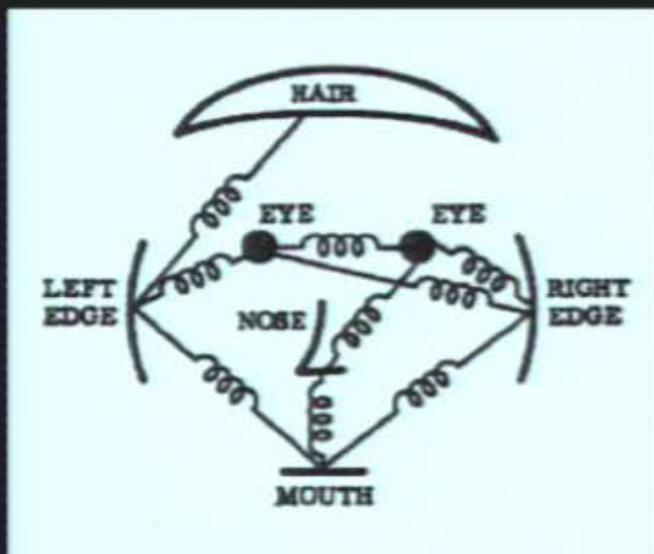


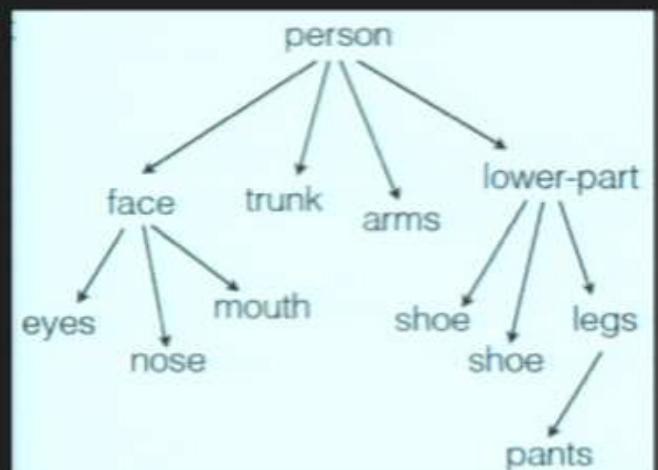
# Deep Learning + Structured Prediction

- ConvNet feature extractor
- Combine with top-down reasoning



[Fischler and R. Elschlager 1973 ]

## Stochastic Grammars



[R. Girshick, P. Felzenszwalb, D. McAllester, Object Detection with Grammar Models, NIPS 2011]

# Summary

---

- Yann LeCun was right!
- Deep ConvNets work well for recognition
  - Quite a bit better than existing vision approaches
- Can be used for many other vision tasks
- But unsupervised learning still an open problem

# DEMO

URL: <http://horatio.cs.nyu.edu>

## Image Classifier Demo



Upload your images to have them classified by a machine! Upload multiple images using the button below or dropping them on this page. The predicted objects will be refreshed automatically or you can manually refresh yourself. Images are resized such that the smallest dimension becomes 256, then the center 256x256 crop is used. More about the demo can be found out [here](#).

[+ Upload Images](#)

[✖ Remove All](#)

### Demo Notes

- The maximum file size for uploads in this demo is 10 MB.
- Only image files (JPEG, JPG, GIF, PNG) are allowed in this demo .
- You can **drag & drop** files from your desktop on this webpage with Google Chrome, Mozilla Firefox and Apple Safari.
- All images from your current session and IP are shown above.
- This demo is powered by research out of New York University. [Click here to find out more](#)

Demo created by: [Matthew Zeiler](#)



NEW YORK UNIVERSITY

# Acknowledgements

- Colleagues in CILVR lab @ NYU:



Matt Zeiler



David Eigen



Pierre Sermanet



Yann LeCun

- Slides: Marc'Aurelio Ranzato
- Funding: NSF, DARPA, ONR  
Microsoft, Google, Facebook

# ICLR 2014 Conference

---

- Deadline: 20<sup>th</sup> December 2013
- Banff, Canada, April 14-16<sup>th</sup> 2014
- Welcome anything to do with representation learning!



