6.819 / 6.869: Advances in Computer Vision

MIT
COMPUTER
VISION

High-level vision:

Object & Scene Recognition: What are the next challenges?

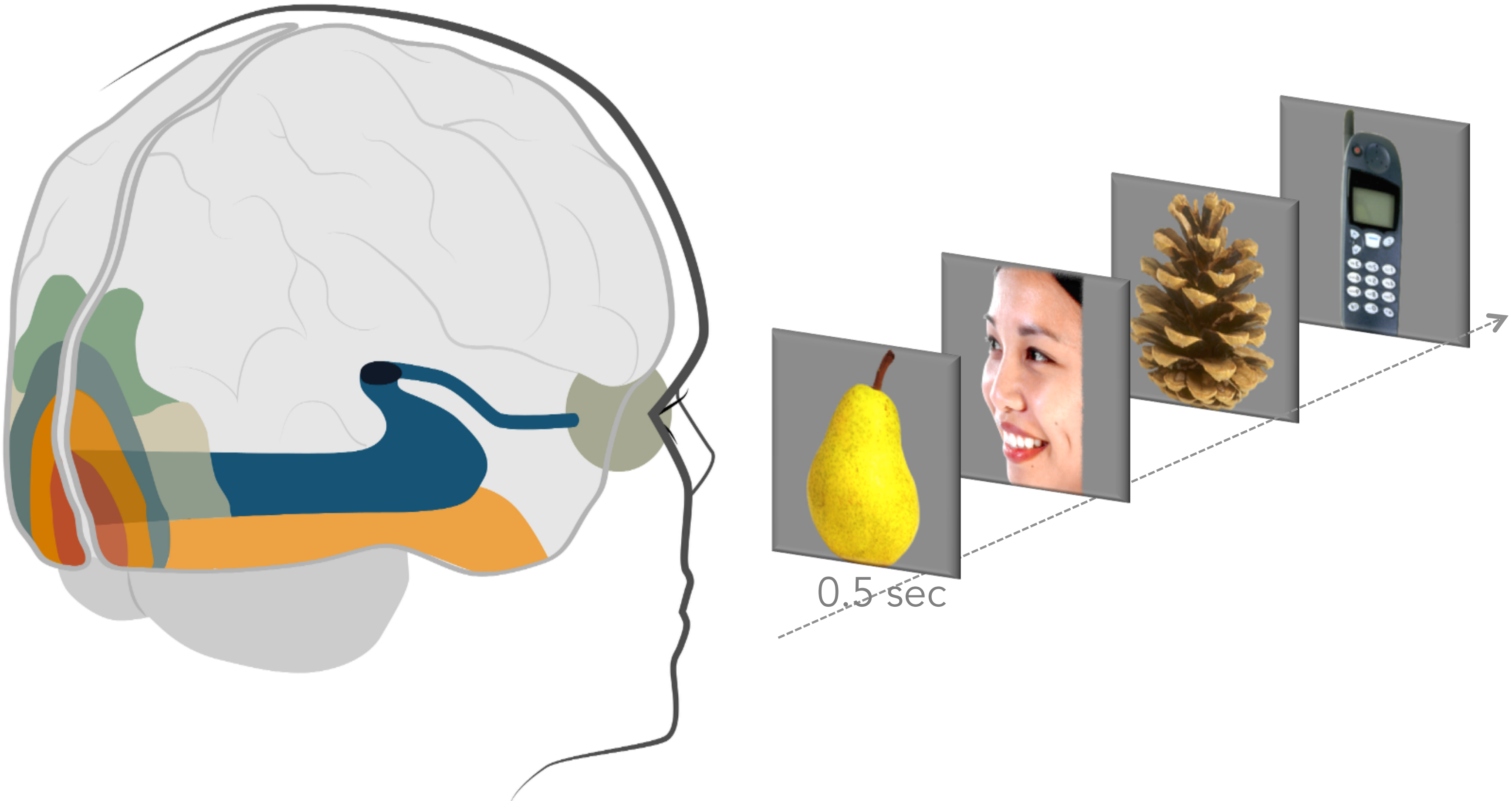
Website:
http://6.869.csail.mit.edu/fa15/

Instructor: Aude Oliva

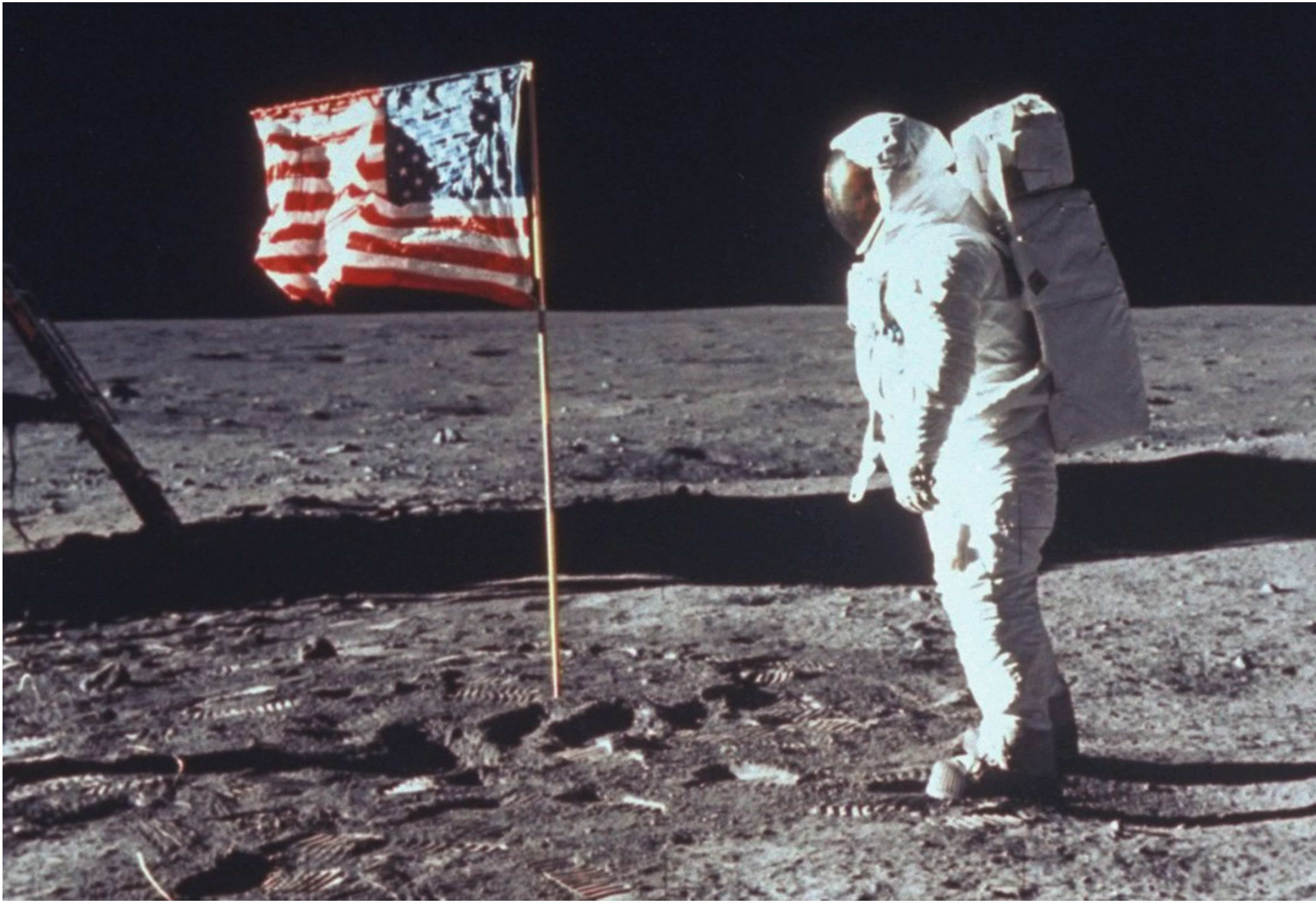
Lecture TR 9:30AM – 11:00AM
(Room 34-101)

