Time, Space and Computation:

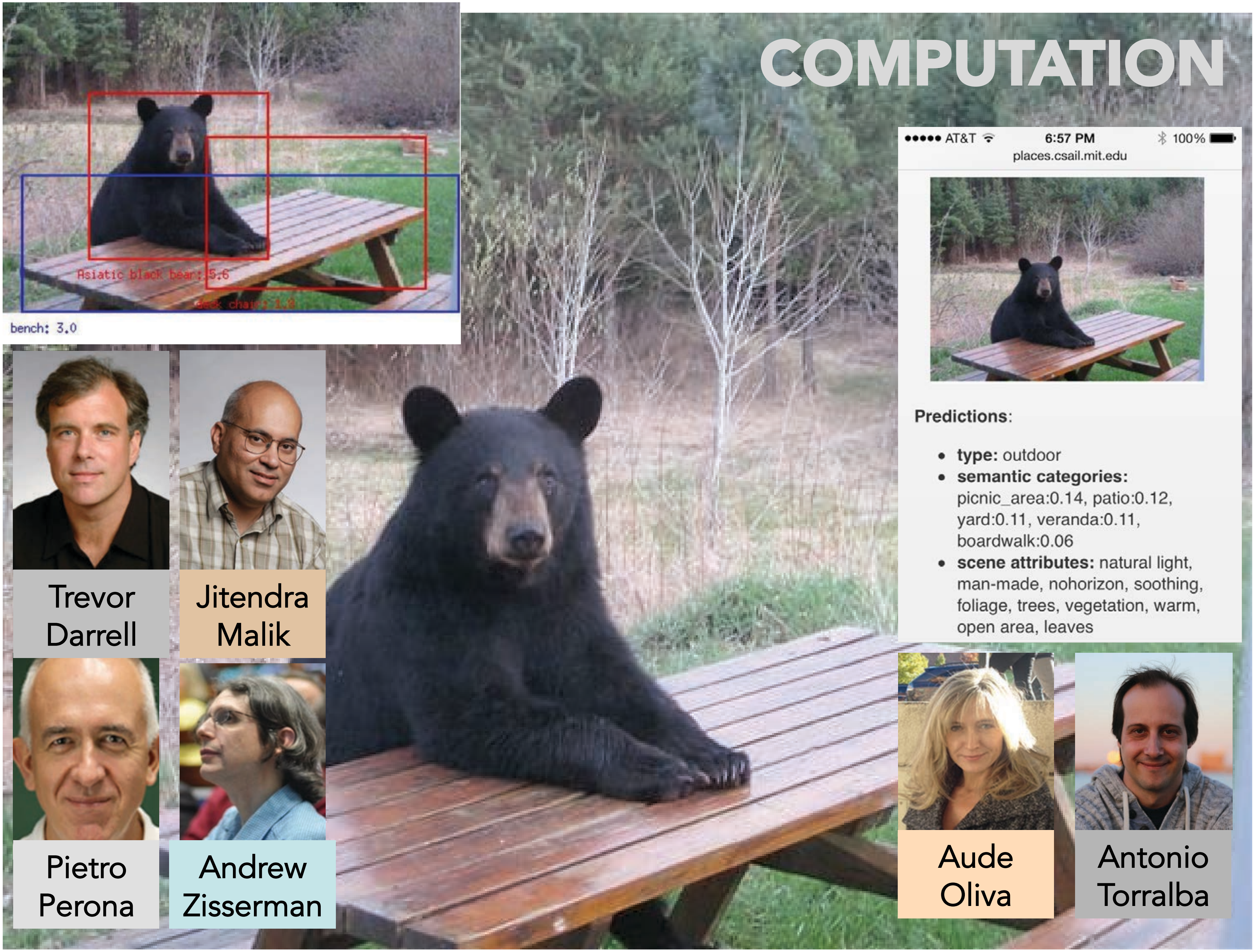
Converging Human Neuroscience

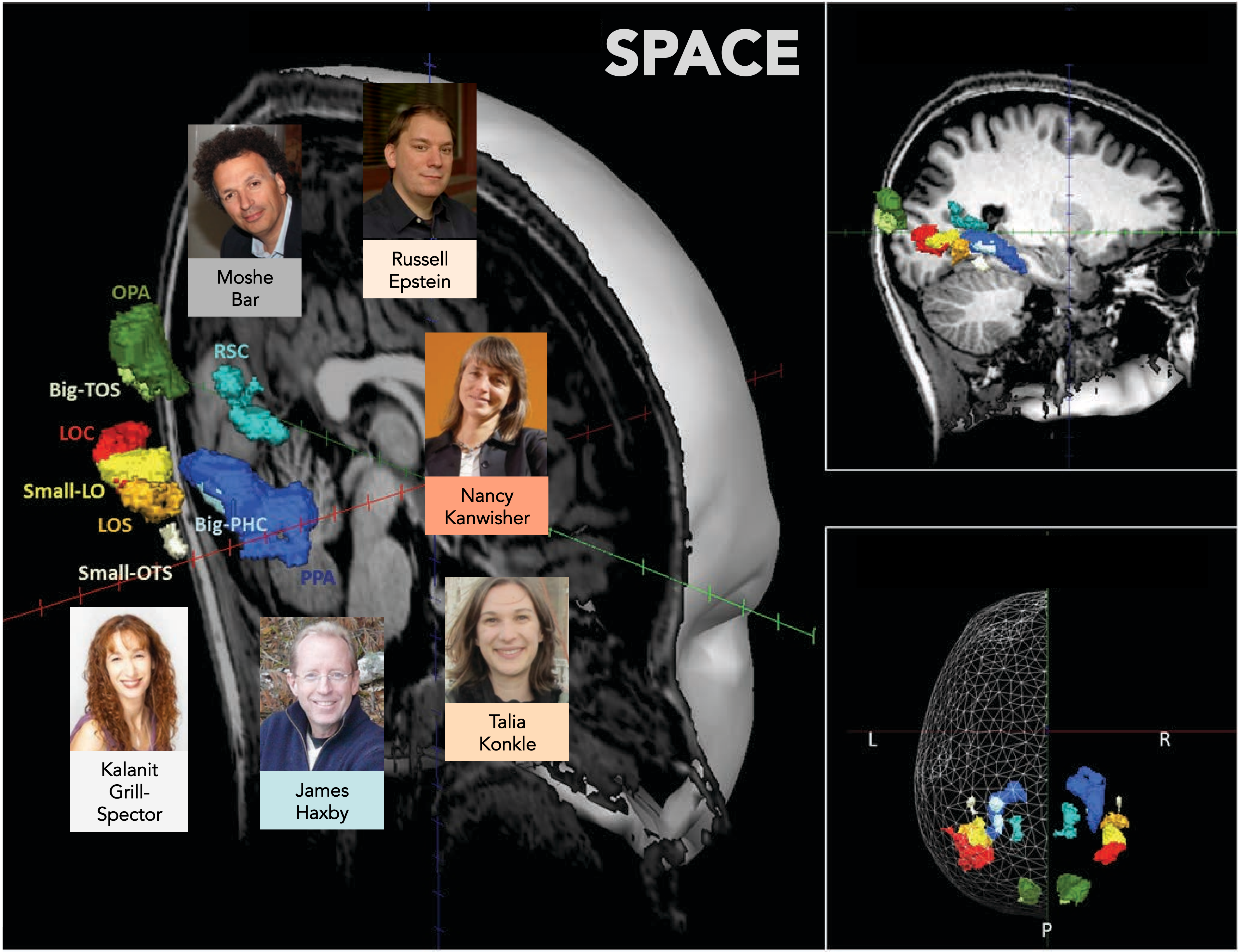
& Computer Science

Aude Oliva

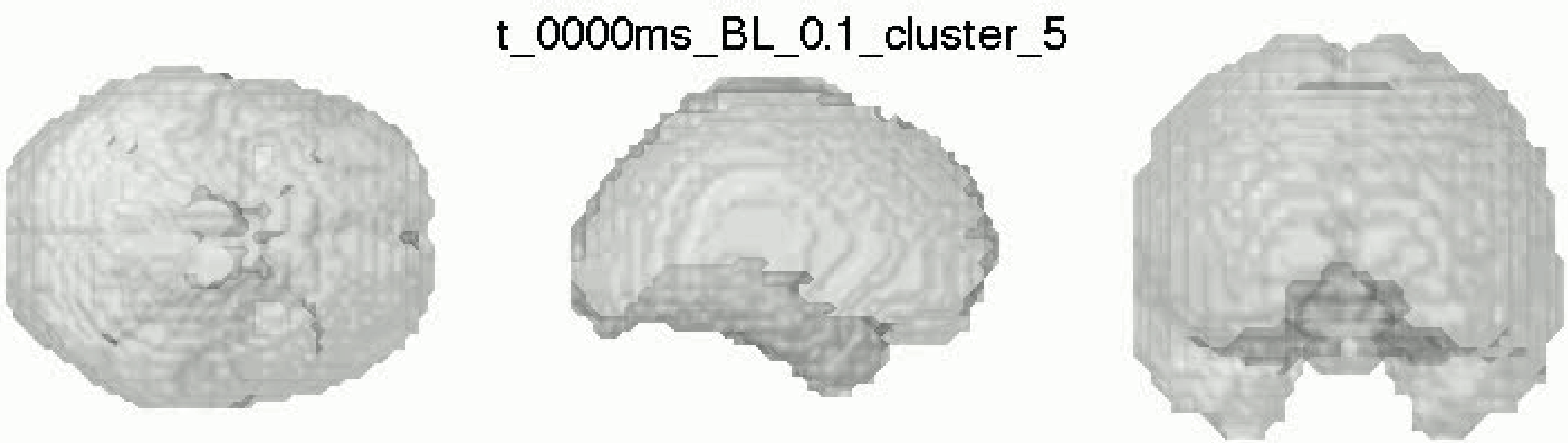
Computer Science and Artificial Intelligence Lab

