Autoencoder

Weizi Li

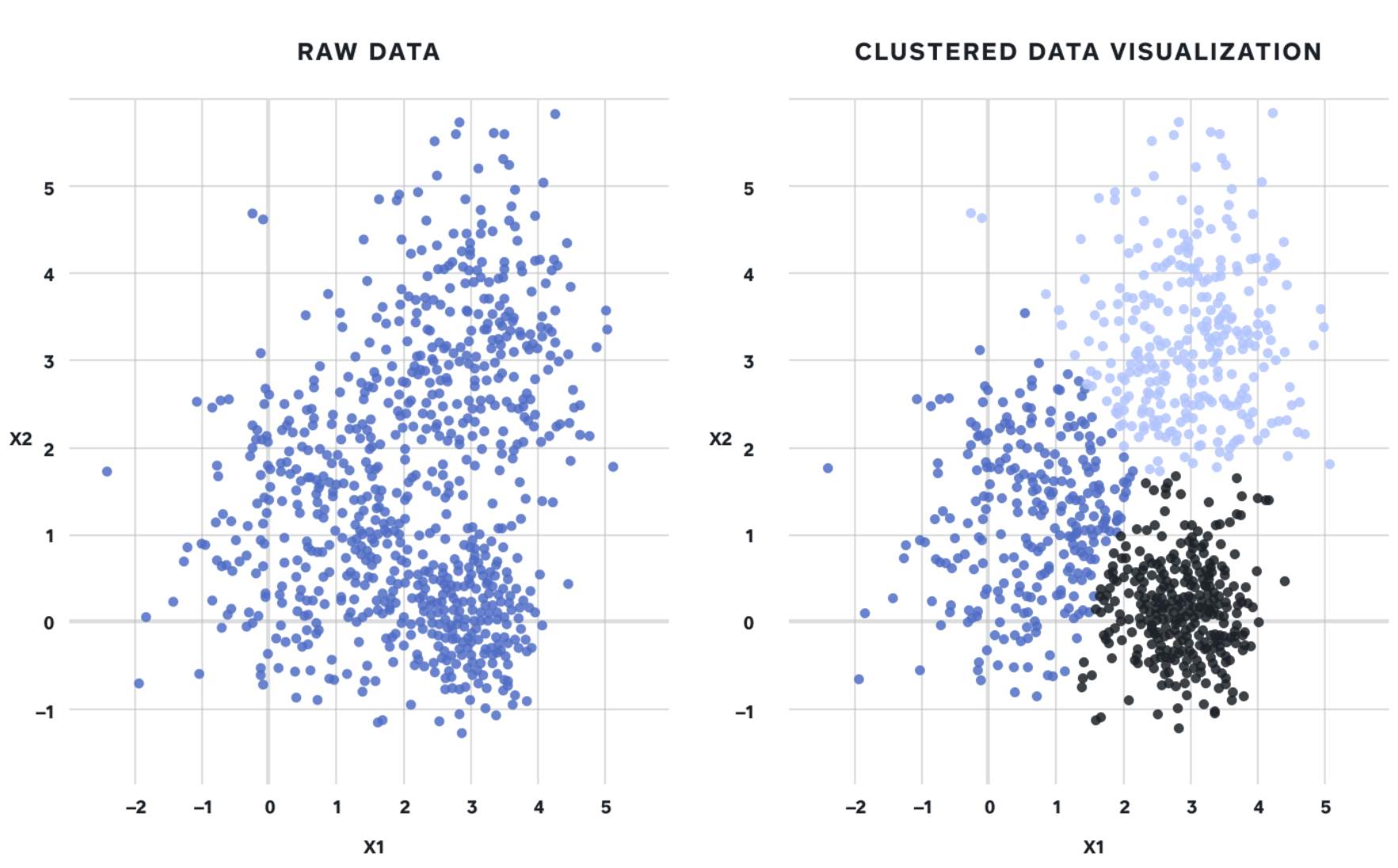
Department of Computer Science
University of Memphis

