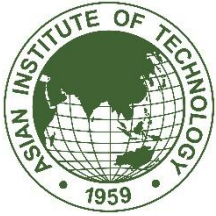


Autoencoders

Dr. Mongkol Ekpanyapong

What is Autoencoder?

Autoencoder

- Auto = self
- Encode = convert into a different form

Autoencoder = a system that teaches itself how to encode information

It is a model that teaches itself how to encode information

Autoencoders

- An unsupervised learning technique that is used as a data representation
- The idea is to use CNN to act as data compression/data encoding by introducing a bottleneck layer
- We must have encoding layers and decoding layers

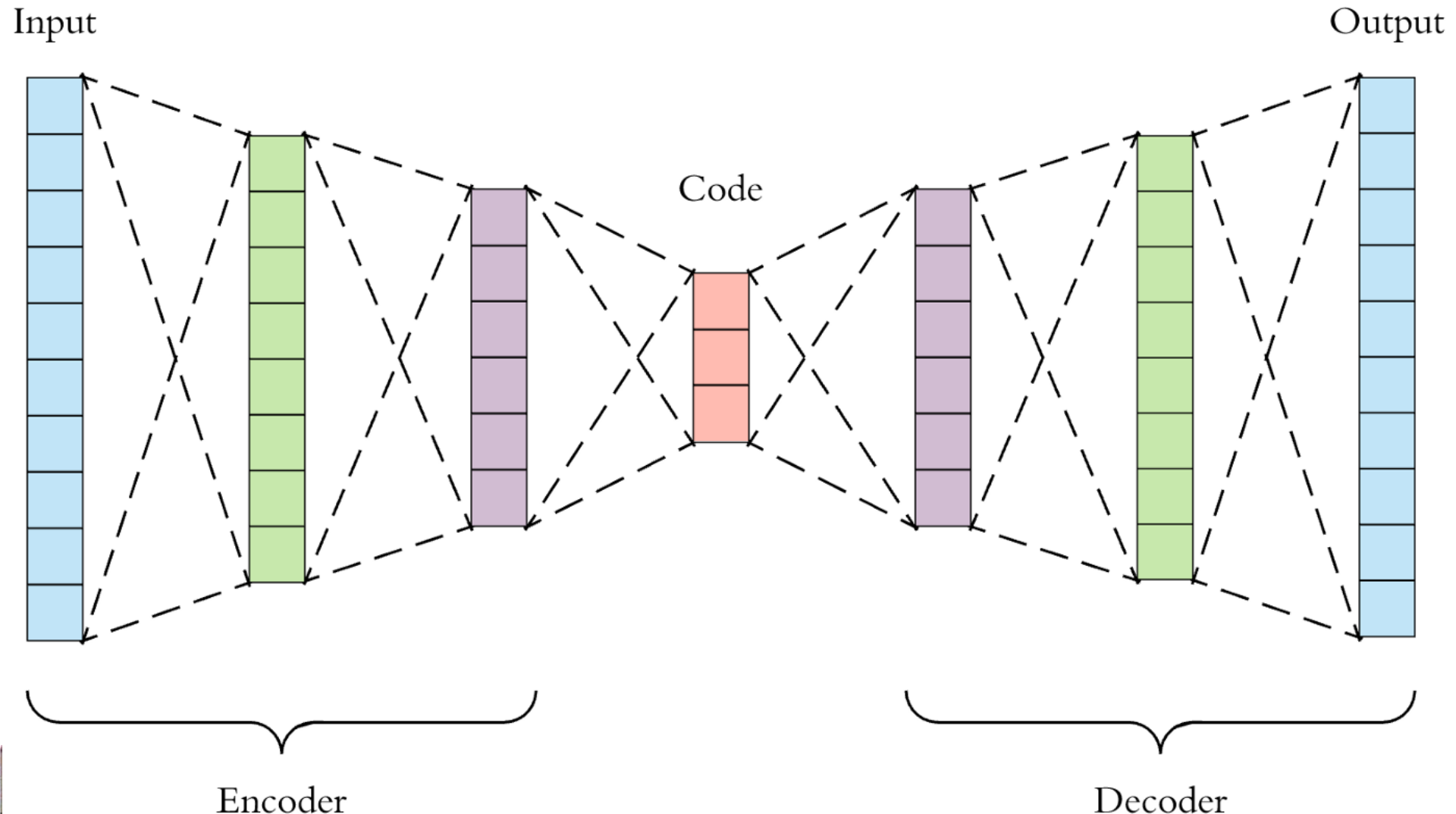
Loss function

$$BCE = -\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{c}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]$$

