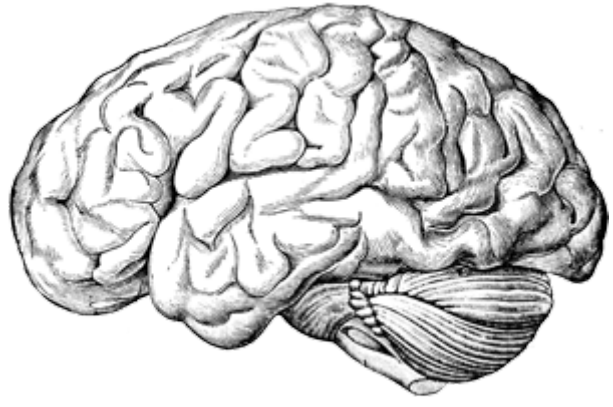
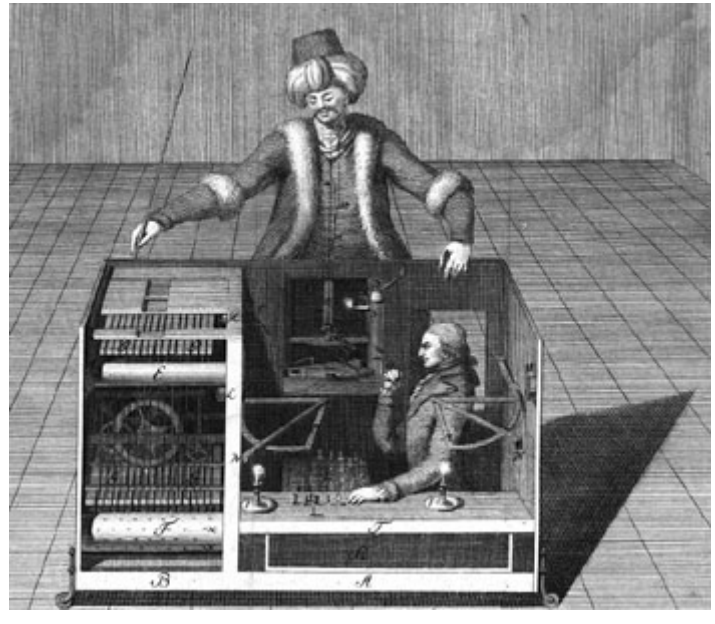


# Using Human Brain Activity to Guide Machine Learning

Paper by Ruth C. Fong, Walter J. Scheirer, David D. Cox

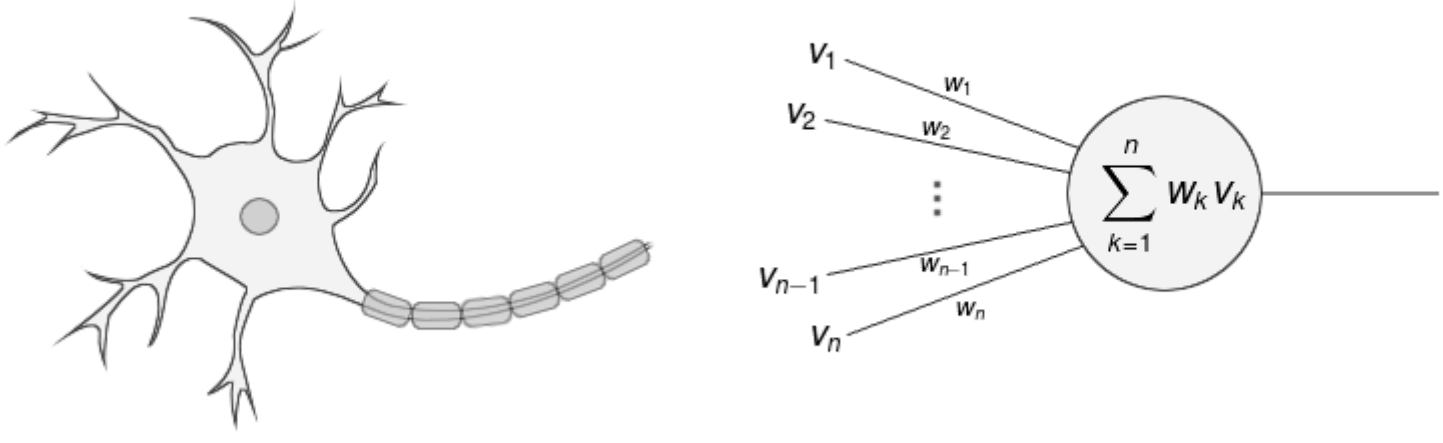
Presented by Chris Foster

What is the relationship between
machine learning and the brain?