Outline

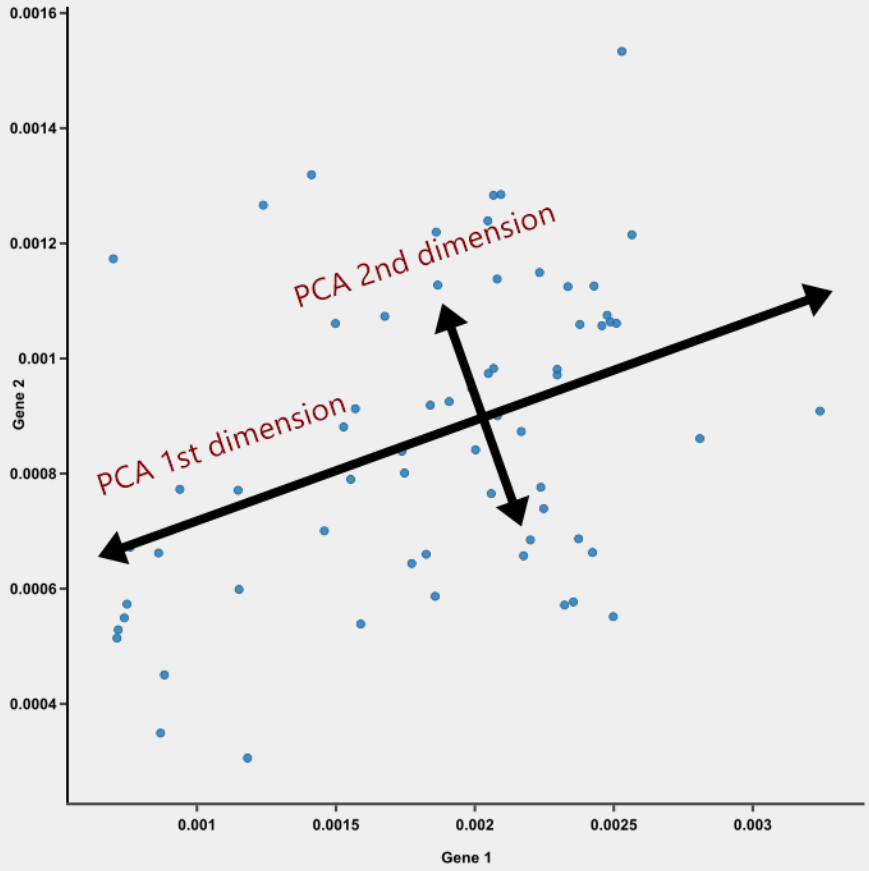
- Unsupervised Learning
- Autoencoder
- Applications

Unsupervised Learning

- Clustering: find underlying structures
  - E.g., K-Means
- Dimension reduction: fast processing, visualization
  - E.g., PCA, Autoencoder
- Generative models: generate synthetic data
  - E.g., Variational Autoencoder, Generative Adversarial Network
- Other applications are possible


RAW DATA
CLUSTERED DATA VISUALIZATION
X2
X1
5
4
3
2
1
0
−1
−2
−1
0
1
2
3
4
5

- N-dimensional data $\rightarrow$ M orthogonal directions in which the data have the most variance