Massachusetts Institute of Technology


COMPUTATION
Asiatic black bear: 5.6
bench: 3.0
AT&T
6:57 PM
100%
places.csail.mit.edu
Predictions:
• type: outdoor
• semantic categories: picnic_area:0.14, patio:0.12, yard:0.11, veranda:0.11, boardwalk:0.06
• scene attributes: natural light, man-made, nohorizon, soothing, foliage, trees, vegetation, warm, open area, leaves
Trevor Darrell
Jitendra Malik
Pietro Perona
Andrew Zisserman
Aude Oliva
Antonio Torralba


SPACE
Moshe Bar
Russell Epstein
Nancy Kanwisher
Talia Konkle
Kalanit Grill-Spector
James Haxby
OPA
RSC
Big-TOS
LOC
Small-LO
LOS
Big-PHC
Small-OTS
PPA
L
R
P

TIME



Radoslaw Cichy

Dimitrios Pantazis

Aude Oliva

Nikolaus Kriegeskorte

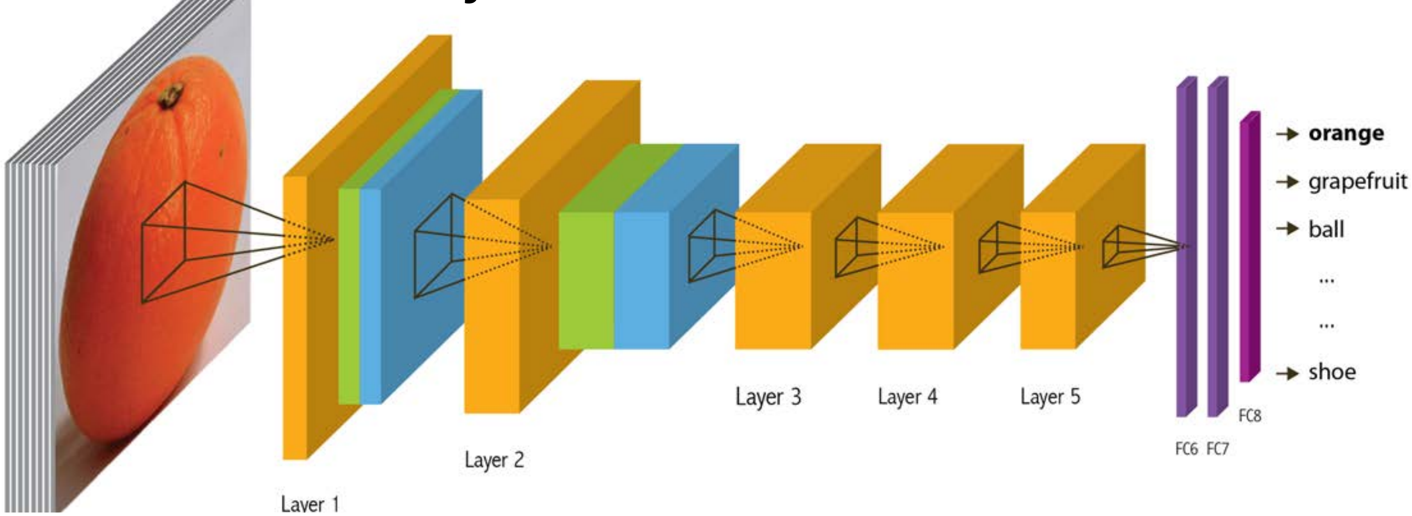
Computation with millions of instances

Deep architectures

Geoffrey Hinton, Yann LeCun

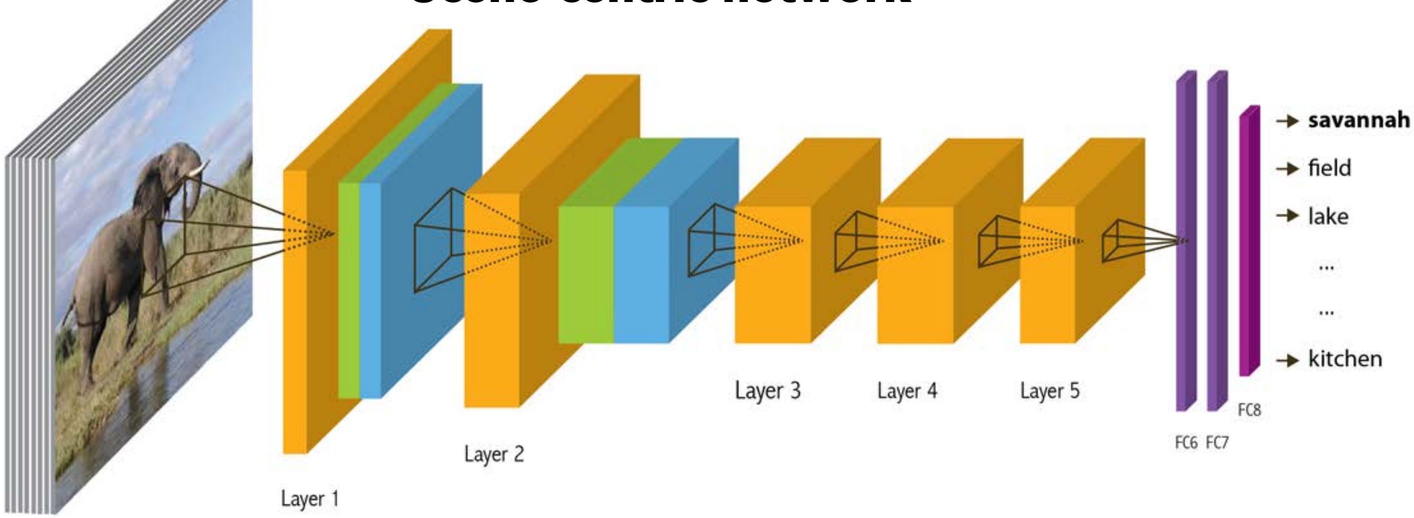
Object-centric network

IMAGENET



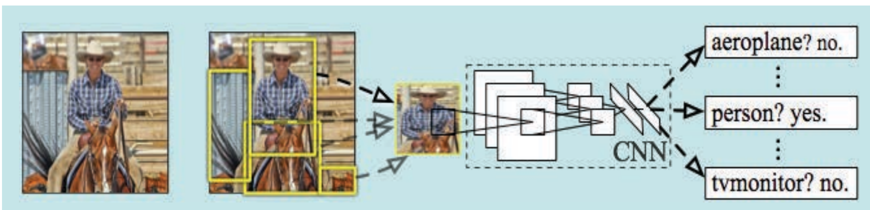
Scene-centric network

PLACES



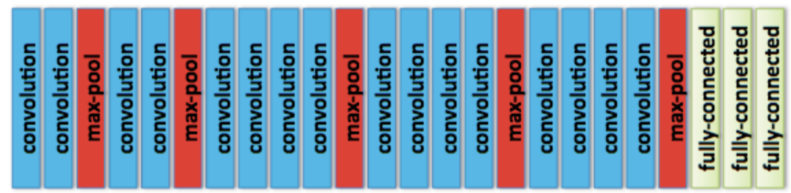
Object-centric deep architectures

R-CNN: Regions with CNN features

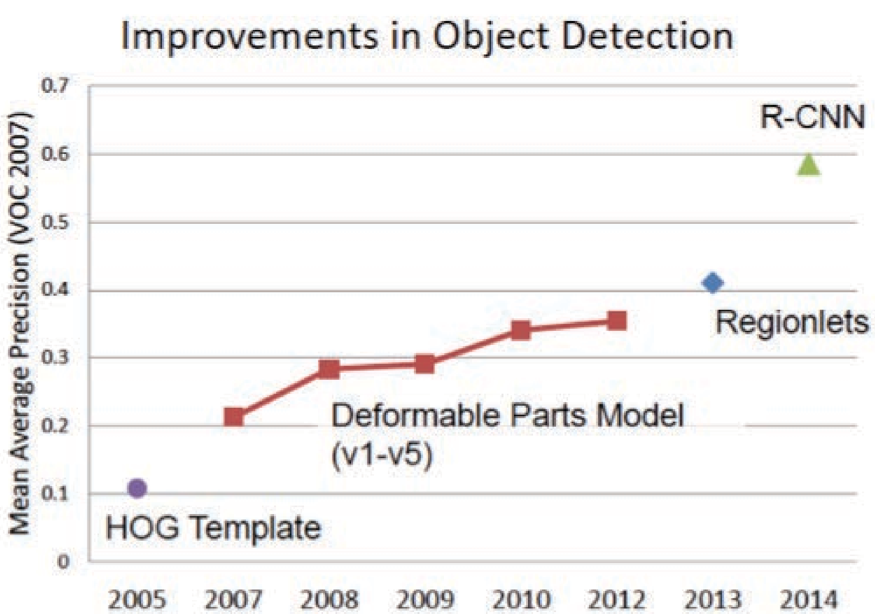


Girshick, Donahue, Darrell & Malik (CVPR 2014)

