Bottom-Up and Top-Down Reasoning with Hierarchical Rectified Gaussians

Peiyun Hu, UC Irvine Deva Ramanan, Carnegie Mellon University

Inspiration from human vision



We explore efficient bidirectional networks that combine bottom-up and top-down feedback

Activations with feedback

Feedforward activations (layer 1)



~ 1ms

Average

Activations with feedback

Feedforward activations (layer 1)



~ 1ms

Average

Feedforward + feedback activations (layer 1)



~ 40ms

Average

Feedback appears to add knowledge about the "hair"



"Fully-convolutional" VGG Long et al, 15









So how do we add feedback to deep models?

CNNs



Past work on CNNs + feedback: Pinherio & Collobert, 14 Cao et al, 15 Gatta et al, 14

So how do we add feedback to deep models?

CNNs



Boltzmann Machines





Insight: Use CNNs to learn Hierarchical Rectified Gaussians

Unroll MAP updates on Rectified Gaussian models into a rectified neural net



Similar architectures:

Autoencoders, DeConvNets, U-Nets, Hourglass Nets, Ladder Networks Past work on unrolling models: Chen et al, 15 Zheng et al, 15 Goodfellow et al, 13

Empirical results: Caltech Occluded Faces



Localization error on occluded points: Bottom-up: 21.3 % Top-down: **15.3** %

Improvement comes "for free" (no increase in parameters)

Take-home messages

- Add top-down feedback into CNNs "for free"
- Unfold inference on rectified probabilistic models into rectified neural nets
- Competitive accuracy on keypoint localization



Caltech Occluded Faces