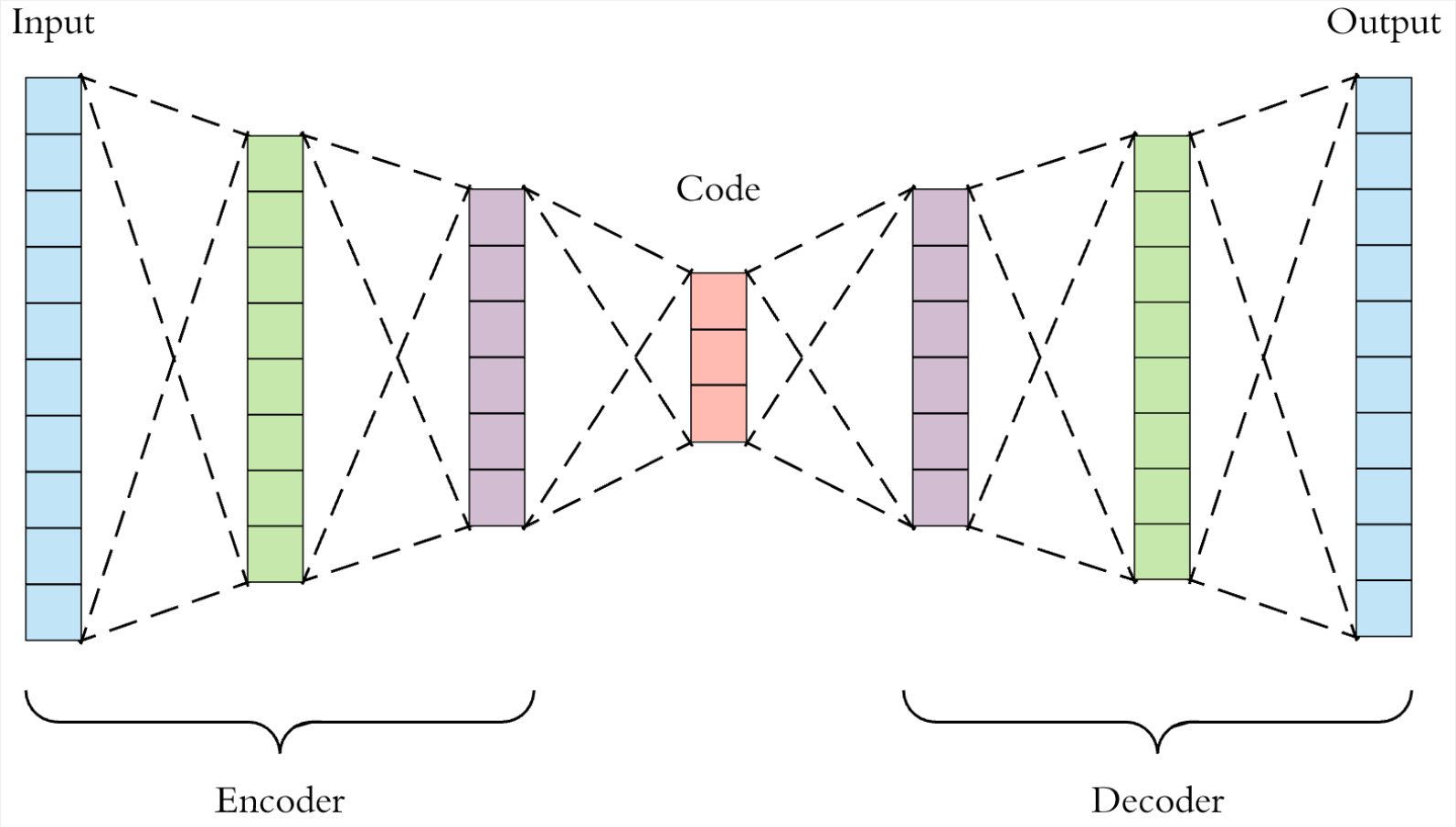


- N-dimensional data $\rightarrow$ M orthogonal directions in which the data have the most variance