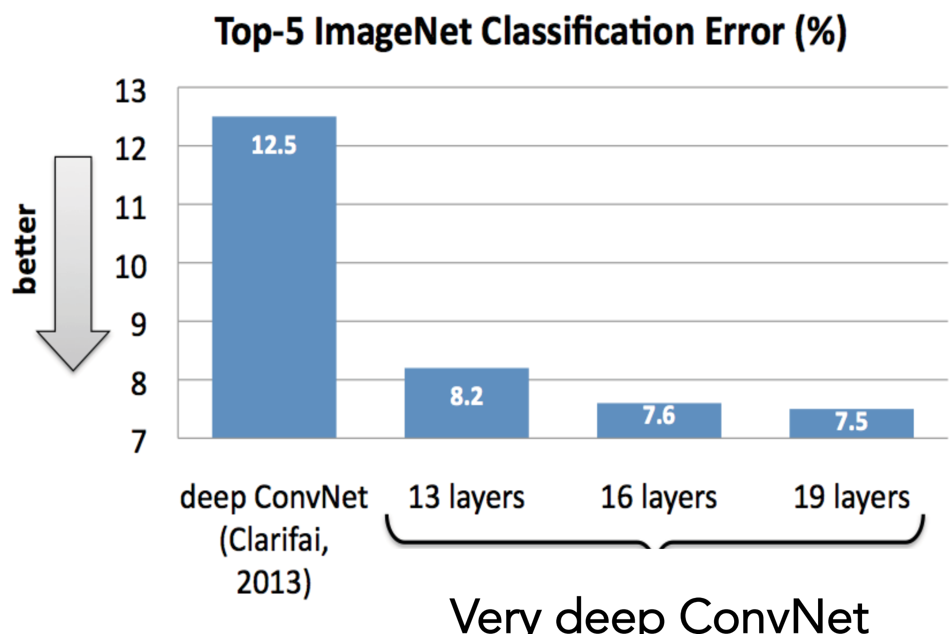
VGGNet: Very deep ConvNet



Simonyan & Zisserman (2014)



(graph by D. Hoeim)



Torralba

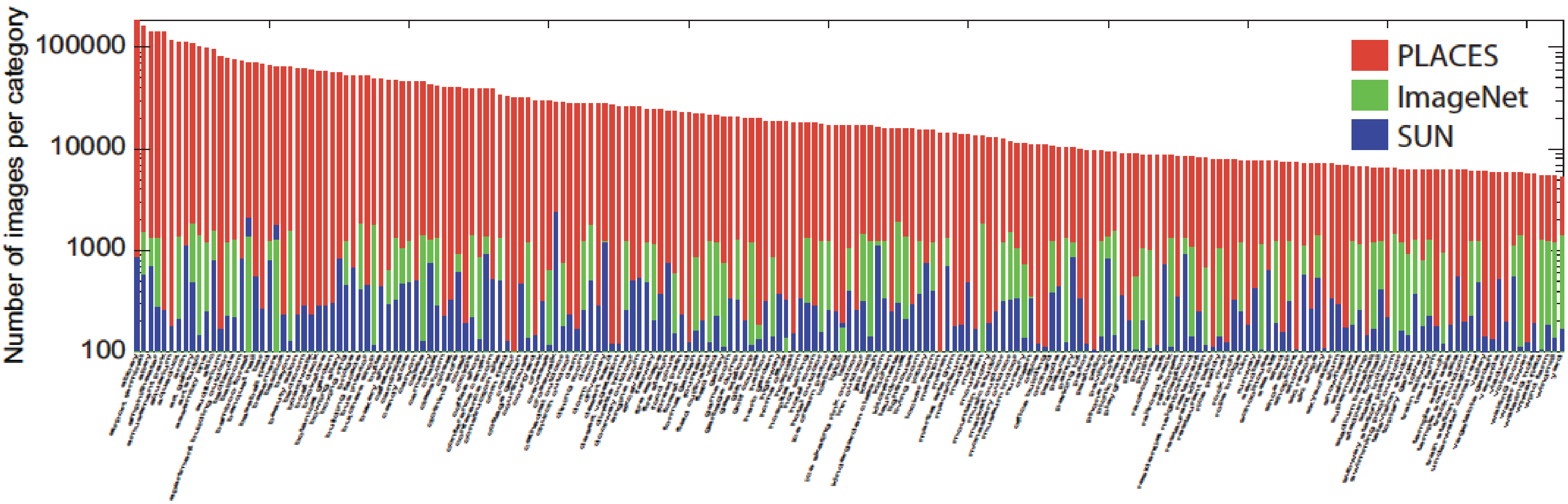
Lapedriza

Places Model

places.csail.mit.edu

Zhou

Xiao



Predictions:

- **type**: indoor
- **semantic categories**: coffee_shop:0.47, restaurant:0.17, cafeteria:0.08, food_court:0.06,

Predictions:

- **type**: indoor
- **semantic categories**: supermarket:0.96,

Predictions:

- **type**: indoor
- **semantic categories**: conference_center:0.51, auditorium:0.12, office:0.08,

Predictions:

- **type**: indoor
- **semantic categories**: bus_interior:0.91,

NIPS 2014 release: 2.5 million images, 205 scene categories

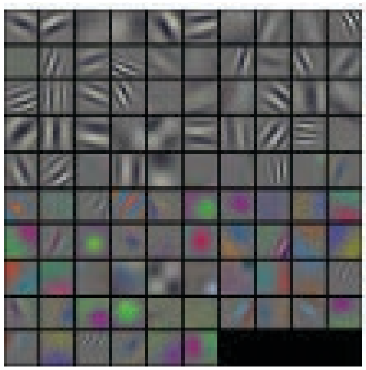
Zhou, Lapedriza, Xiao, Torralba & Oliva (2014), NIPS