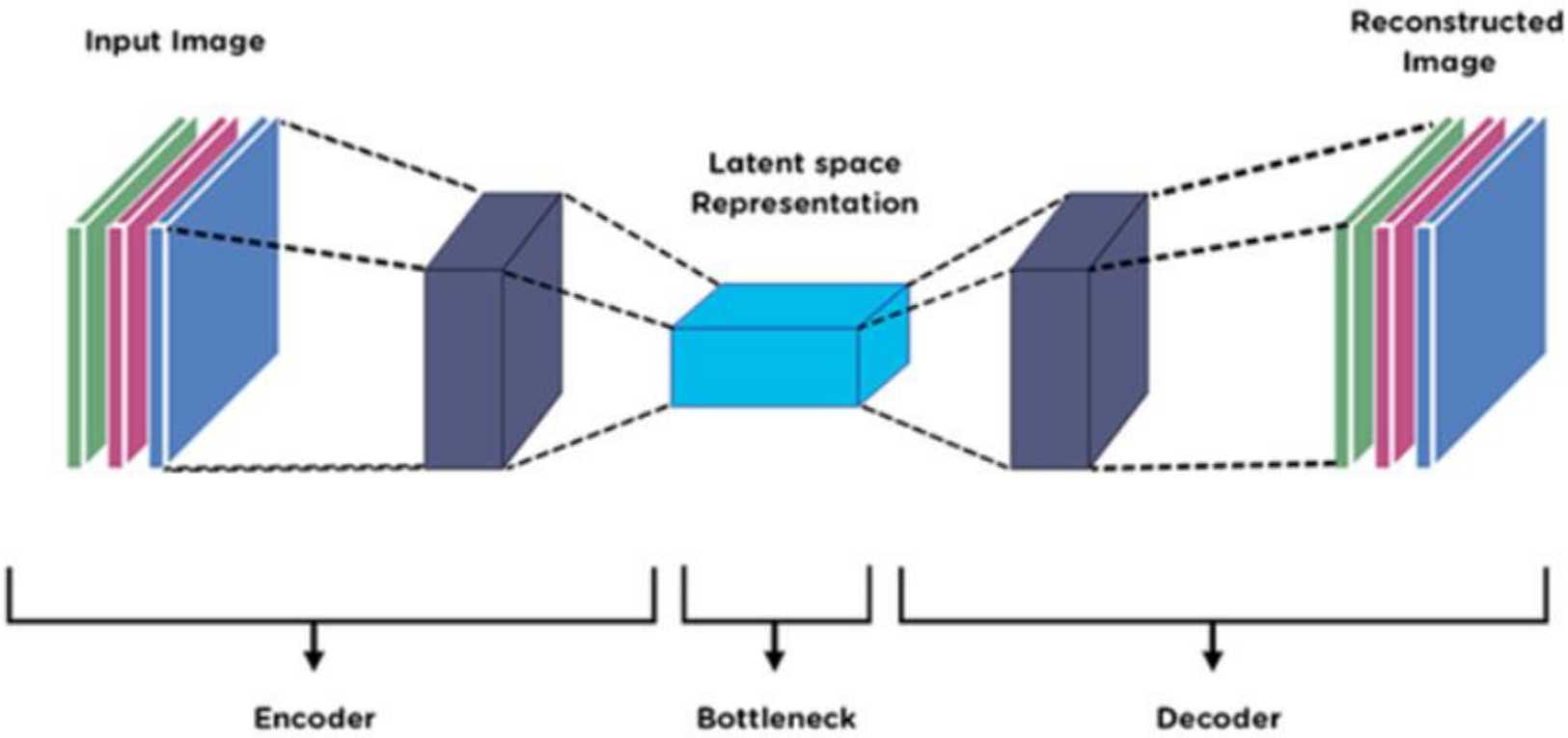
- Or MSE loss

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

Example



CNN Autoencoder

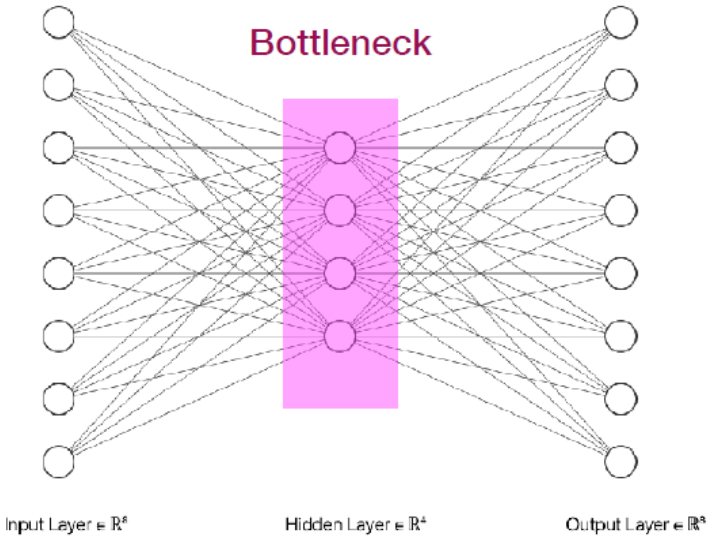


Autoencoder applications

- Denoising
- Fix Image Inpainting
- Information Retrieval
- Anomaly Detection

Autoencoders

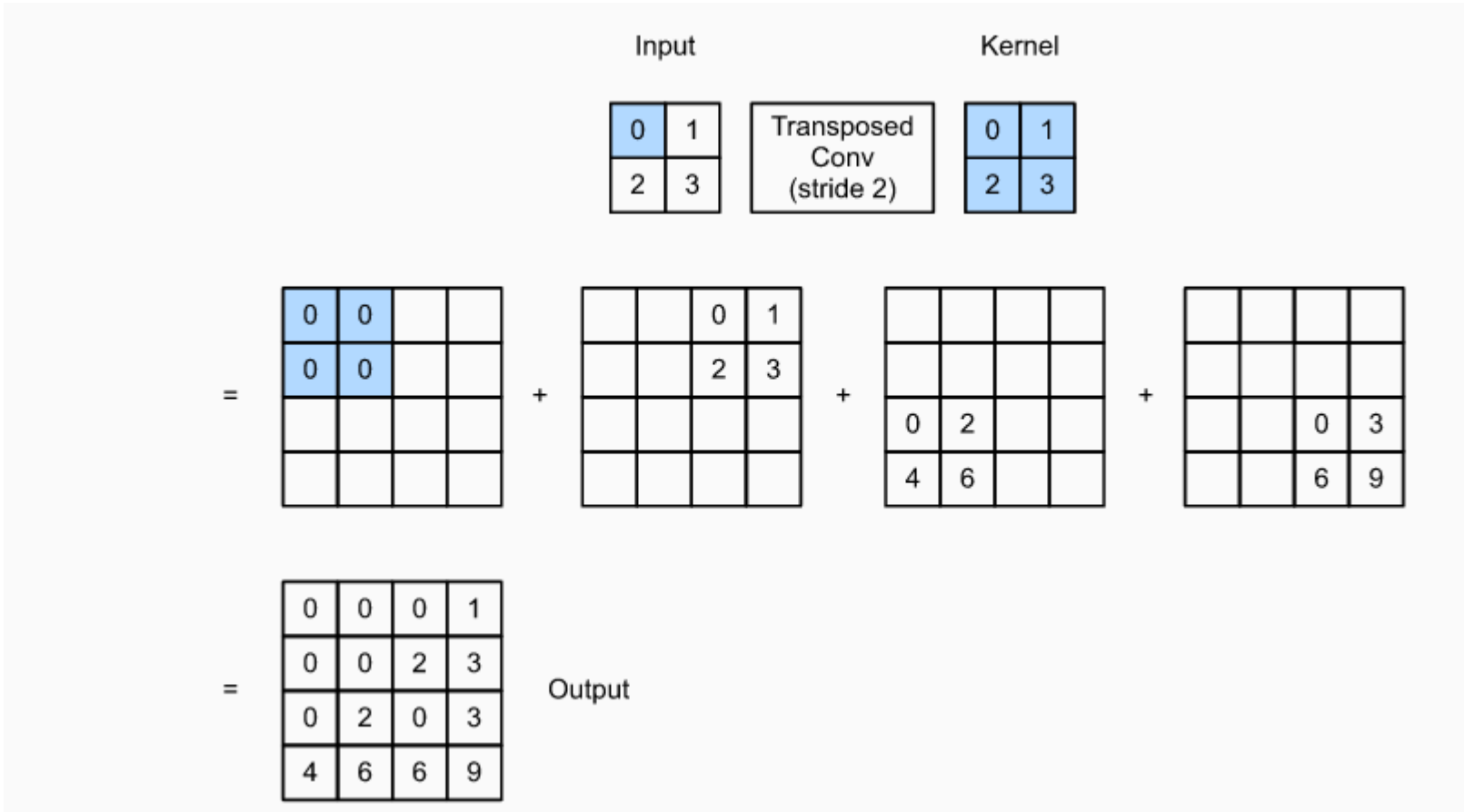
- Introduce a bottleneck layer to compress the data