B
A

ML is used for understanding the brain:

- Brain imagery decoding
- Neural response prediction
- Hierarchical brain modeling




Neural networks are loosely inspired by the brain's neurons...

...but there's still a way to go

- Adversarial failures
- Different error patterns
- Different performance patterns

# The Direct Approach

*Instead of simply conceptually modeling the brain, can we use brain data directly to aid machine learning models?*

# The Dataset

- 1,386 500x500 color images
- Downsampled to 250x250
- Per-pixel object labels
- Cut to single-object categories
- Cut to 20% class minimum
- Cut to 5 categories

# The Categories

- 219 human images
- 180 animal images
- 151 building images
- 59 food images
- 37 vehicle images

Four classification problems:

*Does this image contain <class>?*

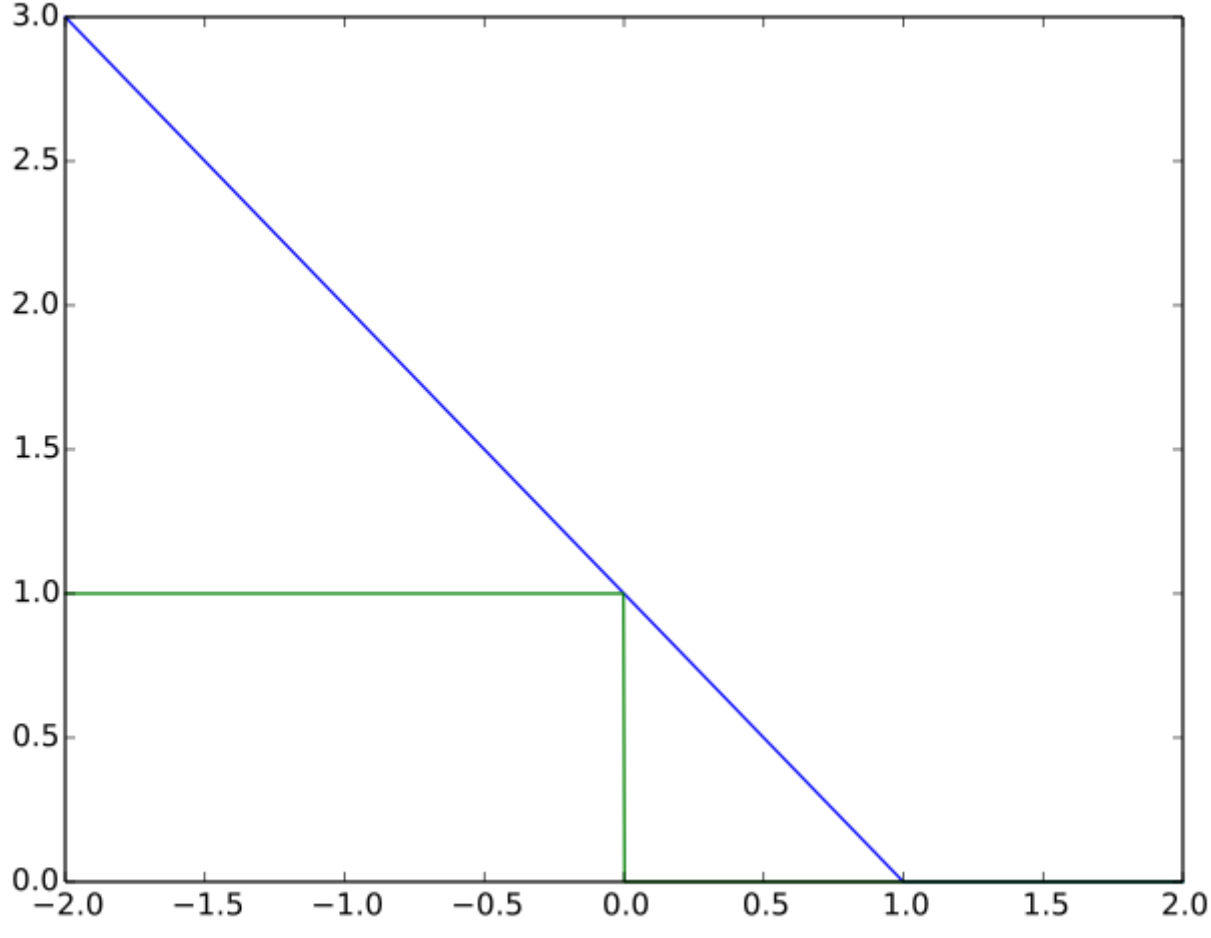
# Prediction Model

- Support Vector Machine
- Radial Basis Function
- Hinge Loss (HL)