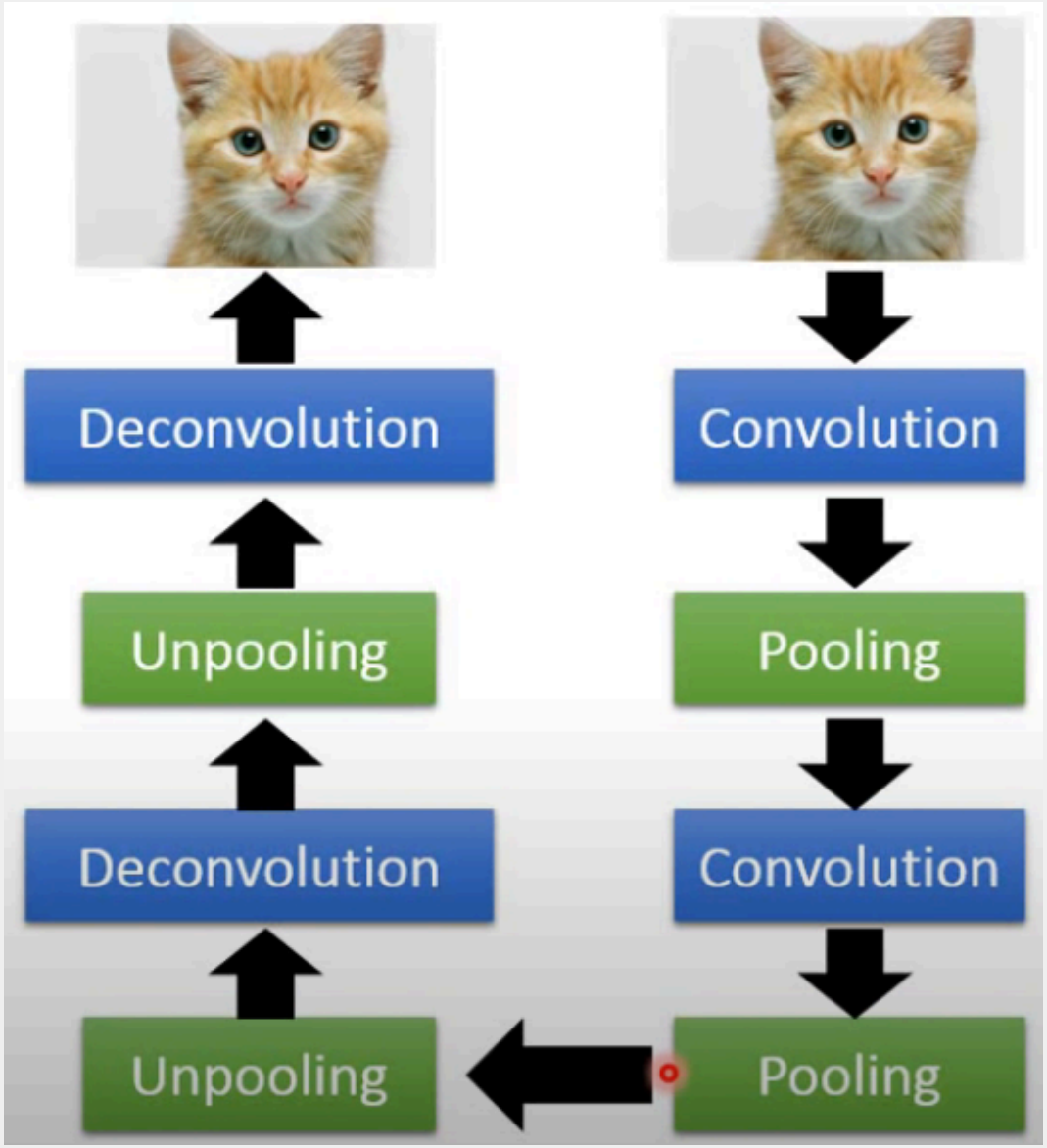
```
from sklearn import decomposition
pca = decomposition.PCA(n_components=2)
pca.fit(diabetes_X)
X_2d = pca.transform(diabetes_X)
```