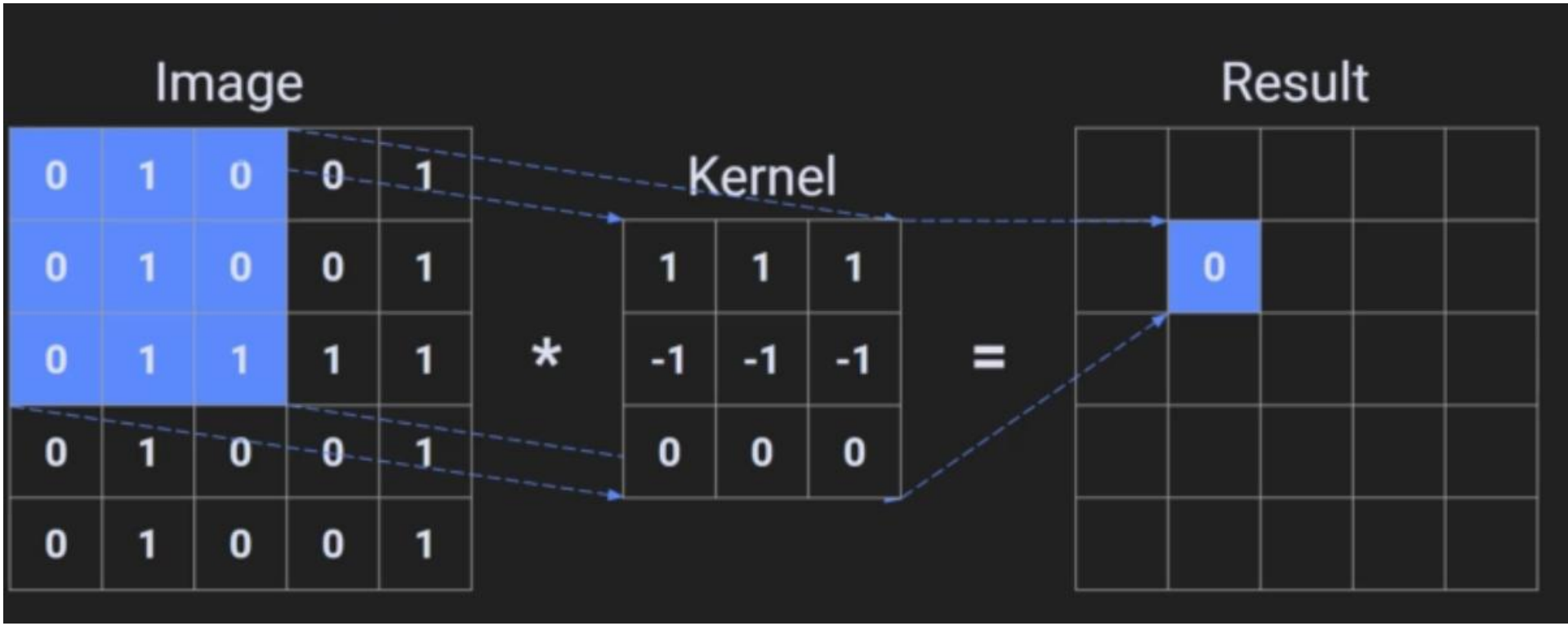
Terminology

- Convolution with stride >= 2 (downsampling)
- Transpose convolution (upsampling)
- Note that convolution and transposed convolution is not inverse operation.
- We use it to create the correct matrix dimension and back propagation to identify the value

