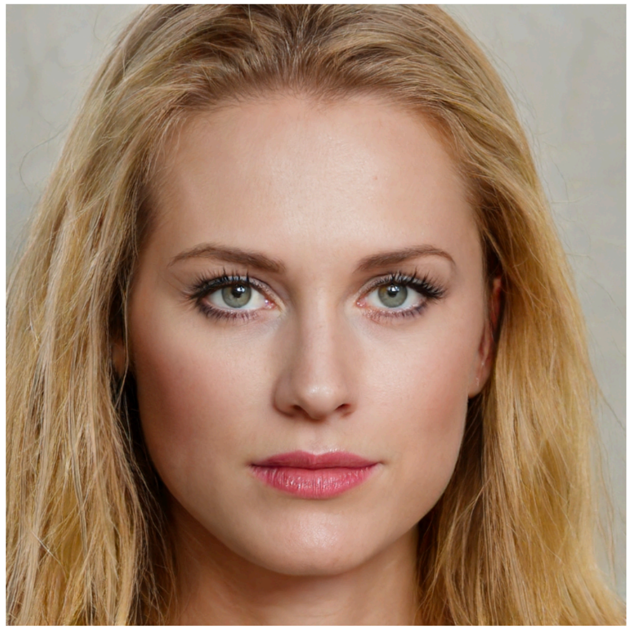
Generative Models: Variational Autoencoders

Foundations of Data Analysis

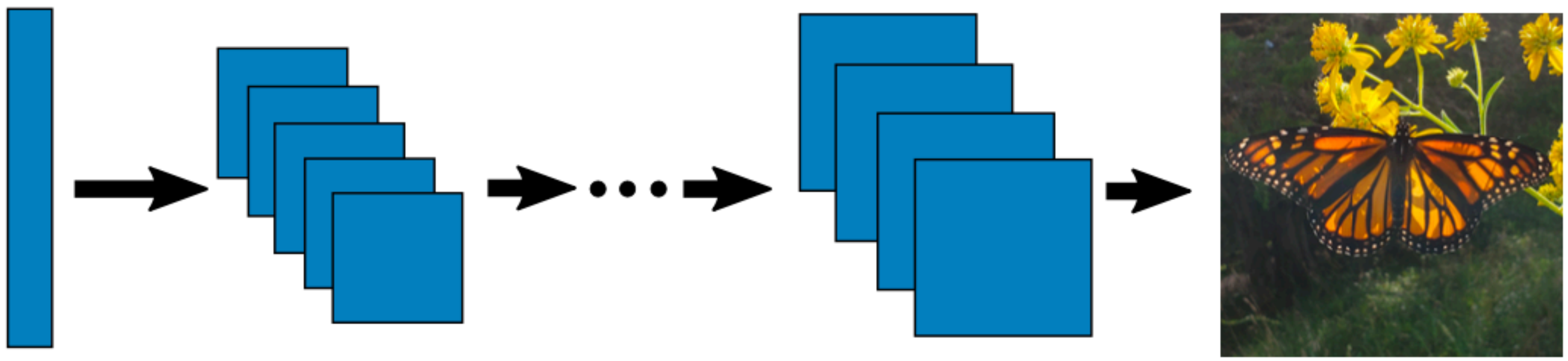
April 26, 2023

These are not real people

Karras et al., CVPR 2020, and thispersondoesnotexist.com

Deep Generative Models



Input:

$$z \in \mathbb{R}^d$$

$$z \sim N(0, I)$$

$$\xrightarrow{g = g_L \circ g_{L-1} \circ \cdots \circ g_1}$$

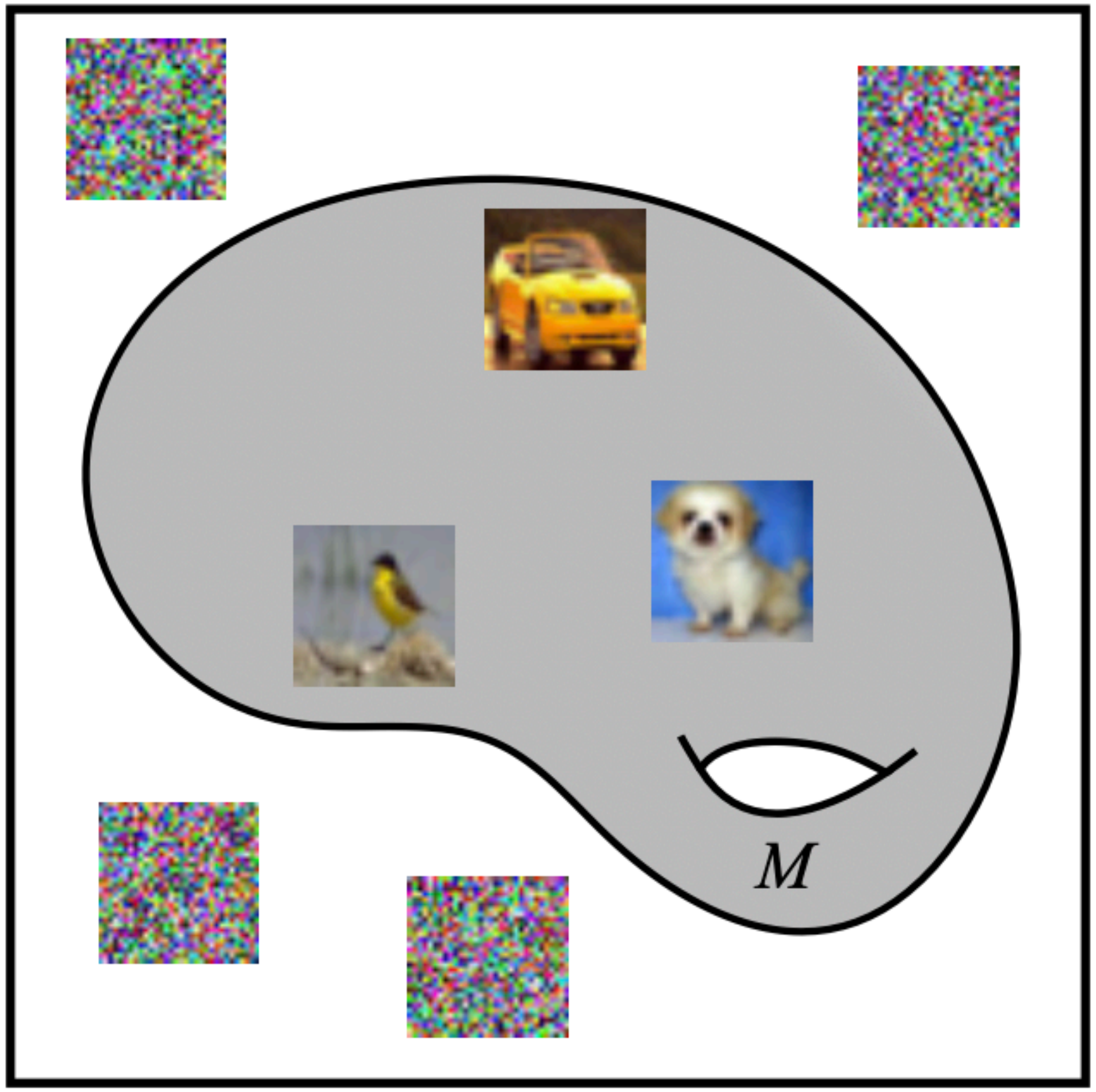
Output:

$$x \in \mathbb{R}^D$$

$$d << D$$

Manifold Hypothesis

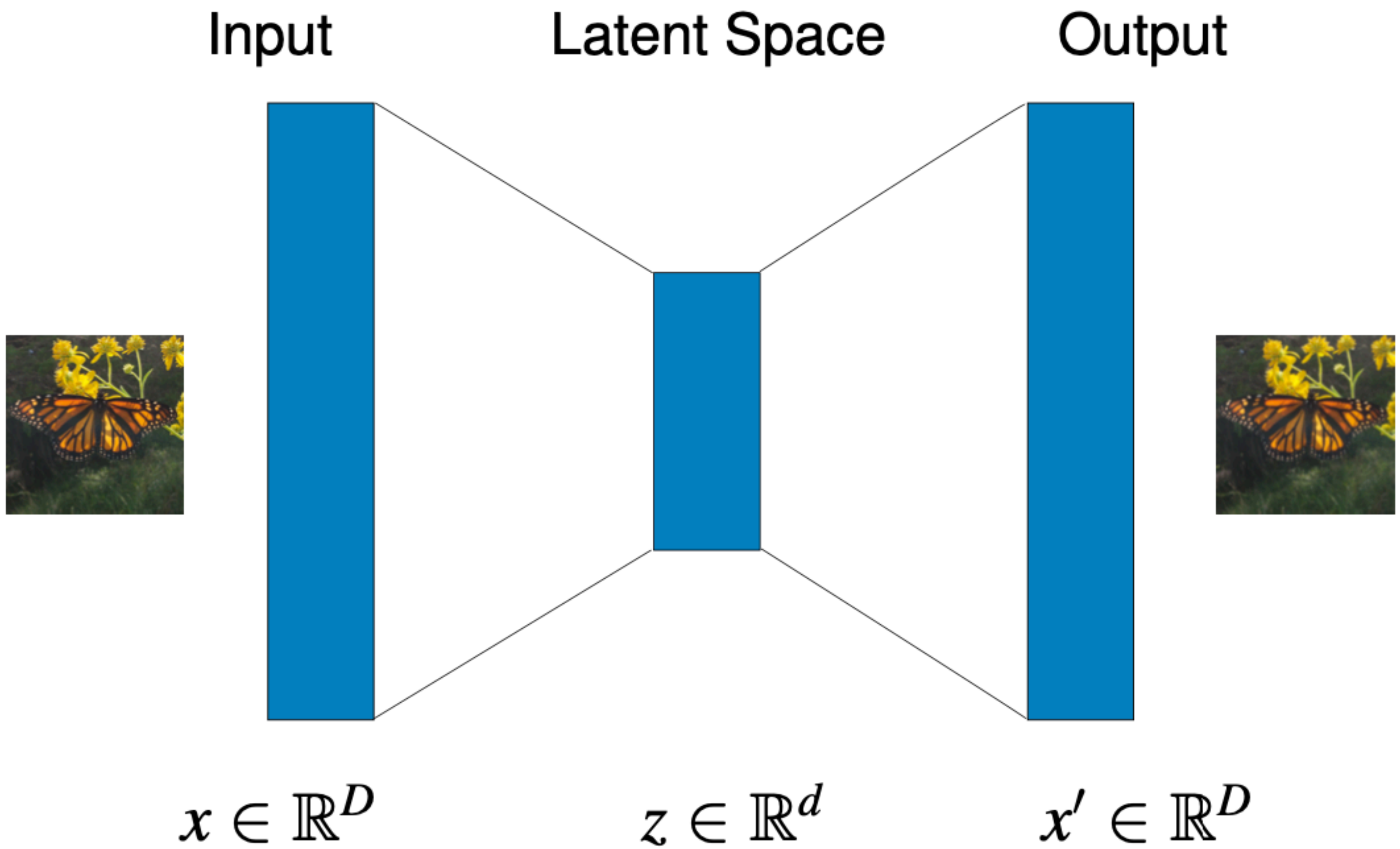
Real data lie near lower-dimensional manifolds



Talking about this paper:

Diederik Kingma and Max Welling, Auto-Encoding Variational Bayes, In *International Conference on Learning Representation (ICLR)*, 2014.