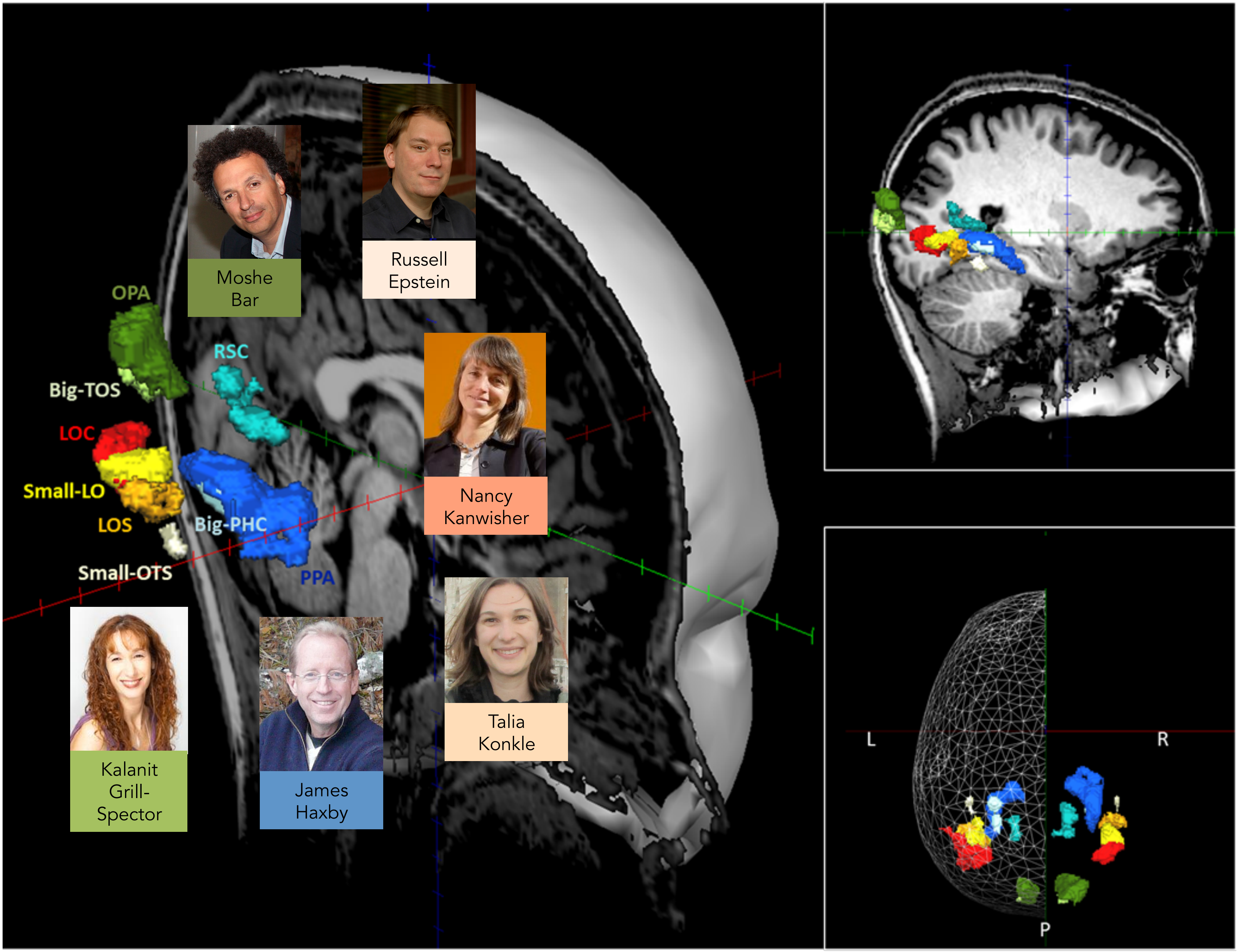
Asiatic black bear: 5.6
bench: 3.0
AT&T
6:57 PM
100%
places.csail.mit.edu
Predictions:
type: outdoor
semantic categories: picnic_area:0.14, patio:0.12, yard:0.11, veranda:0.11, boardwalk:0.06
scene attributes: natural light, man-made, nohorizon, soothing, foliage, trees, vegetation, warm, open area,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

High Computing Visual Engine: Object recognition



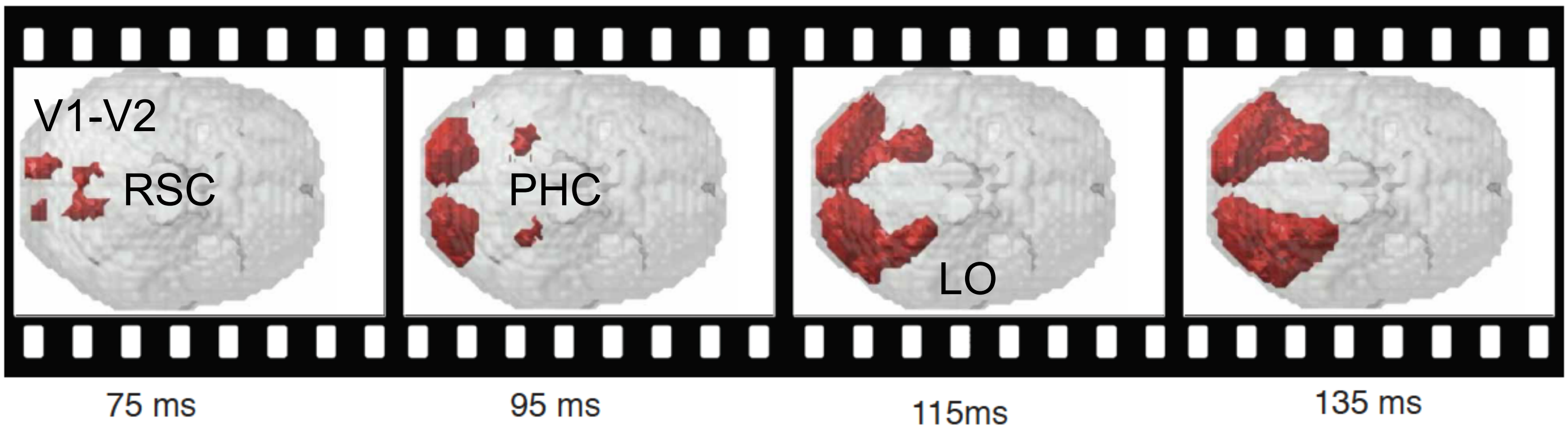

BOND





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

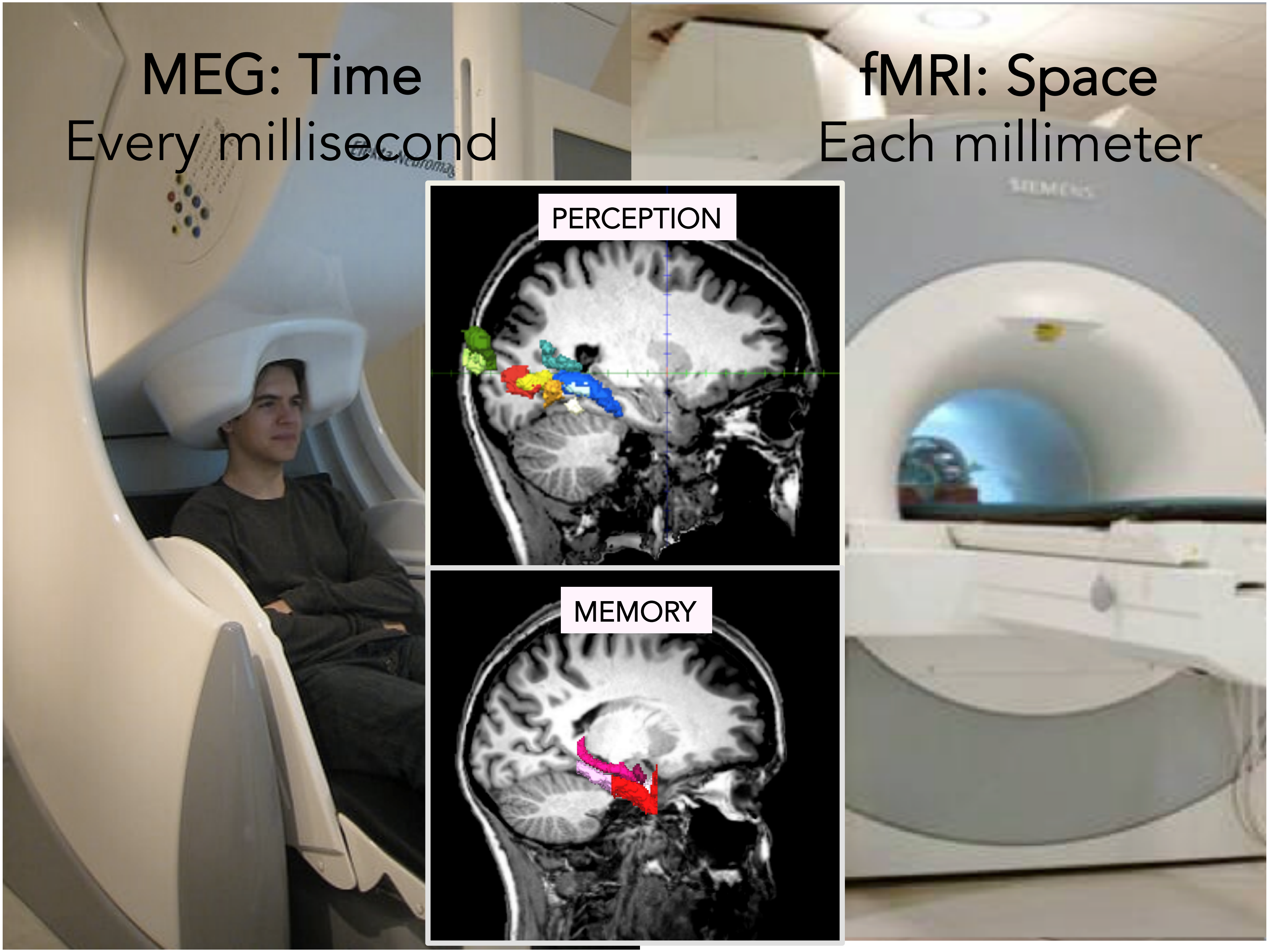
Spatiotemporal map of correlations between MEG and fMRI