International Conference on  
Learning Representations 2014

# Summary

---

- Yann LeCun was right!
- Deep ConvNets work well for recognition
  - Quite a bit better than existing vision approaches
- Can be used for many other vision tasks
- But unsupervised learning still an open problem

# Removing Artifacts

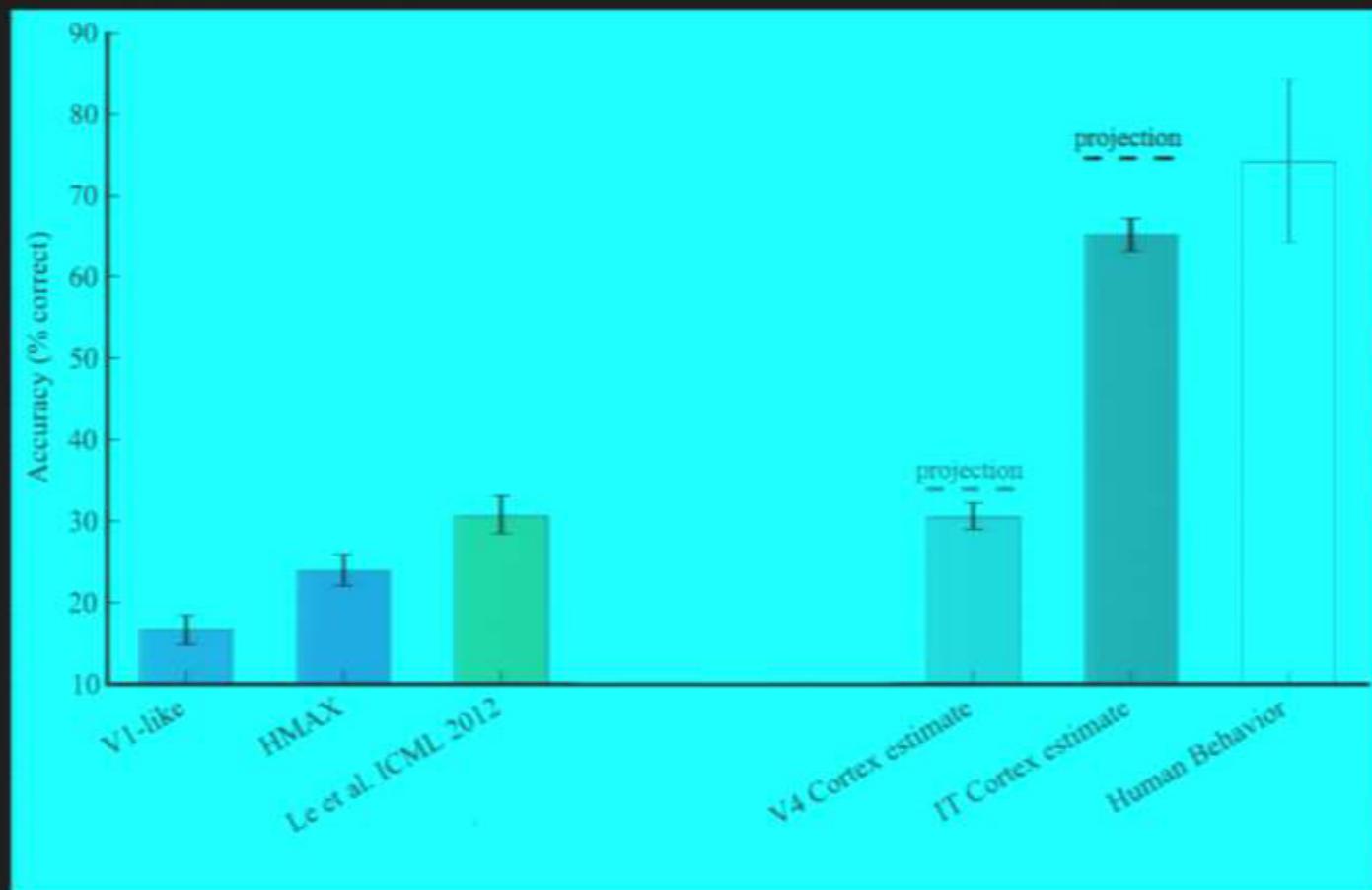
[Eigen et al. "Restoring an Image Taken Through a Window Covered with Dirt or Rain" ICCV 2013]