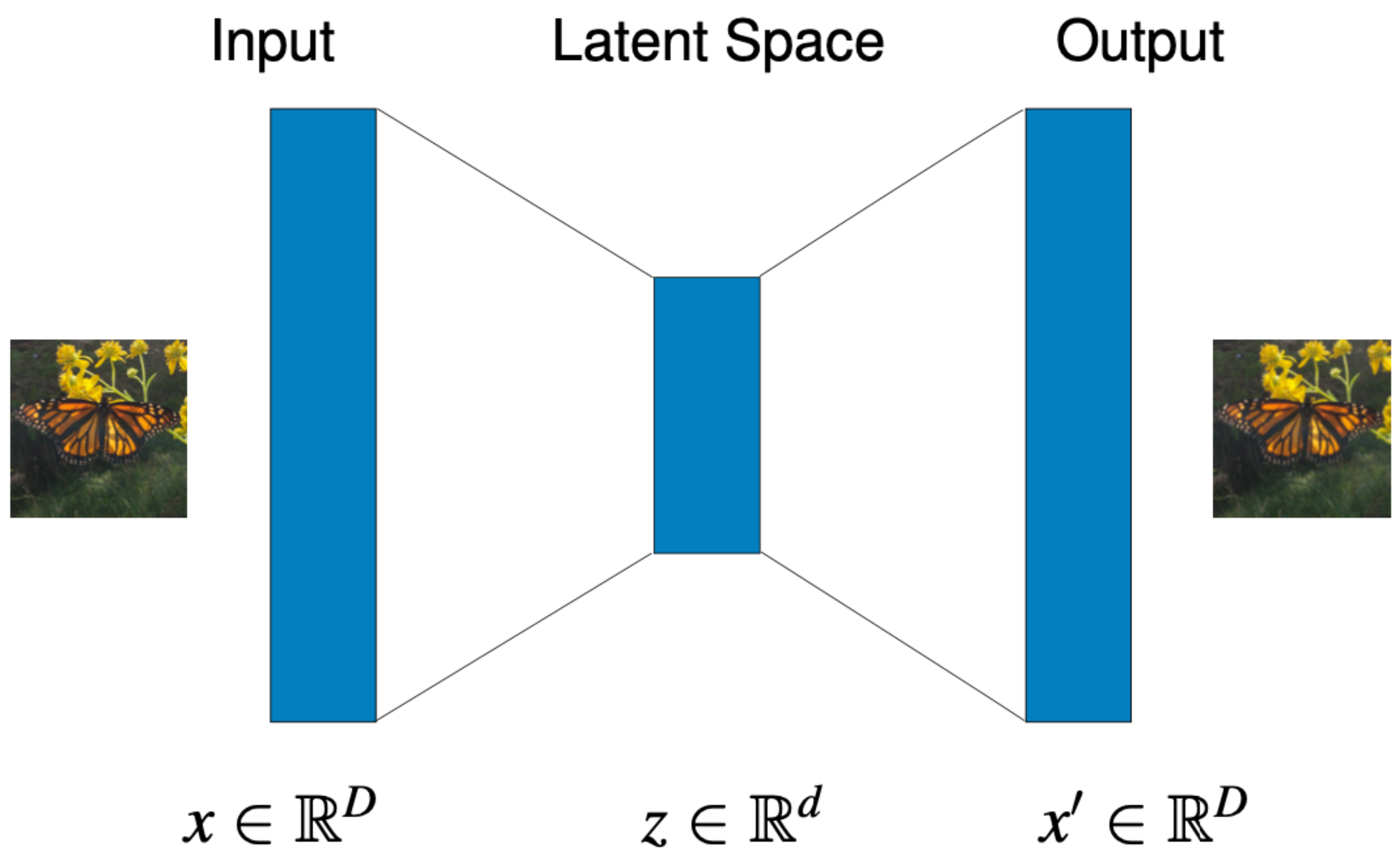
AutoEncoders



$$d << D$$

AutoEncoders

- Linear activation functions give you PCA



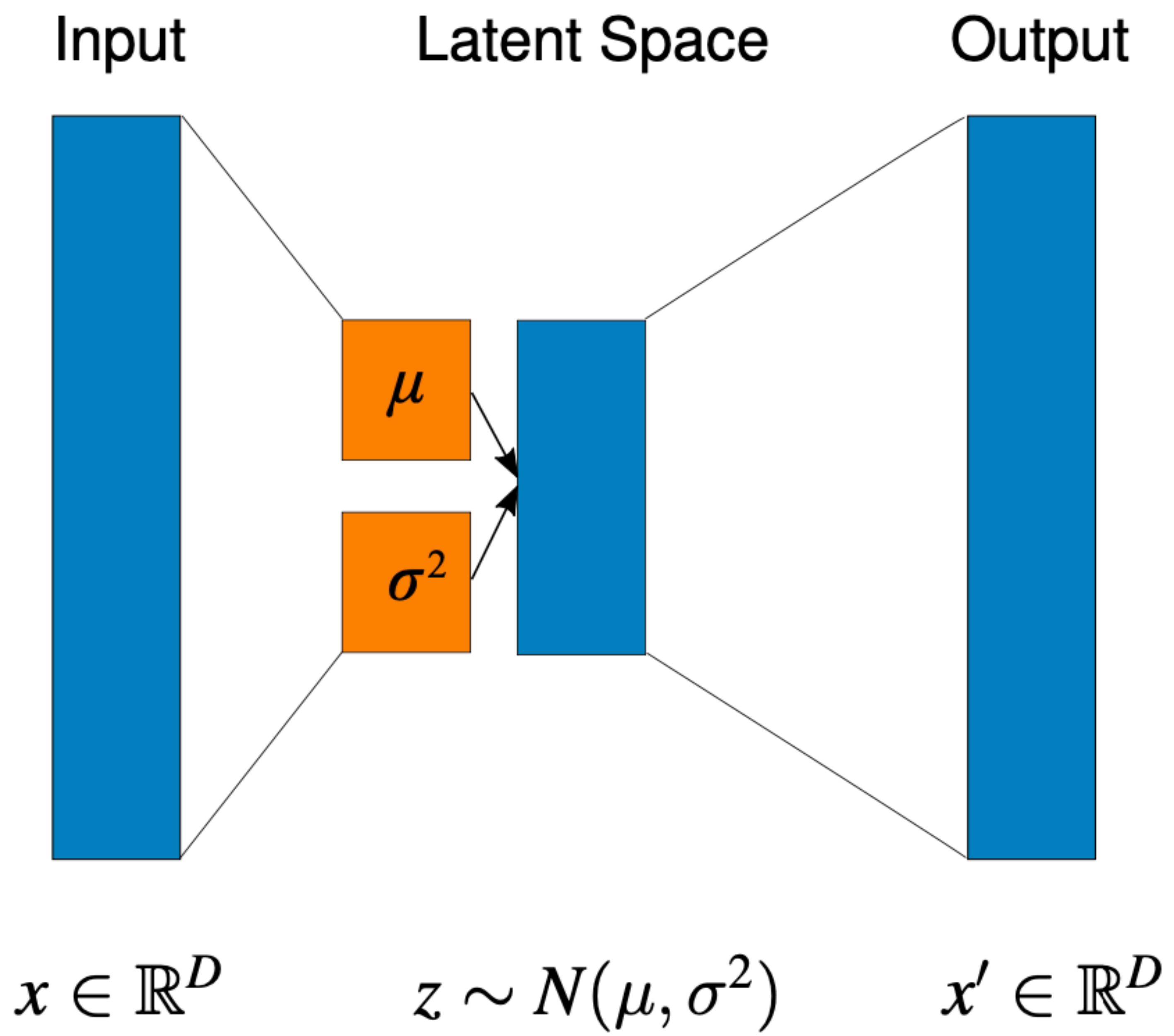
AutoEncoders

- Linear activation functions give you PCA
- Training:
  1. Given data x, feedforward to x′ output
  2. Compute loss,e.g., $L(x, x') = \|x - x'\|^2$
  3. Backpropagate loss gradient to update weights

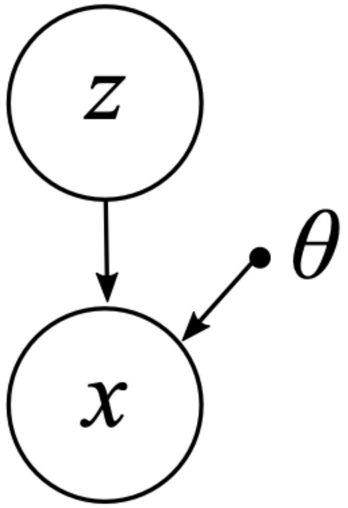
AutoEncoders

- Linear activation functions give you PCA
- Training:
  1. Given data x, feedforward to x′ output
  2. Compute loss,e.g.,$L(x, x') = ||x - x'||^2$
  3. Backpropagate loss gradient to update weights
- **Not** a generative model!