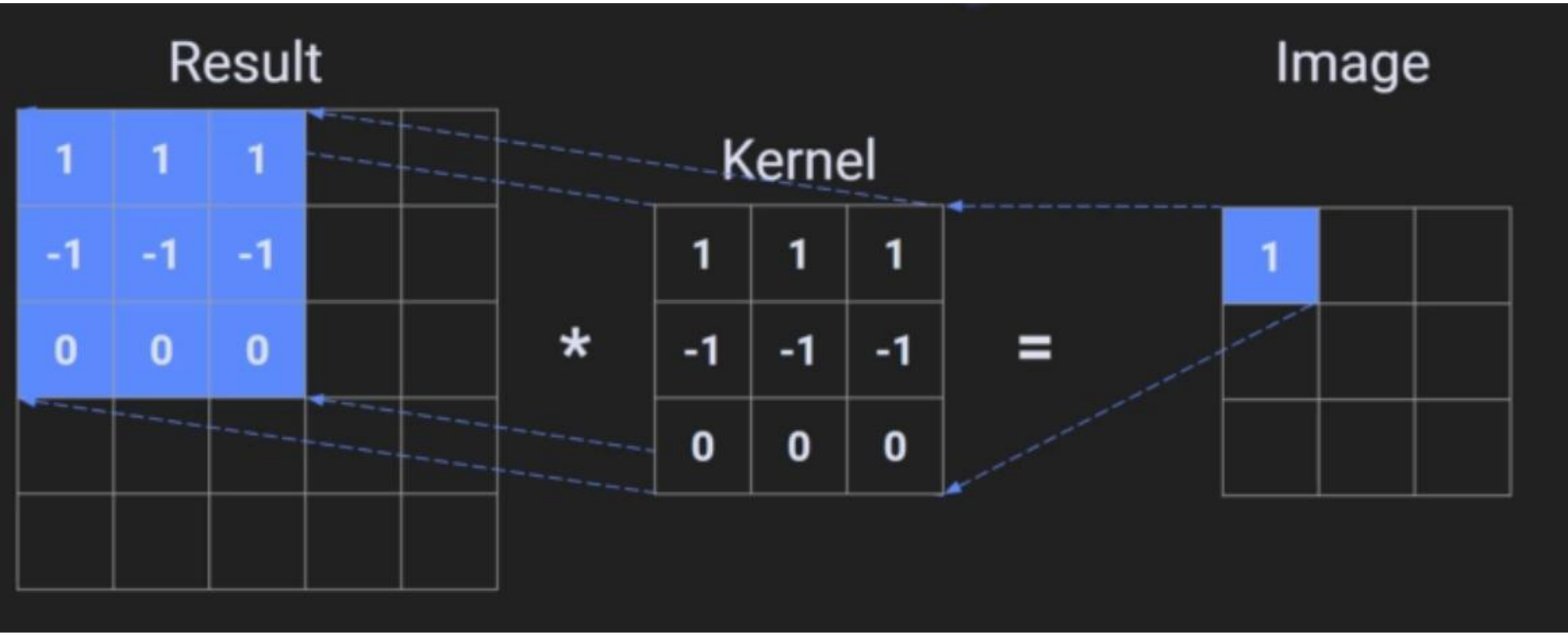
Convolution/Transposed Convolution



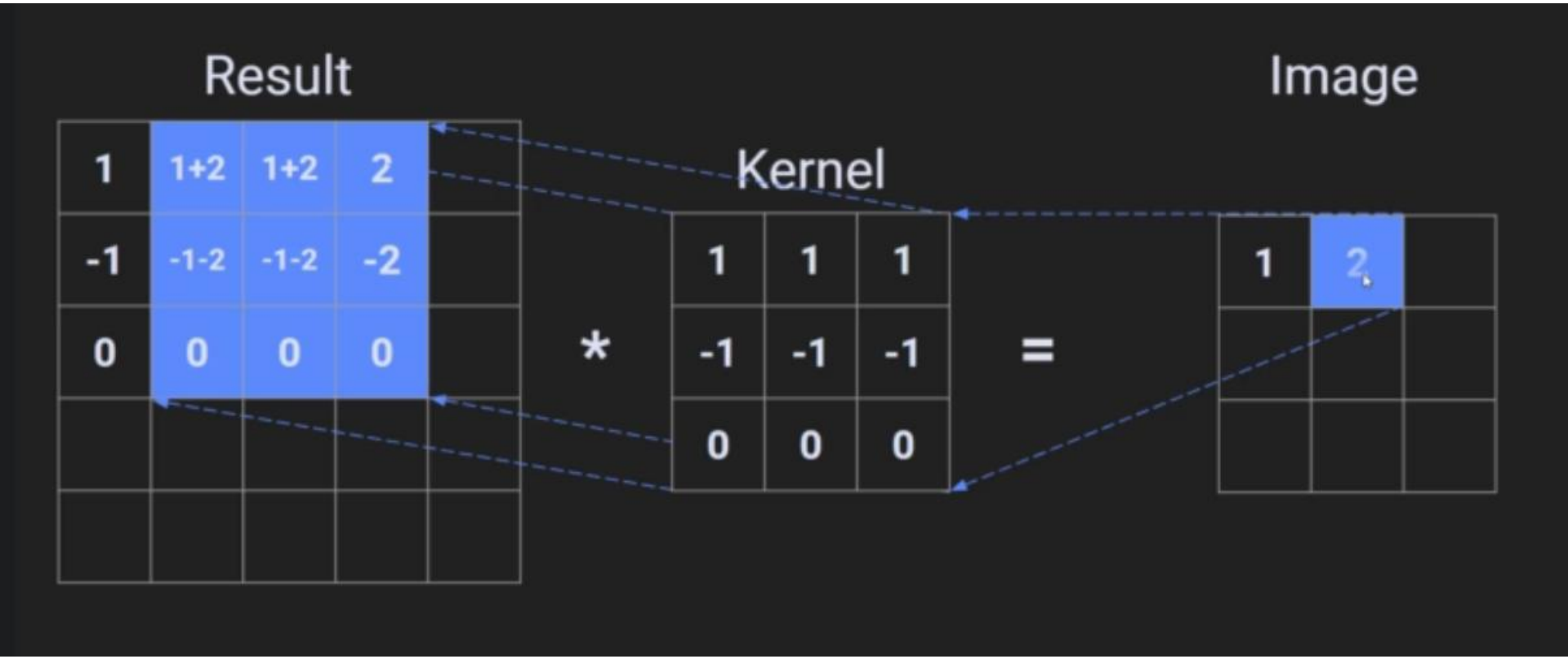
Convolution



Transpose Convolution



Next Step



Question

- Please complete the previous example when the input of the transposed convolution is:

1	2	1
0	1	0
1	-1	1

Transpose Convolution Size

$$N_h = s_h (M_h - 1) + k - 2p$$

h is for height

M is input

N is output

k is kernel

p is padding

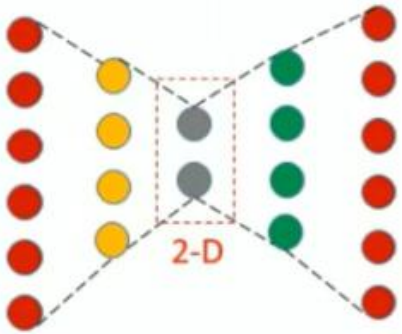
Example

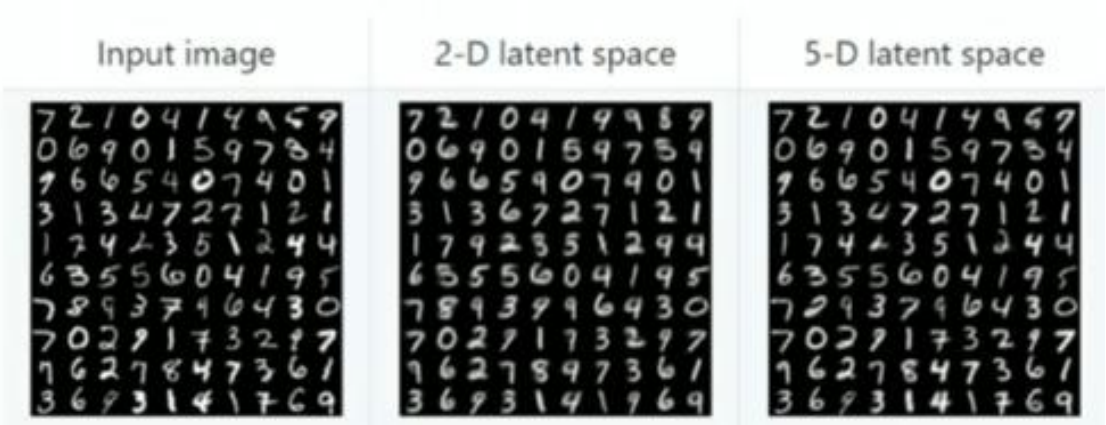
$$N_h = s_h (M_h - 1) + k - 2p$$

$$N = 1(3 - 1) + 3 - 2 \text{ x}0$$

The number of latent variables

- The number of latent variables (the number of output neurons/dimension) of the last encoder layer/the compressing layer (the dimensionality of the compressed representation) matters. The more, the better:





Training an Autoencoder

- The label data is the same as training data
- The loss function can be binary cross entropy or MSE
- Autoencoder are lossy

```
autoencoder.fit(x_train, x_train,
                epochs=50,
                batch_size=256,
                shuffle=True,
                validation_data=(x_test, x_test))
```