Deep architectures

A Visualization of the learned representation for each unit

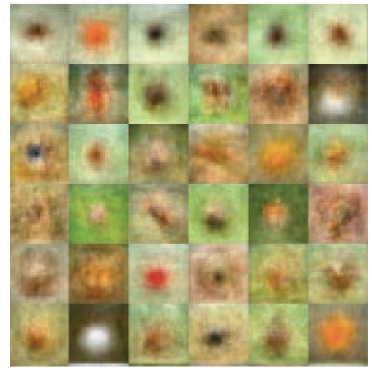
Object-centric CNN

C1 filters

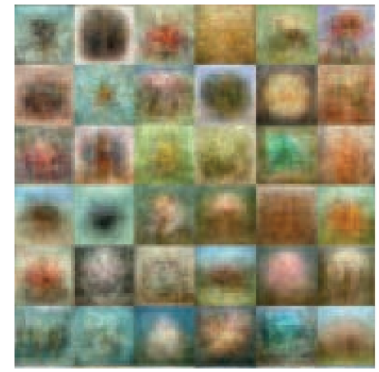


C2 feature maps

C5 feature maps



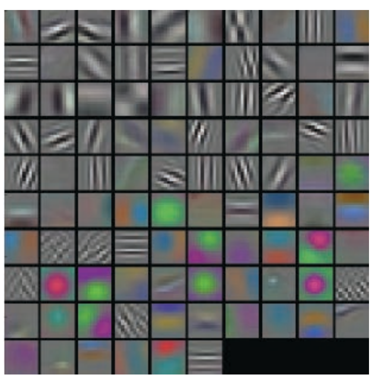
C7 feature maps



Object like shapes

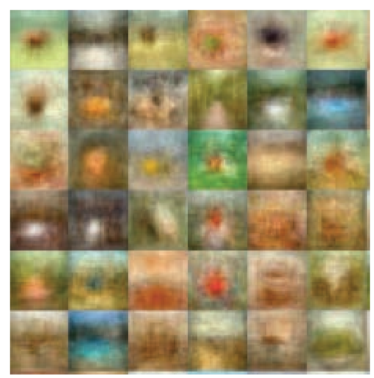
Scene-centric CNN

C1 filters

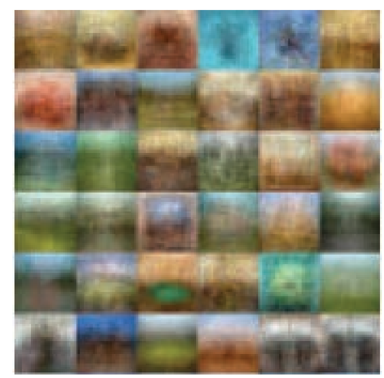


C2 feature maps

C5 feature maps



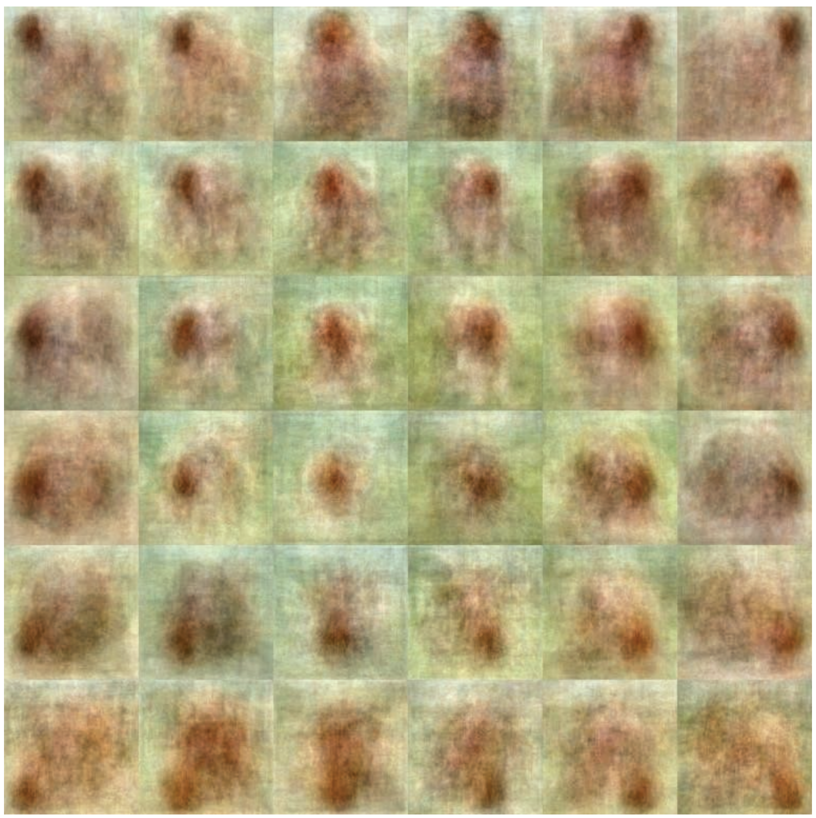
C7 feature maps

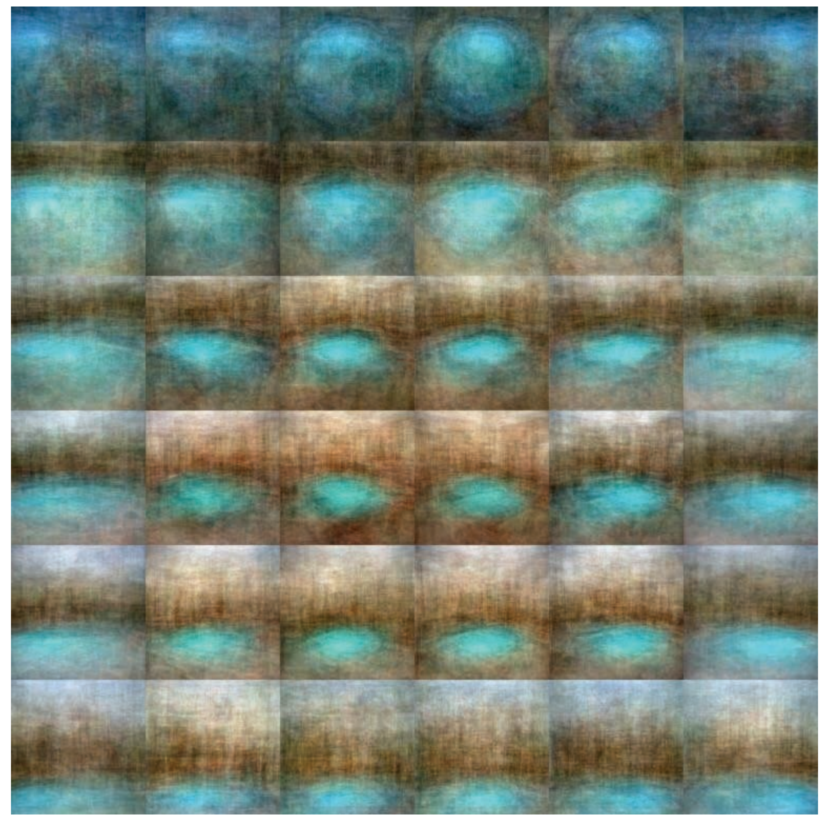


Space like shapes

Zhou, Lapedriza, Xiao, Torralba & Oliva (2014), NIPS

Layer 5: Artificial Receptive fields





Zhou, Lapedriza, Xiao, Torralba & Oliva (2014), NIPS

Dimitrios
Pantazis

Non invasive neuro-imaging techniques

Radoslaw
Cichy

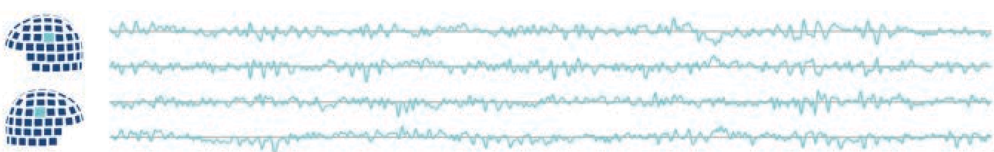
MEG
(msec-resolution)

Magneto encephalography

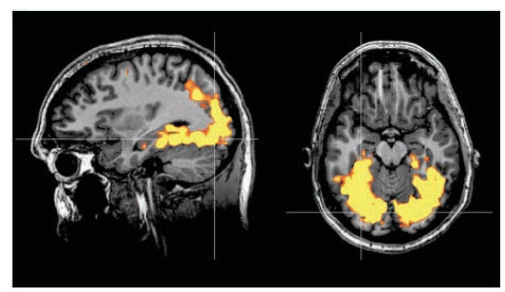
fMRI
(mm-resolution)

Functional Magnetic Resonance Imaging

?



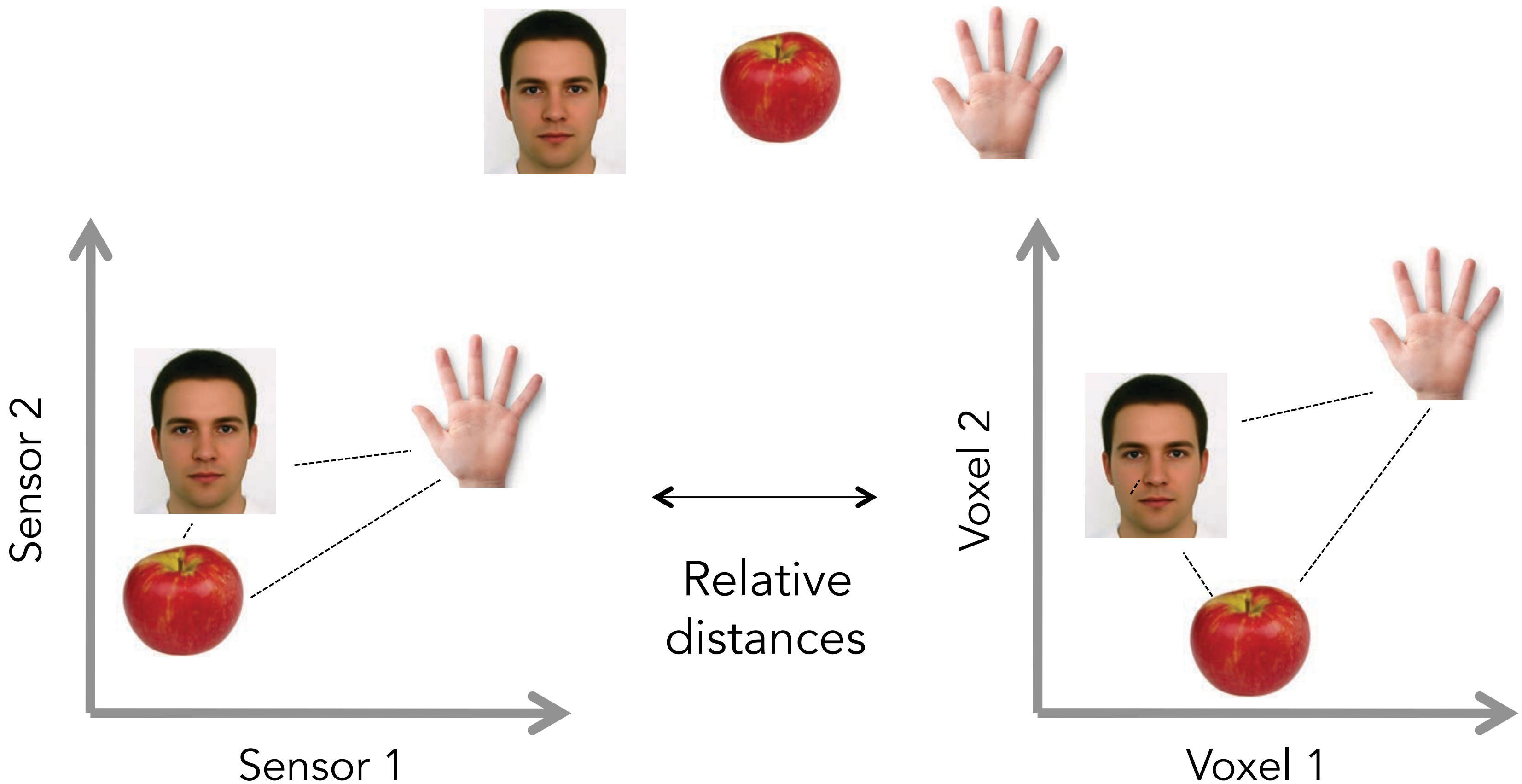
Sensors (time)