Variational AutoEncoders



Generative Models

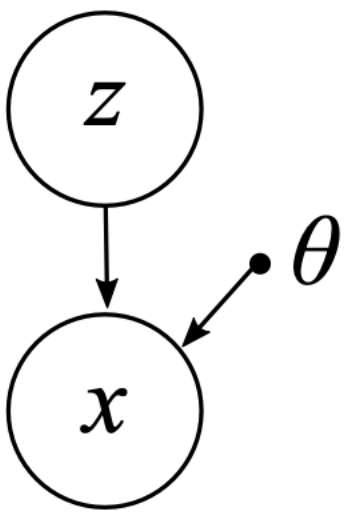


Sample a new x in two steps:

Prior: $p(z)$

Generator: $p_\theta(x \mid z)$

Generative Models



Sample a new x in two steps:

Prior: $p(z)$

Generator: $p_\theta(x \mid z)$

Now the analogy to the "encoder" is:

Posterior: $p(z \mid x)$

Bayesian Inference

Posterior via Bayes' Rule:

$$p(z \mid x) = \frac{p_\theta(x \mid z)p(z)}{p(x)}$$

$$= \frac{p_\theta(x \mid z)p(z)}{\int p_\theta(x \mid z)p(z)dz}$$

Integral in denominator is (usually) intractable!

Variational Inference

Approximate intractable posterior $p(z\,|\,x)$ with a manageable distribution $q(z)$