Example

- See Jupiter Notebook

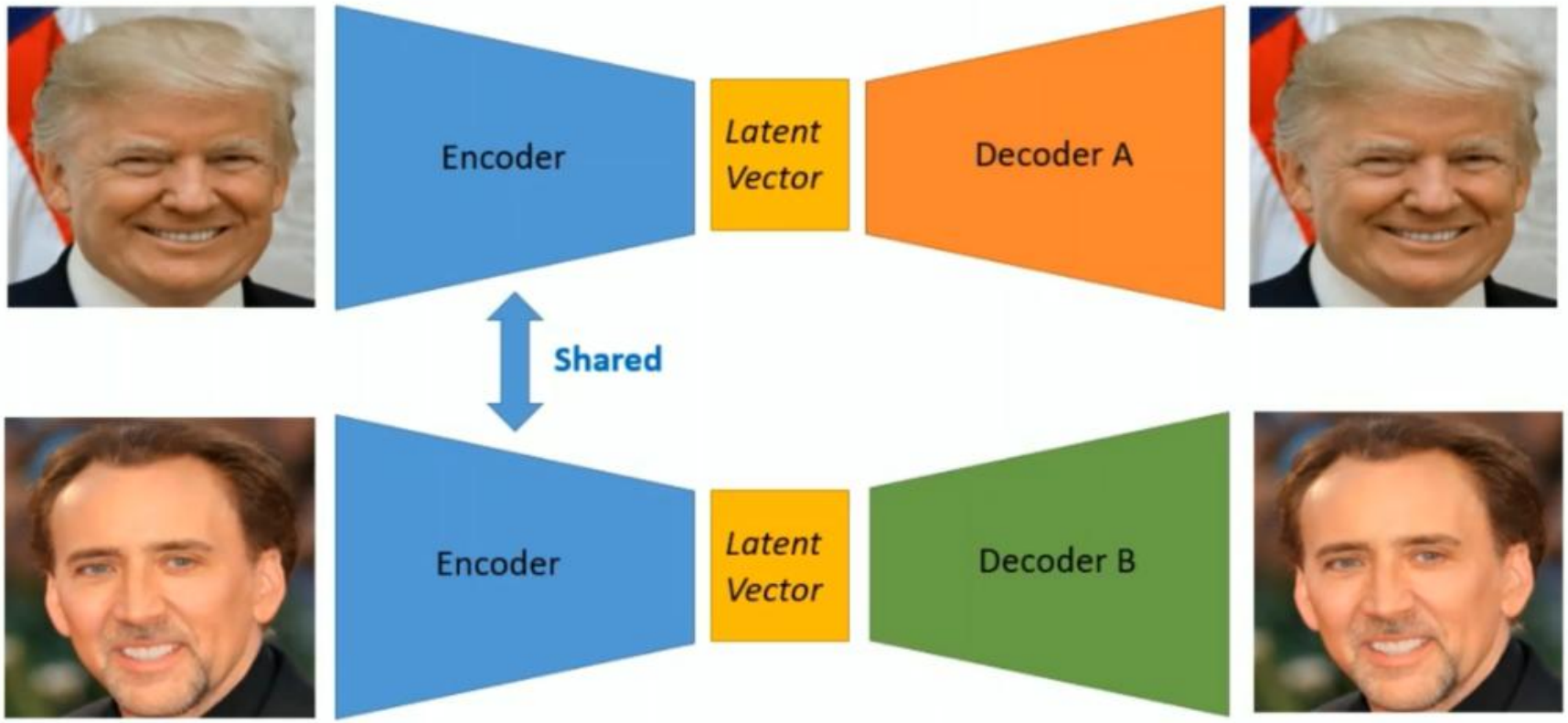
Deep Fake with AutoEncoder

Donald Trump → Mr. Bean

Home Alone → Home Stallone

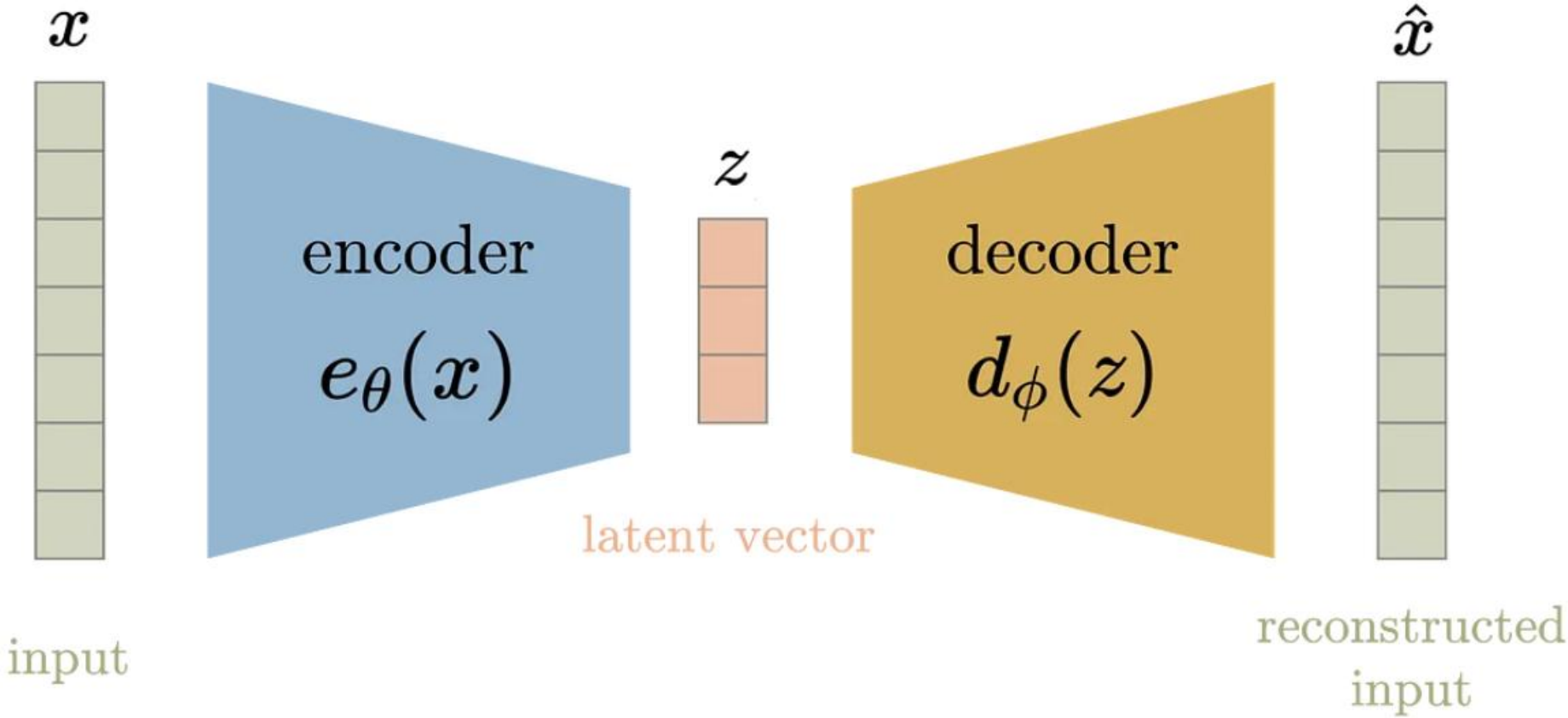
Training Phase

The **Decoder A** is only trained with faces of A; the **Decoder B** is only trained with faces of B. However, all latent faces are produced by the **same Encoder**.



