Deep Belief Networks

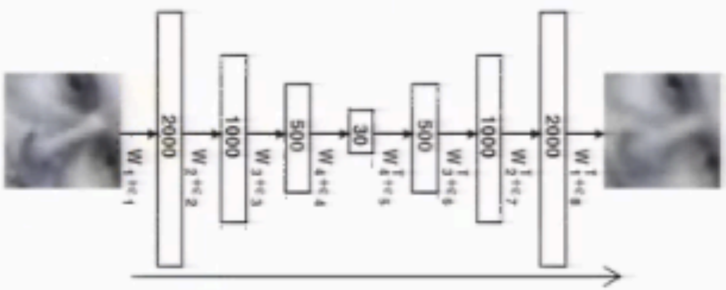
Deep Belief Networks

[Hinton & Salakhutdinov, *Science*, 2006]

- Problem: training networks with many hidden layers doesn't work very well
  - local minima, very slow training if initialize with zero weights

- Deep belief networks
  - autoencoder networks to learn low dimensional encodings

  - but more layers, to learn better encodings



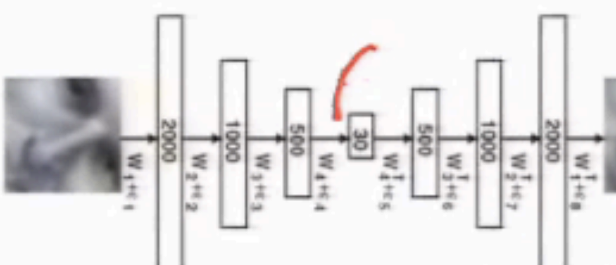
Deep Belief Networks

[Hinton & Salakhutdinov, 2006]