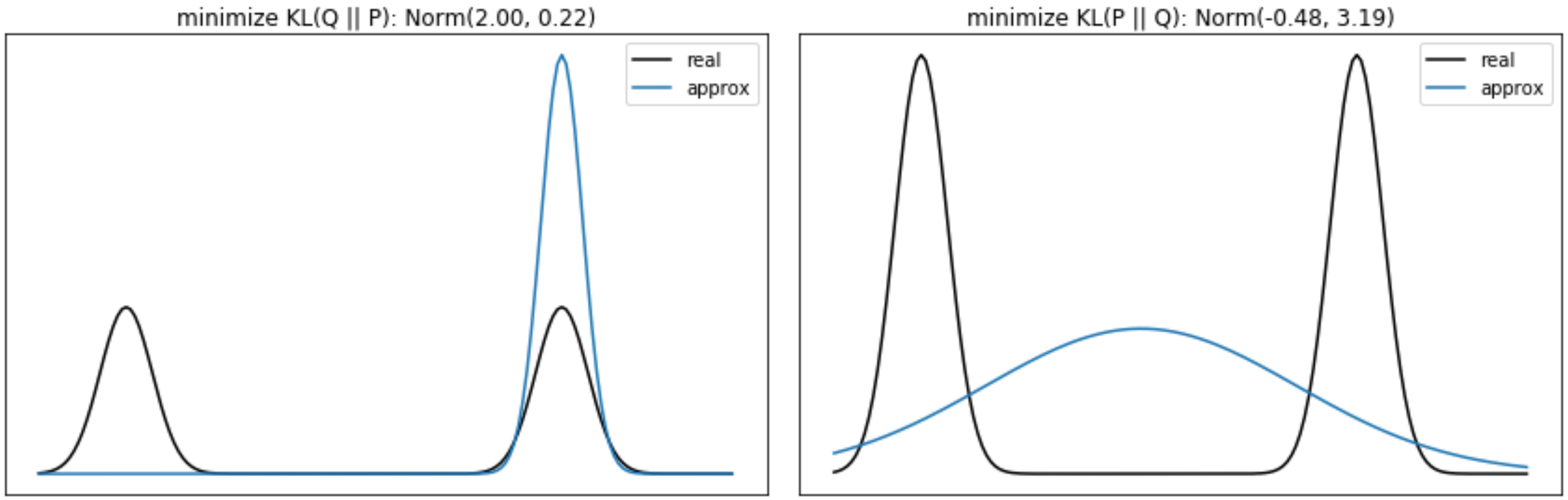
Minimize the KL divergence: $D_{KL}(q(z)||p(z\,|\,x))$