RSC = Retrosplenial cortex

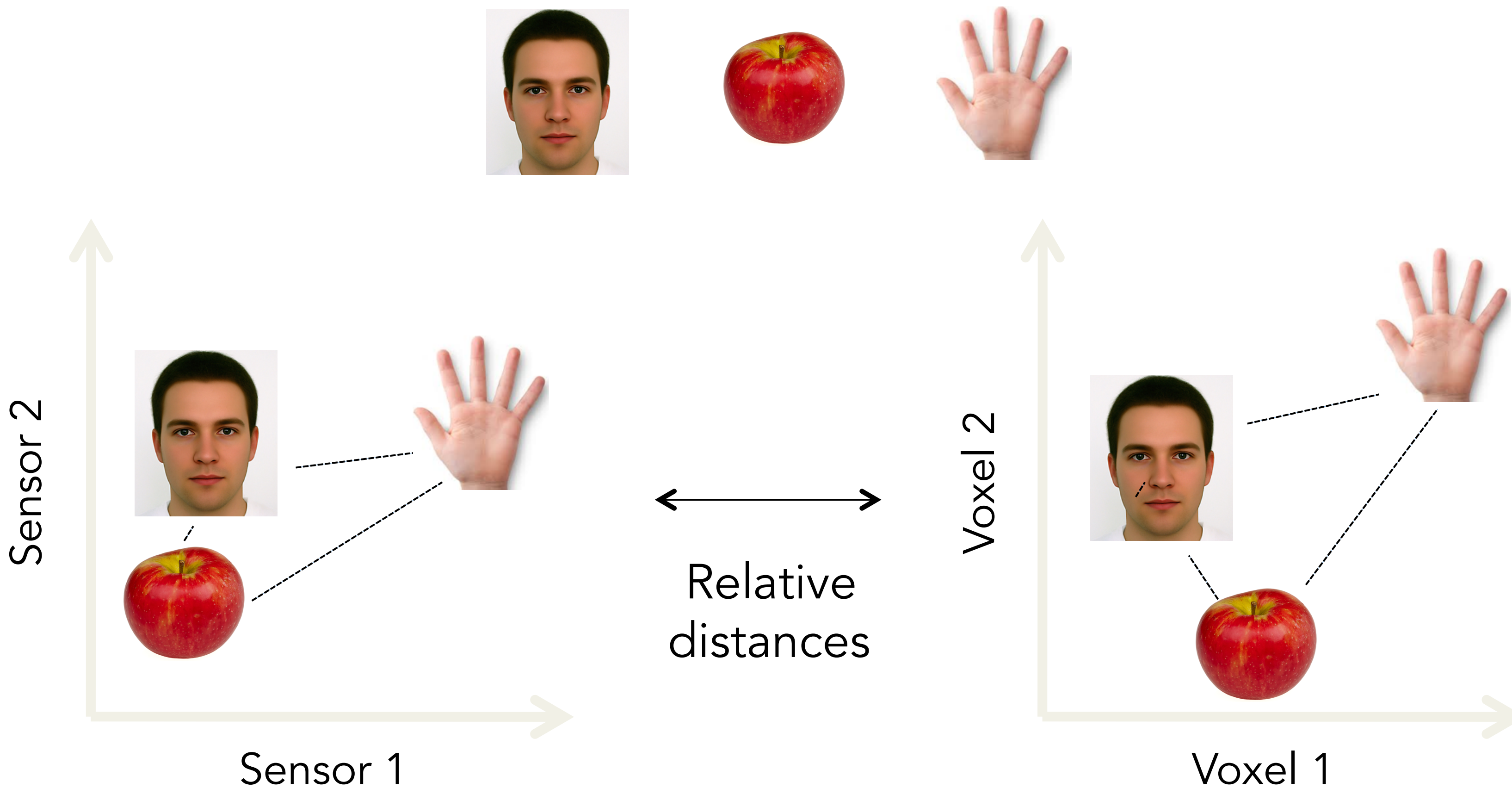
PHC = Parahippocampal cortex

LO = Lateral Occipital cortex


MEG: Time
Every millisecond
fMRI: Space
Each millimeter
PERCEPTION
MEMORY

Representational Geometry

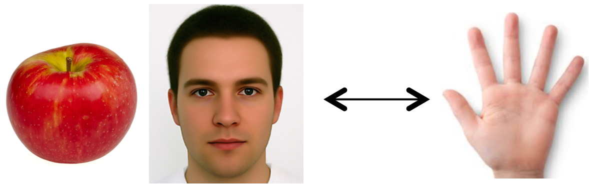
Nikolaus Kriegeskorte (2008)



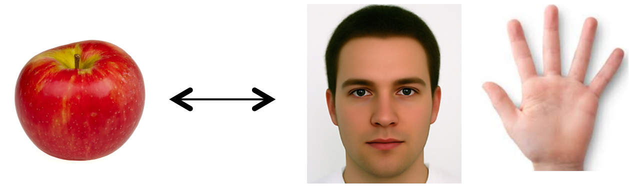
Shepard et al., 1980; Kruskal and Wish., 1978; Edelman et al. 1998; Kriegeskorte et al., 2008; Mur et al., 2009; Liu et al., 2013

Representational Geometry

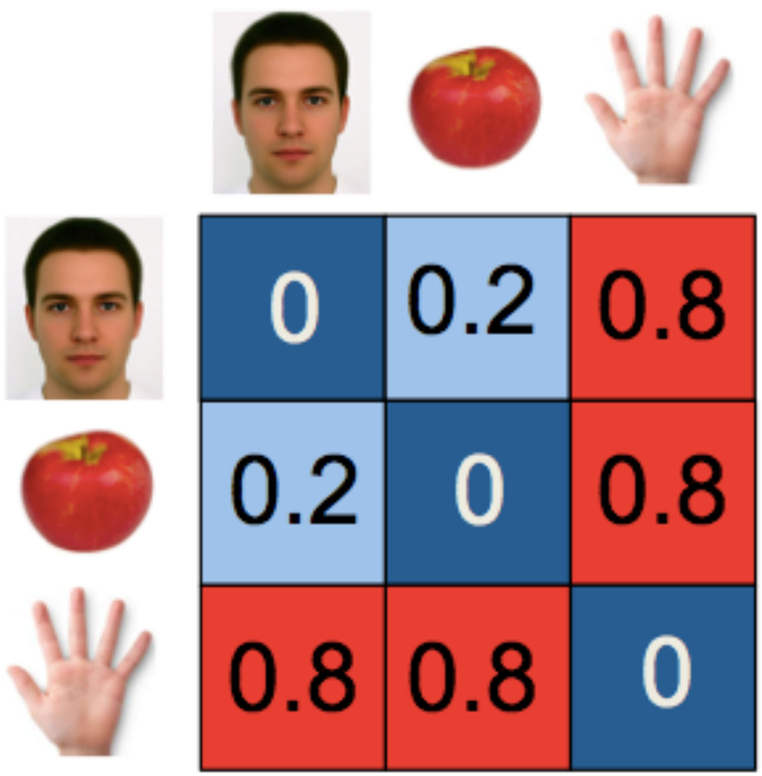
Nikolaus Kriegeskorte (2008)