Autoencoder

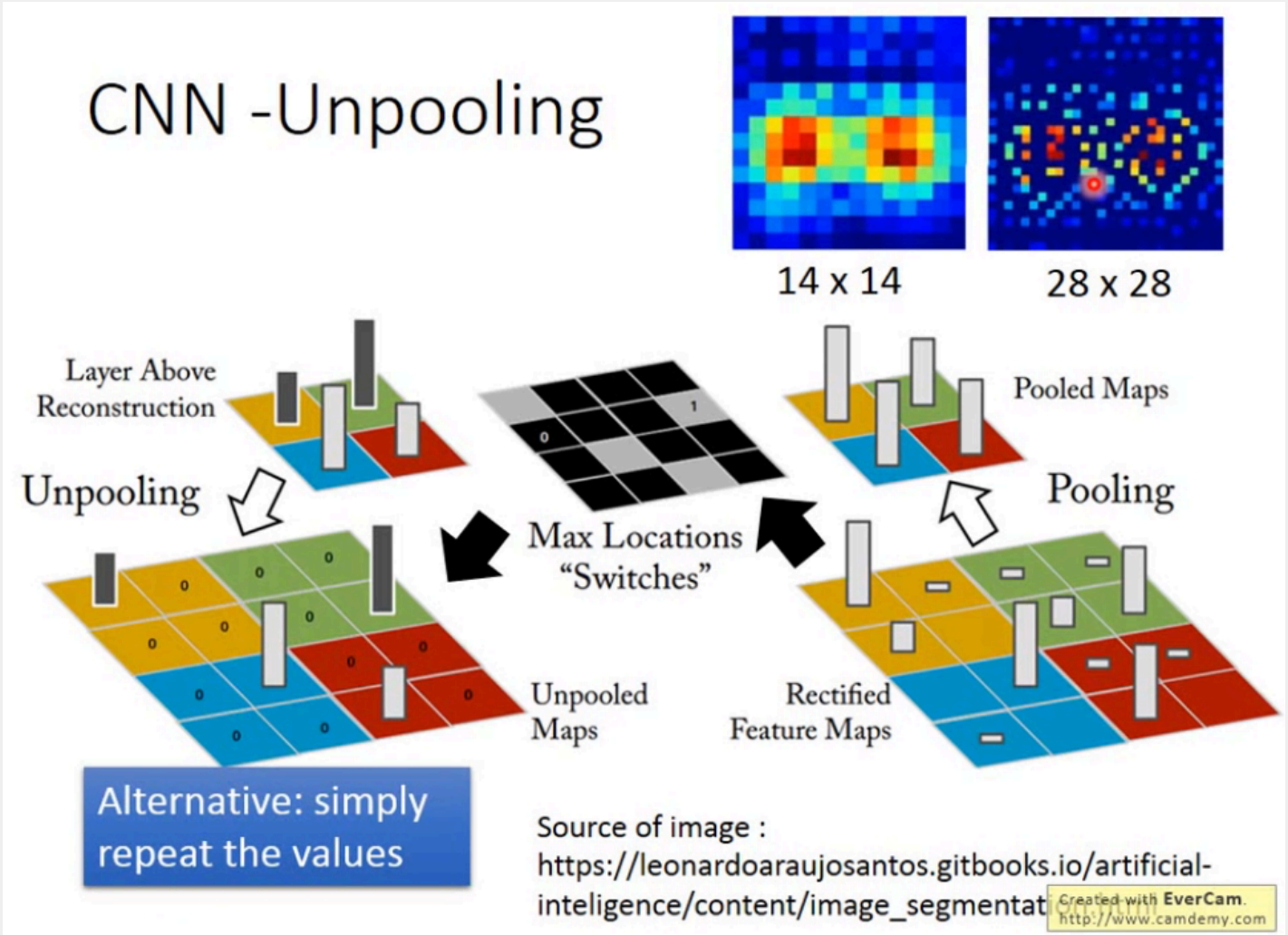
- Goal: minimize $\|input - output\|$



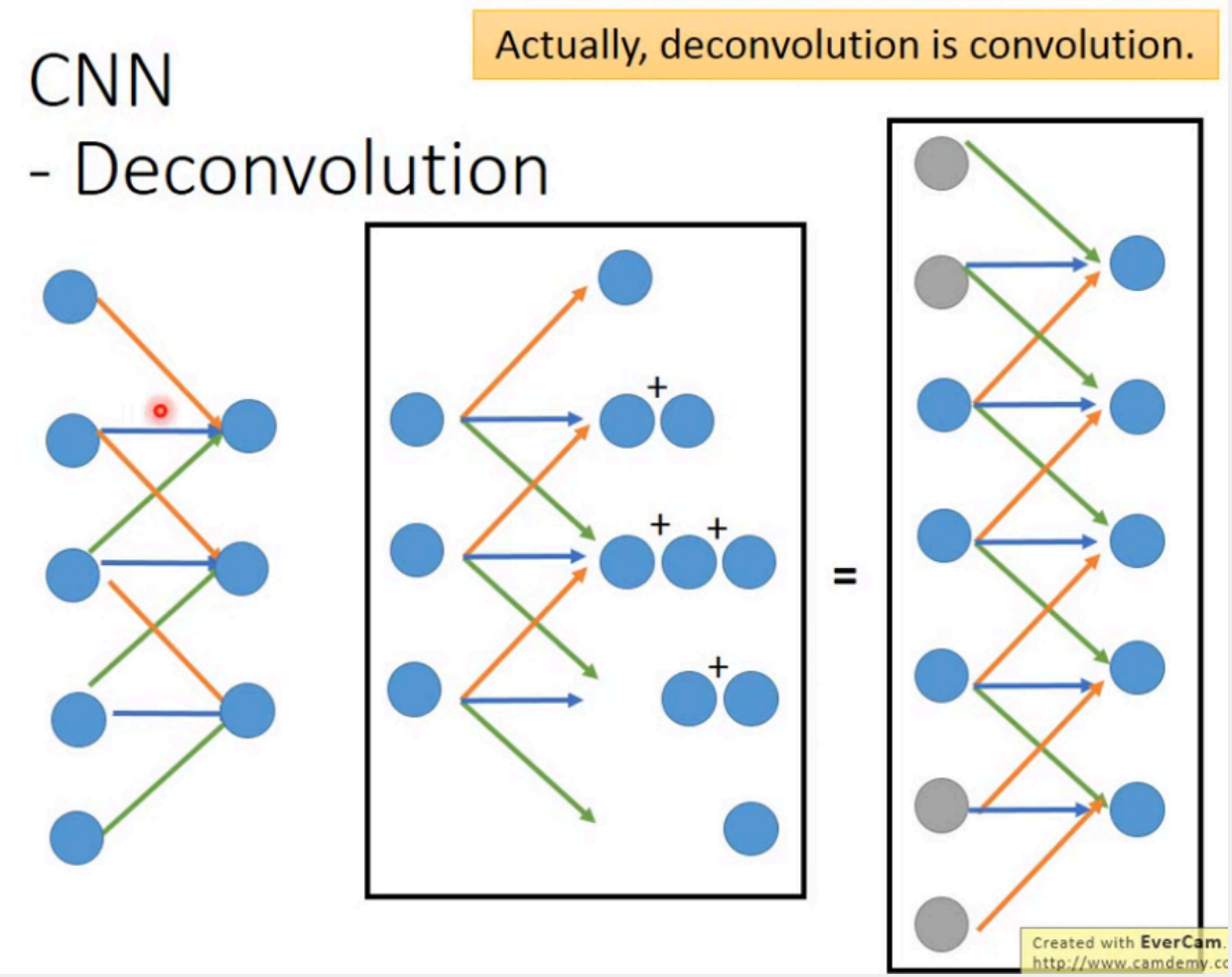
- PCA is a linear operation and can be realized using Autoencoder (w/ linear activation)
- Autoencoder can achieve non-linear dimension reduction through non-linear activation functions
- Hinton and Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks, 2006 (link)


Deconvolution
Convolution
Unpooling
Pooling
Deconvolution
Convolution
Unpooling
Pooling

- E.g., Max unpooling: keep the "max" and fill the rest with 0 or just repeat the "max" value



- Deconvolution is just convolution



Applications

- Denoising
  - input: noisy data
  - output: clean data
  - Goal: minimize $\|input - output\|$
- Vincent et al., Extracting and Composing Robust Features with Denoising Autoencoders, 2008 (influential paper on denoising AE)
- Example code: link

