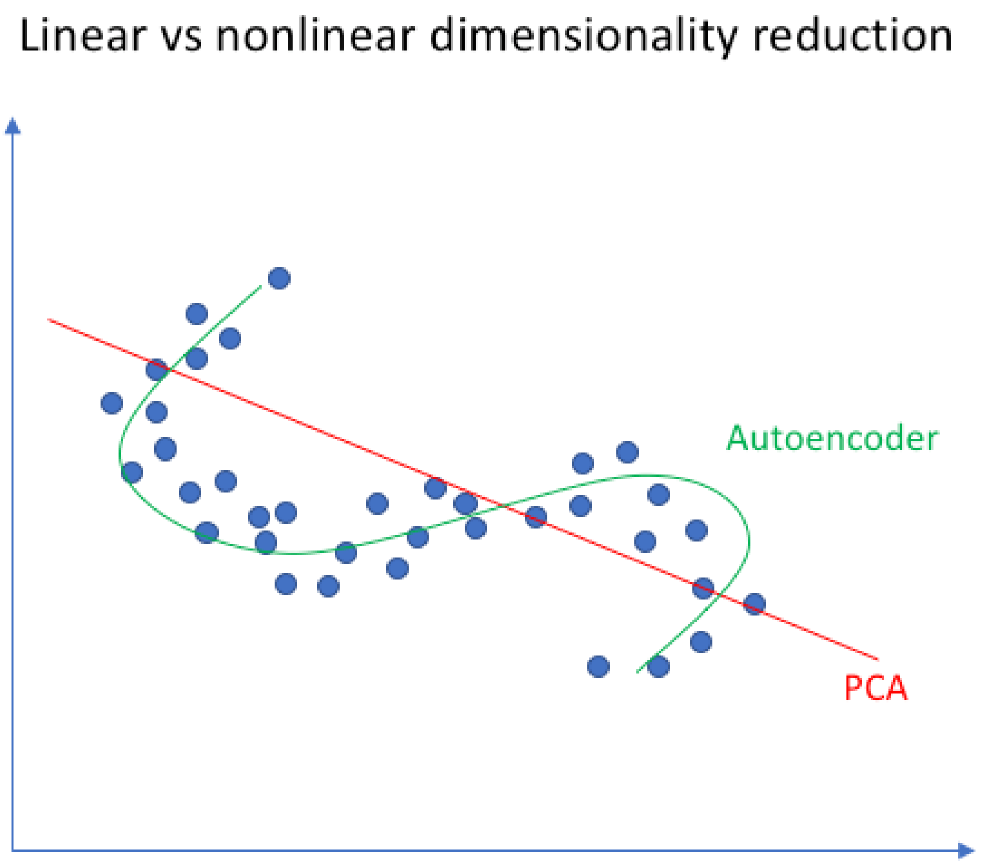
$$U\Sigma = XW$$

$$U\Sigma W^T = X$$

PCA is a linear autoencoder

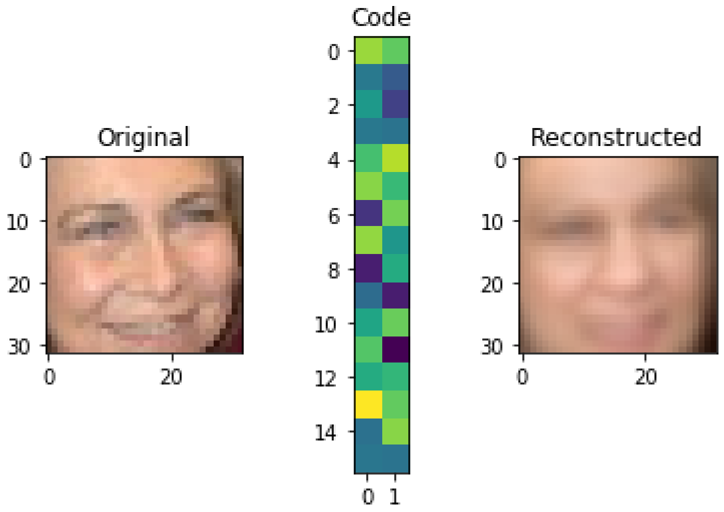


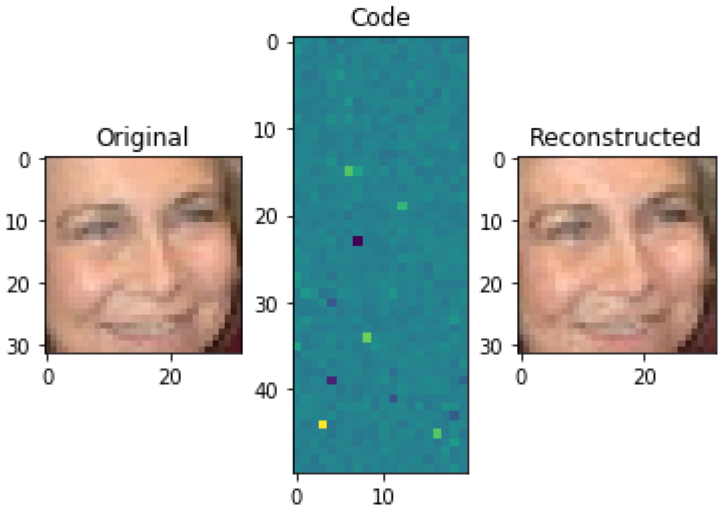
Why non-linear autoencoders?



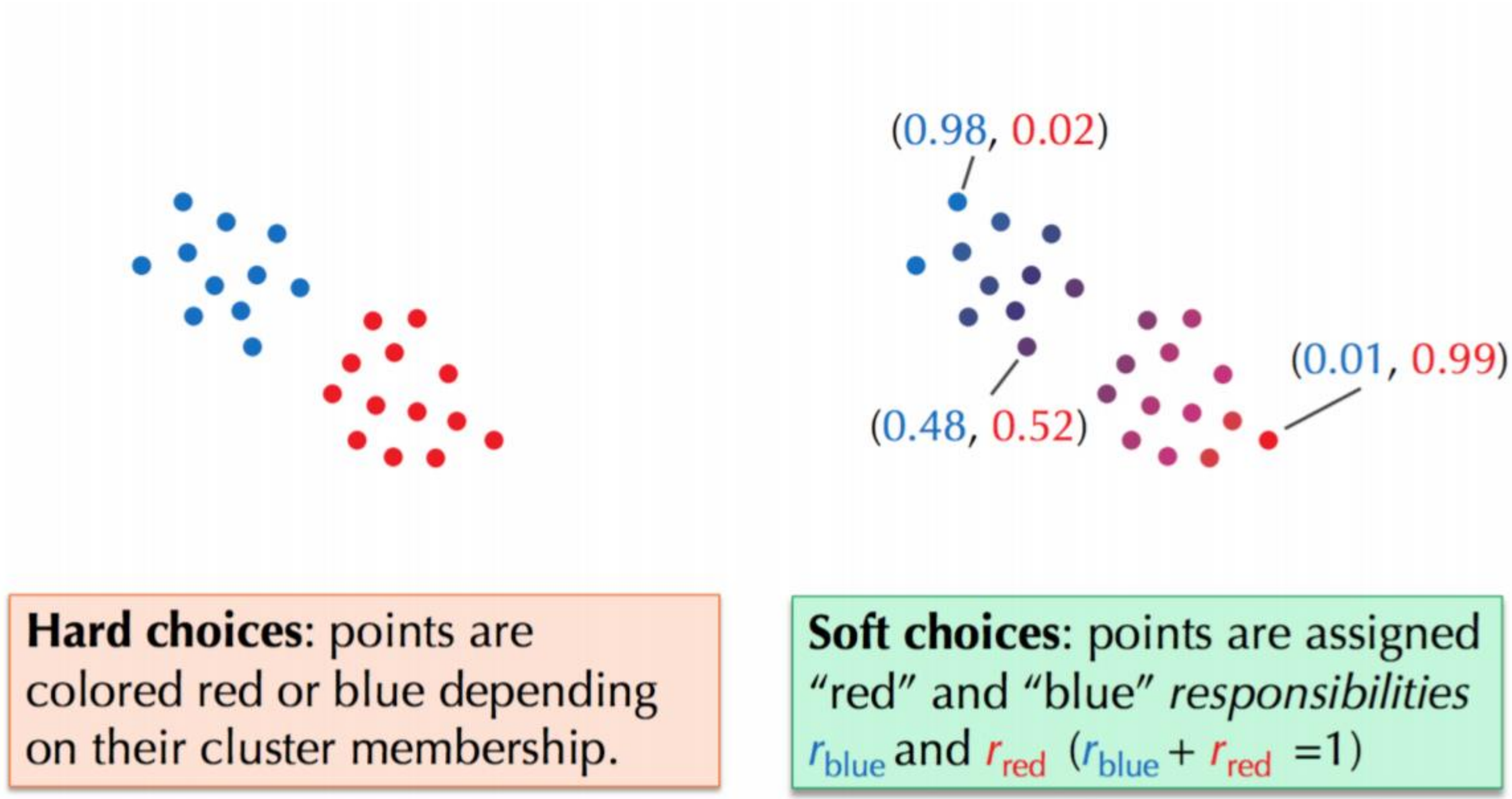
Neural network autoencoder

- The size of the embedding will affect how much the data is compressed, and how good is the reconstruction error