Kullback-Leibler Divergence

$$D_{KL}(q||p) = -\int q(z)\log\frac{p(z)}{q(z)}dz$$
$$= E_q(-\log\frac{p}{q})$$

The average *information gained* from moving from q to p

Kullback-Leibler Divergence

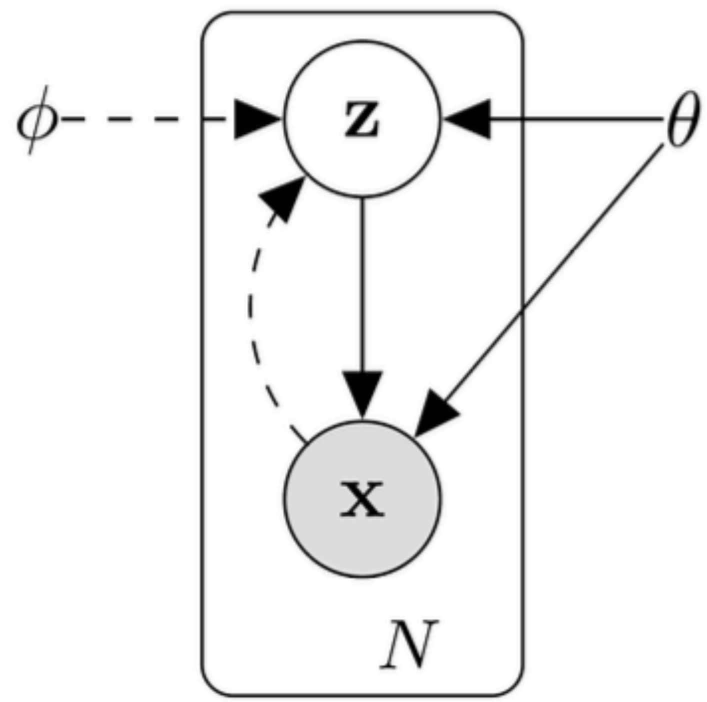


The average *information gained* from moving from q to p

Variational Autoencoder

$q_\phi(z \mid x)$

Encoder Network

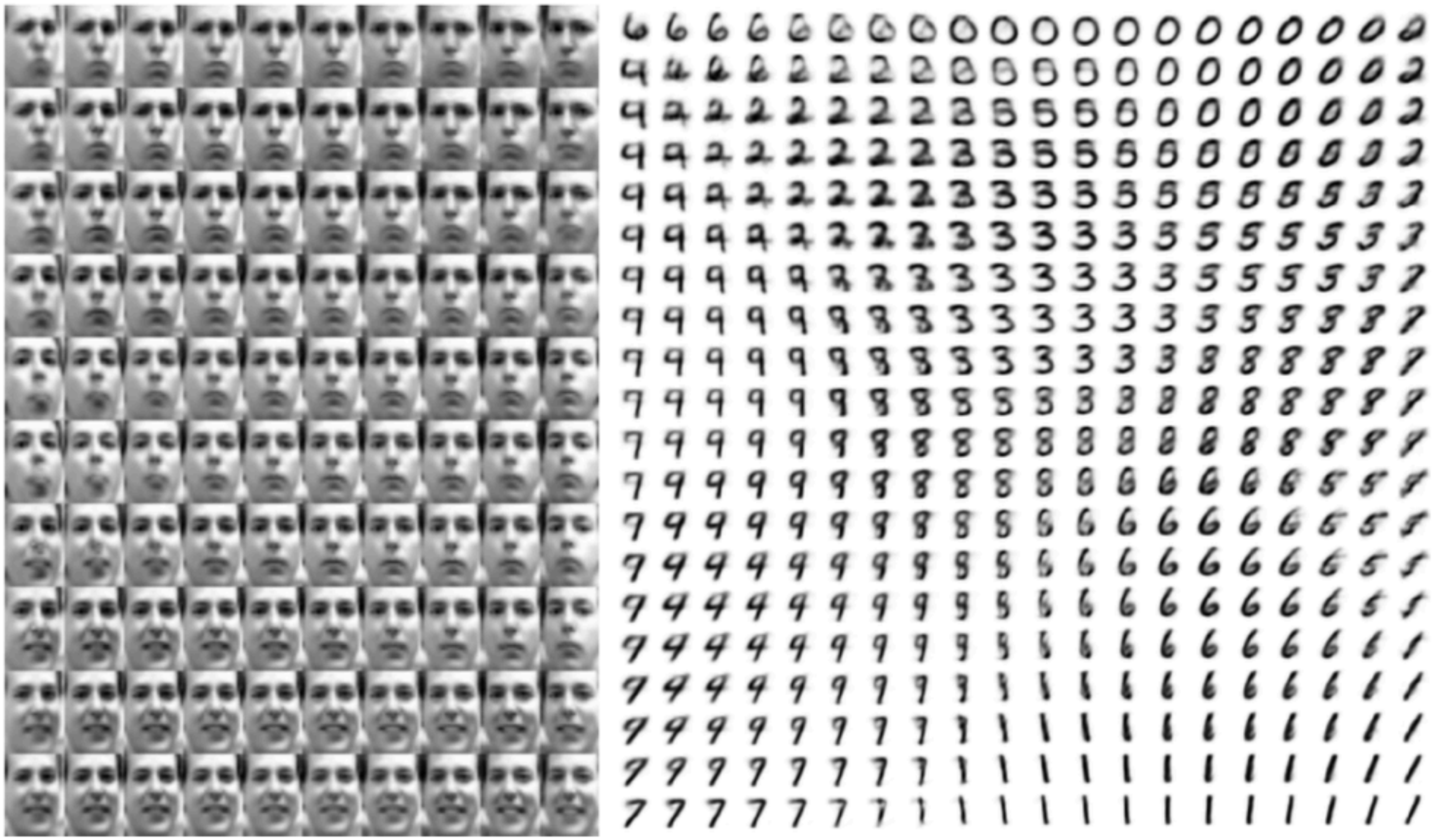


$p_\theta(x \mid z)$

Decoder Network

Minimize the KL divergence: $D_{KL}(q(z) \| p(z \mid x))$.