Voxels (space)

Representational Geometry

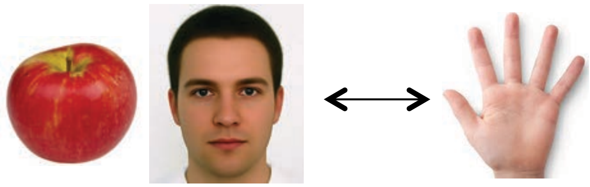
Nikolaus Kriegeskorte (2008)



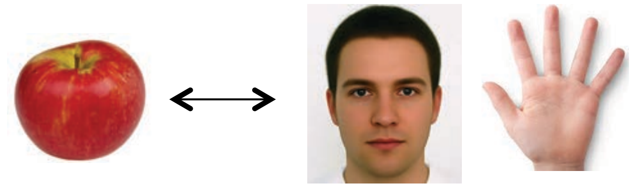
Shepard et al., 1980; Kruskal and Wish., 1978; Edelman et al. 1998; Kriegeskorte et al., 2008; Mur et al., 2009; Liu et al., 2013

Representational Geometry

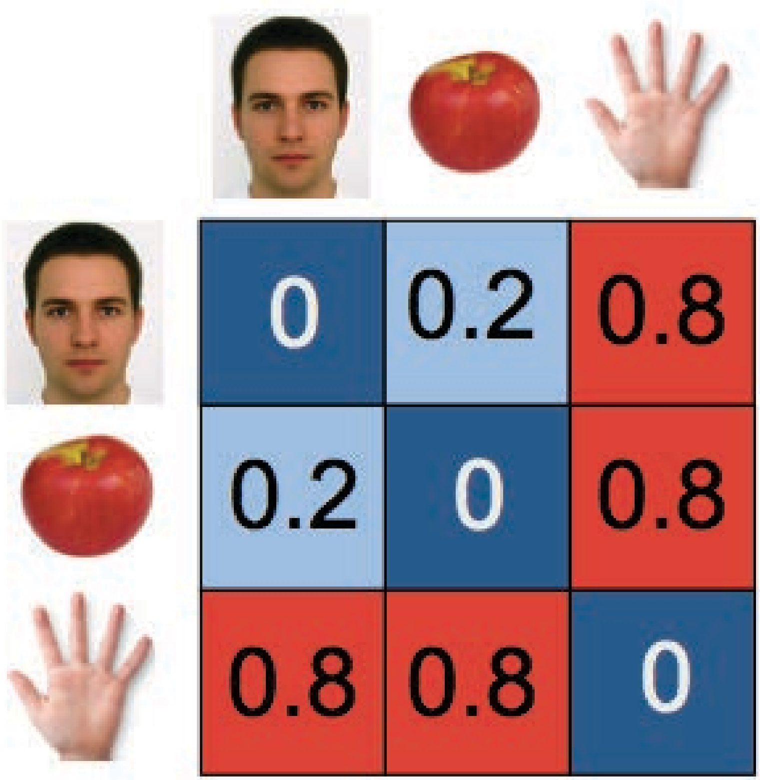
Nikolaus Kriegeskorte (2008)