K-Means vs Soft K-Means



https://www.cs.cmu.edu/~02251/recitations/recitation_soft_clustering.pdf

K-Means vs Soft K-Means

$$X \approx QP$$

Cluster centroid

Cluster Assignment

K-Means $Q \in \mathbb{B} \quad \sum_j Q_{ij} = 1$ One-hot cluster indicator

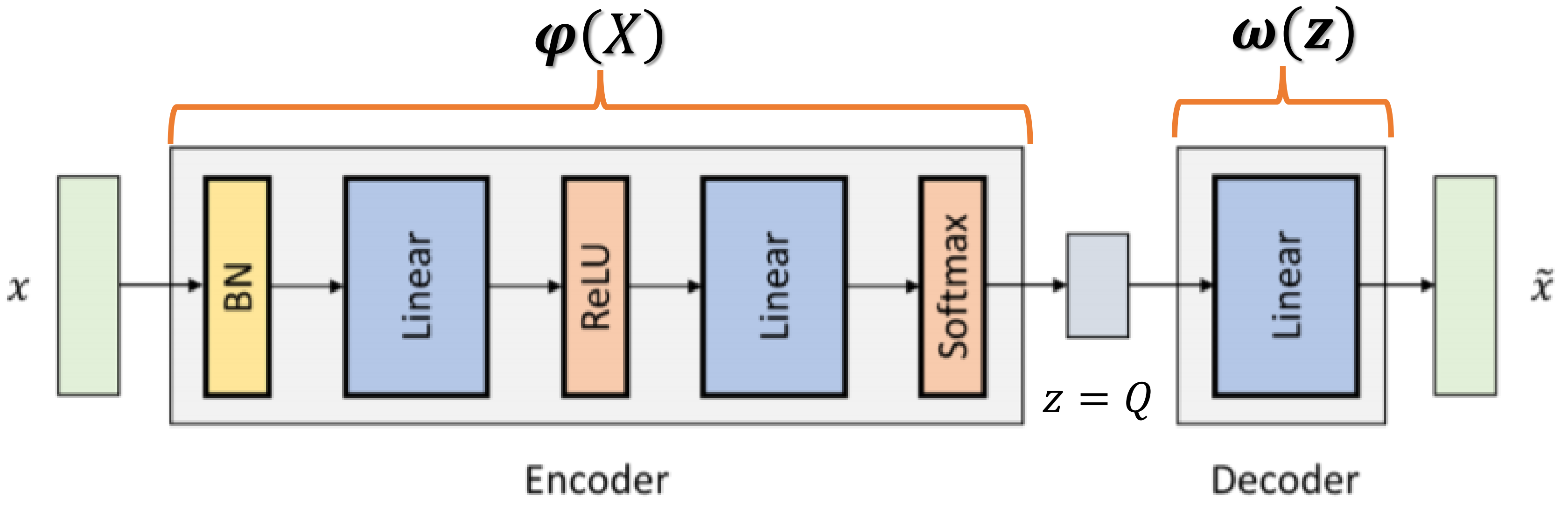
Soft K-Means $Q > 0 \quad \sum_j Q_{ij} = 1$ Cluster percentages

ADMIXTURE/Soft K-means as an Autoencoder

- ADMIXTURE is a likelihood approach typically used in population genetics similar to **Soft K-Means**

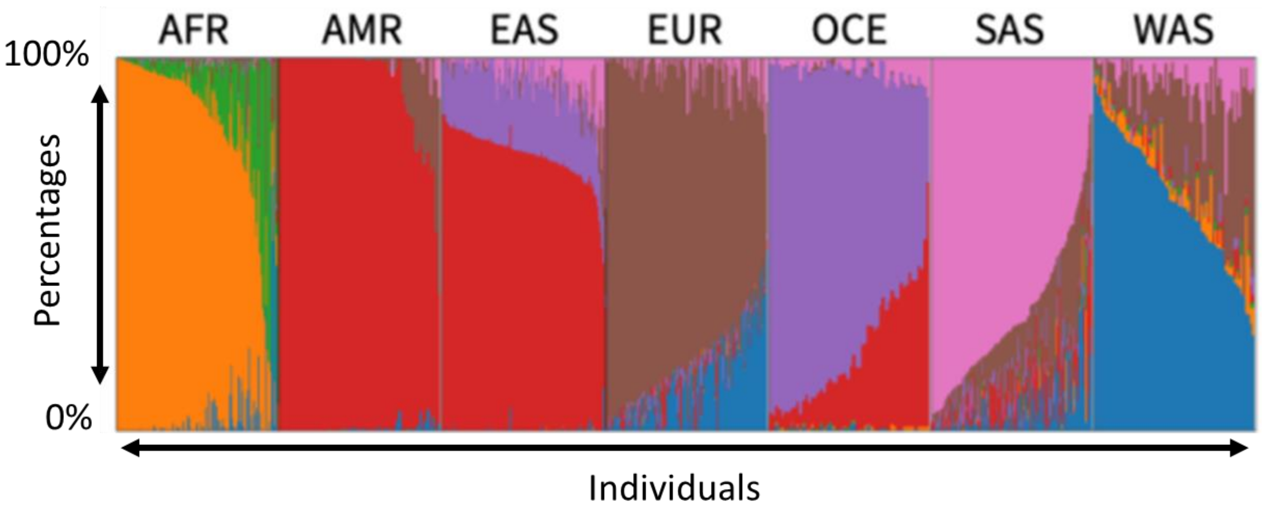
$$X \approx QP = \boldsymbol{\varphi}(X)P$$

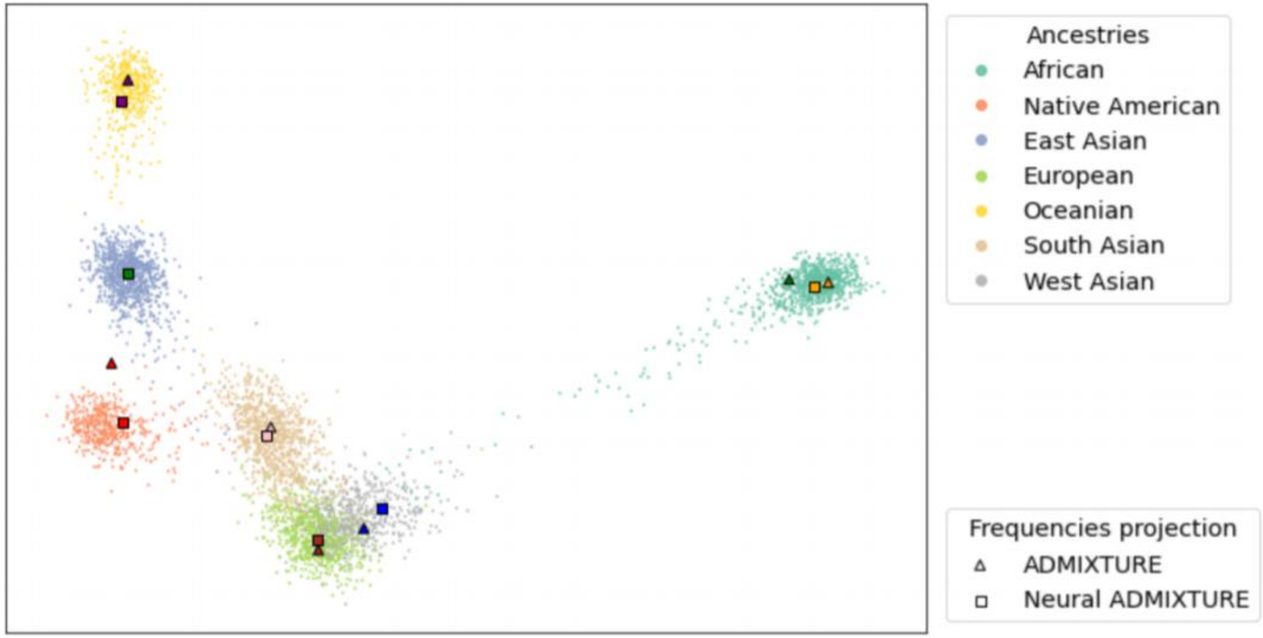
$$z = Q = \boldsymbol{\varphi}(X)$$
$$\tilde{\boldsymbol{x}} = \boldsymbol{\omega}(\boldsymbol{z}) = QP$$



https://www.biorxiv.org/content/10.1101/2021.06.27.450081v3.full.pdf

ADMIXTURE as an Autoencoder



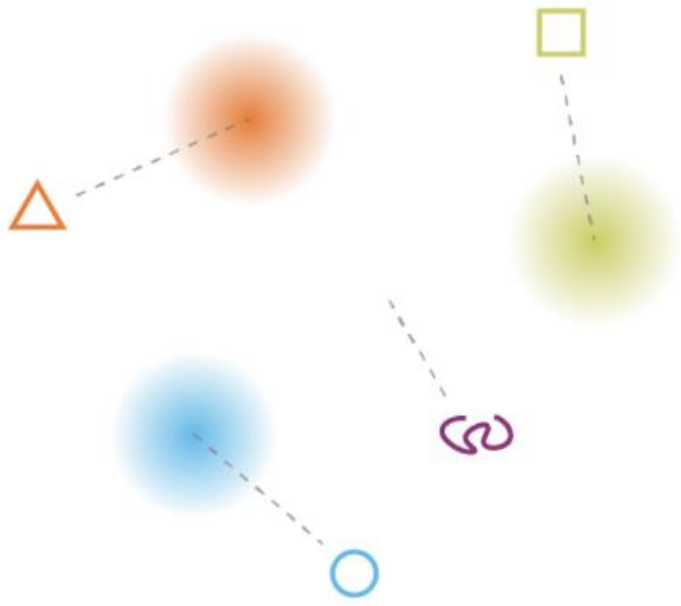


https://www.biorxiv.org/content/10.1101/2021.06.27.450081v3.full.pdf

Variational Autoencoders

Variational Autoencoder

- What if we inject a random vector into the decoder?
- How do we know which type of vector we need to input in order to get a good output?