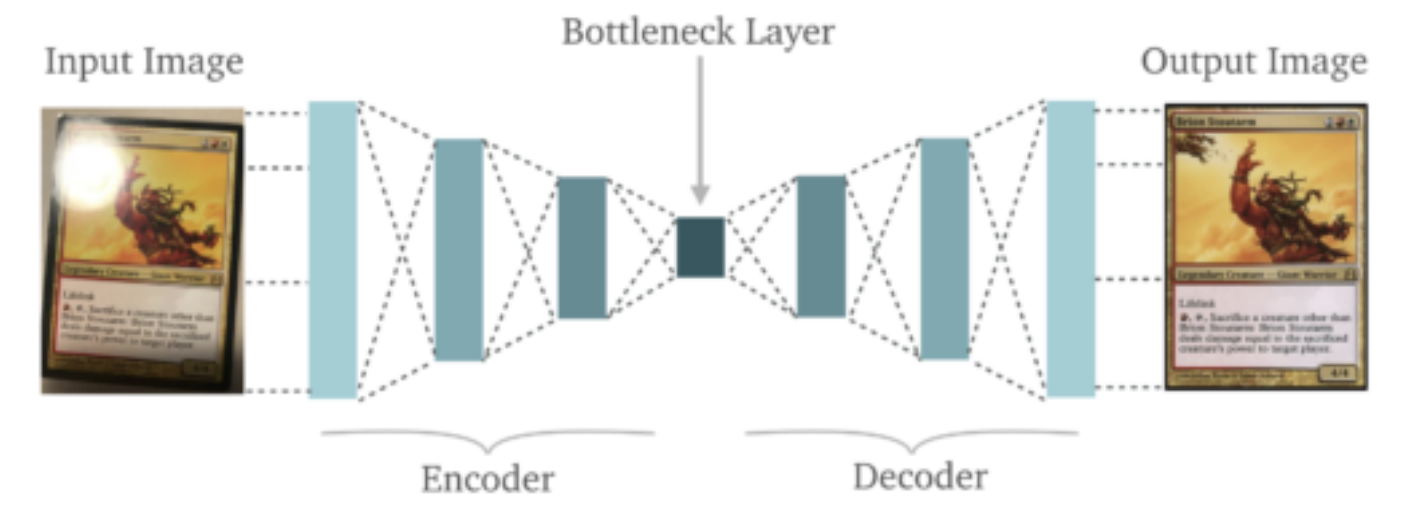
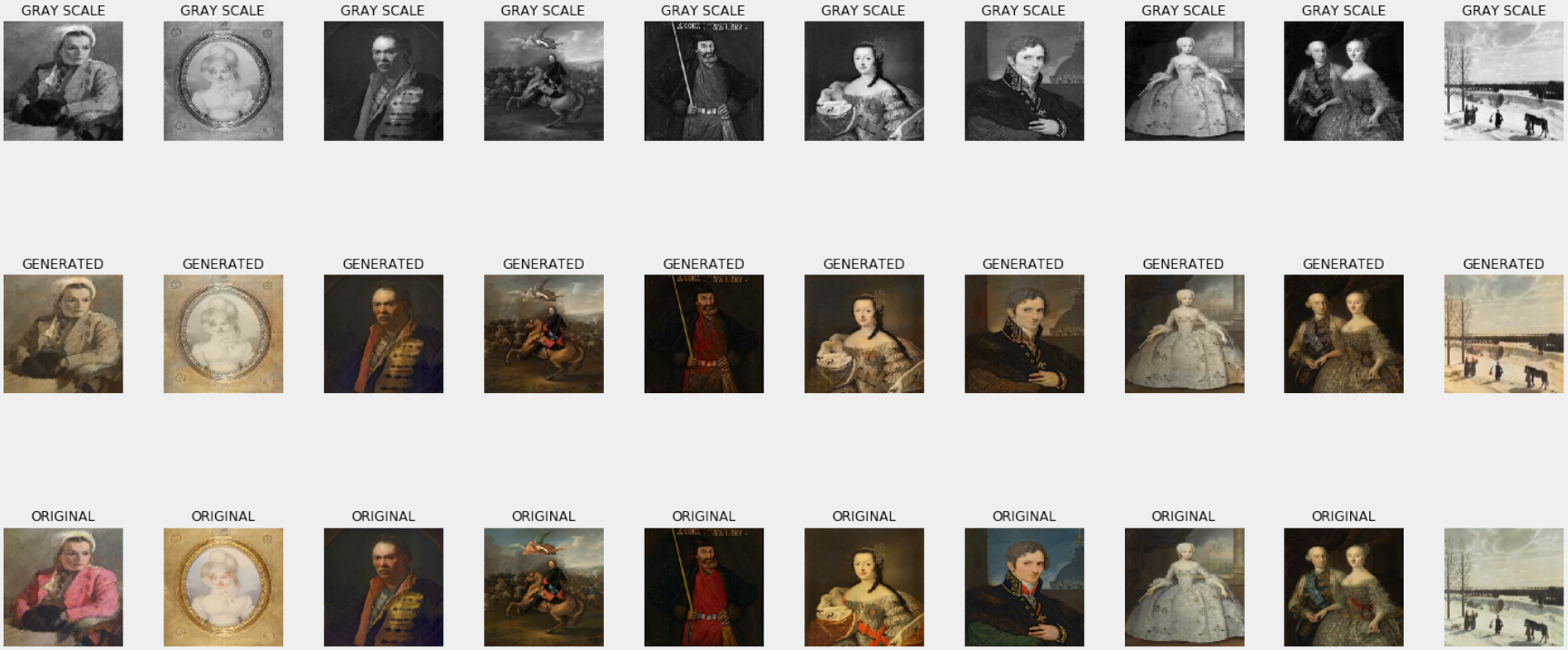

Input Image
Bottleneck Layer
Output Image
Encoder
Decoder

Image Coloring



- link