# Hinge Loss

$$\phi_h(z) = \max(0, 1 - z)$$

# Hinge Loss



# Histogram of Oriented Gradients

Handpicked computer vision features
Designed for object detection
60-65% accuracy

Input image

Histogram of Oriented Gradients

# Convolutional Neural Network

Automatically learned features
Trained on ILSVRC 2012
AlexNet architecture
80-85% accuracy

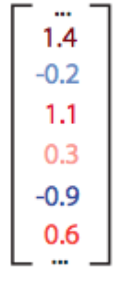
# Brain Imagery

- Subjects viewed the 1,386 images
- 3T Siemens Tim Trio MR Scanner
- Reduced to voxels in visual ROIs
- *EBA, FFA, LO, OFA, PPA, RSC, TOS*
- 2.25x2.25x2.99mm voxel size

# Stimulus Capture

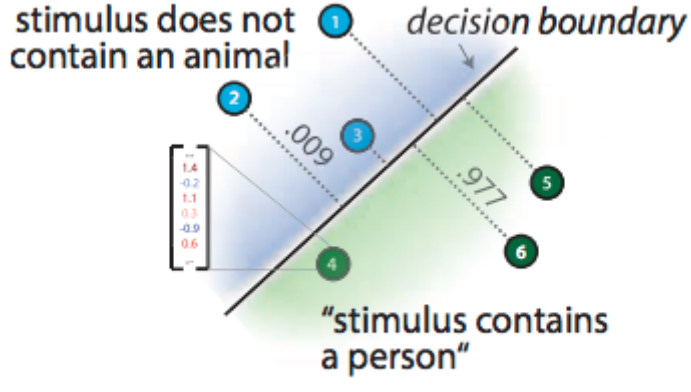
*Stimulus*

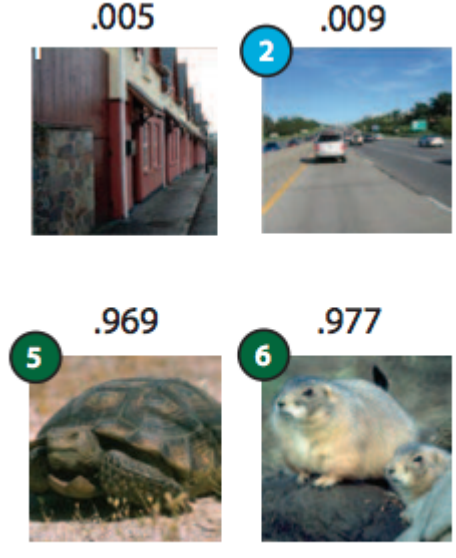
*fMRI Images*




*Activity Vector*

# SVM Classifier



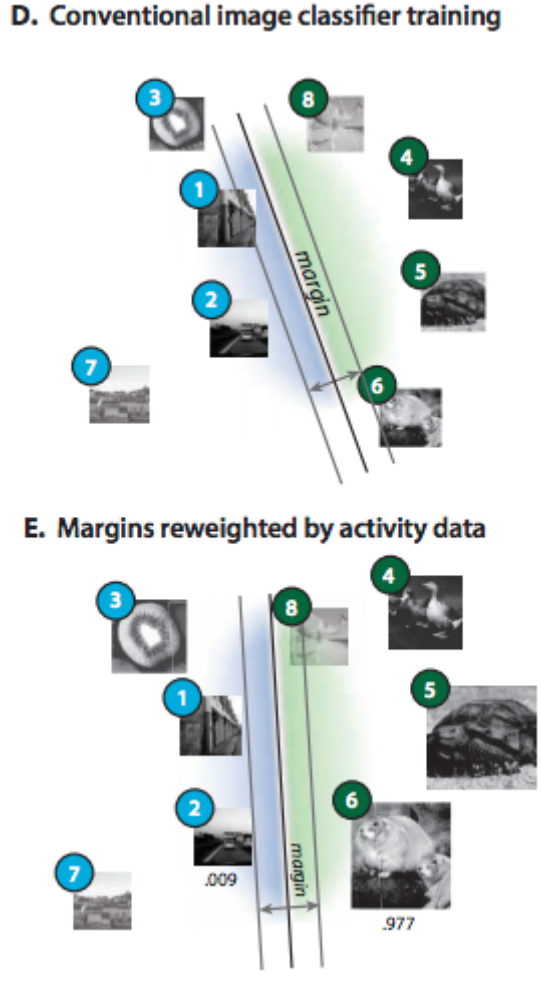



# Activity Weighted Loss

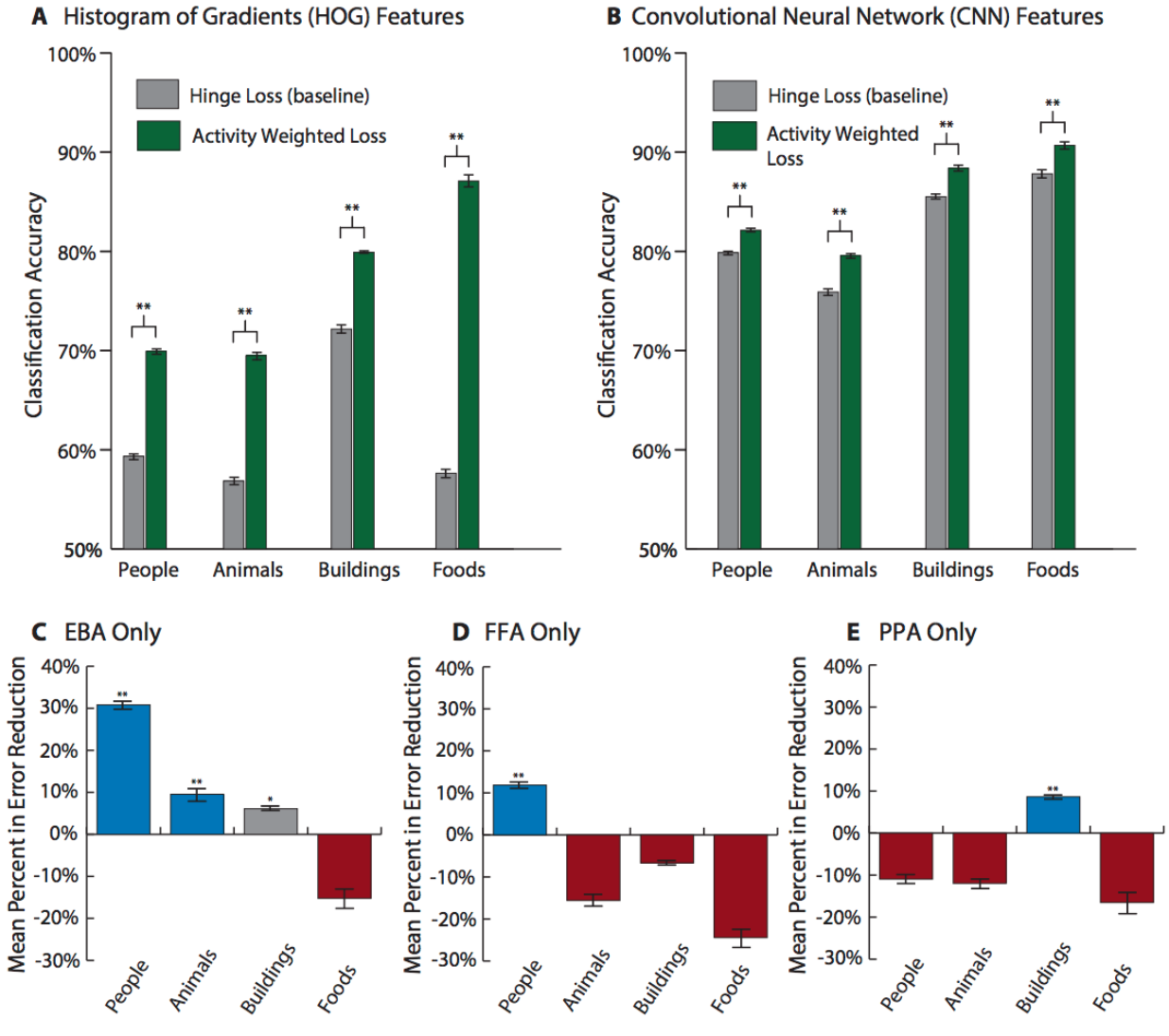
$$\phi_\psi(x, z) = \max(0, (1 - z) \cdot M(x, z))$$

$$M(x, z) = \begin{cases} 1 + c_x, & \text{if } z < 1 \\ 1, & \text{otherwise} \end{cases}$$