https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73

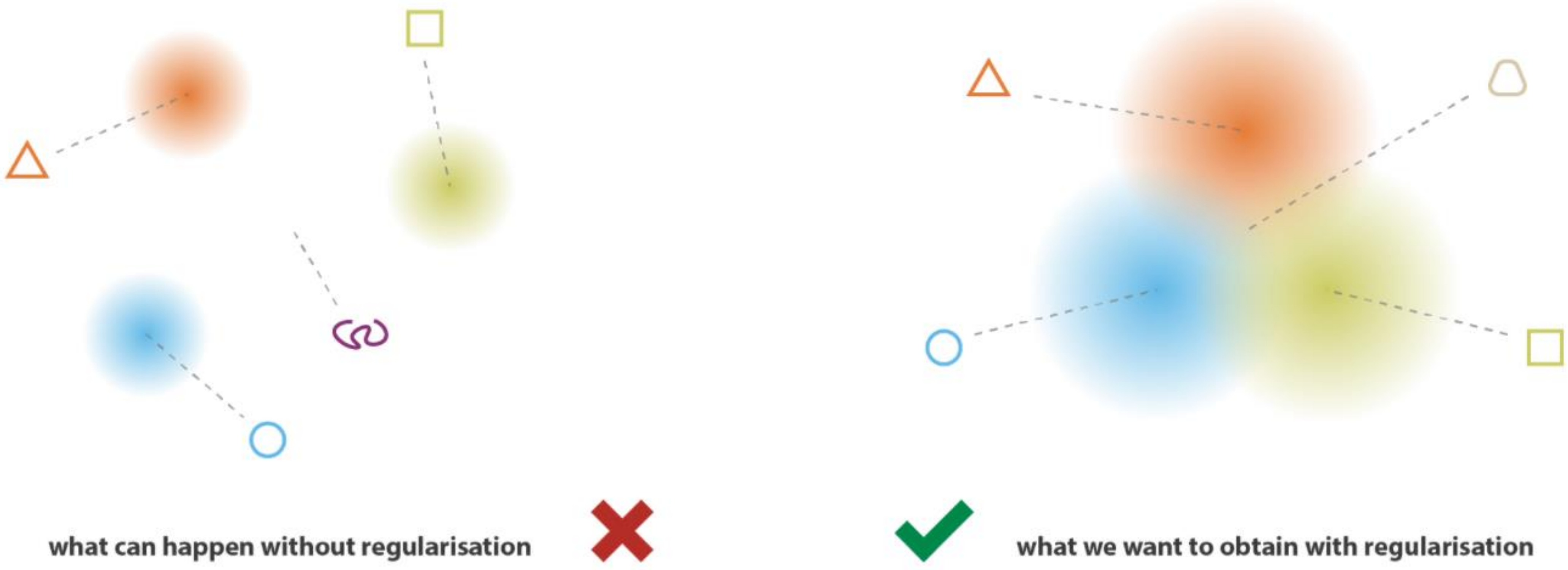
Variational Autoencoder

- We would like the embedding vectors to follow a know statistical distribution (e.g., **a gaussian**)

- If the latent vectors (z) follow a Gaussian distribution, they will be:
  - Centered and constrained: points closer to the origin will provide good simulations

  - Smooth: neighboring points will provide similar simulations

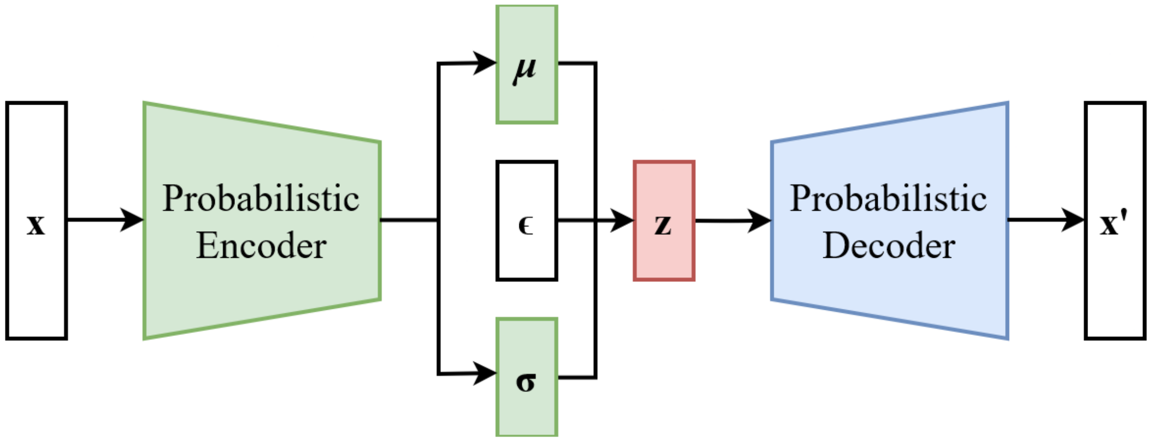
Variational Autoencoder



https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73

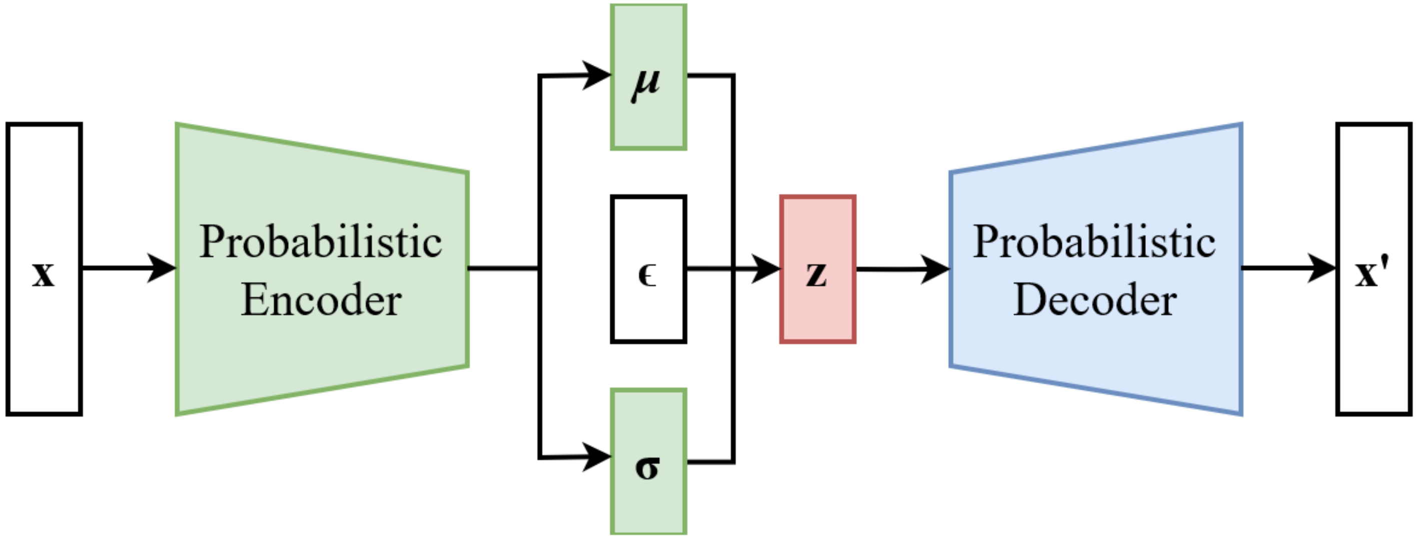
Variational Autoencoder

- **Centered and constrained:** an L2 regularization is applied to the mean of the latent code → larger values are penalized
- **Smooth:** small gaussian noise is applied to the latent code → reparameterization trick
- **Gaussian:** the KL Divergence between the z and a gaussian is applied



Kingma, D.P.; Welling, M. *Auto-encoding variational bayes*. arXiv preprint arXiv:1312.6114, 2013.

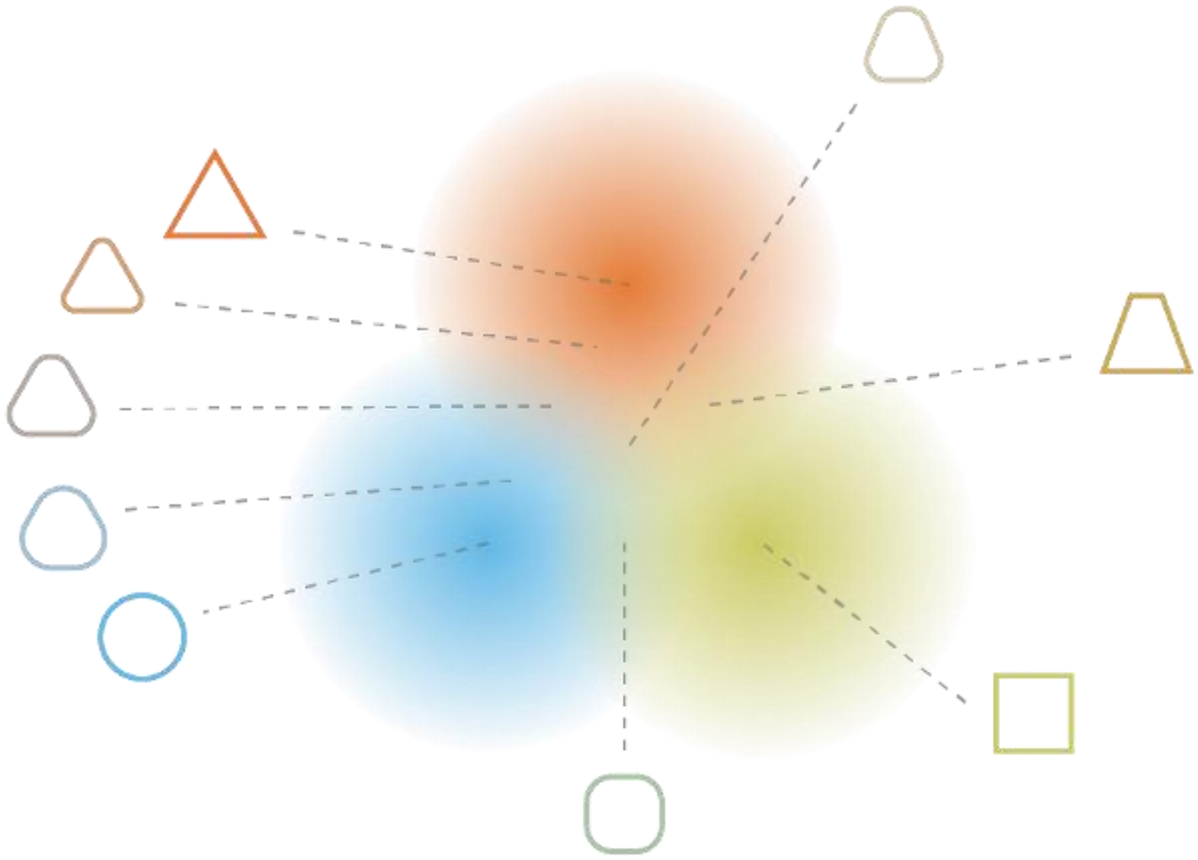
Variational Autoencoder



$$\mathbf{z} \sim q_\phi(\mathbf{z} \mid \mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \qquad \mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \boldsymbol{\varepsilon}.$$