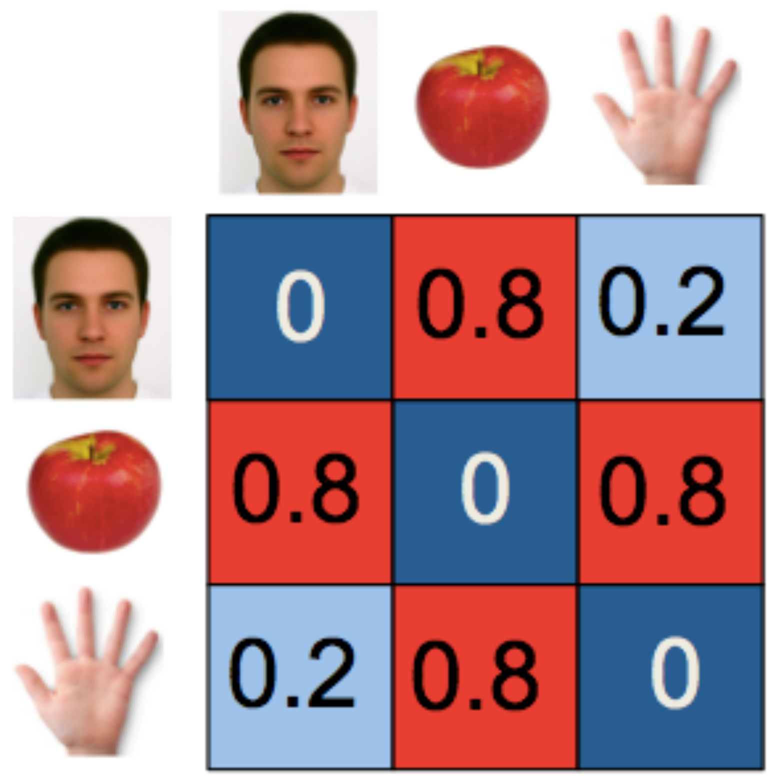


Round shape



Body part

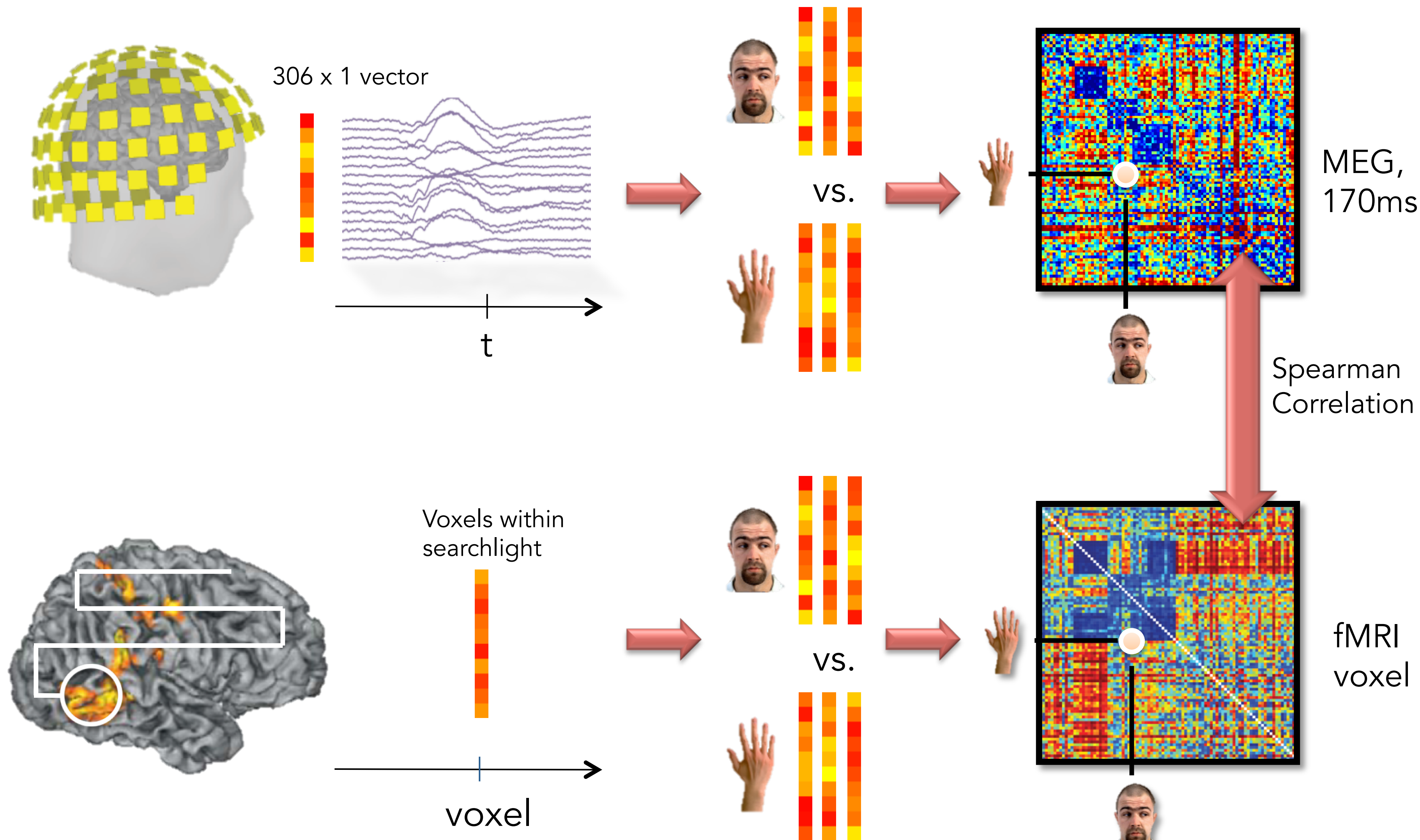




"RDMs as a hub to relate different representations across sensors and models"

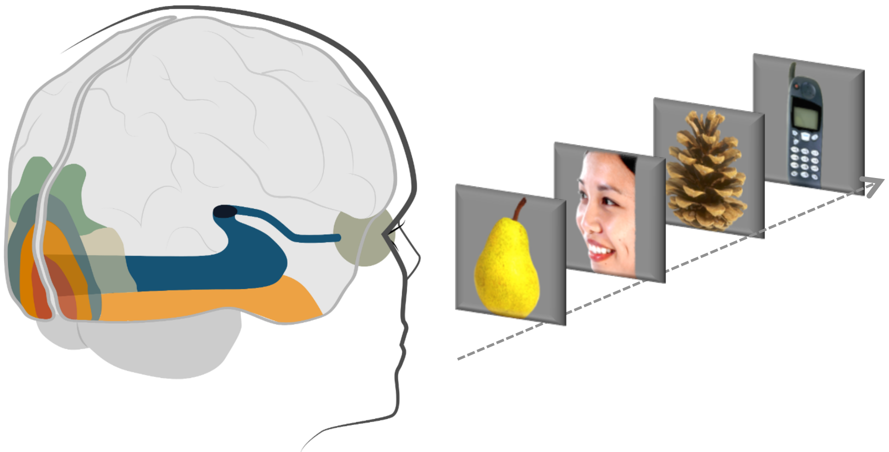
Time-specific fMRI searchlight analysis

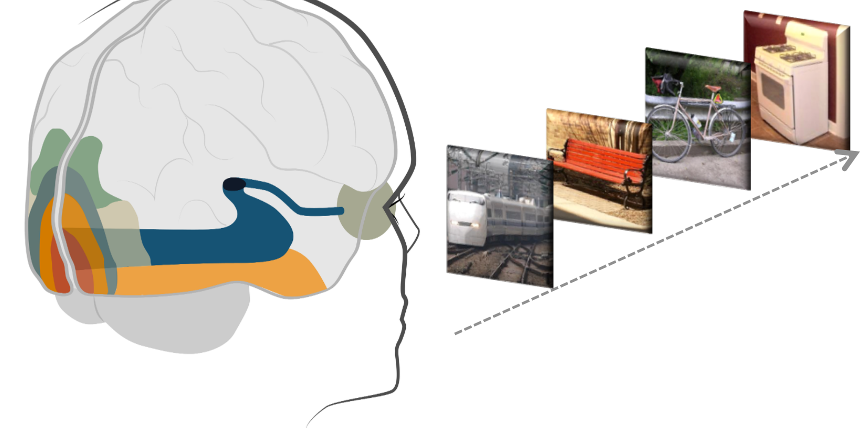
A spatially unbiased view of the relations in similarity structure between MEG and fMRI