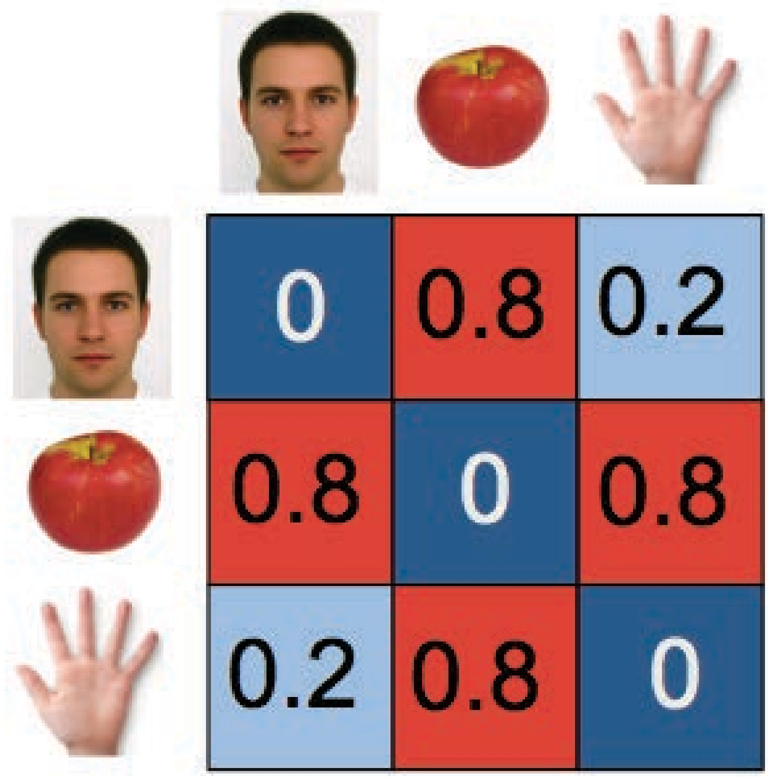


Round shape



Body part

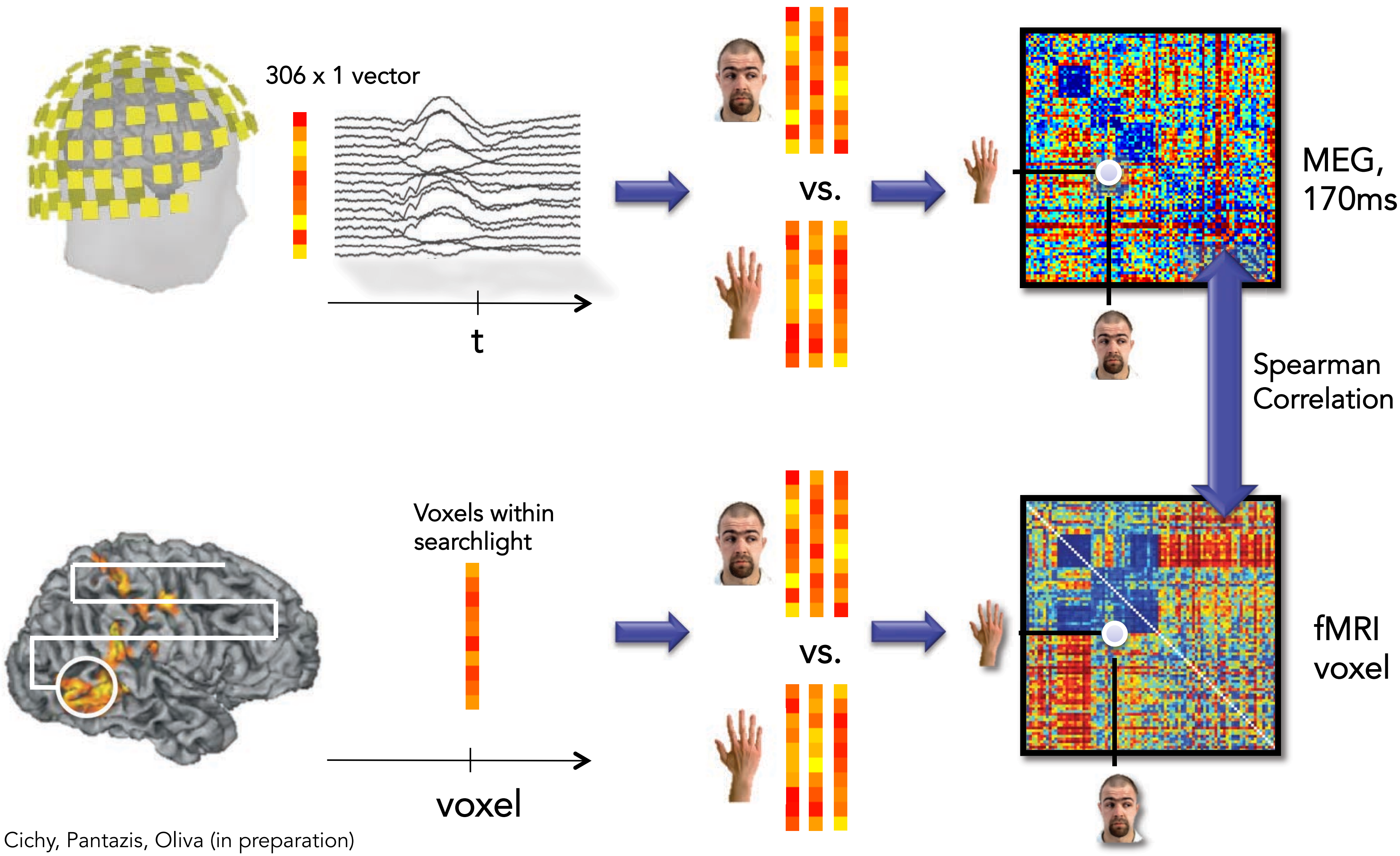




"RDMs as a hub to relate different representations across sensors and models"

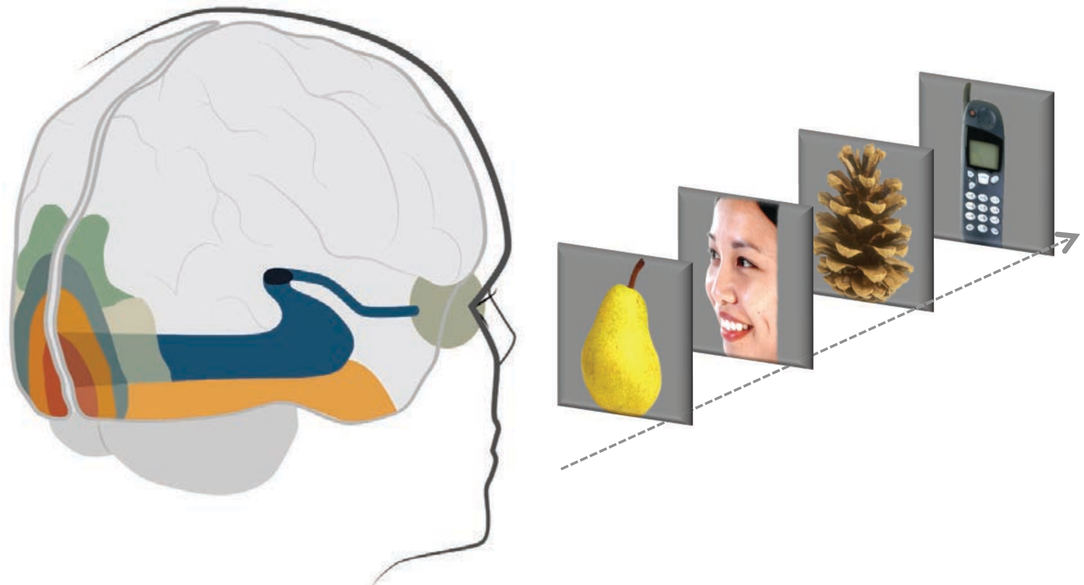
Time-specific fMRI searchlight analysis

A spatially unbiased view of the relations in similarity structure between MEG and fMRI

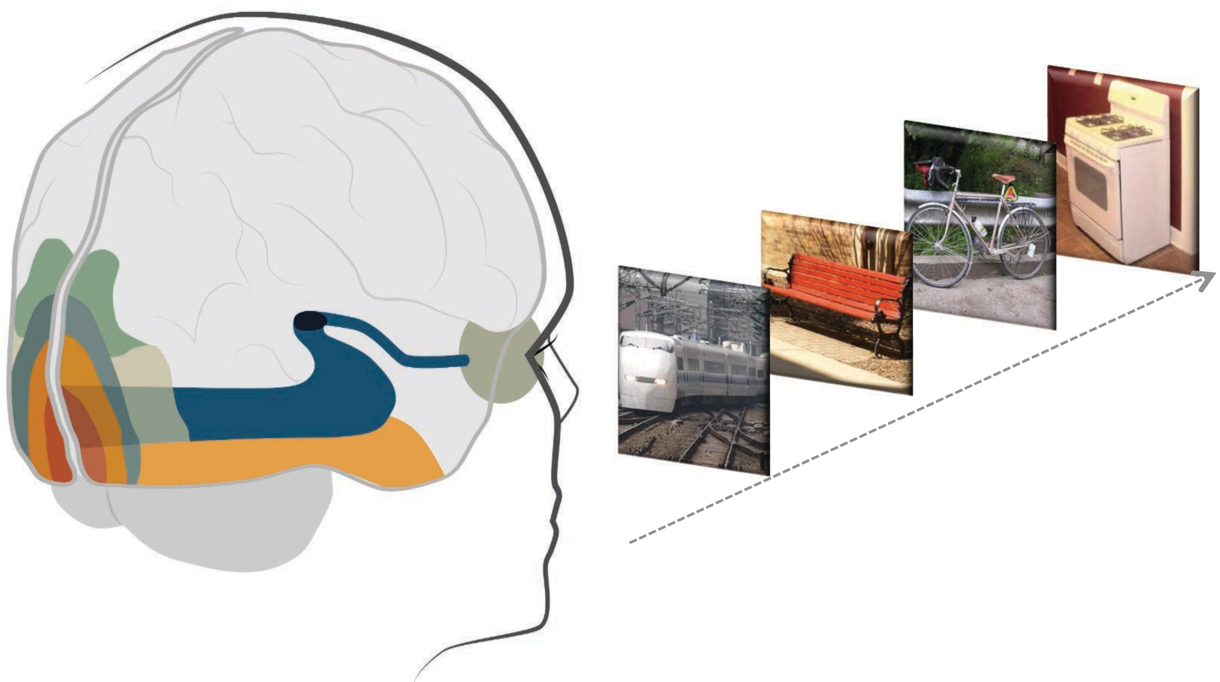


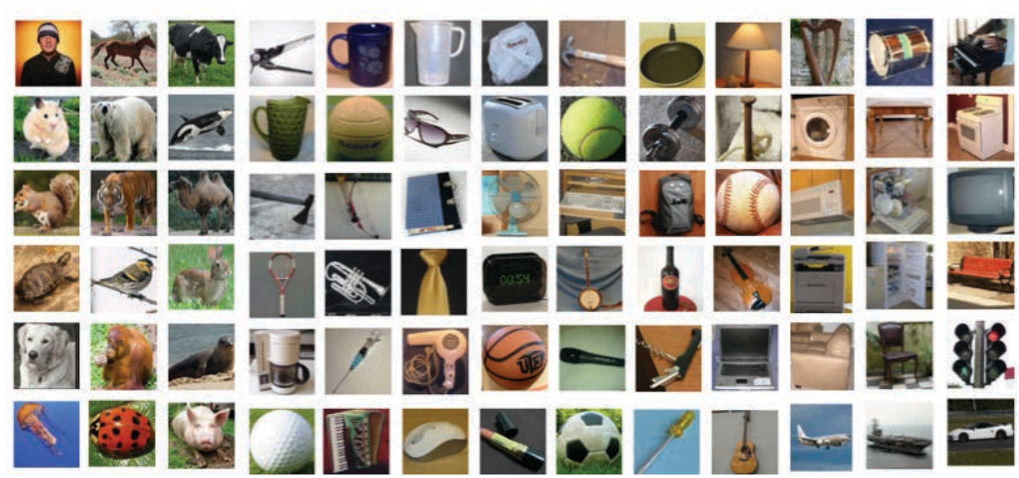
Cichy, Pantazis, Oliva (in preparation)