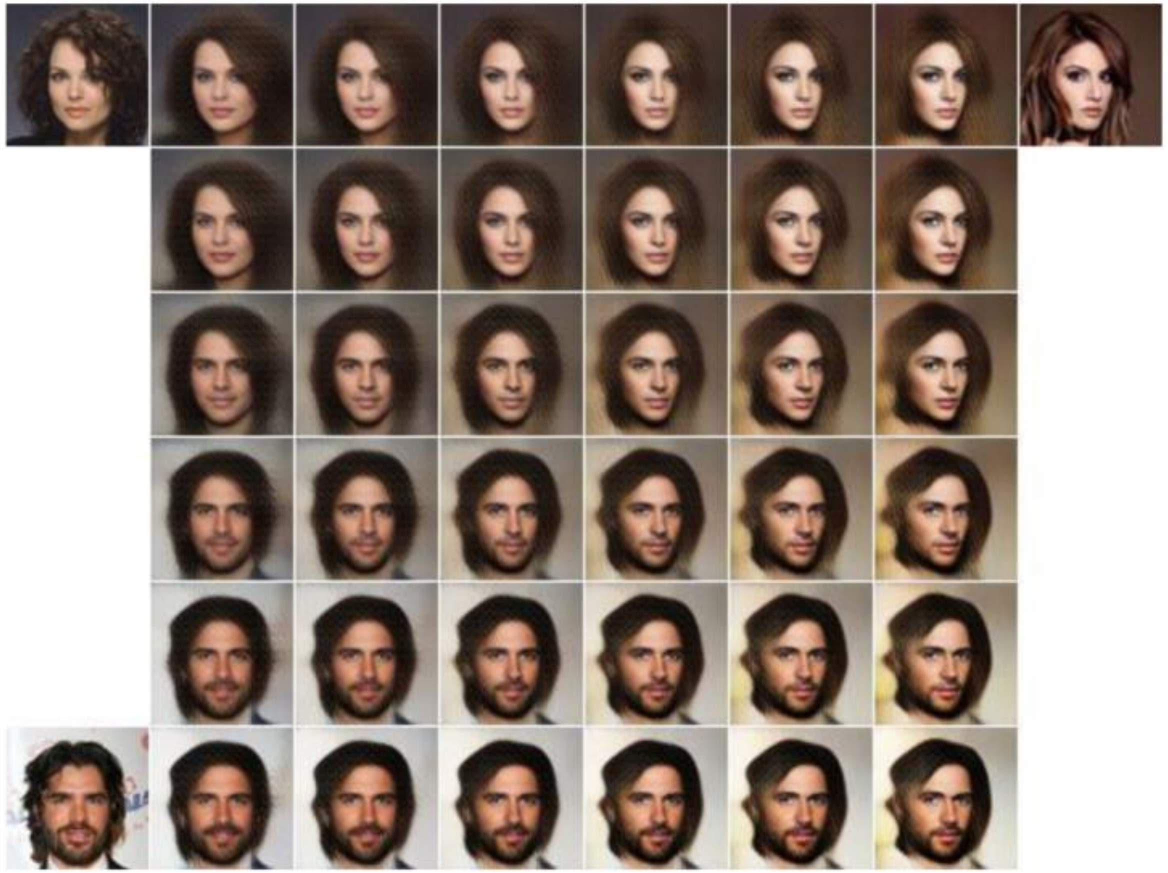
$$\mathcal{L} = -\sum_{j=1}^{J} \frac{1}{2}\left[1 + \log\left(\sigma_i^2\right) - \sigma_i^2 - \mu_i^2\right] - \frac{1}{L}\sum_{l} E_{\sim q_\theta(z|x_i)}\left[\log p(x_i|z^{(i,l)})\right]$$

Variational Autoencoder



https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73

Variational Autoencoder



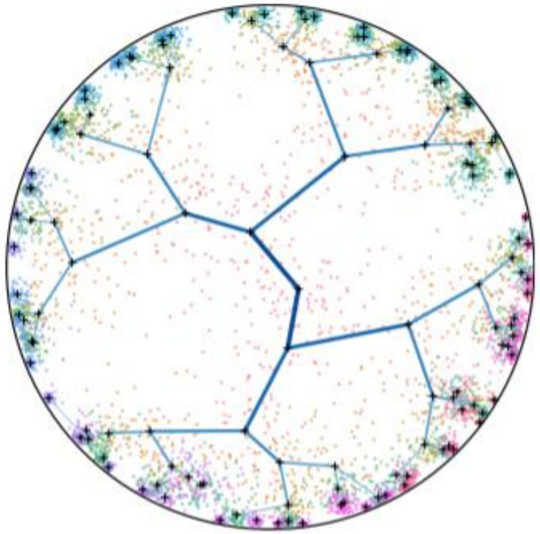
3-way **Latent space** interpolation for **faces**

https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73

Other autoencoders

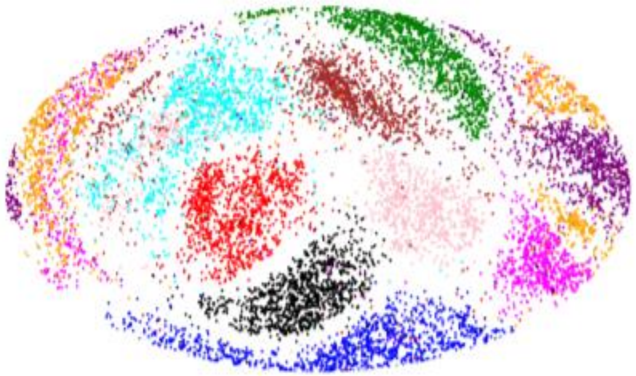
- Besides Gaussian VAE, there are many flavors:

Hyperbolical



https://arxiv.org/pdf/1901.06033.pdf

Spherical

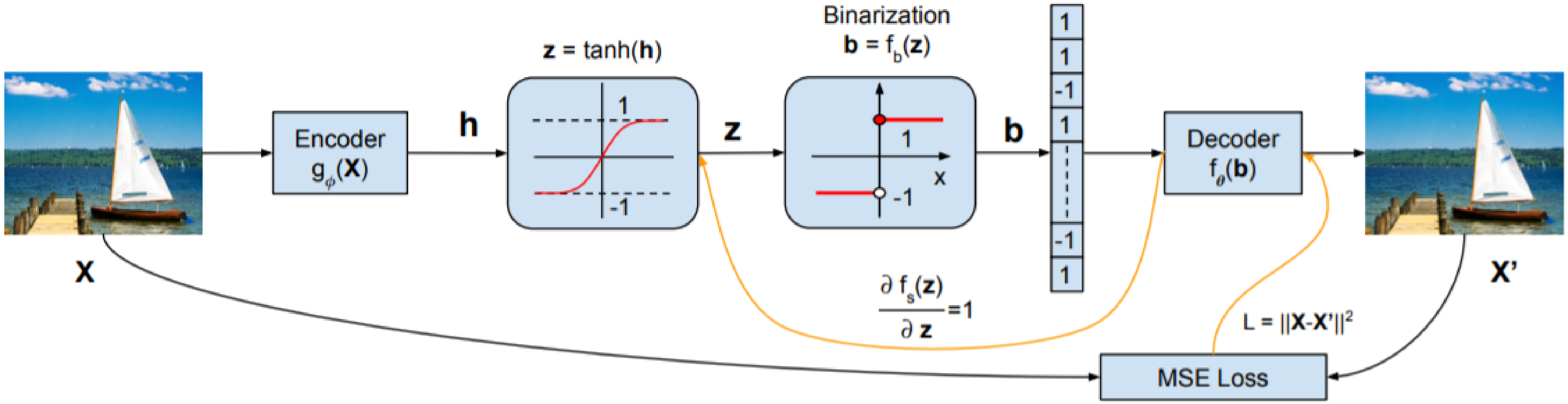


https://arxiv.org/pdf/1804.00891.pdf

Other autoencoders - Bernoulli

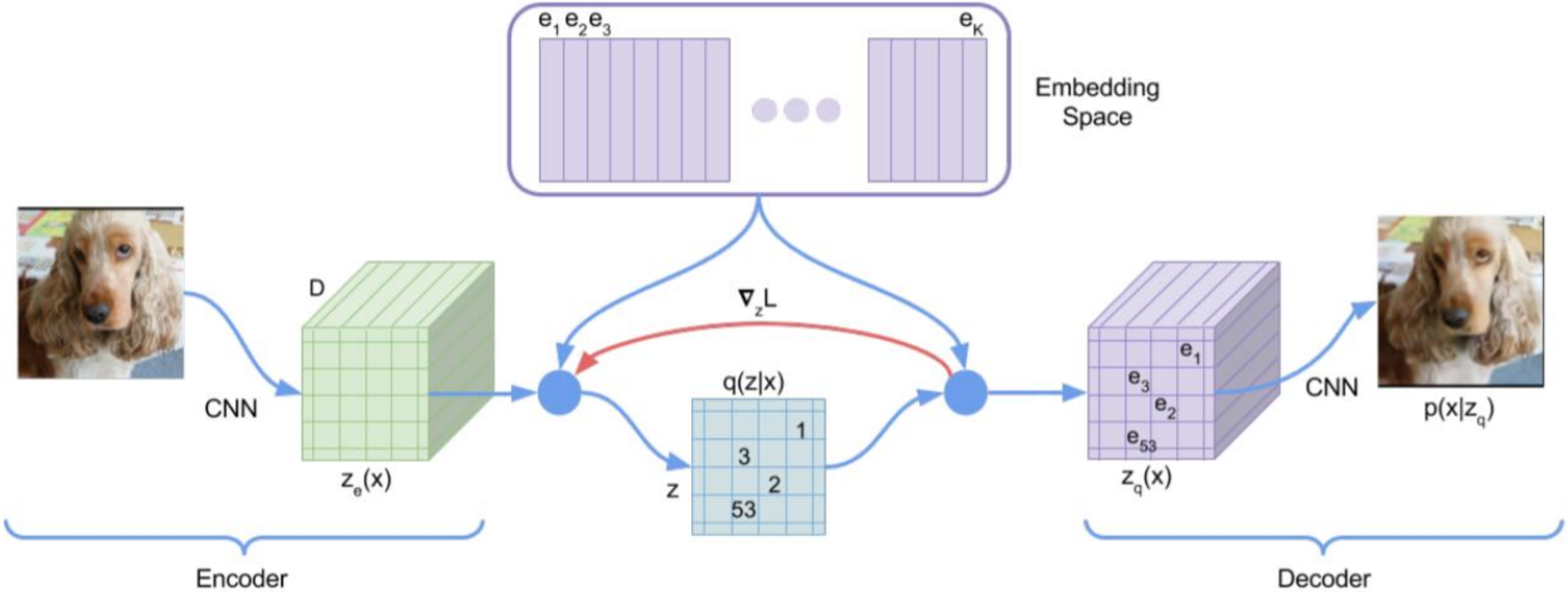
- Besides Gaussian VAE, there are many flavors:

Bernoulli



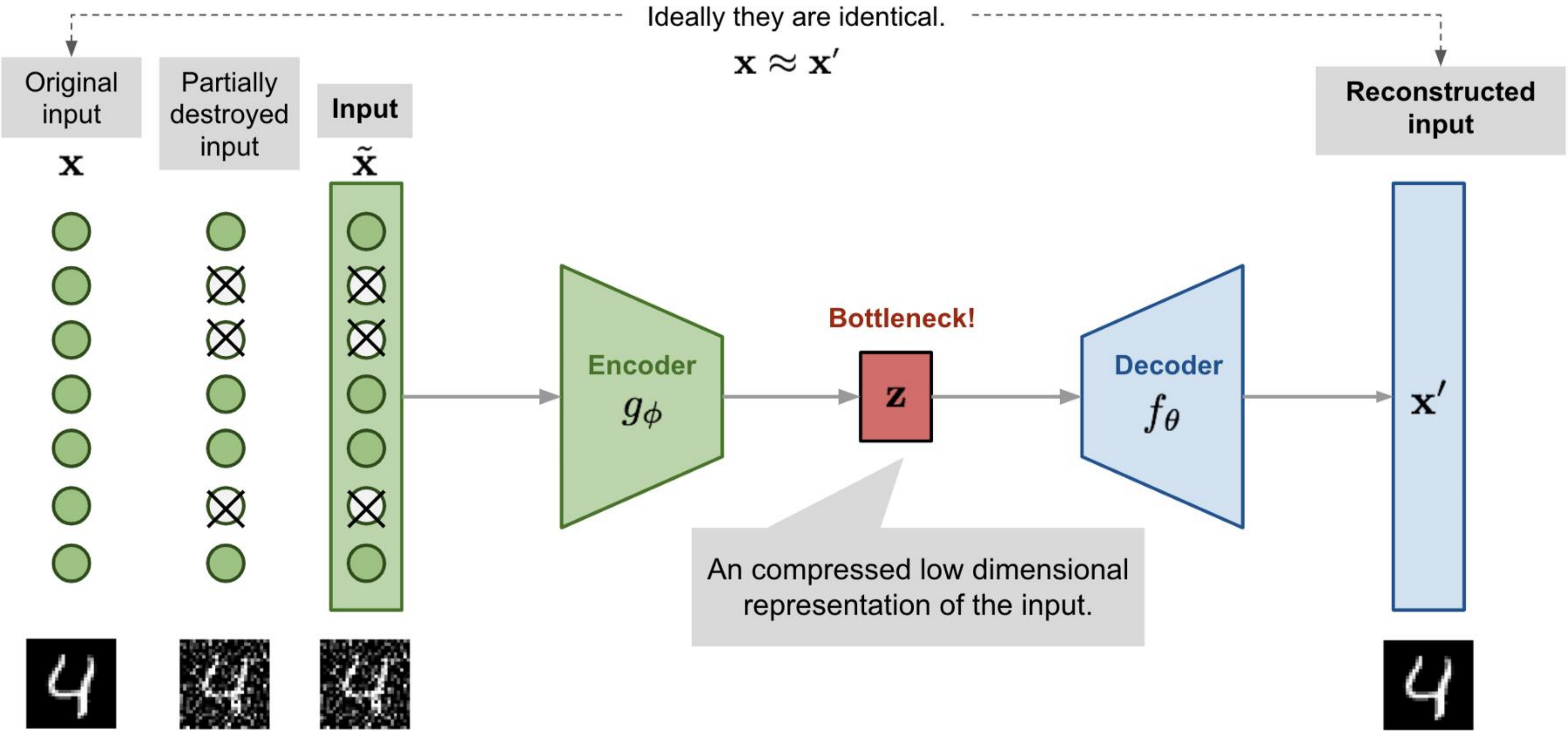
http://proceedings.mlr.press/v119/fajtl20a/fajtl20a.pdf

Other autoencoders - VQ-VAE

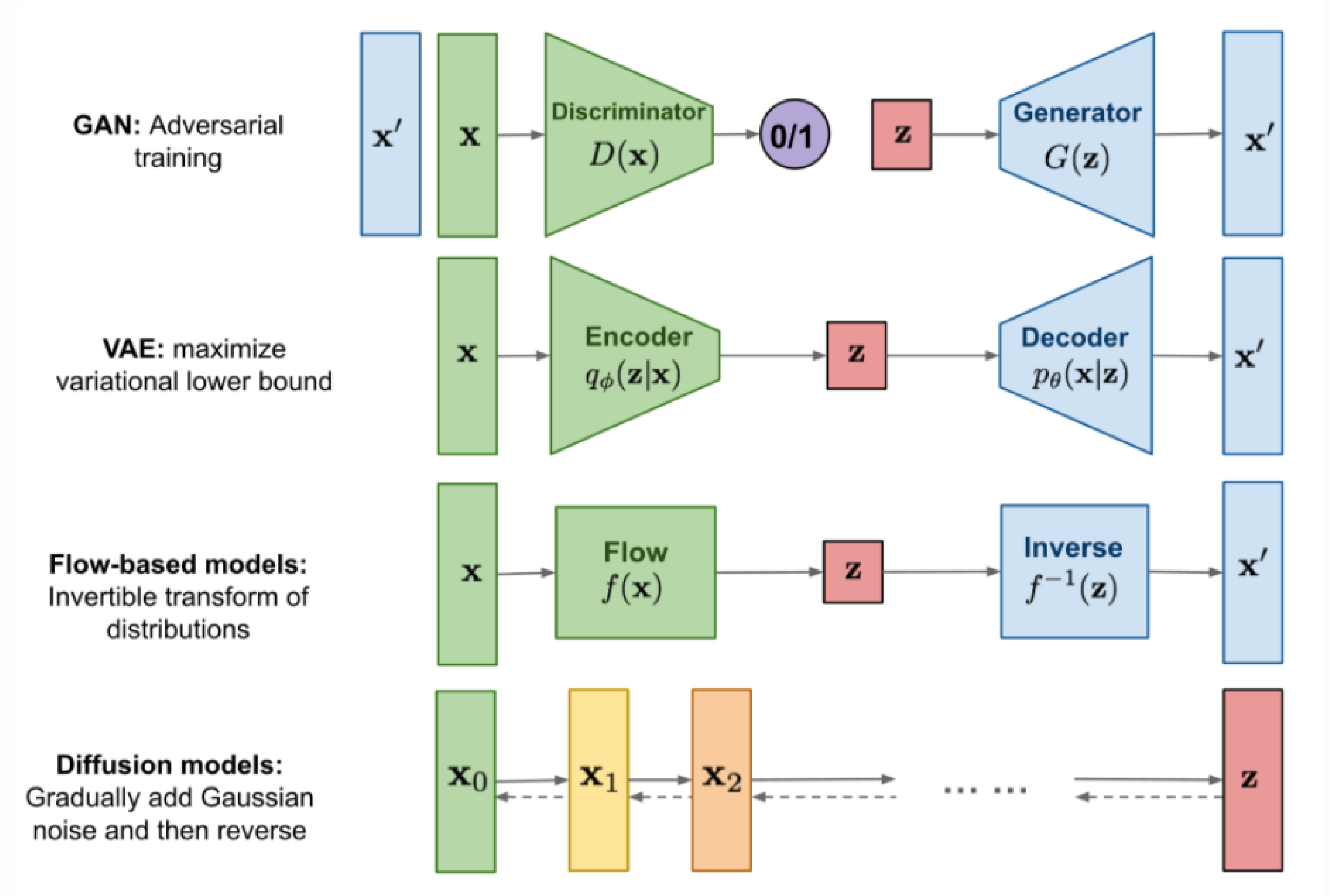


https://arxiv.org/abs/1711.00937v2

Other autoencoders - Denoising Autoencoder



Generative Networks



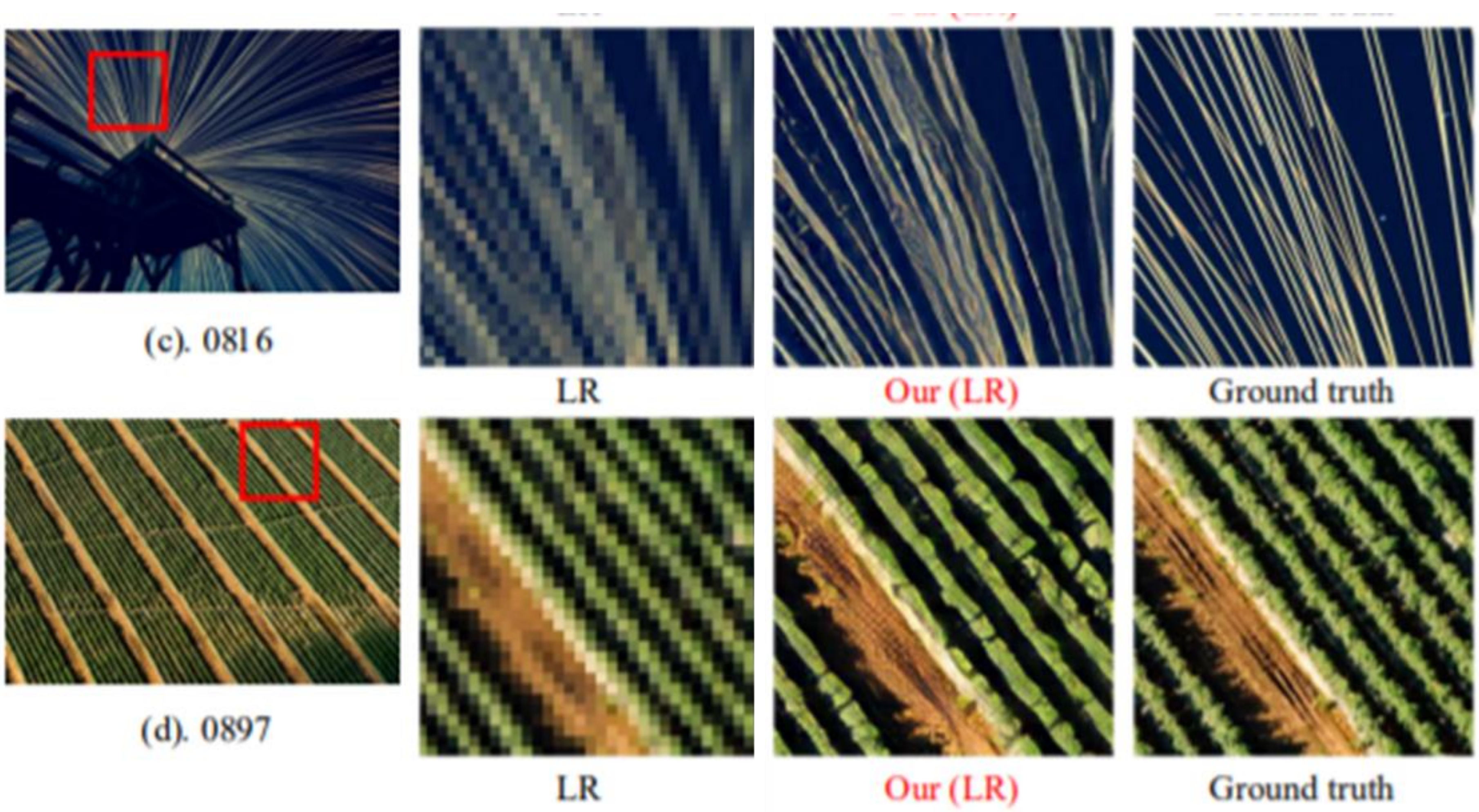
https://lilianweng.github.io/lil-log/2021/07/11/diffusion-models.html

Cool Applications!