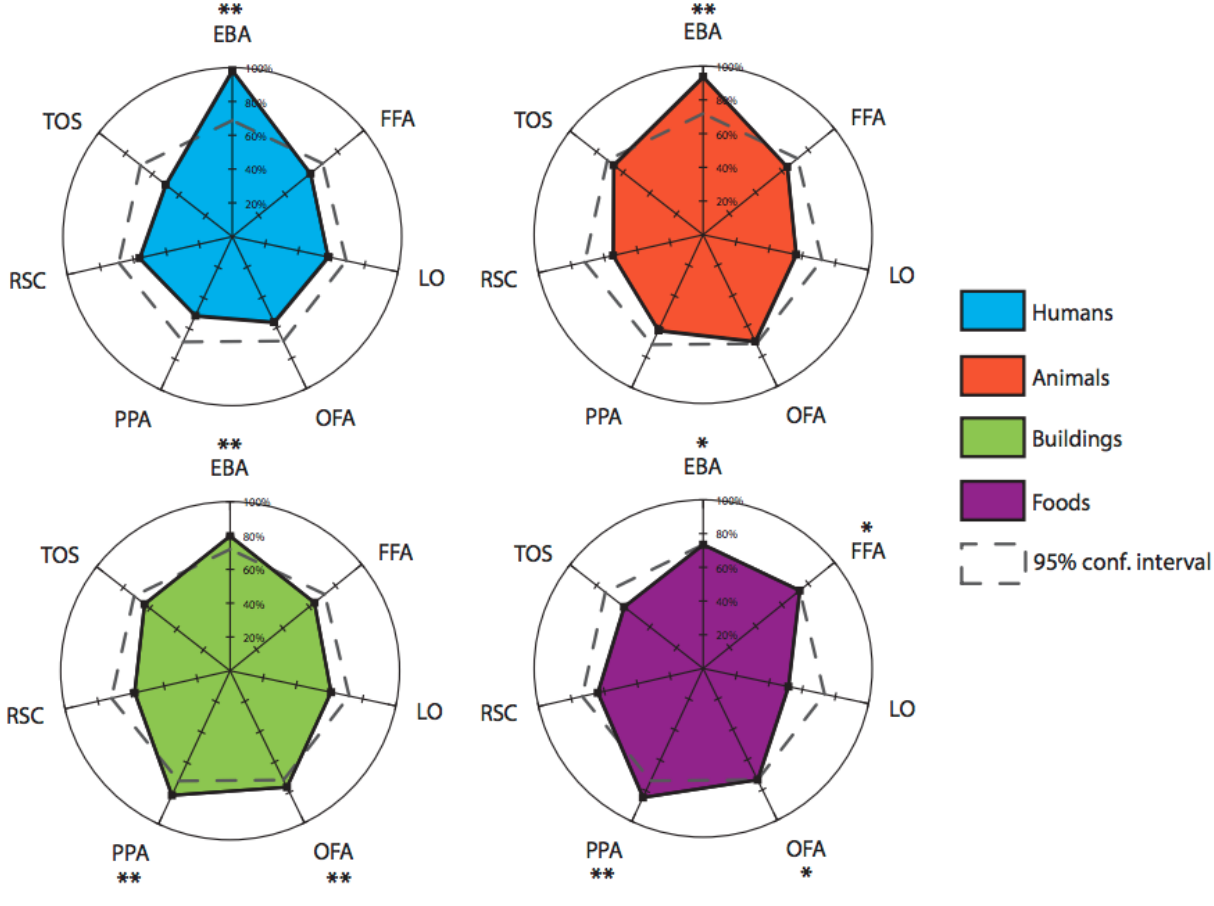
# Margins


D. Conventional image classifier training
3
8
4
1
margin
5
2
7
6
E. Margins reweighted by activity data
3
8
4
1
5
2
6
7
.009
margin
.977

# Results


A Histogram of Gradients (HOG) Features
B Convolutional Neural Network (CNN) Features
Hinge Loss (baseline)
Activity Weighted Loss
Classification Accuracy
100%
90%
80%
70%
60%
50%
People
Animals
Buildings
Foods
**
C EBA Only
D FFA Only
E PPA Only
Mean Percent in Error Reduction
40%
30%
20%
10%
0%
-10%
-20%
-30%
*

# Analysis of ROIs



# Summary

- Brain data can improve machine learning classifiers
- Data is only required at training time, not test time
- Training can still be done when brain data is missing
- This is an early application and has room to experiment

Thanks!