original image

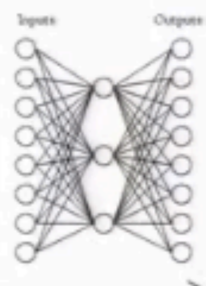
reconstructed from 2000-1000-500-30 DBN

reconstructed from 2000-300, linear PCA



logistic transformations

versus



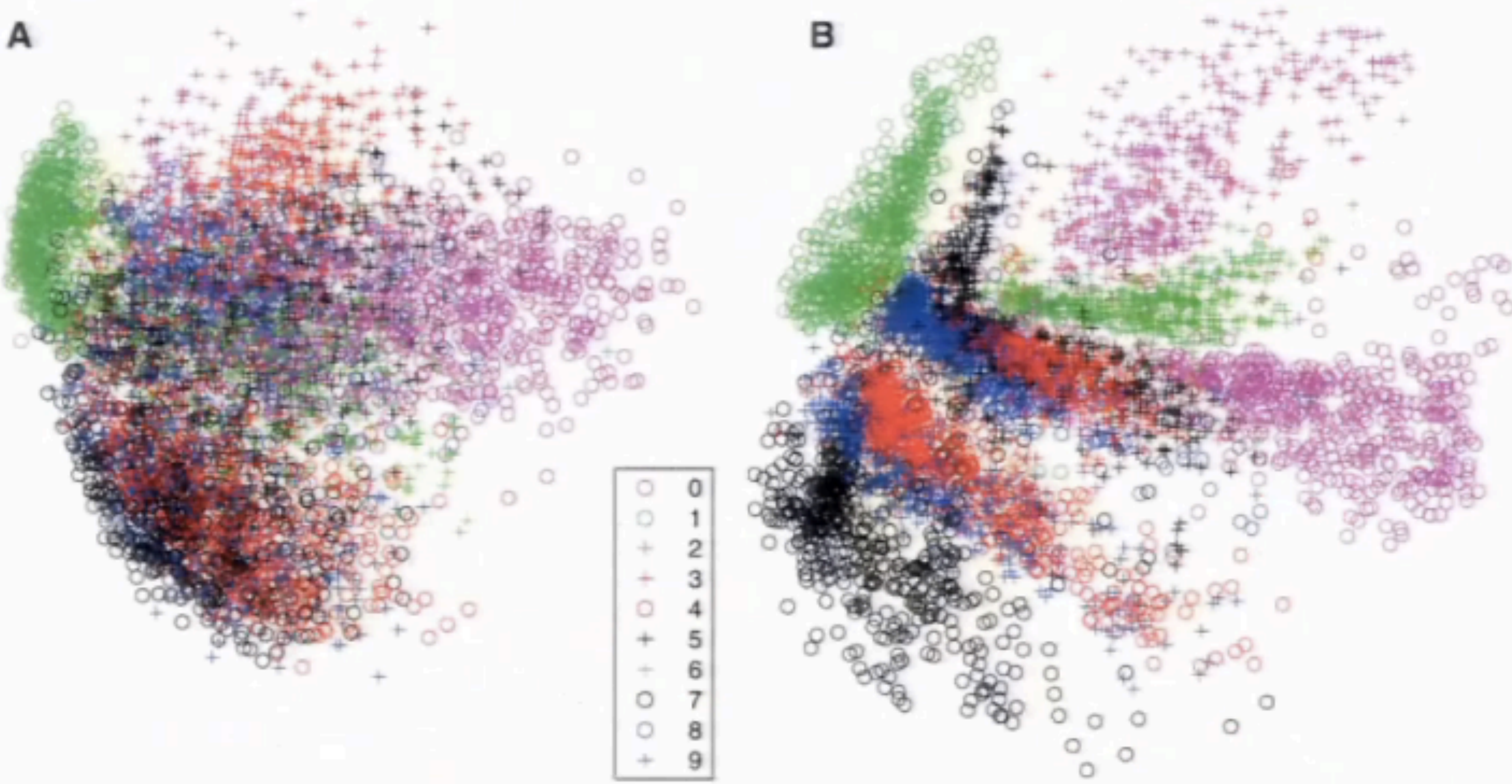
linear transformations

Encoding of digit images in two dimensions

[Hinton & Salakhutdinov, 2006]

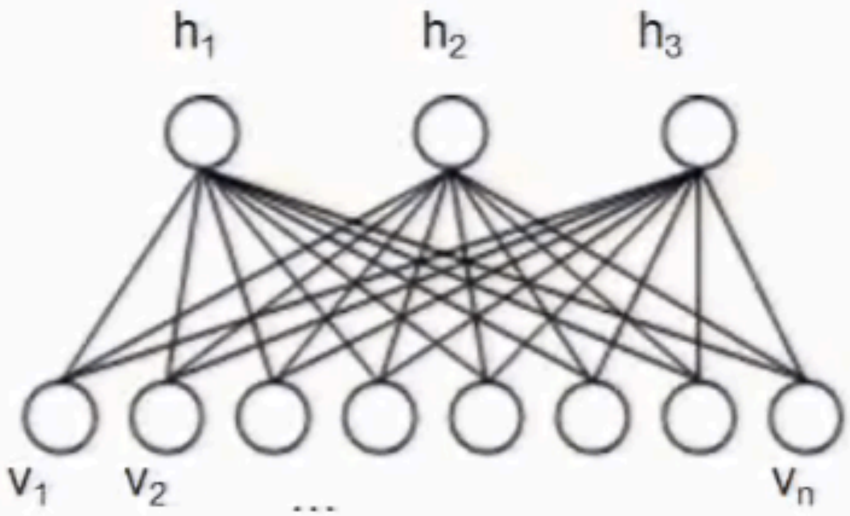
784-2 linear encoding (PCA)

784-1000-500-250-2 DBNet



Restricted Boltzman Machine

- Bipartite graph, logistic activation
- Inference: fill in any nodes, estimate other nodes
- consider v_i, h_j are boolean variables