Inpainting Autoencoder

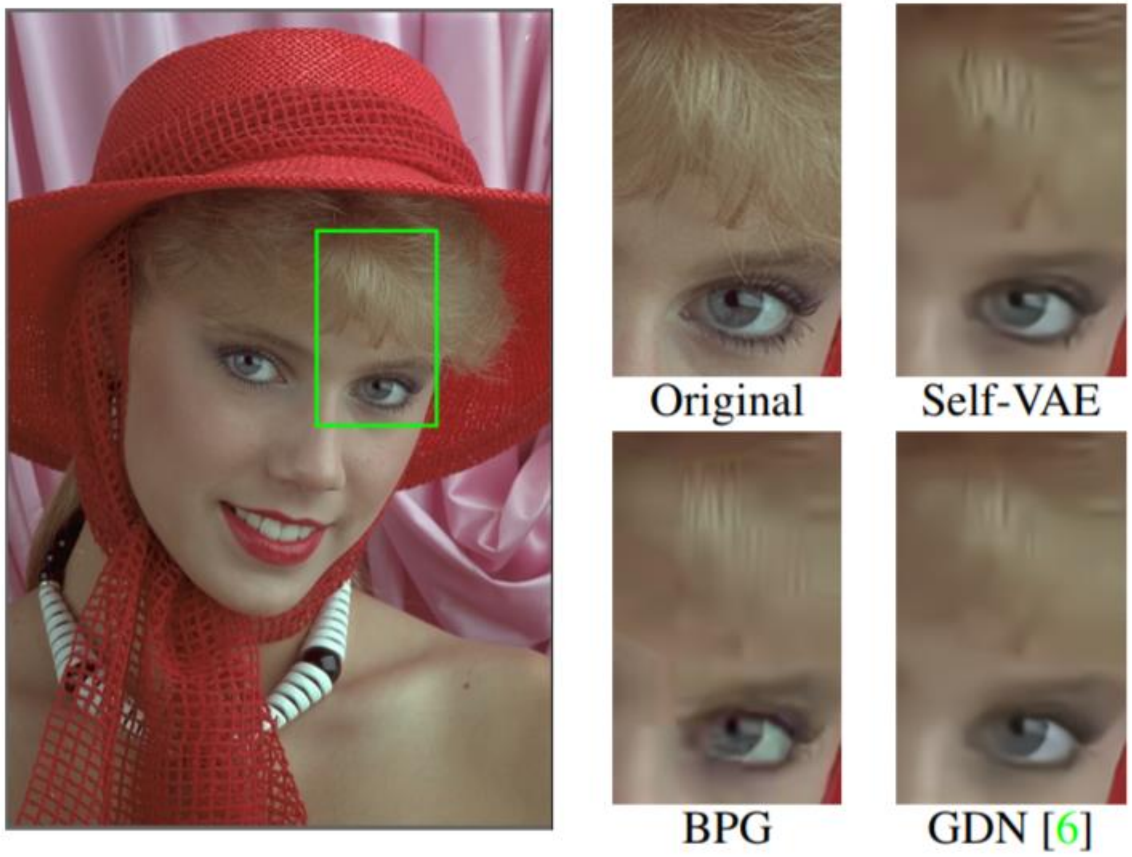
https://vincentcartillier.github.io/papers/variational-image-inpainting.pdf