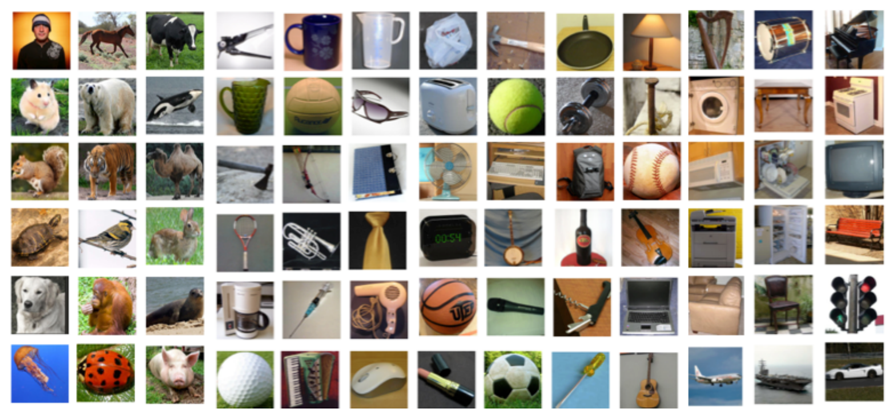


Cichy, Pantazis, Oliva (submitted)

Object recognition



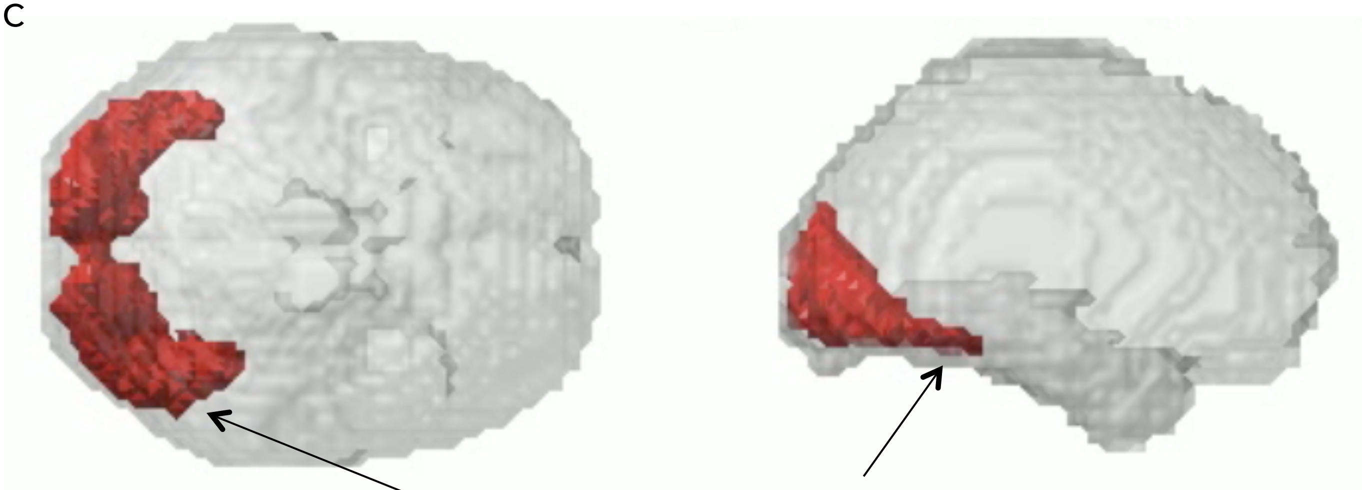




Spatiotemporal maps of correlations between MEG and fMRI

100 msec

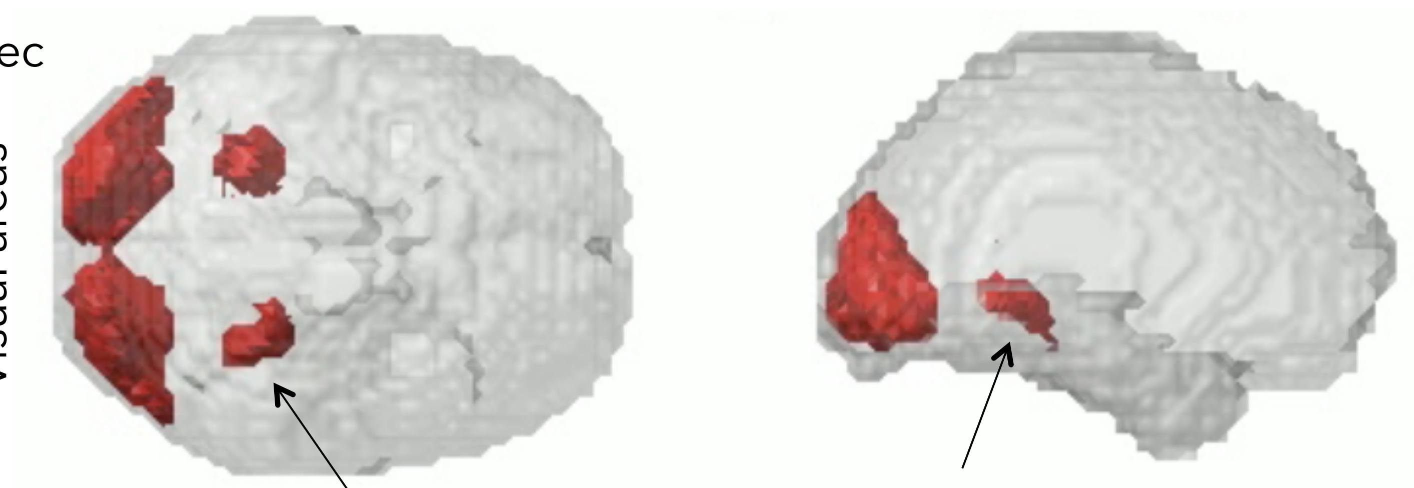
Visual areas