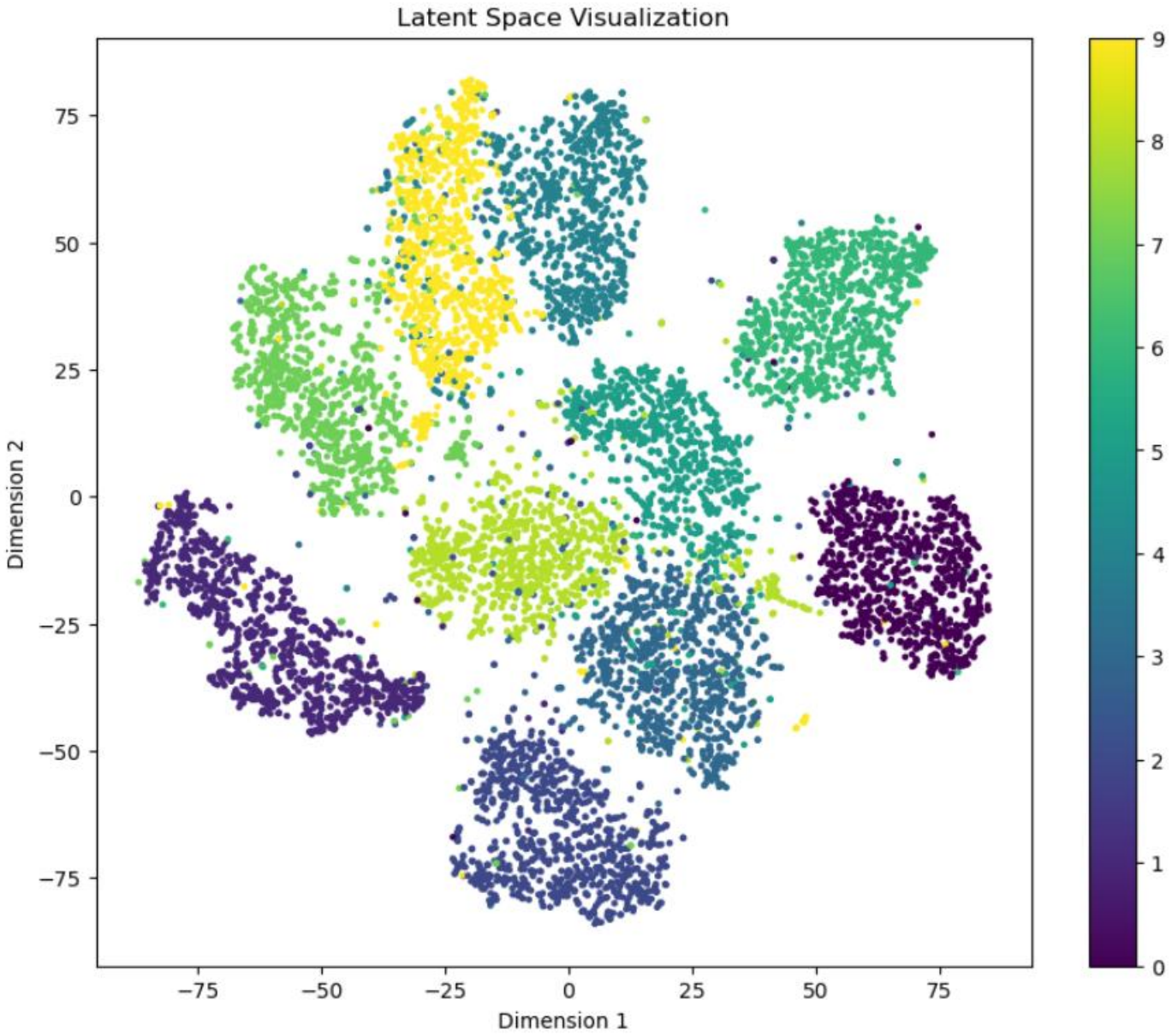
Variation AutoEncoder vs. Auto Encoder

AutoEncoder

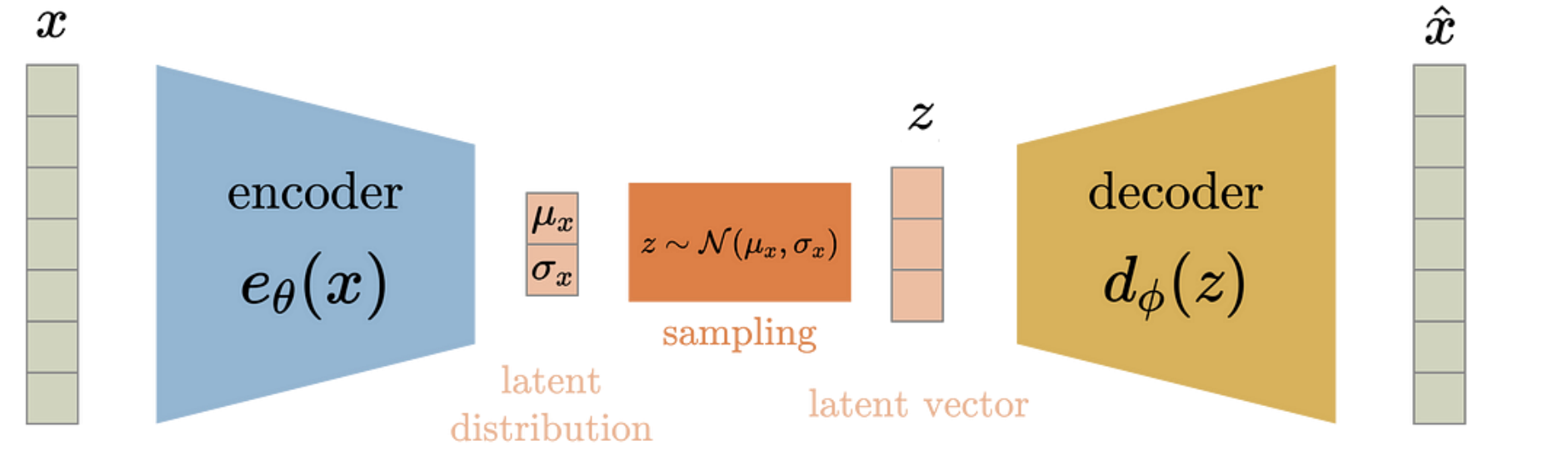


$$loss = \|x - \hat{x}\|_2 = \|x - d_\phi(z)\|_2 = \|x - d_\phi(e_\theta(x))\|_2$$

image by Aqeel Anwar


Latent Space Visualization
Dimension 1
Dimension 2
75
50
25
0
−25
−50
−75
9
8
7
6
5
4
3
2
1
0

Variational AutoEncoder



$$\text{reconstruction loss} = \|x - \hat{x}\|_2 = \|x - d_\phi(z)\|_2 = \|x - d_\phi(\mu_x + \sigma_x \epsilon)\|_2$$

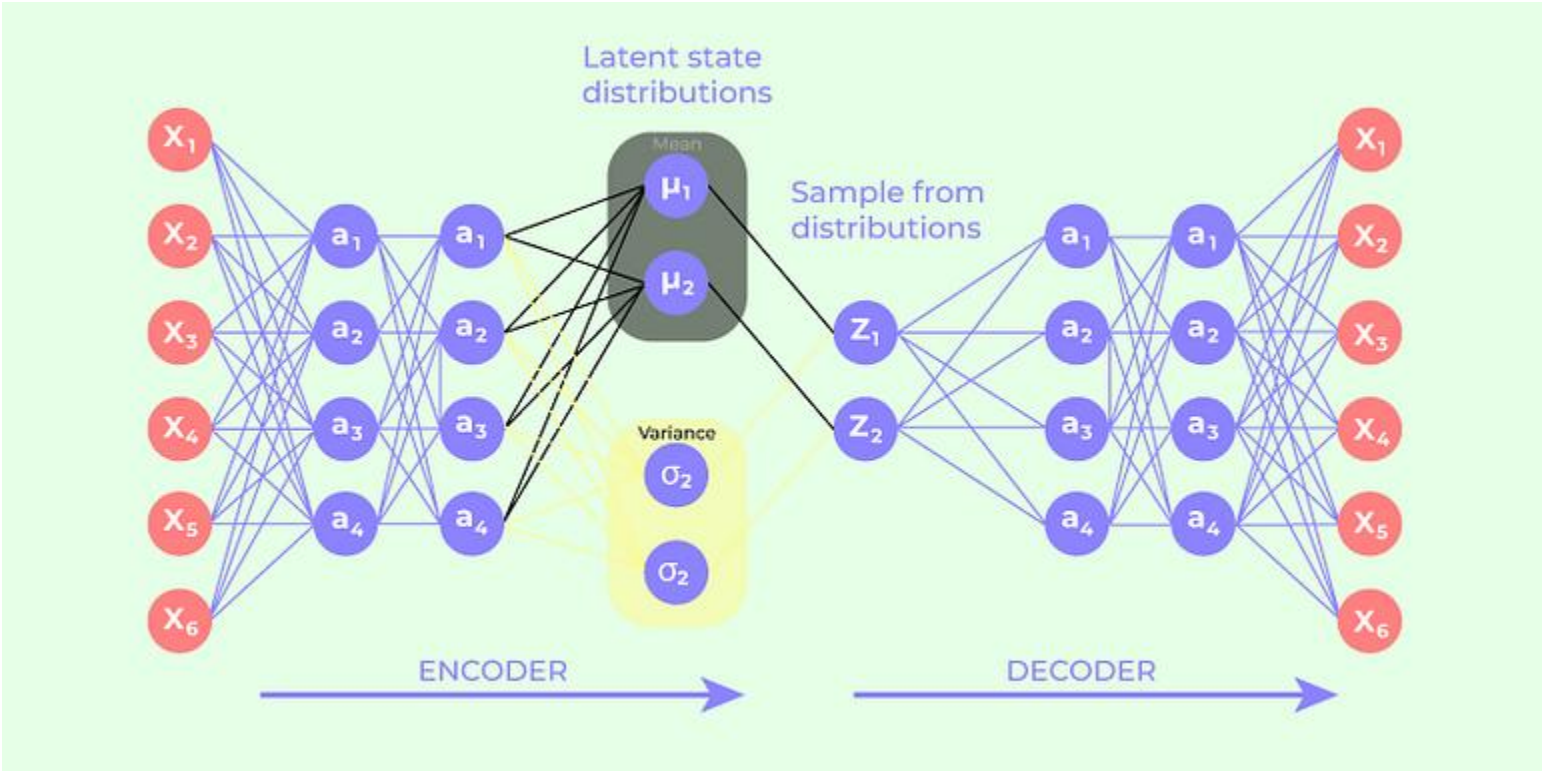
$$\mu_x, \sigma_x = e_\theta(x), \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