Super-resolution



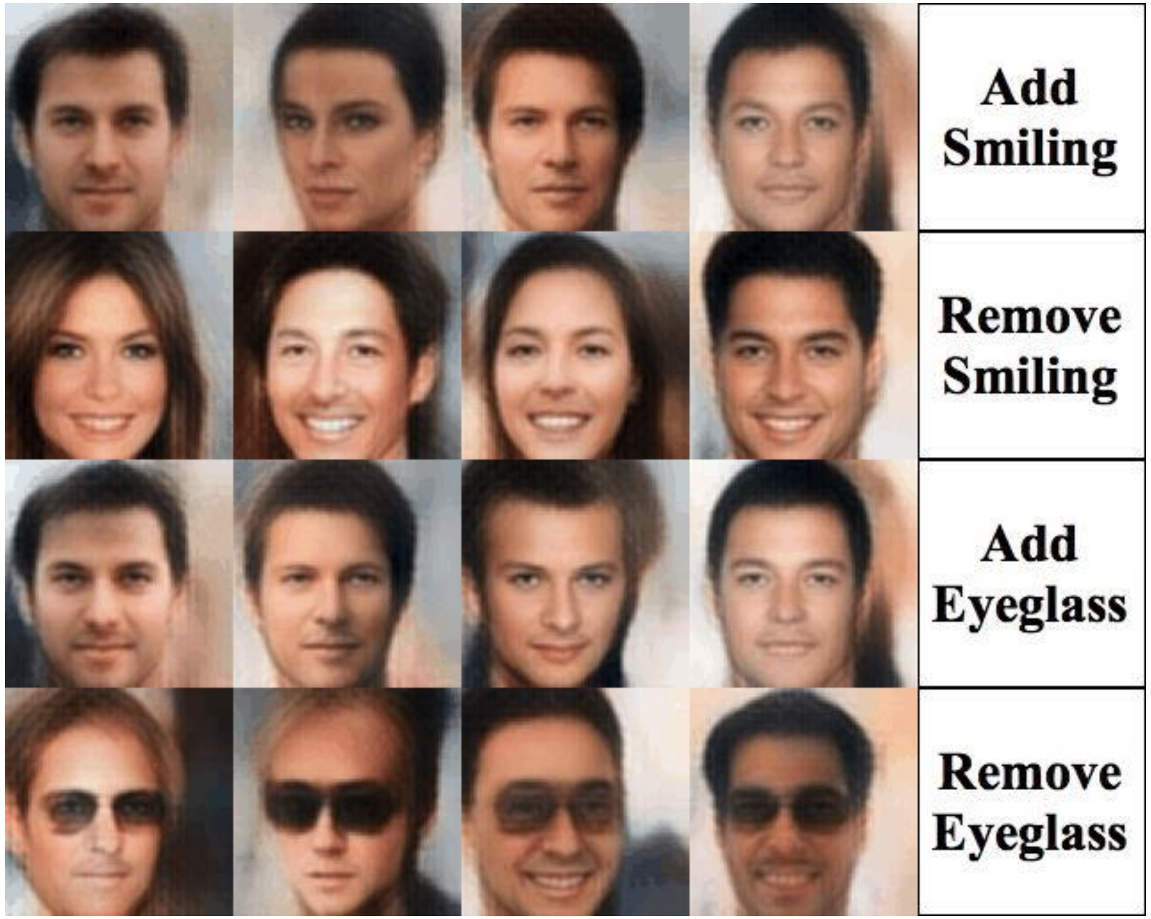
https://arxiv.org/pdf/2006.05218.pdf

Image Compression



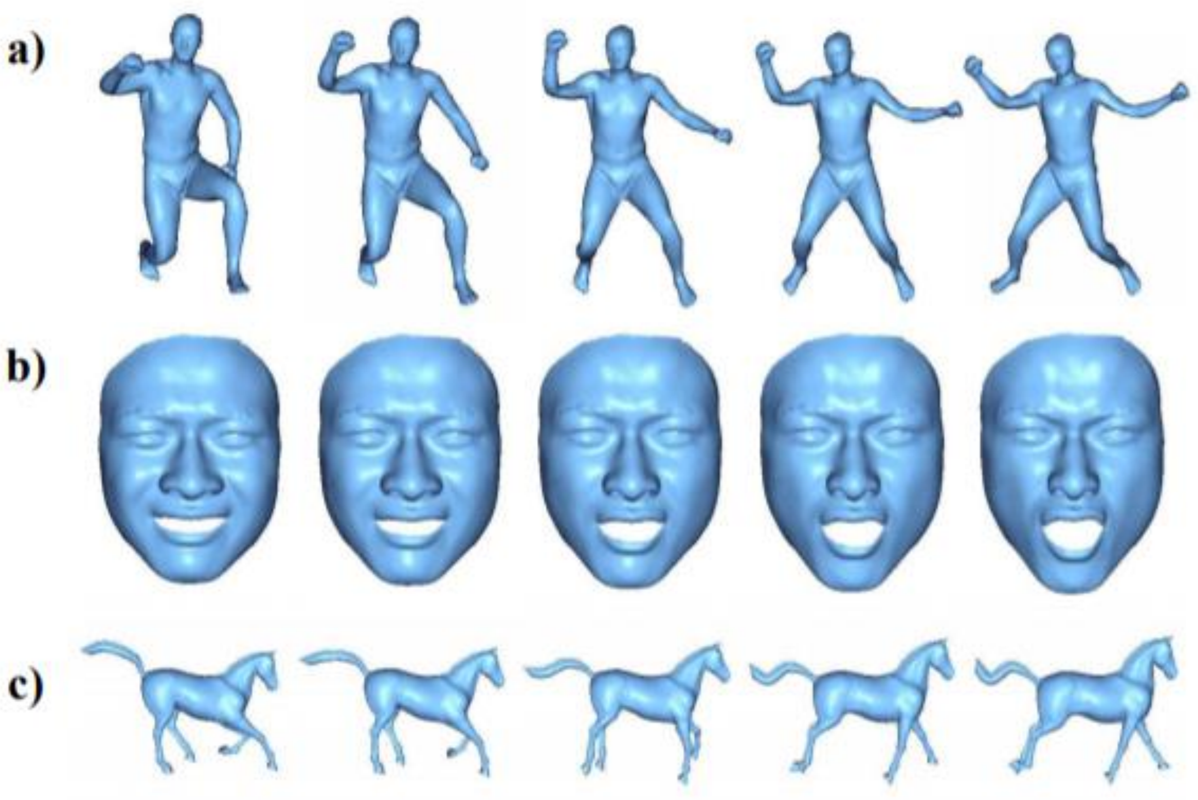
https://arxiv.org/pdf/2105.12107.pdf

Simulation and Interpolation



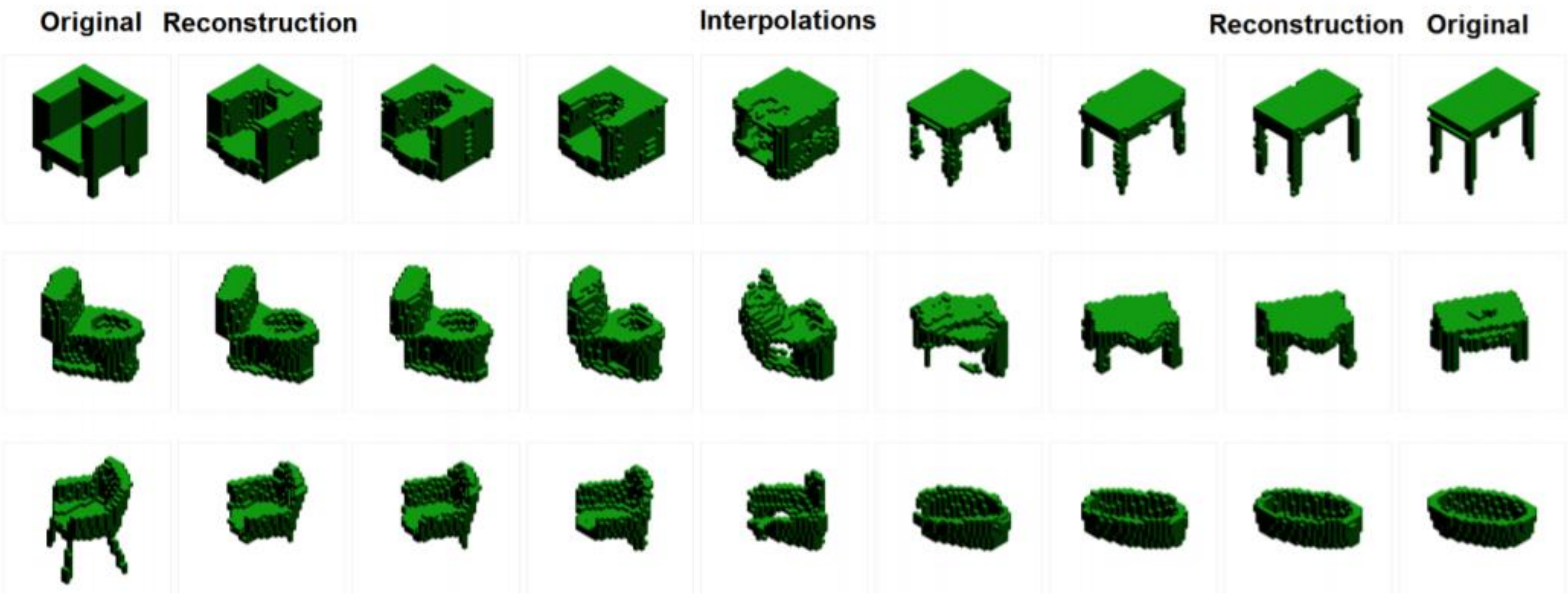
https://houxianxu.github.io/assets/project/dfcvae (animated gif)

3D Mesh Modeling



https://openaccess.thecvf.com/content_cvpr_2018/papers/Tan_Variational_Autoencoders_for_CVPR_2018_paper.pdf

3D Voxel Modeling



https://arxiv.org/pdf/1608.04236.pdf

Deepfakes

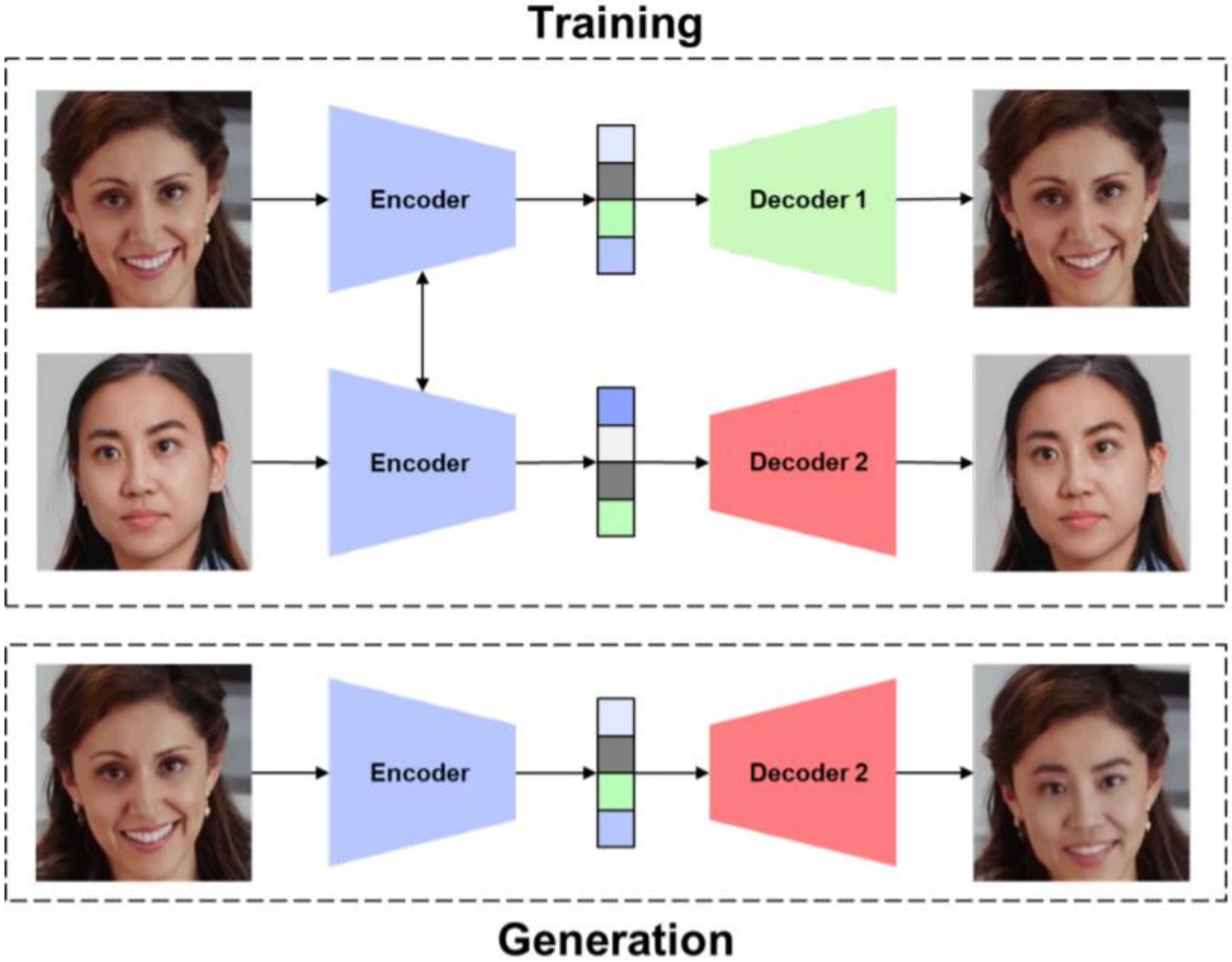
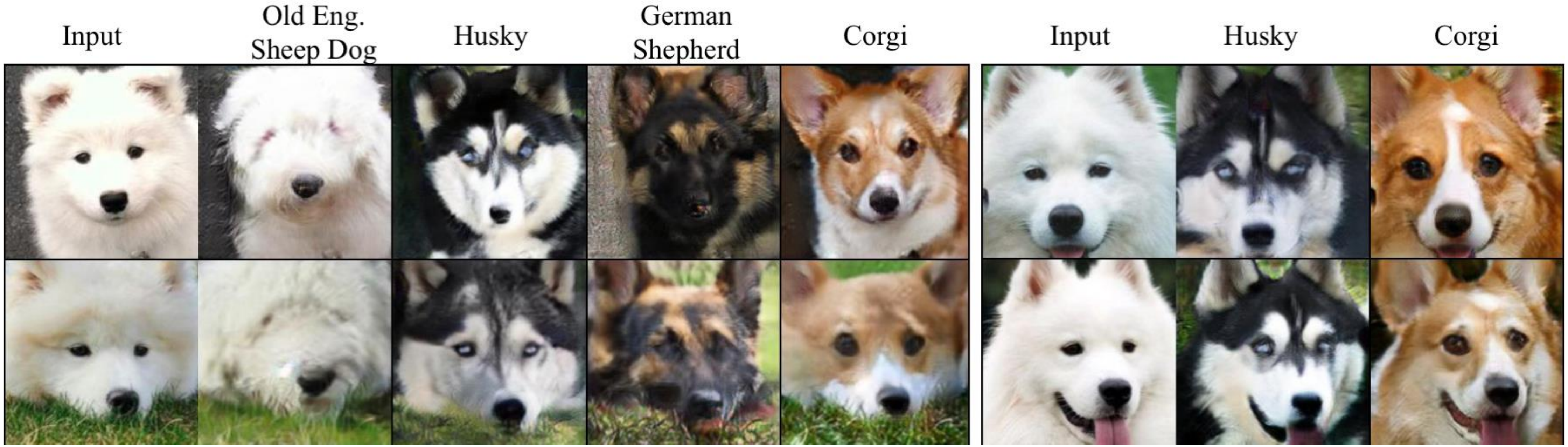