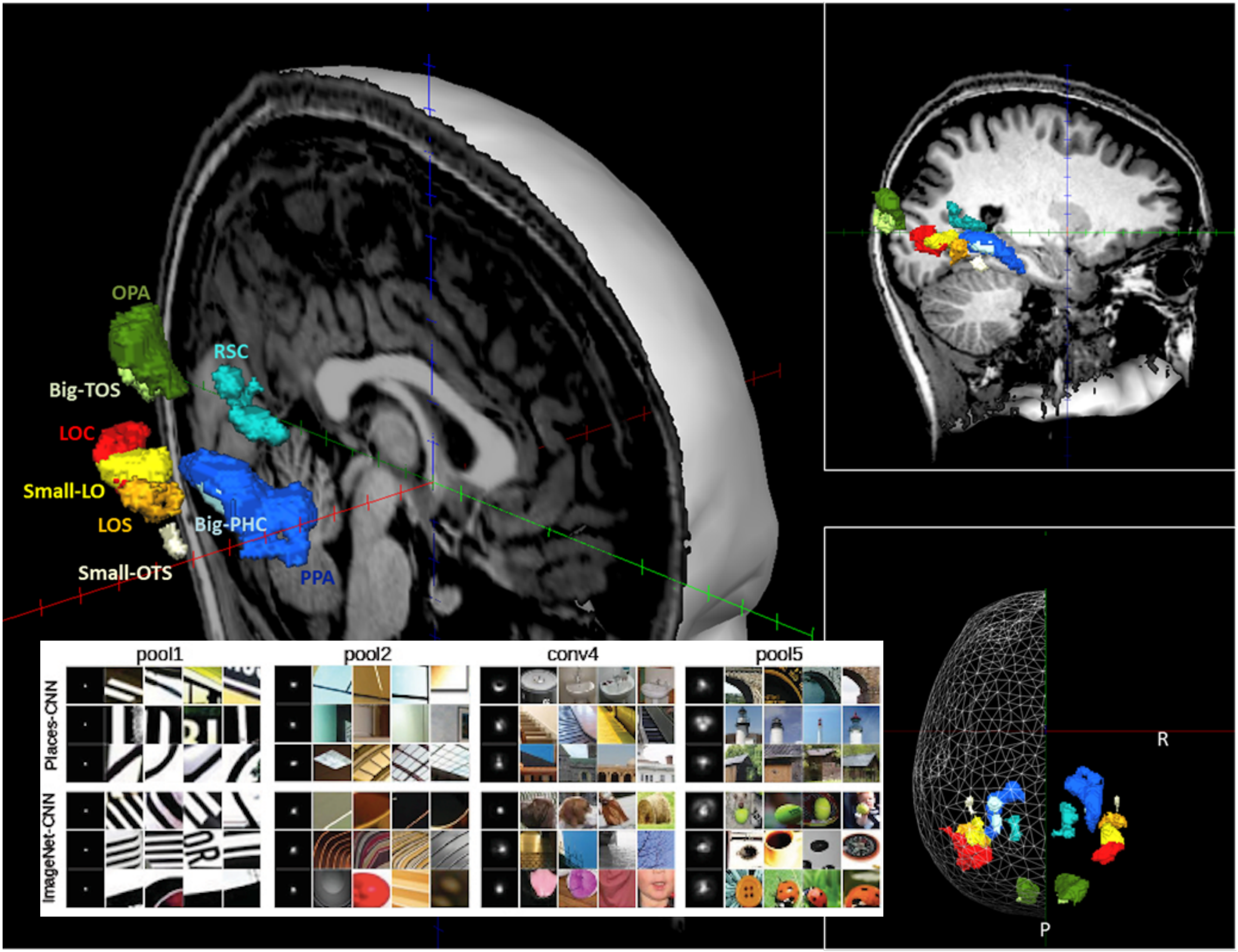
Inferior-temporal cortex

100 msec

Visual areas

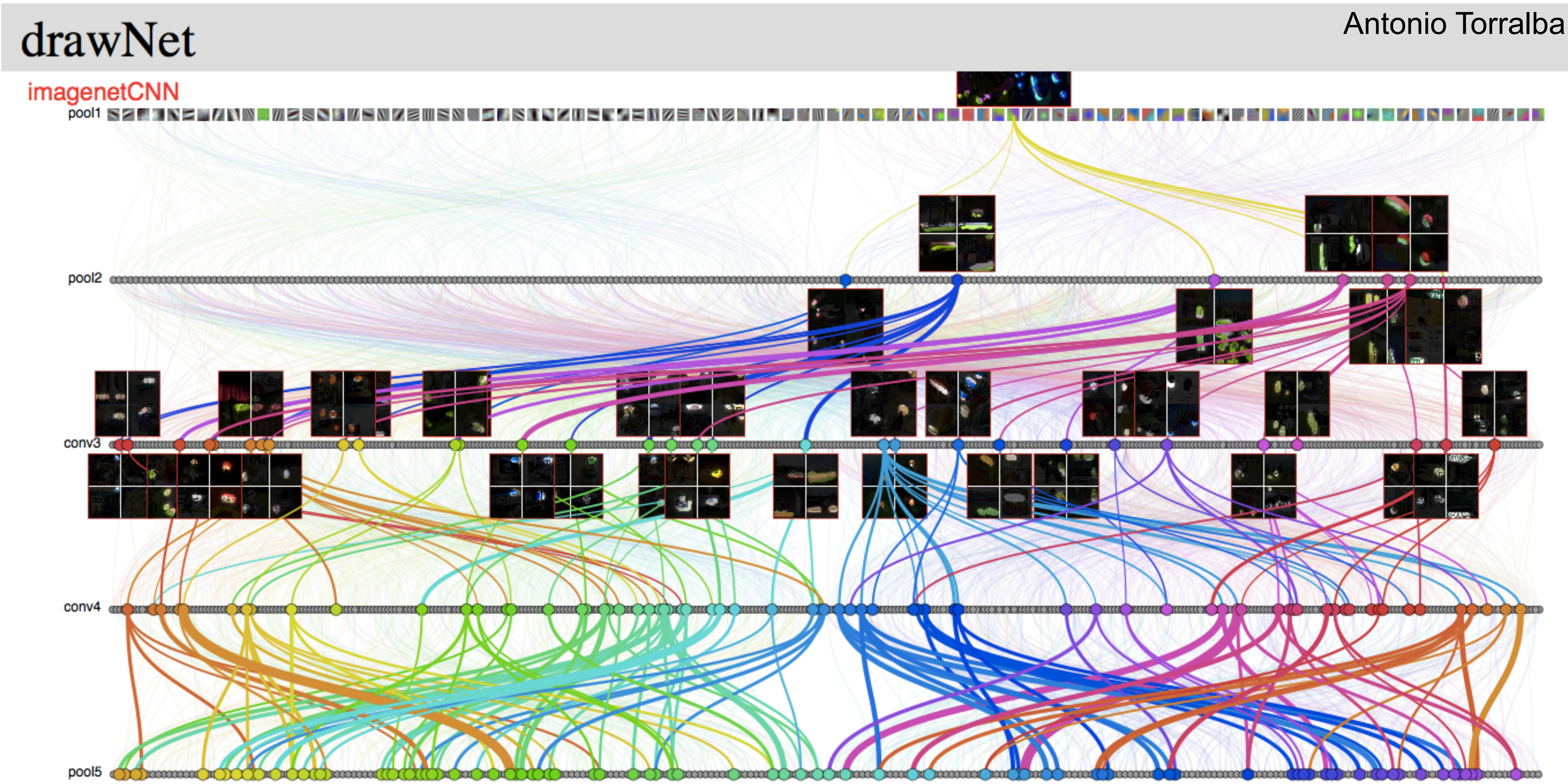


Parahippocampal cortex


OPA
RSC
Big-TOS
LOC
Small-LO
LOS
Big-PHC
Small-OTS
PPA
R
P
pool1
pool2
conv4
pool5
Places-CNN
ImageNet-CNN