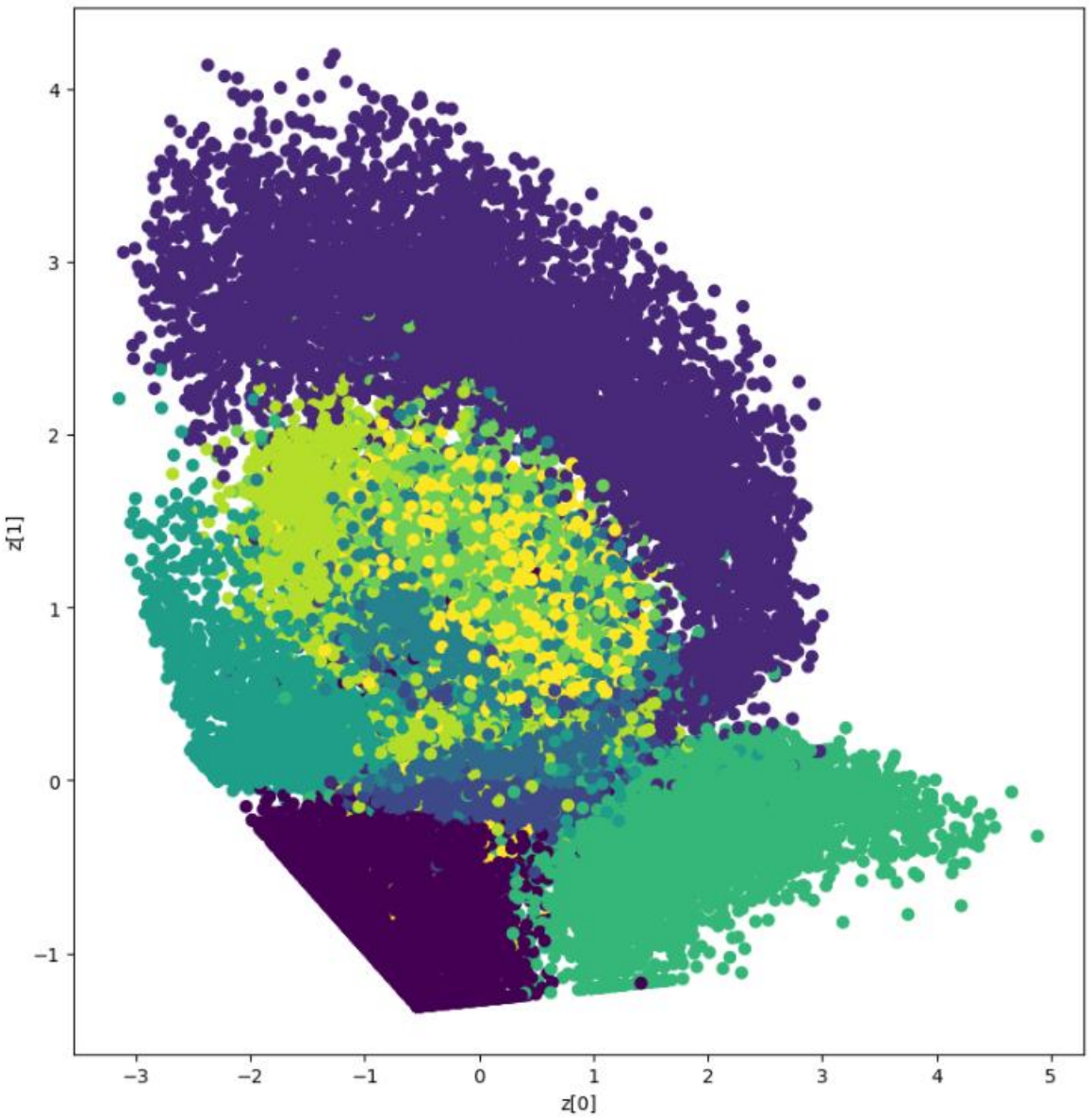
$$similarity\ loss = KL\ Divergence = D_{KL}(\mathcal{N}(\mu_x, \sigma_x) \parallel \mathcal{N}(\mathbf{0}, \mathbf{I}))$$

$$loss = reconstruction\ loss + similarity\ loss$$

image by Aqeel Anwar

Latent dimension =4




z[1]
z[0]
4
3
2
1
0
−1
−3
−2
−1
0
1
2
3
4
5
9
8
7
6
5
4
3
2
1
0

VAE

https://www.youtube.com/watch?v=sV2FOdGqlX0

Questions?

Homework

- Denoising Occluded Mnist

Modify noisy data with this occlude function instead

```
import random
def occlude(array):
"""    Adds occlusion.    """
  new_array = copy.deepcopy( array )
  print(new_array.shape)
  for k in range( len(new_array)):
        x = random.randint(0,25)
        new_array[k,x:x+2,:] = 1.0
return new_array
```

Variational Auto Encoder