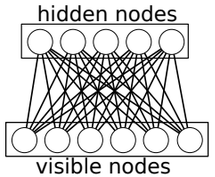
(Charlie) Yichuan Tang
Ruslan Salakhutdinov
Geoffrey Hinton
University of Toronto, Canada

Robust Boltzmann Machines
for Recognition and Denoising

Introduction

- Allows Boltzmann Machines to be robust to noise and occlusions
- Multiplicative gating allows for the presence of novel occluders with exotic appearances
- The structure of occluders and noise statistics can be learned from the data in an unsupervised fashion
- Completely automated image inpainting and denoising correspond to posterior inference in the model

Gaussian Restricted Boltzmann Machine



$$p(h_j = 1|\mathbf{v}) = \frac{1}{1+\exp(-\sum_i W_{ij}v_i - c_j)}$$

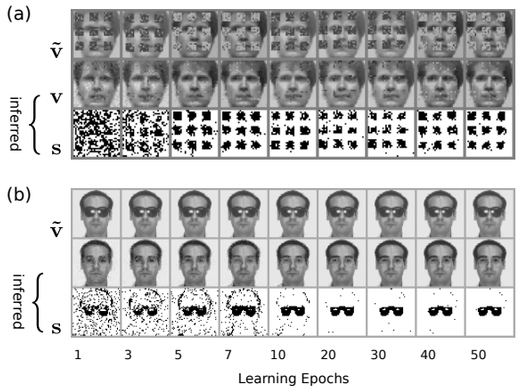
$$p(v_i|\mathbf{h}) = \mathcal{N}(v_i|\mu_i, \sigma_i^2)$$

$$\mu_i = b_i + \sigma_i^2 \sum_j W_{ij} h_j$$

$$E_{GRBM}(\mathbf{v},\mathbf{h}) = \frac{1}{2}\sum_i \frac{(v_i - b_i)^2}{\sigma_i^2} - \sum_j c_j h_j - \sum_{ij} W_{ij} v_i h_j$$

Learning Noise Structure

- A binary RBM component within the RoBM learns the structure of the occluders
- A learned occluder model allows for improved denoising and recognition



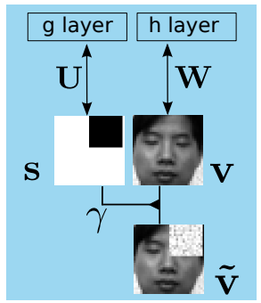
RoBM

$$E_{RoBM}(\mathbf{v},\tilde{\mathbf{v}},\mathbf{s},\mathbf{h},\mathbf{g}) = \frac{1}{2}\sum_i \frac{\gamma_i^2}{\sigma_i^2} s_i (v_i - \tilde{v}_i)^2 + \frac{1}{2}\sum_i \frac{(\tilde{v}_i - \tilde{b}_i)^2}{\tilde{\sigma}_i^2}$$

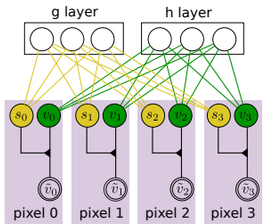
$$- \sum_i d_i s_i - \sum_k e_k g_k - \sum_{i,k} U_{ik} s_i g_k$$

$$+ \frac{1}{2}\sum_i \frac{(v_i - b_i)^2}{\sigma_i^2} - \sum_j c_j h_j - \sum_{ij} W_{ij} v_i h_j$$

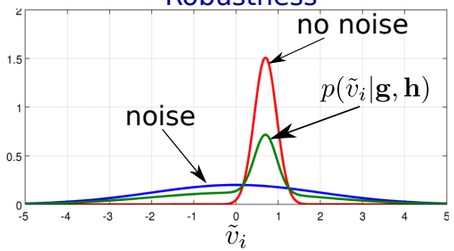
Idea



Model Diagram



Robustness



Inference using Gibbs Sampling:

$p(\mathbf{v},\mathbf{s},\mathbf{h},\mathbf{g}|\tilde{\mathbf{v}})$

1) Sample from $p(\mathbf{s}|\mathbf{h},\mathbf{g},\tilde{\mathbf{v}})$

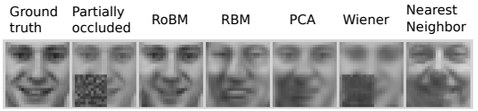
2) Sample from $p(\mathbf{v}|\mathbf{s},\mathbf{h},\tilde{\mathbf{v}})$

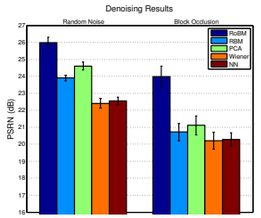
3) Sample from $p(\mathbf{h},\mathbf{g}|\mathbf{v},\mathbf{s})$

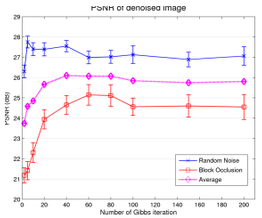
Block Occlusion Denoising

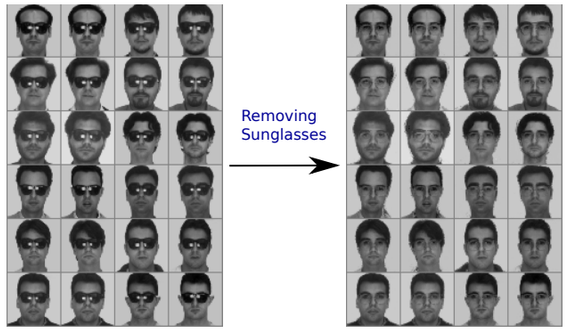
Salt and Pepper Denoising

Denoising

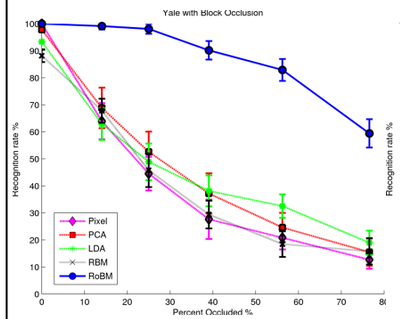


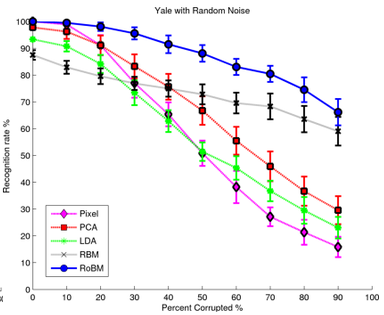






Face Recognition - Yale Database





AR faces - Inference

AR faces - Recognition

Algorithms	Sunglasses	Scarf
RoBM	84.5 %	80.7 %
RBM	61.7 %	32.9 %
Eigenfaces	66.9 %	38.6 %
LDA	56.1 %	27.0 %
Pixel	51.3 %	17.5 %