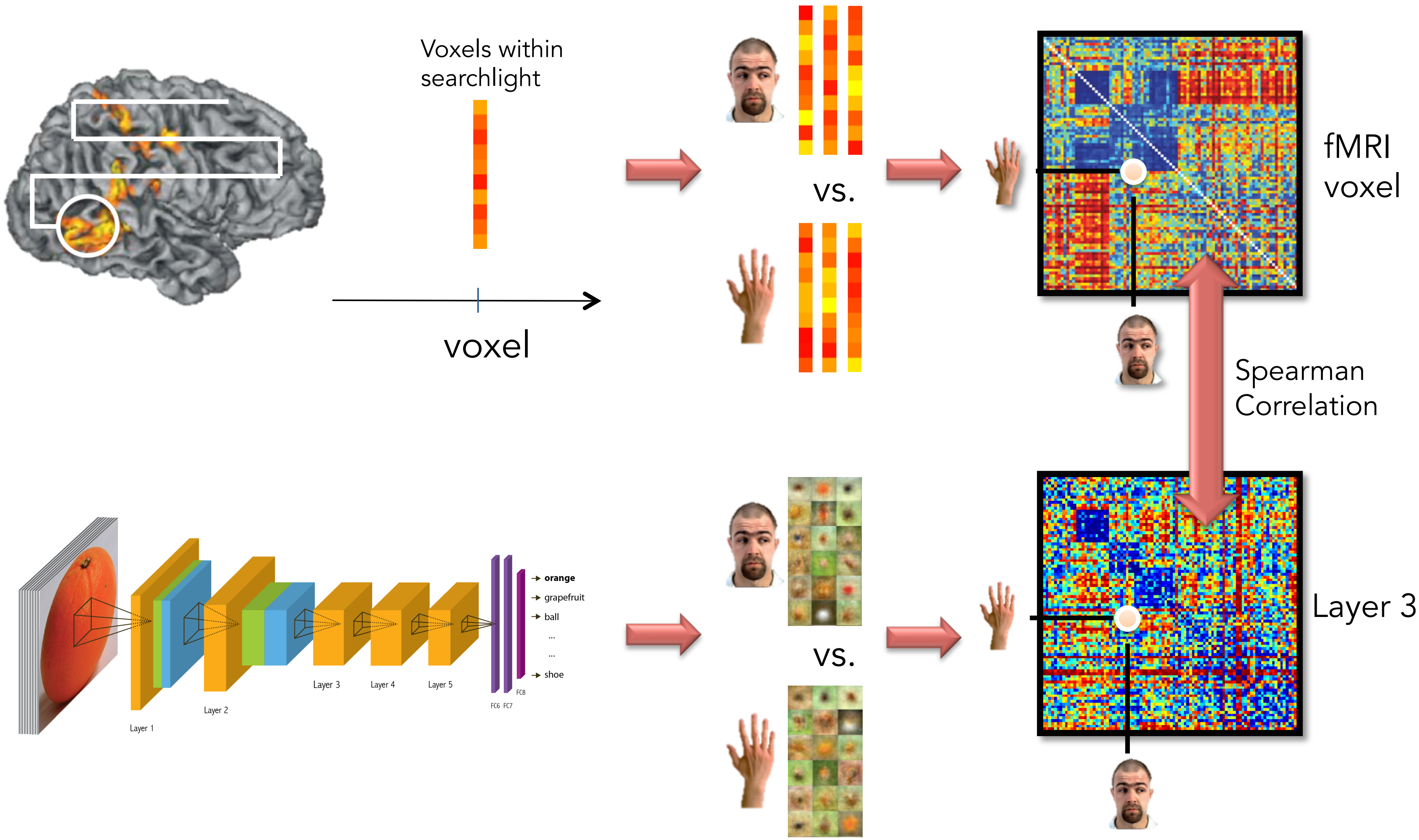
Visualizing model RFs & connections

Antonio Torralba



Algorithmic-specific fMRI searchlight analysis

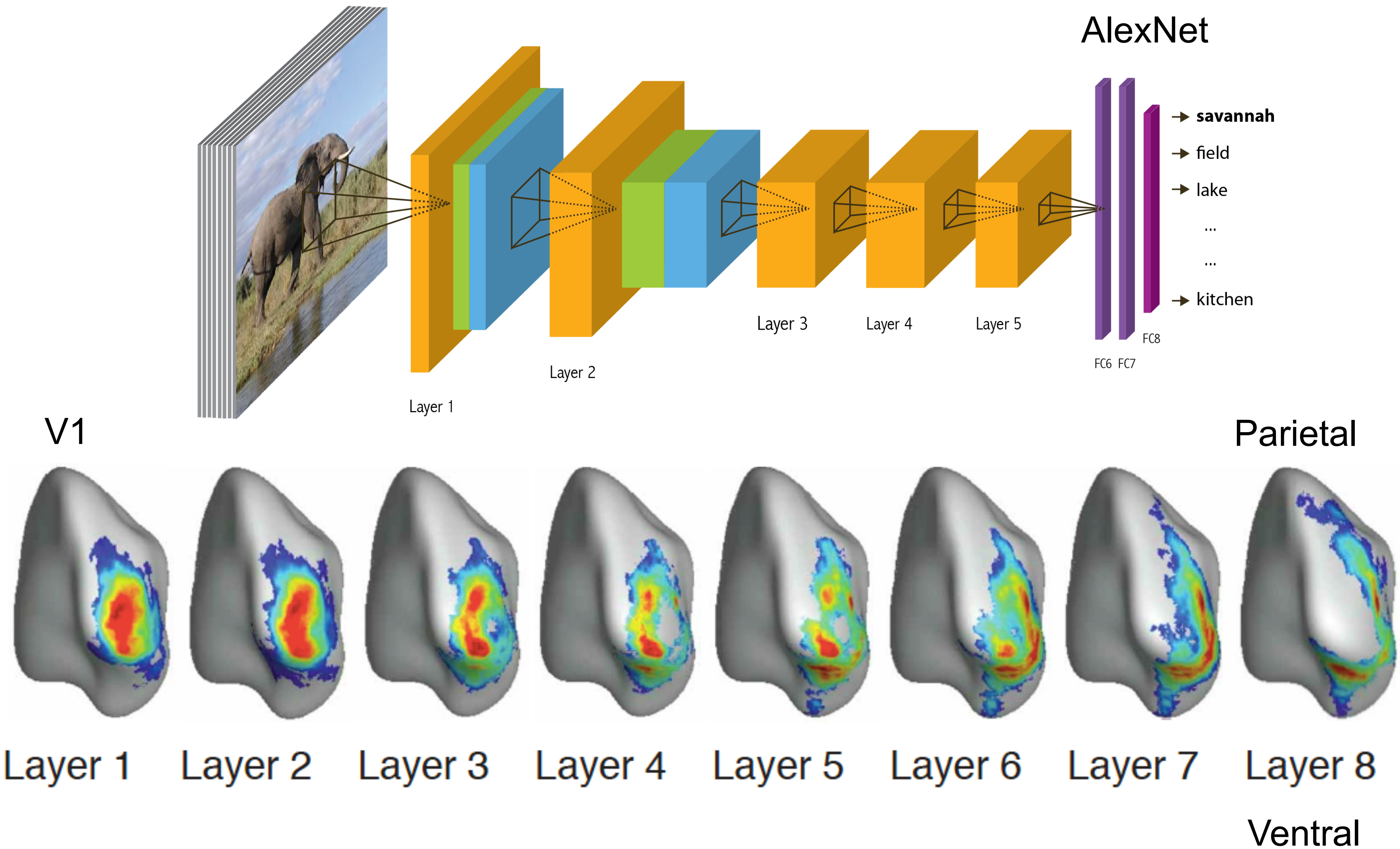
A spatially unbiased view of the relations in similarity structure between deep architectures and fMRI



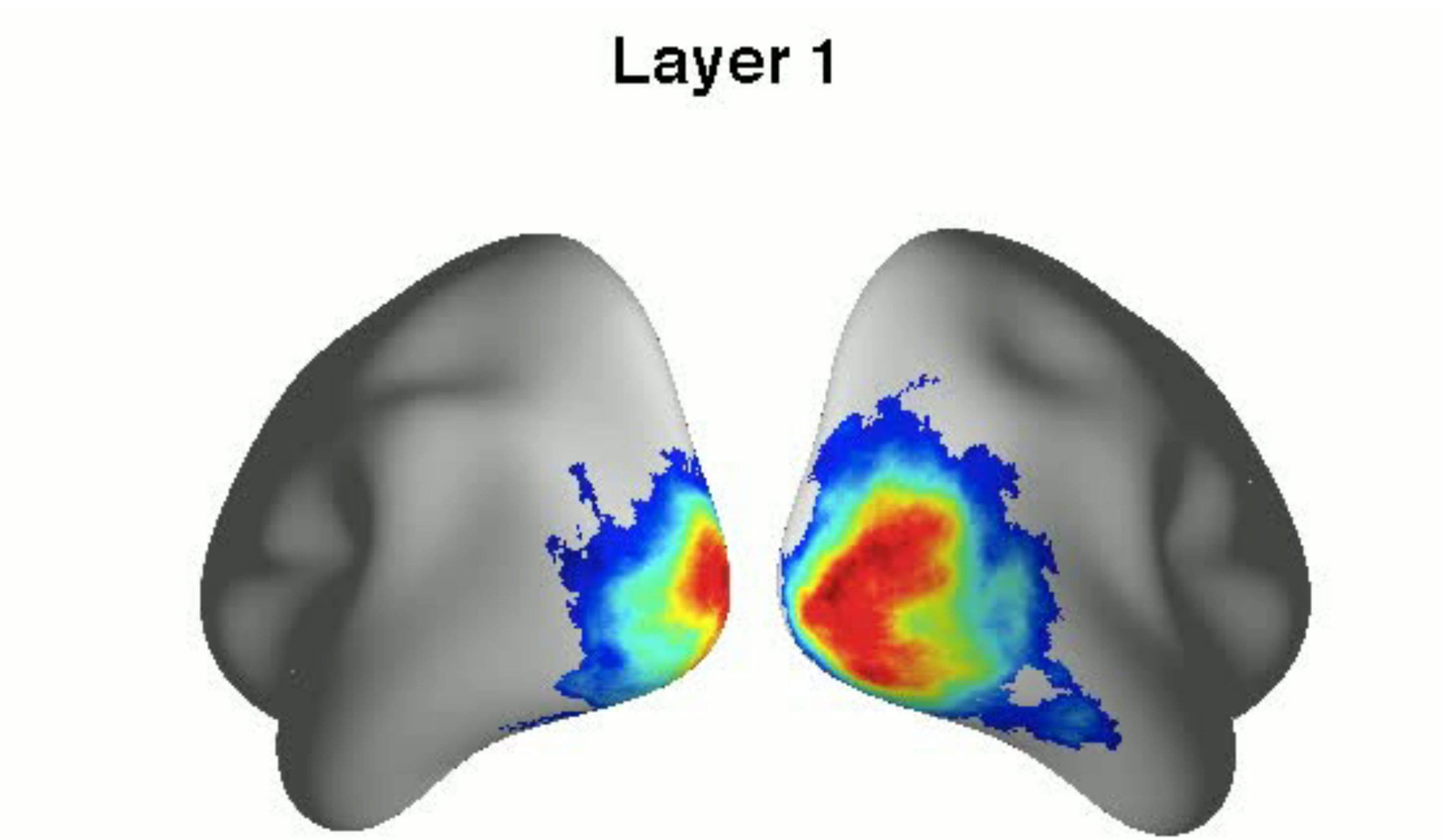
Cichy, Khosla, Pantazis, Torralba, Oliva (submitted)