$$P(h_j = 1|\mathbf{v}) = \frac{1}{1 + \exp(\sum_i w_{ij} v_i)}$$

$$P(v_i = 1|\mathbf{h}) = \frac{1}{1 + \exp(\sum_j w_{ij} h_j)}$$

Deep Belief Networks: Training

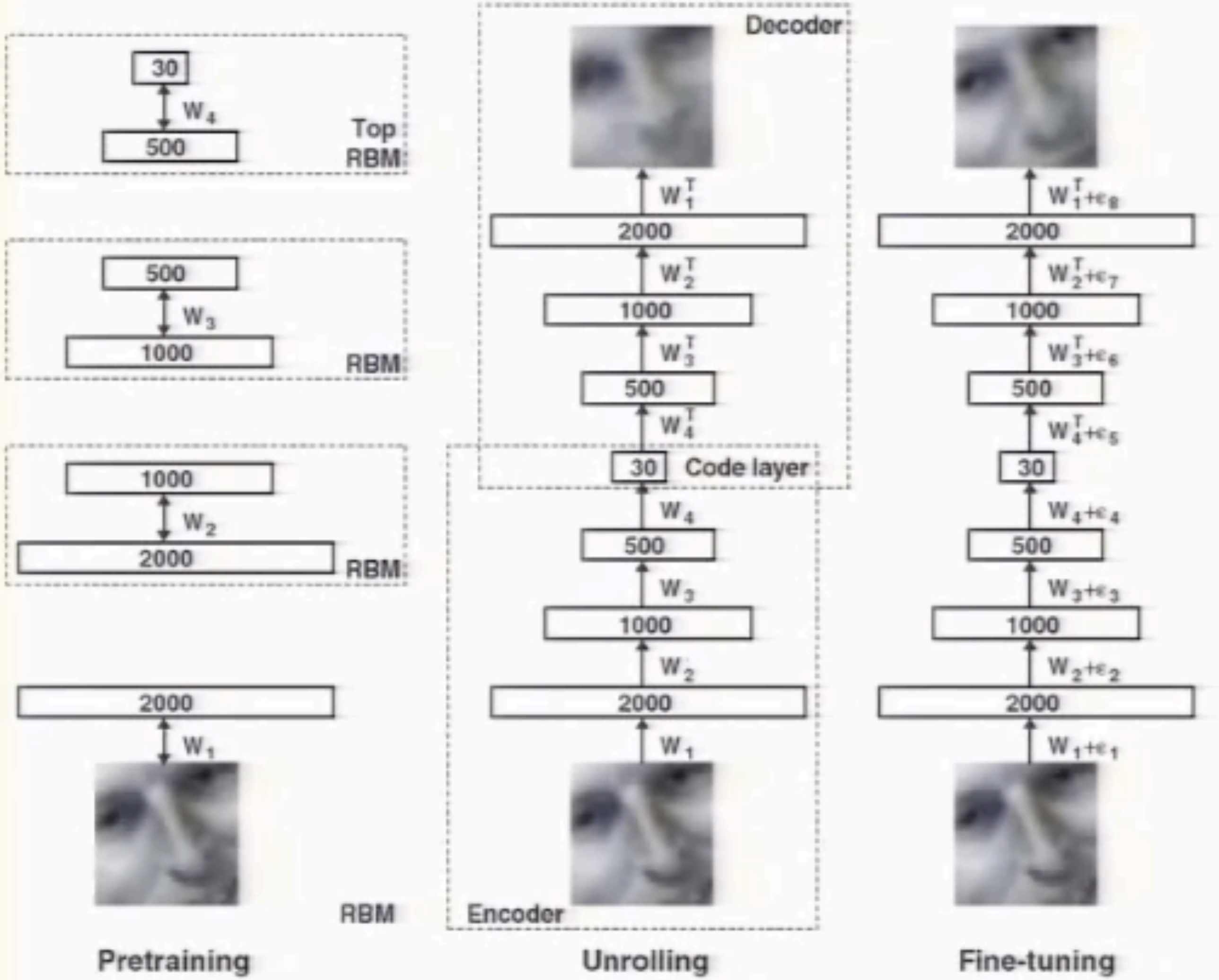
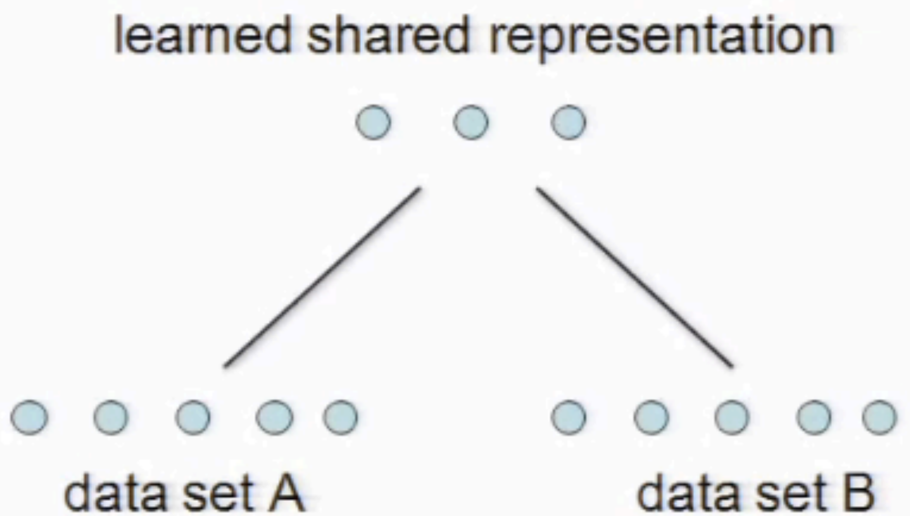


Fig. 1. Pretraining consists of learning a stack of restricted Boltzmann machines (RBMs), each having only one layer of feature detectors. The learned feature activations of one RBM are used as the "data" for training the next RBM in the stack. After the pretraining, the RBMs are "unrolled" to create a deep autoencoder, which is then fine-tuned using backpropagation of error derivatives.