# **Other Vision Applications**

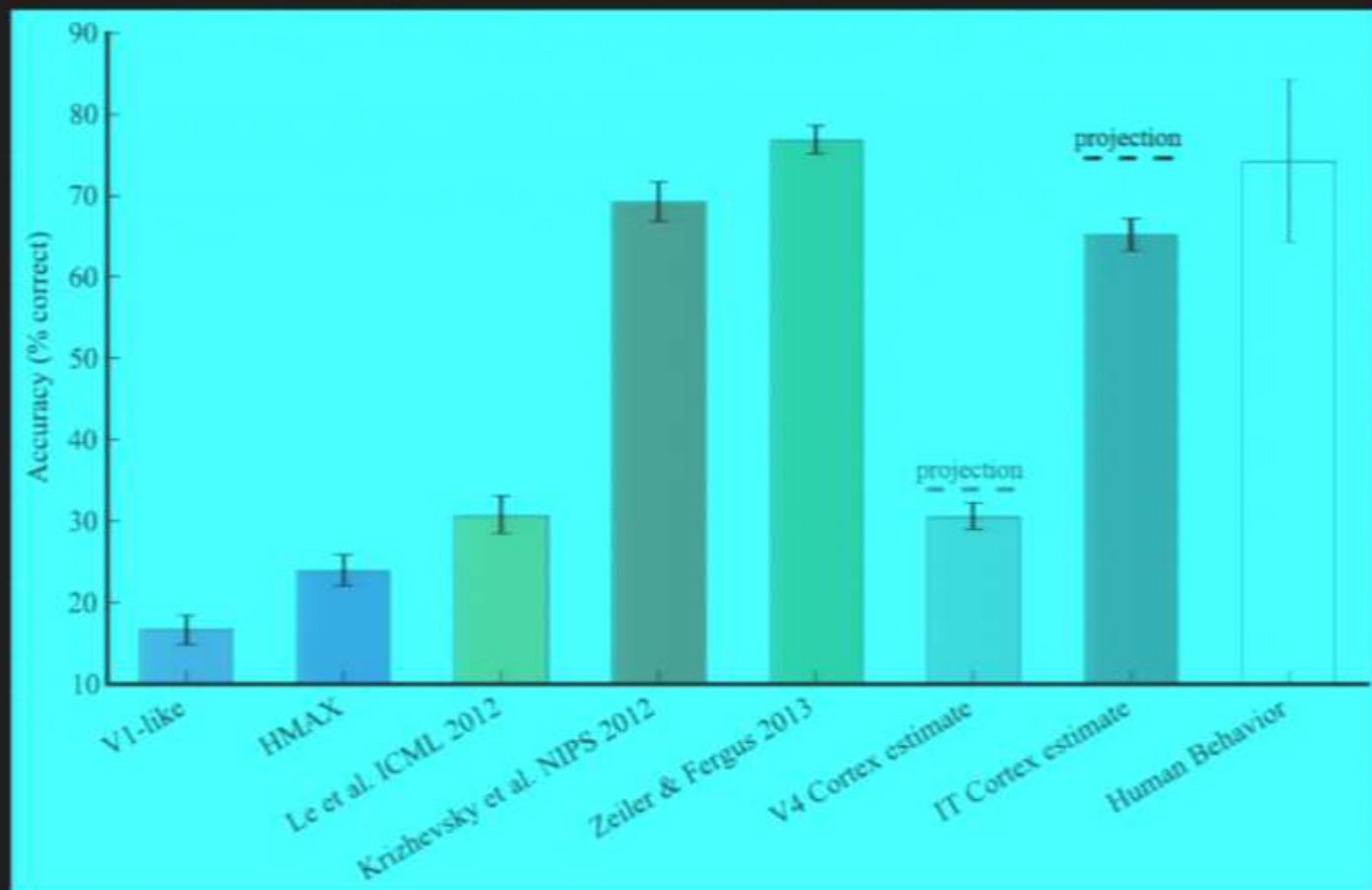
# Deep Nets vs Monkey vs Humans [Cadieu et al.]