Training
Encoder
Decoder 1
Encoder
Decoder 2
Encoder
Decoder 2
Generation

Image Translation



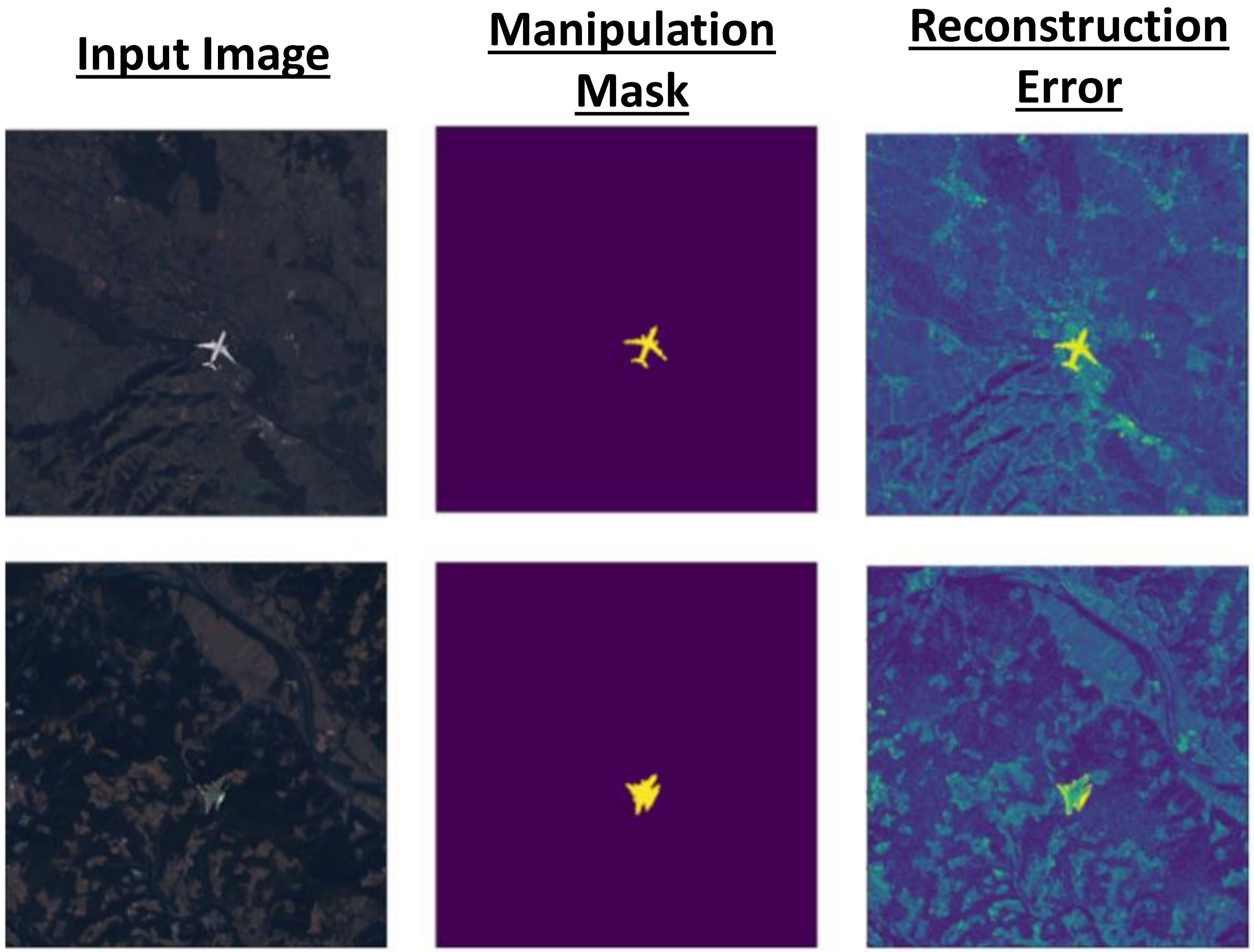
https://arxiv.org/pdf/1703.00848.pdf

Image Translation

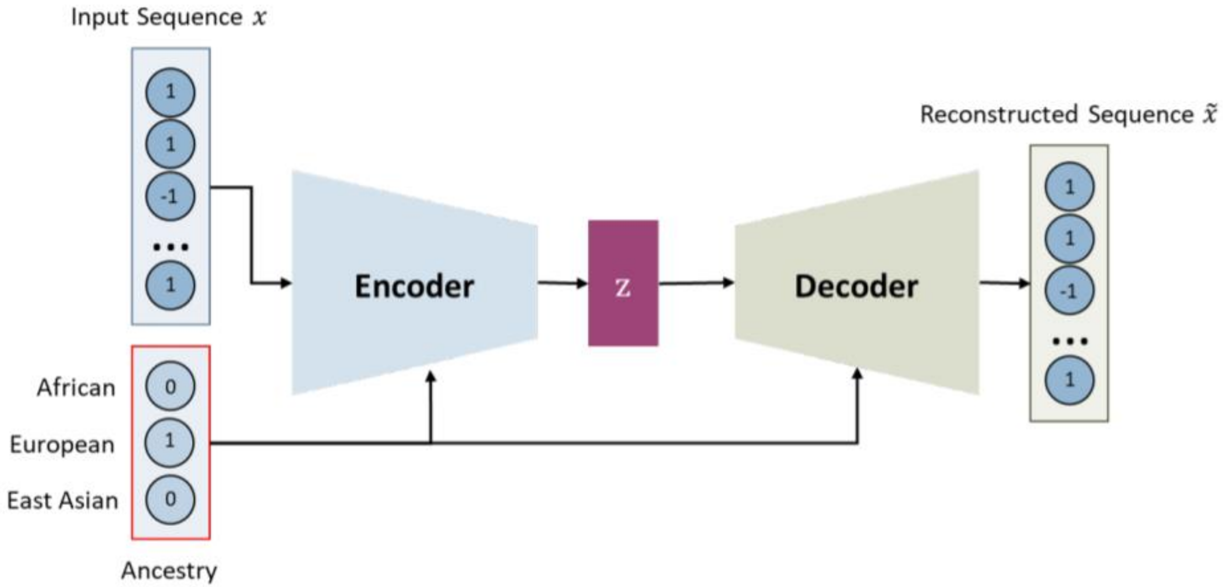
Clothing Simulation

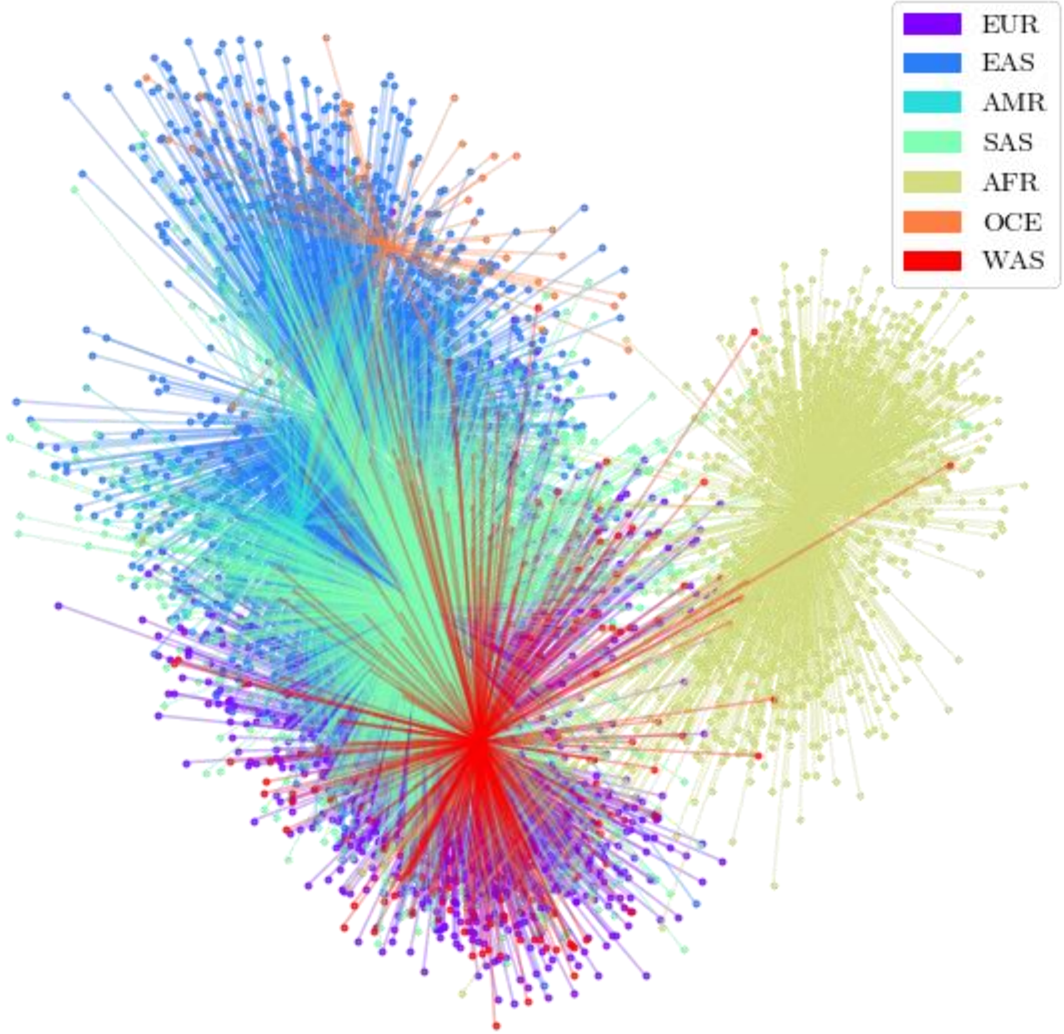
https://arxiv.org/pdf/1901.02284.pdf

Anomaly Detection



DNA Simulation





https://arxiv.org/pdf/1911.13220.pdf

DALL-E

TEXT PROMPT

an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES

Edit prompt or view more images↓

TEXT PROMPT

an armchair in the shape of an avocado [...]

AI-GENERATED IMAGES

Edit prompt or view more images↓

https://openai.com/blog/dall-e/

Thank you!

dmasmont@stanford.edu