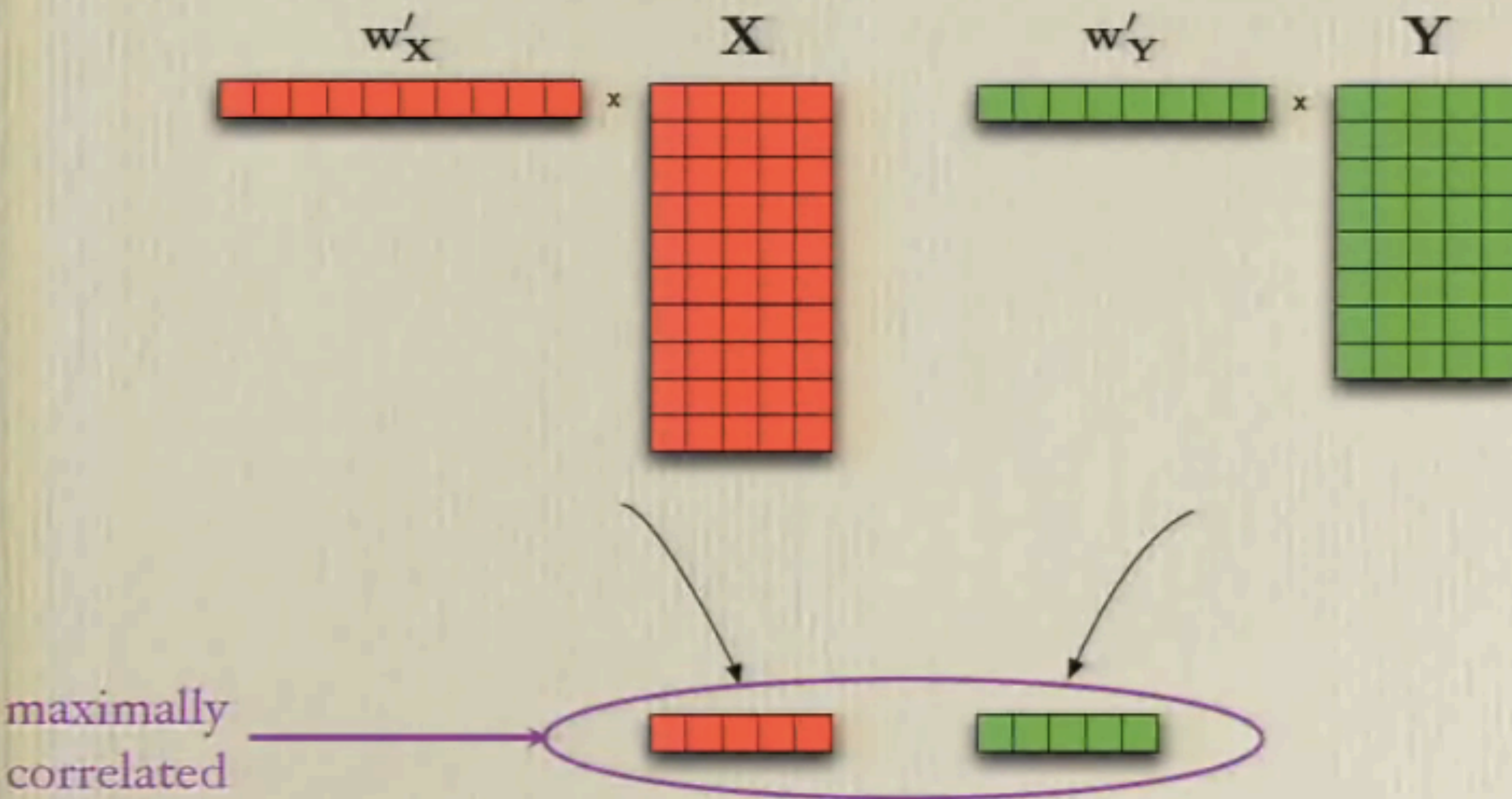
Dimensionality reduction across multiple datasets

- Given data sets A and B, find linear projections of each into a <u>common</u> lower dimensional space!
  - Generalized SVD: minimize sq reconstruction errors of both
  - Canonical correlation analysis: maximize correlation of A and B in the projected space



Canonical correlation analysis

$$Corr(A,B) = \frac{1}{N}\sum_{i=1}^{N}\frac{(A_i-\bar{A})}{\sigma_A}\;\frac{(B_i-\bar{B})}{\sigma_b}$$



$\mathbf{w}'_{\mathbf{X}}\mathbf{X}$ $\mathbf{w}'_{\mathbf{Y}}\mathbf{Y}$

[slide courtesy of Indra Rustandi]