- Rapid presentation experiments (100ms)
- Feed-forward processing only in monkey/humans



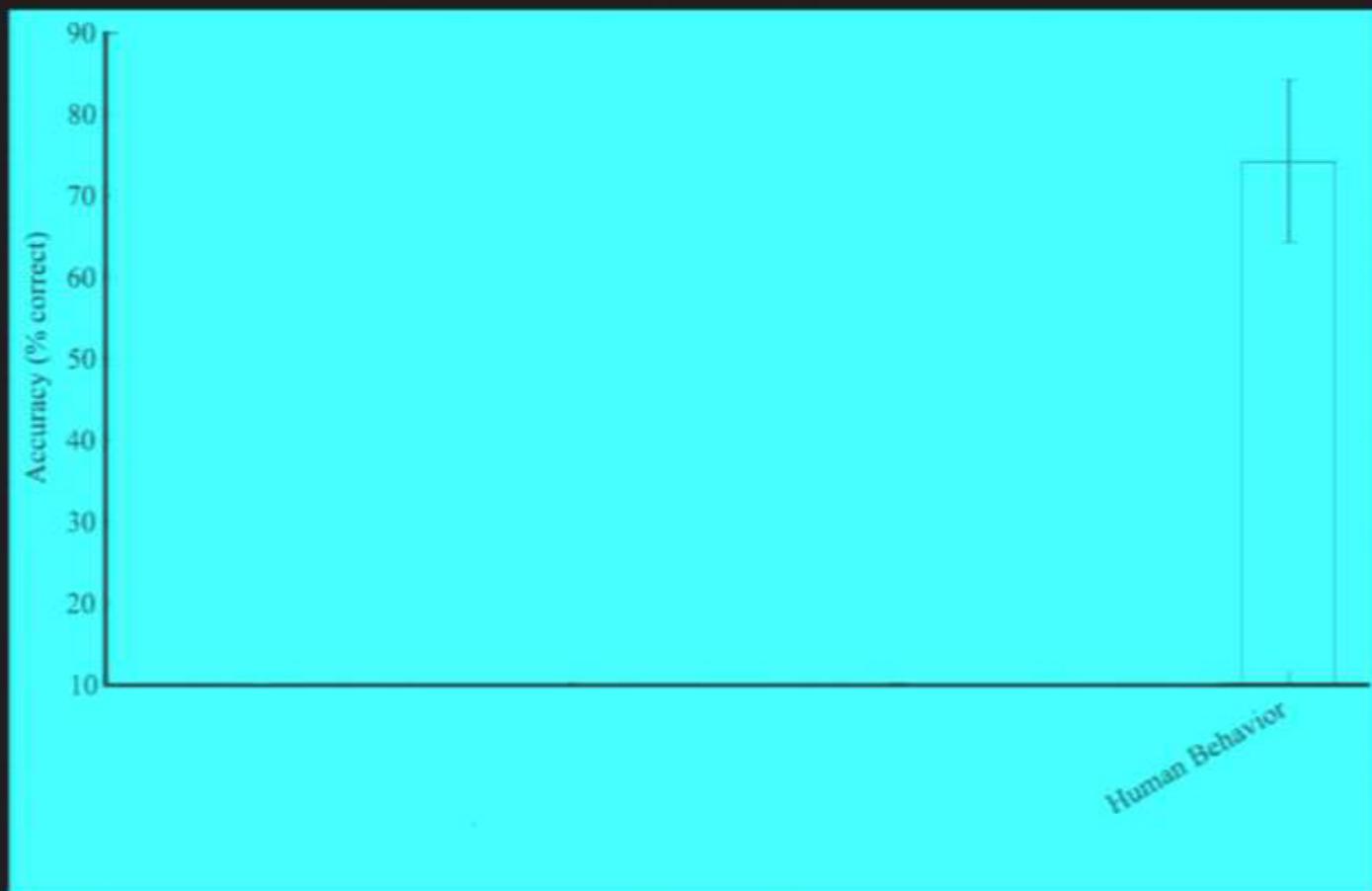
# Deep Nets vs Monkey vs Humans [Cadieu et al.]

