Introduction
- Deep Belief Nets are excellent models of high dimensional visual data
- However, they are not robust to simple noise not in the training set
- We introduce two ways to boost recognition:
  1. Sparsely connected first layer
  2. Probabilistic denoising on $h^2$

MNIST test error for standard DBN
[784 500 500 2000 10]

1.03% 66.14% 33.78% 79.83%

Denoising
- Hypothetical state space with the noisy images and their $\log p(h^2)$
- $Z$ added

Denoising Details
- $b_j = \phi(v)$
- $\text{bottom up input, computed using recognition weights}$
- $\text{gradient ascent on unnormalized log p}$
- $h^{b}_j$ is the set of all nodes in $h^1$ except $h_j$
- $h^{b}_j$ is $h^1$ with the $j$th node replaced by $p(h_j) = 1 - 1b_j$
- $p(h_j) = \frac{\exp(\sum h^b_j W_{bh} + b_j)}{Z(h^b)}$
  - where $\phi = (W^+_b)^t b_j + b$
  - determine which hidden nodes to unclamp
- $p(s_j | h_j) = \sigma(\sum W_{sh} h_j + b_j)$
- $h_{j+1} = \sigma(\sum W_{sh} h_j + b_j)$
  - block Gibbs sampling
- $h_{j+1} = \phi$
  - update activation for next timestep

Combining With Bottom Up
- Attenuate noisy parts of $V$, with attention-like gating
- Neurophysiological evidence for attentional modulation early in visual processing pathway

$sRBm$ Receptive Fields 7x7

Table 1. Sparse RBM and sparse DBN evaluations

<table>
<thead>
<tr>
<th>RF size</th>
<th>hidden</th>
<th>log probability</th>
<th>sDBN error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$7x7$</td>
<td>950</td>
<td>-94.62</td>
<td>1.19%</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-92.50</td>
<td>1.30%</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>-91.77</td>
<td>1.69%</td>
</tr>
<tr>
<td>$10x10$</td>
<td>950</td>
<td>-92.53</td>
<td>1.17%</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-90.36</td>
<td>2.14%</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>-89.78</td>
<td>1.55%</td>
</tr>
<tr>
<td>$15x15$</td>
<td>950</td>
<td>-90.30</td>
<td>1.68%</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-89.72</td>
<td>1.63%</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>-89.56</td>
<td>1.63%</td>
</tr>
</tbody>
</table>

Conclusions
- Sparse connections and denoising improves recognition
- Knowledge of the occluder greatly facilitates denoising
- In many cases, less error with denoising than after supervised training including noisy images
- Tradeoff between speed and accuracy (feedforward only or with feedback)