Denoising Gated Boltzmann Machines

Charlie Tang

Centre for Theoretical Neuroscience
University of Waterloo

August 7, 2010

Introduction

- Visual recognition in the real world is noisy (occlusions, shadows, etc...)
- For discriminative classifiers, we can only improve recognition by adding examples to the training set or play with model architecture
- RBM and DBM are generative models, we should use them to denoise *before* recognition

MNIST Recognition With Noise

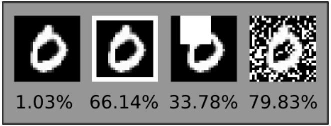

Figure: Standard DBN is not robust when facing noise during recognition.

(Tang & Eliasmith, ICML '10)

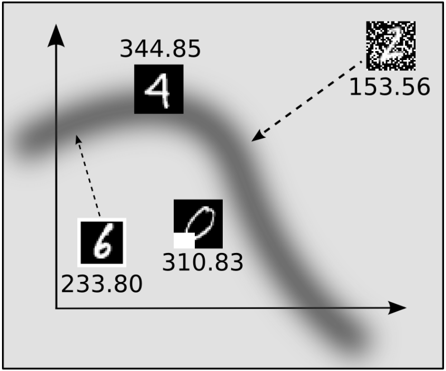

Figure: Network with attentional feedback modulation.

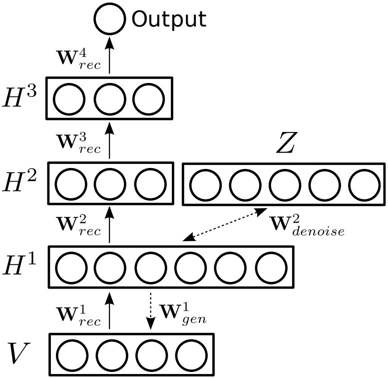

Figure: Network with attentional feedback modulation.

Results

Table: Summary of recognition results

Network	clean	border	block	random
After denoising	**1.24%**	**1.29%**	19.09%	**3.83%**
Trained w/ noise	1.61%	1.77%	**8.39%**	6.64%

- Denoising algorithm is computationally expensive, complicated
- Top-down and bottom-up inputs are combined in an rather ad hoc way

A Better Denoising BM

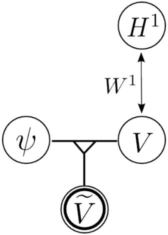

Figure: "Denoising gated Boltzmann machine".

- Adding $\tilde{v}$ layer allows a buffer between noisy input ($\tilde{v}$) and clean images (v)
- Use ψ to explain the noise and occlusion
- Top down modulation of ψ can help denoising
- Denoising is easy by running Gibbs sampling on conditional $p(\mathbf{v}|\tilde{\mathbf{v}})$

DGBM Energy

$$E(\mathbf{v}, \tilde{\mathbf{v}}, \psi, \mathbf{h}^1) = -\mathbf{b}^\top \mathbf{v} - \mathbf{c}^\top \mathbf{h}^1 - \mathbf{v}^\top W^1 \mathbf{h}^1 - \mathbf{a}^\top \boldsymbol{\psi} + \sum_i \gamma_i \psi_i \log(1 + \eta_i (v_i - \tilde{v}_i)^2))$$

Noise likelihood:

- let $d_i = |v_i - \tilde{v}_i|$
- $p(d_i | \psi_i = 1) \propto (1 + \eta_i d_i^2)^{-\gamma_i}$
- $p(d_i | \psi_i = 0) \propto constant$

Learning:

- max $\frac{1}{N} \log p(\mathbf{v}, \tilde{\mathbf{v}}, \psi)$ using Persistent CD
- +ve phase samples $p(\mathbf{h} | \mathbf{v})$
- -ve phase samples $p(\mathbf{h} | \mathbf{v})$ and $p(\mathbf{v}, \tilde{\mathbf{v}}, \psi | \mathbf{h})$

DGBM Learning