- Rapid presentation experiments (100ms)
- Feed-forward processing only in monkey/humans



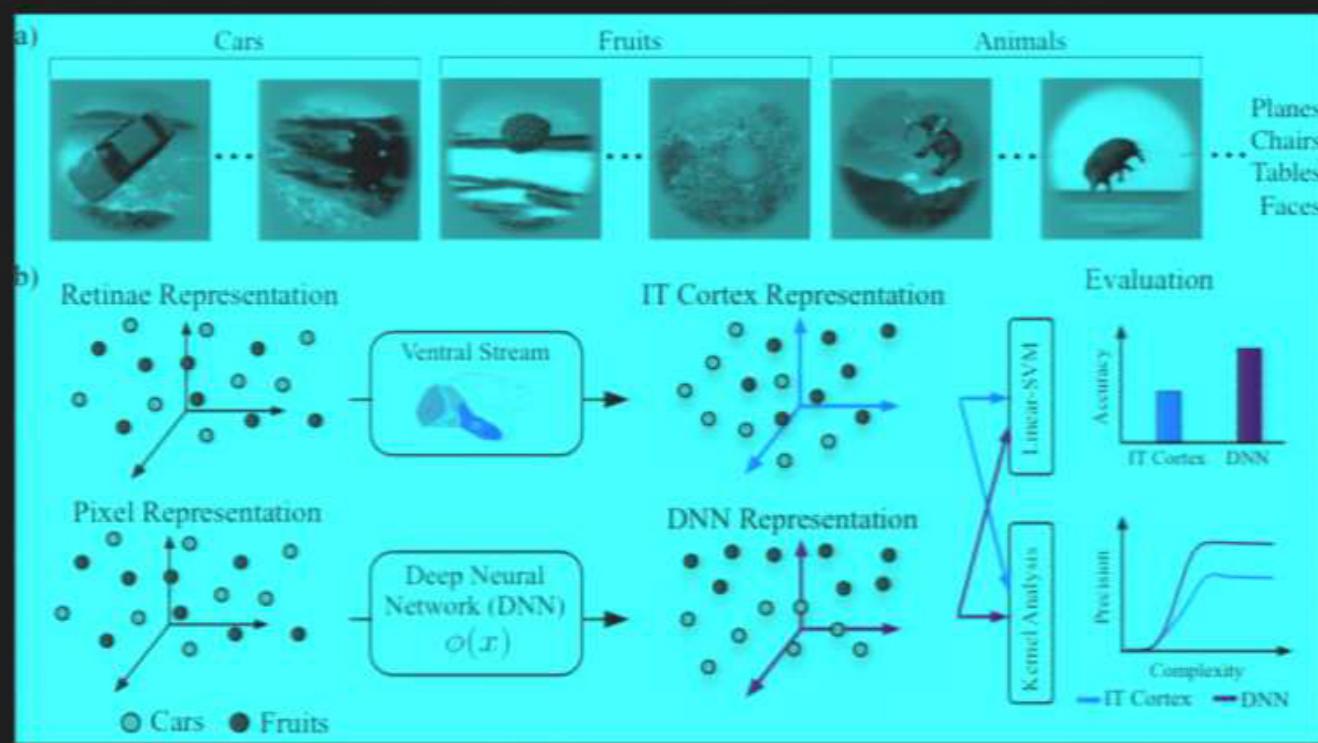
# Deep Nets vs Monkey vs Humans [Cadieu et al.]

- Rapid presentation experiments (100ms)
- Feed-forward processing only in monkey/humans



# Deep Nets vs Monkey vs Human

C.F. Cadieu, H. Hong, D. Yamins, N. Pinto, E.A. Solomon, N.J. Majaj, and J.J. DiCarlo. *Deep Neural Networks Rival the Object Recognition Performance of the Primate Visual System.* (PLOS One Biology, in submission, 2013).



# Deep Learning

# Convnet Successes

- Handwritten text/digits
  - MNIST (0.17% error [Ciresan et al. 2011])
  - Arabic & Chinese [Ciresan et al. 2012]
- Simpler recognition benchmarks
  - CIFAR-10 (9.3% error [Wan et al. 2013])
  - Traffic sign recognition
    - 0.56% error vs 1.16% for humans [Ciresan et al. 2011]





# Deep Learning for Computer Vision

*NIPS 2013 Tutorial*

Rob Fergus

Dept. of Computer Science  
New York University

# Unpooling Operation