Results from Kingma & Welling



Applications of Autoencoder / VAE Models

Image-to-Image Networks

Instead of trying to reconstruct the original input:

1. Encode input: $z = \text{encode}(x)$
2. Decode **derived** output: $y = \text{decode}(z)$

Image-to-Image Networks

Instead of trying to reconstruct the original input:

1. Encode input: $z = \text{encode}(x)$
2. Decode **derived** output: $y = \text{decode}(z)$

Example: Image Segmentation

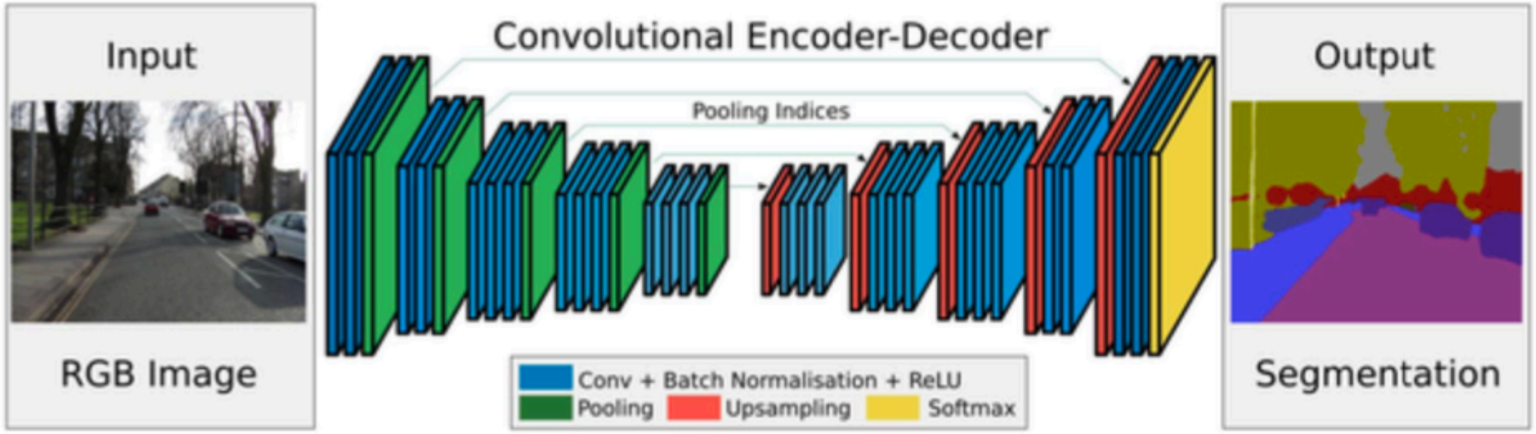
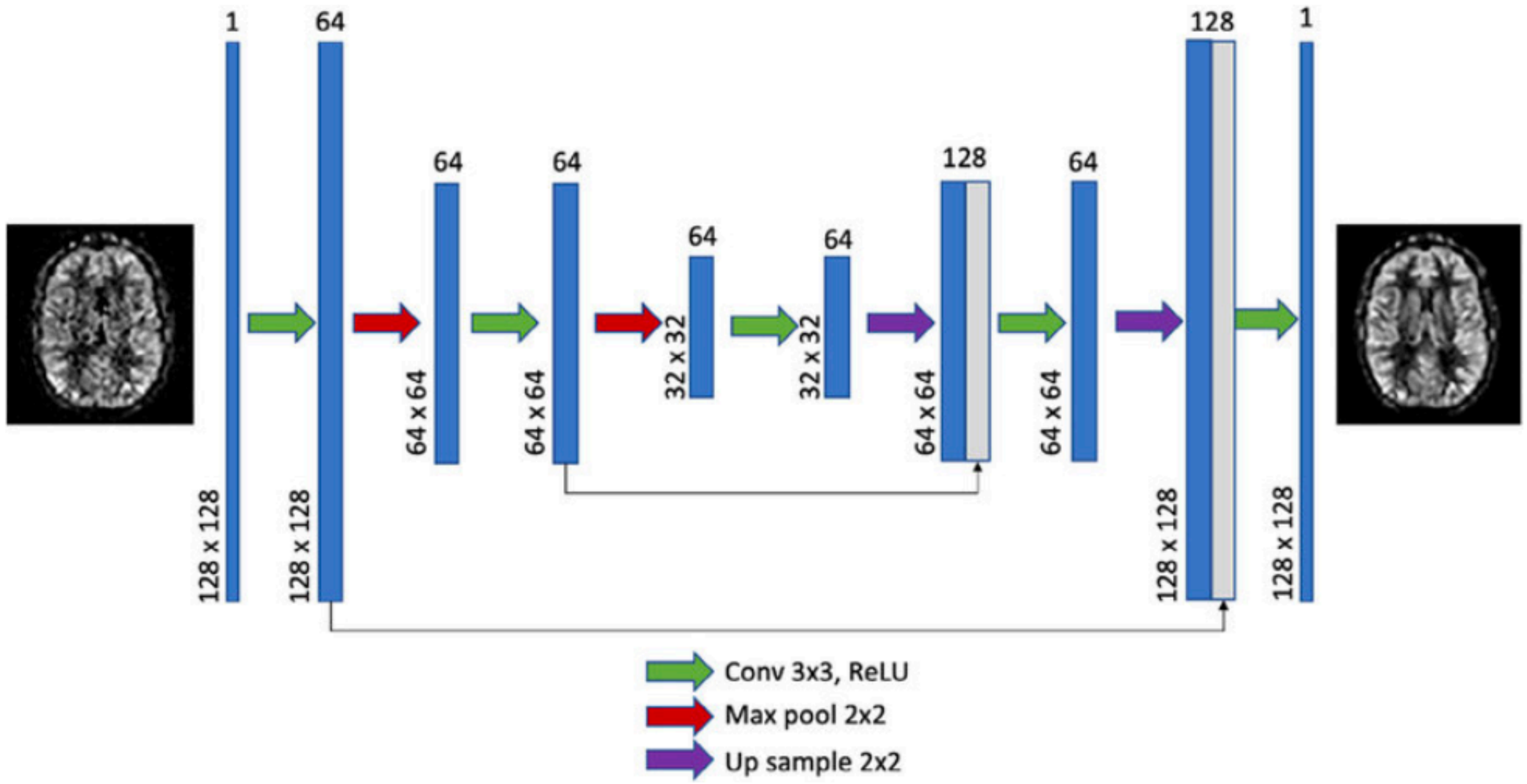