The dynamics of object recognition

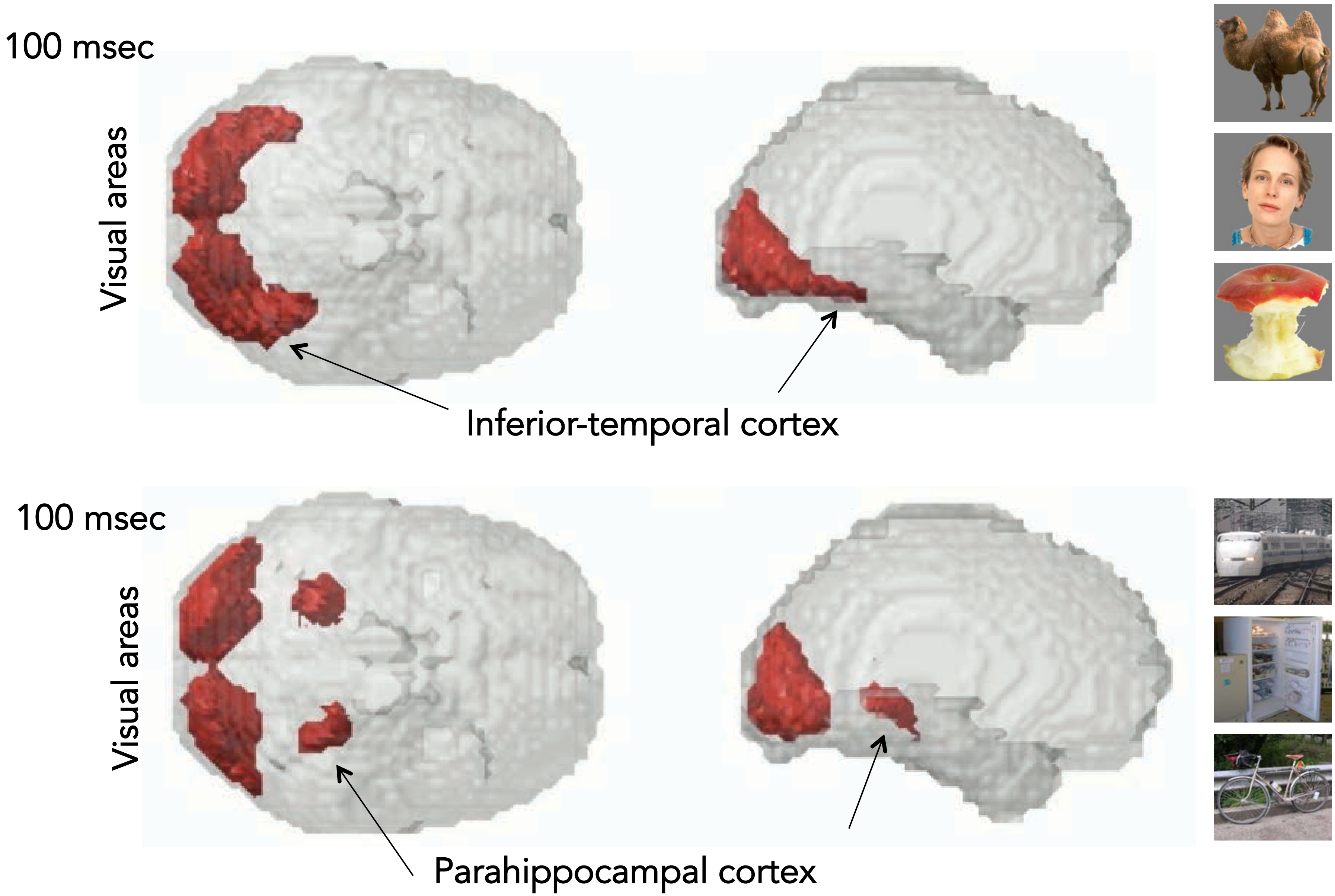


Object recognition in context



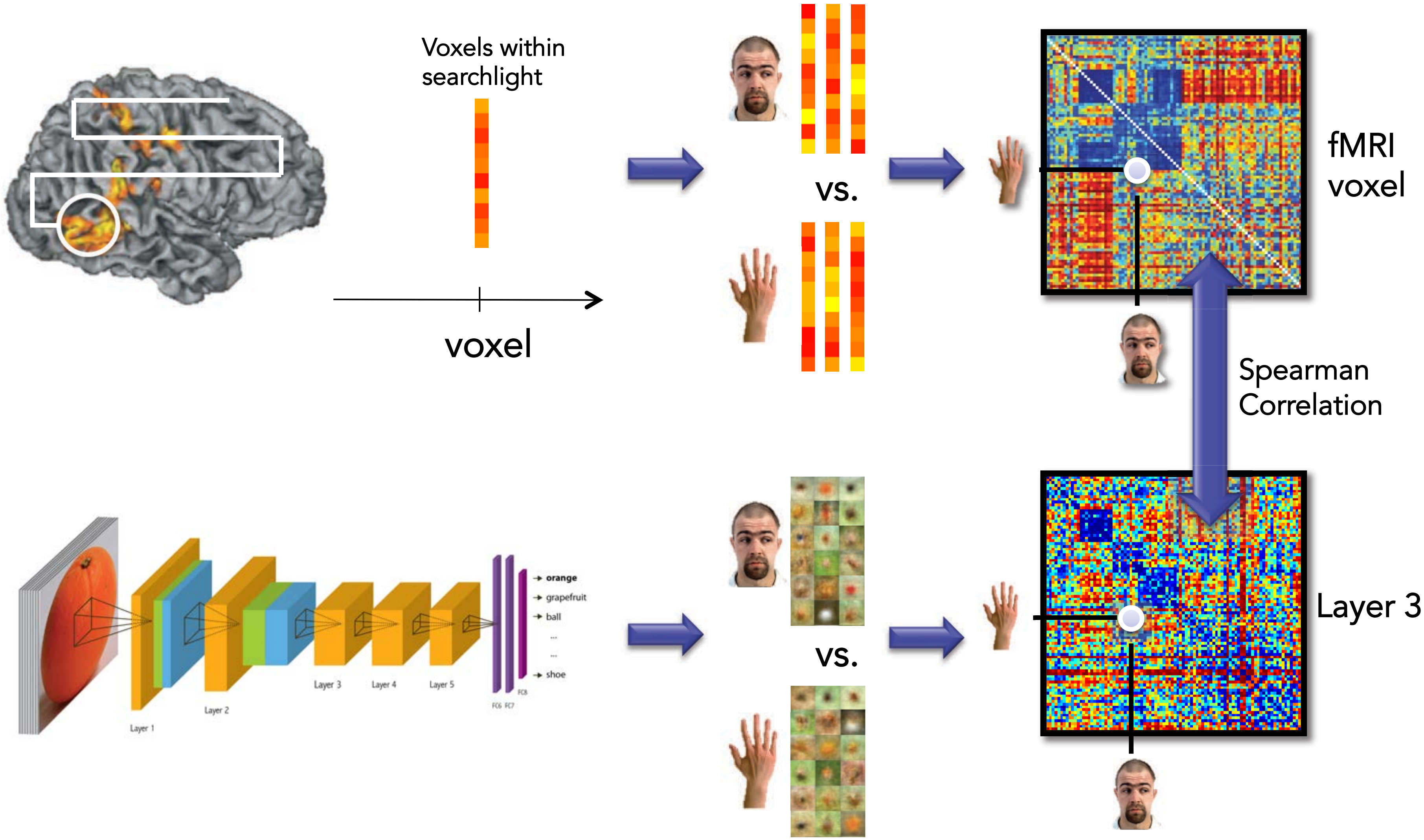


Spatiotemporal maps of correlations between MEG and fMRI



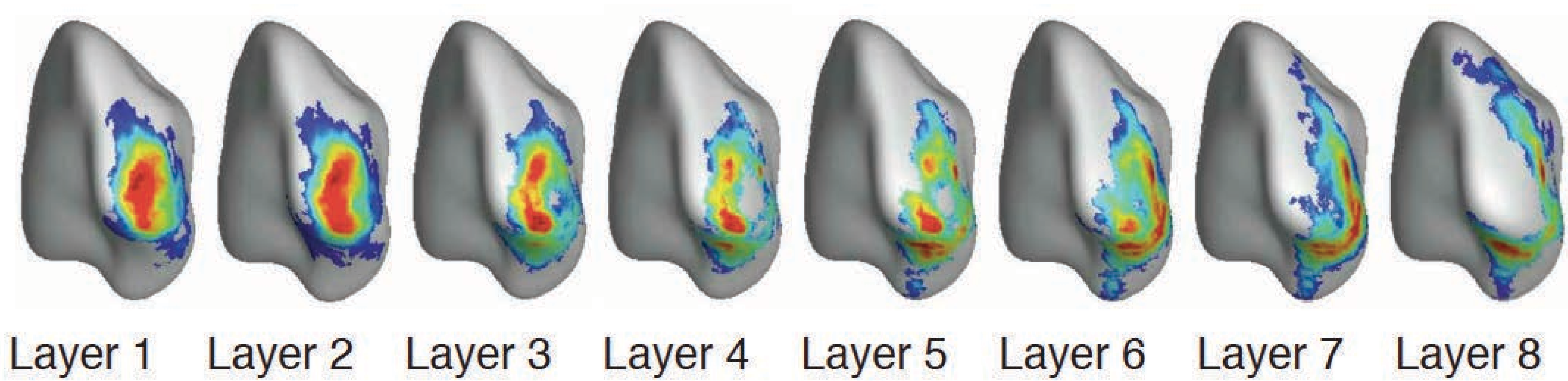
Algorithmic-specific fMRI searchlight analysis

A spatially unbiased view of the relations in similarity structure between deep architectures and fMRI



Cichy, Khosla, Pantazis, Torralba, Oliva (in preparation)   See also Kaligh-Razavi & Krigeskorte (2014)