Image Denoising

Learning mapping from noisy inputs → clean outputs



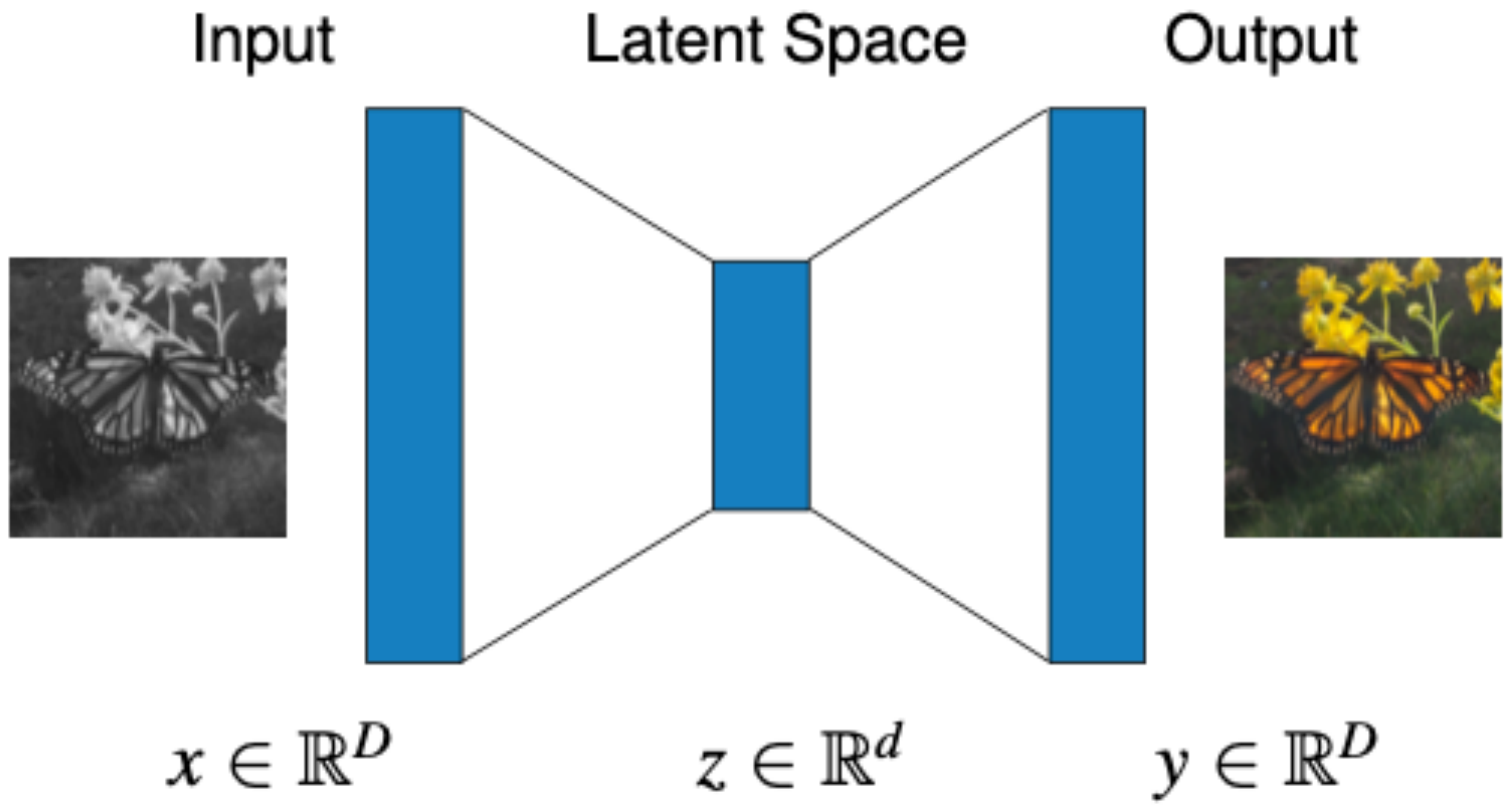
Hales et al., JMRI 2020

Image Super-resolution

Learning mapping from low-res inputs→ high-res outputs

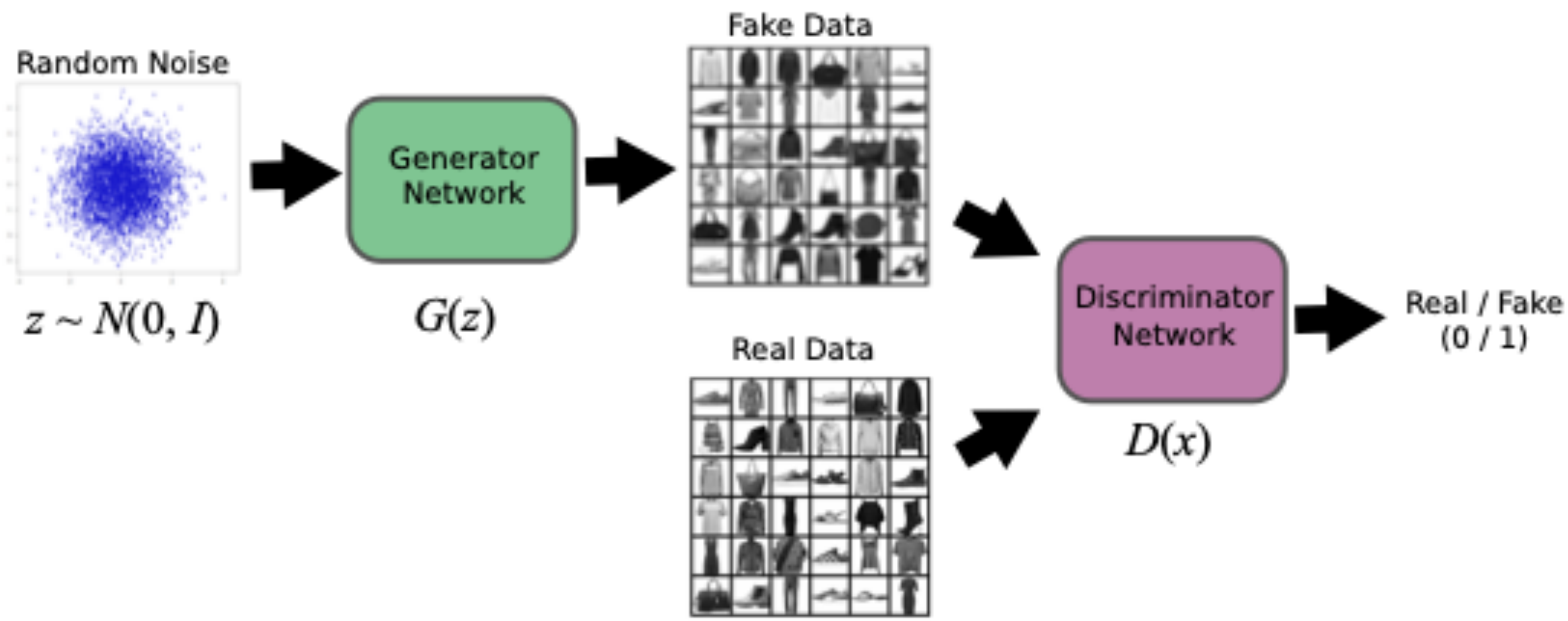
From: https://towardsdatascience.com/deep-learning-based-super-resolution-without-using-a-gan-11c9bb5b6cd5

Image Colorization



Generative Adversarial Networks (GANs)

Generative Adversarial Network



GAN Game Theory

GAN training is framed as a competition where:

1. Discriminator is trying to maximize its reward
2. Generator is trying to minimize it

$$\min_{G} \max_{D} V(D, G)$$

$$V(D, G) = E_{x \sim p(x)}[\log D(x)] + E_{z \sim N(0,I)}[\log 1 - D(G(z))]$$

Original GAN Faces (2014)

Goodfellow et al., NeurIPS 2014

Final Exam - Take Home

- This is an open book exam, but requires independent work. No team work is allowed!
- Exam Hours: 3:30pm - 11:59pm on May 1st.
- The exam will cover all topics discussed in FoDA class.
- Programming is not involved.

Final Exam - Take Home

- Bonus questions will be provided.

ELECTRICAL AND COMPUTER ENGINEERING

Date: Friday, April 28, 2023

Time: 2:00 PM

Location: Thornton Hall E316, 2-3pm EST [in person seminar]

ADD TO CALENDAR

RSVP TO THIS EVENT

Polina Golland

Principal investigator in the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL)

Title: Learning to read xray: applications to heart failure monitoring

Abstract: We propose and demonstrate a novel approach to training image classification models based on large collections of images with limited labels. We take advantage of availability of radiology reports to construct joint multimodal embedding that serves as a basis for classification. We demonstrate the advantages of this approach in application to assessment of pulmonary edema severity in congestive heart failure that motivated the development of the method.