Spatiotemporal map of correlations between fMRI and model layers

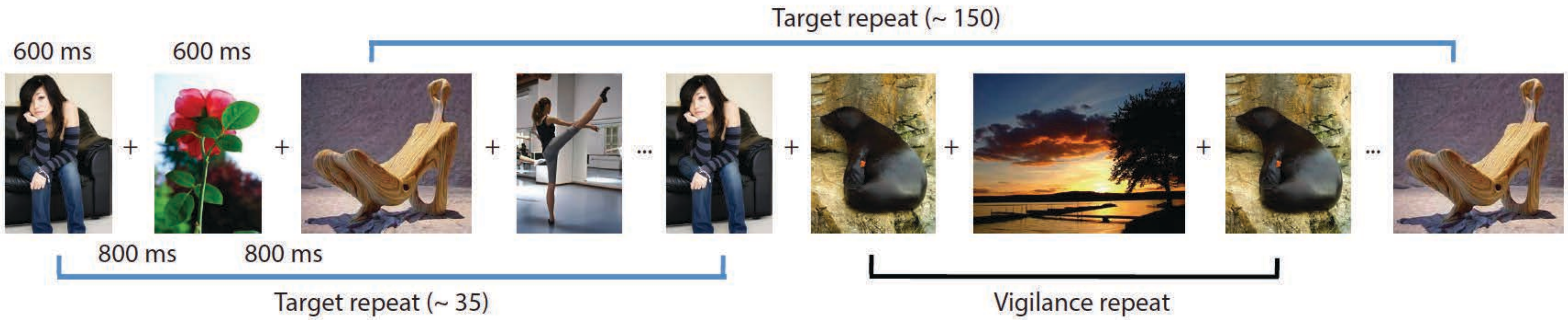


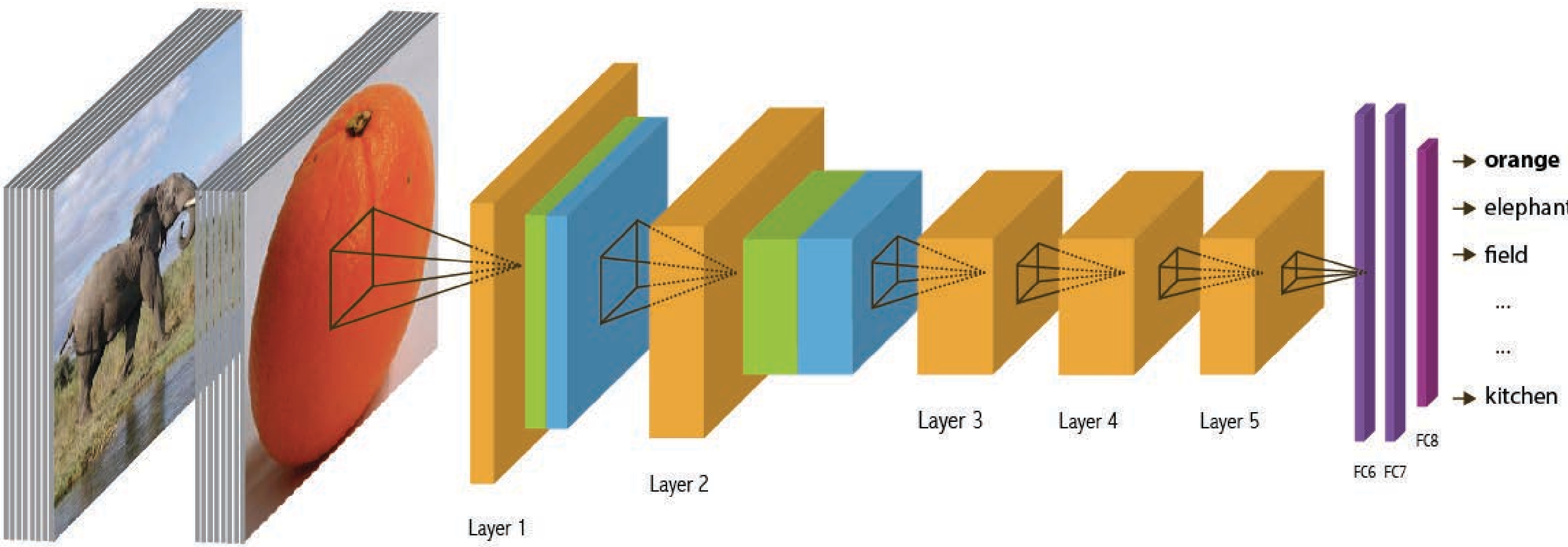
Can we predict which images are memorable ?

Predicting Visual Memorability

~ 60,000 photographs with Memorability scores

Aditya Khosla

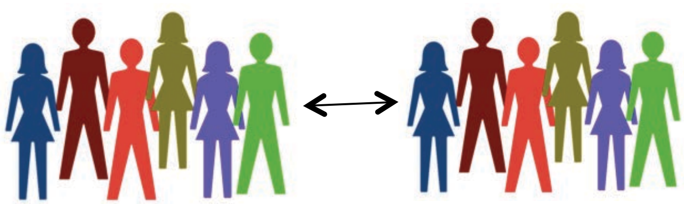


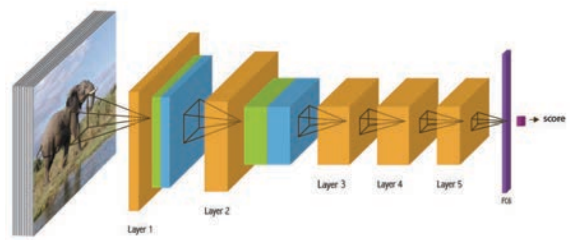


Predicting Visual Memorability

~ 60,000 photographs with Memorability scores

Most memorable | Less memorable





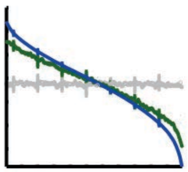
Human ρ = 0.68

Deep feature ρ = 0.64

Cognitive-level Algorithms

Memorability: metric of the utility of information

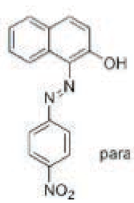
Understand human memory



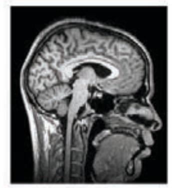
Data Visualization

Logos
Slogans
- words-

Education
-Individual difference



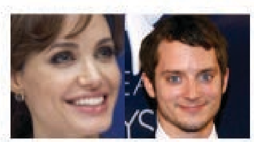
Diagnose memory problems



Mobile applications

Social Networking

Face Memorability



Design mnemonic aids

"heavy"

"lourds"

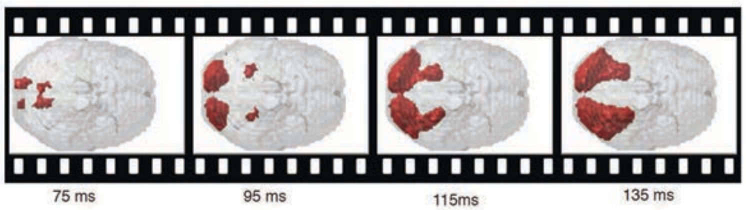
Retrieve better images from search

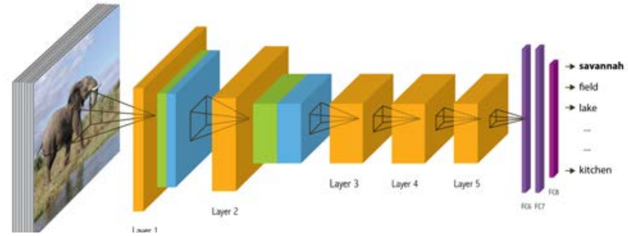
Computer Graphics
- *cognitive* saliency

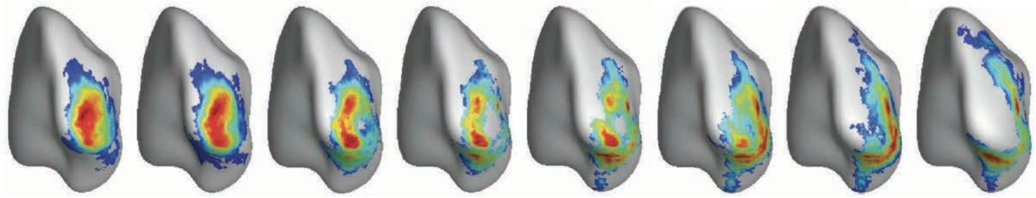
Summarize Bigdata – images, videos

Time, Space and Computation

A converging framework for hypothesis testing







Power of Prediction

Comparing large-scale processing between natural and artificial systems will not only allow us to understand why biological systems have implemented a certain mechanism, but will allow

- Studying the strategies that work best for performing specific tasks
- Characterizing the operations when the system is broken
- Exploring the alternatives biological systems have not taken

A.I "Alien" Intelligence *(Kevin Kelly, Wired magazine)*

NSF CISE, RI: 1016862 National Eye Institute NATIONAL INSTITUTES OF HEALTH R01-EY020484 Google