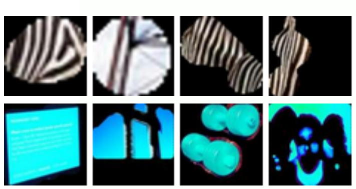
See also Kaligh-Razavi & Krigeskorte (2014)

Spatiotemporal map of correlations between human brain and model layers

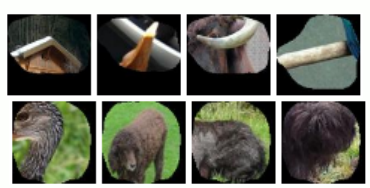


Spatiotemporal maps of correlations between human brain and CNN layers

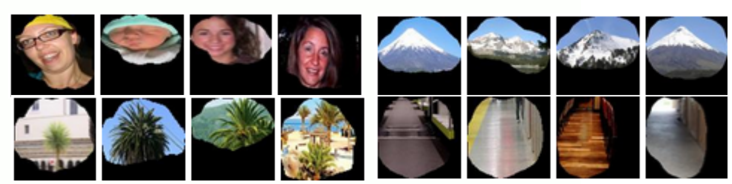




Layers 1-2



Layers 2-4

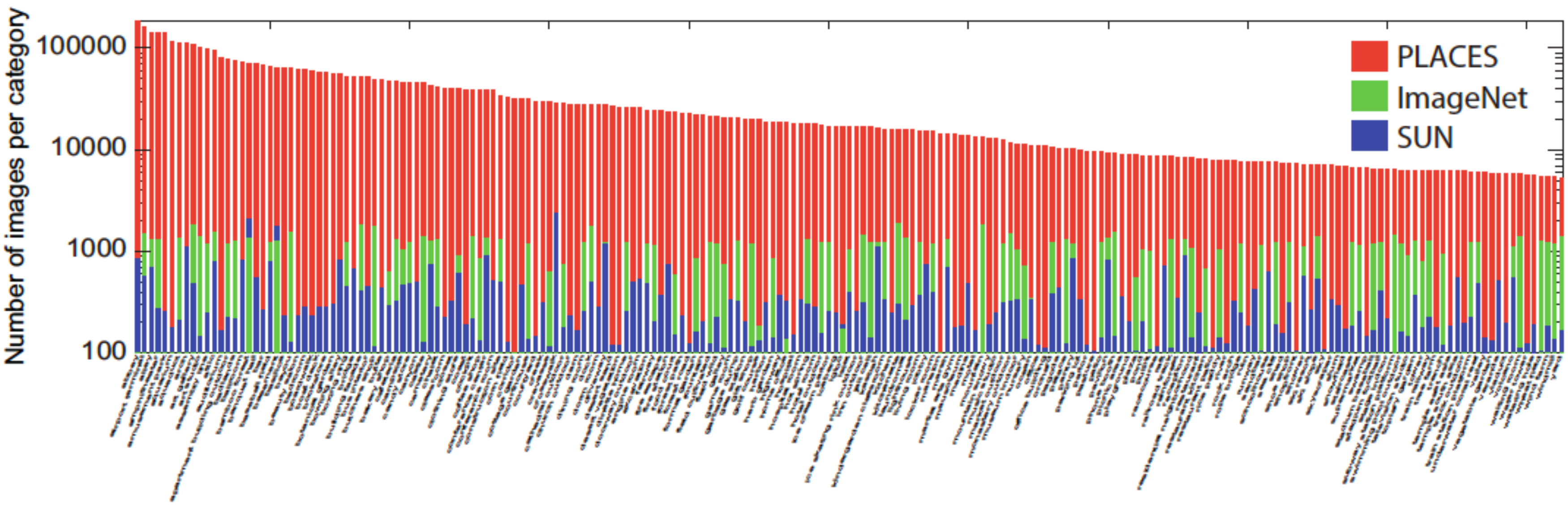


Layers 5-8

places

400 Categories, 10 M images

places.csail.mit.edu



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,

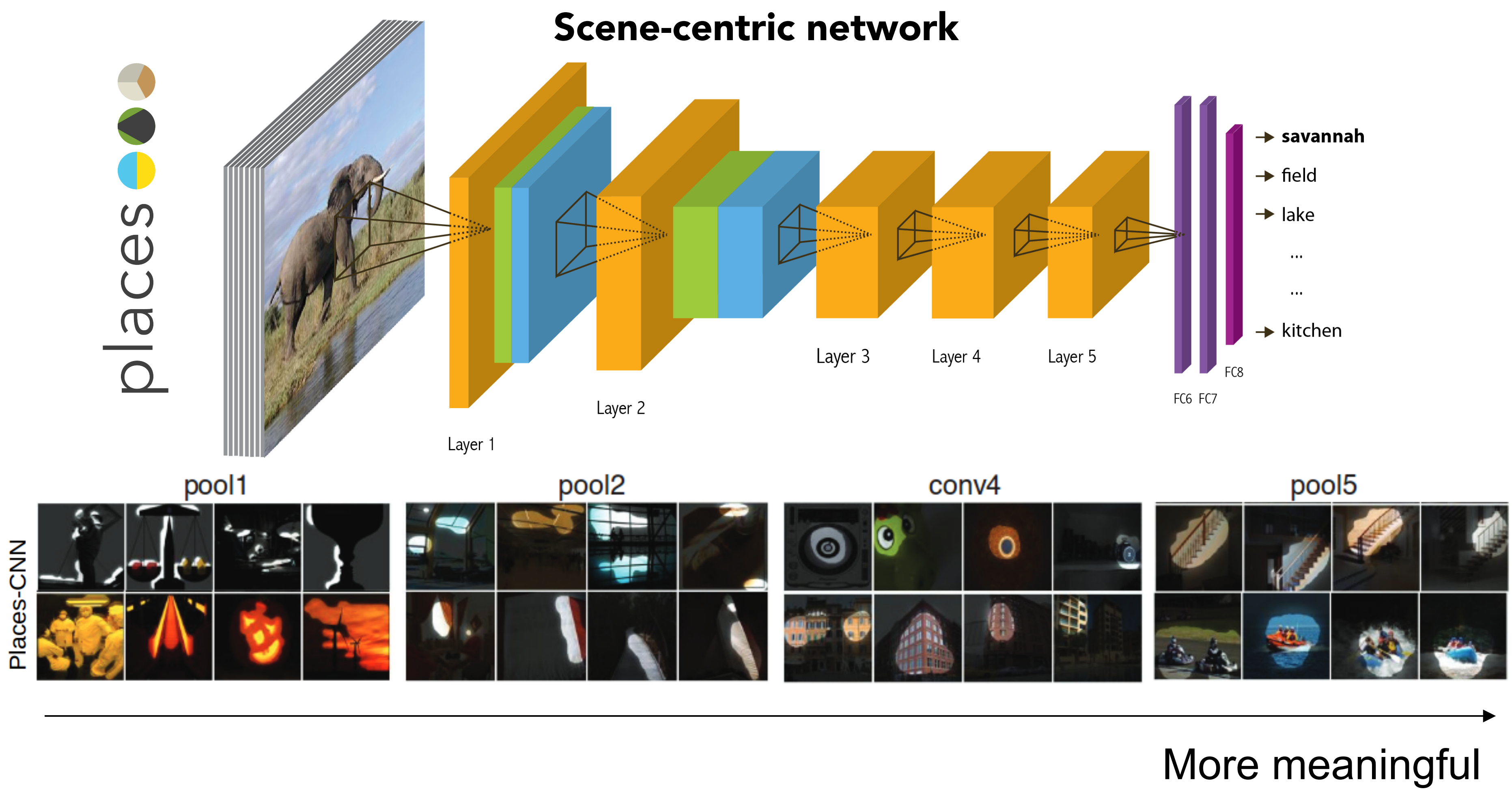
Predictions:

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

NIPS 2014 release: 2.5 million images, 205 scene categories

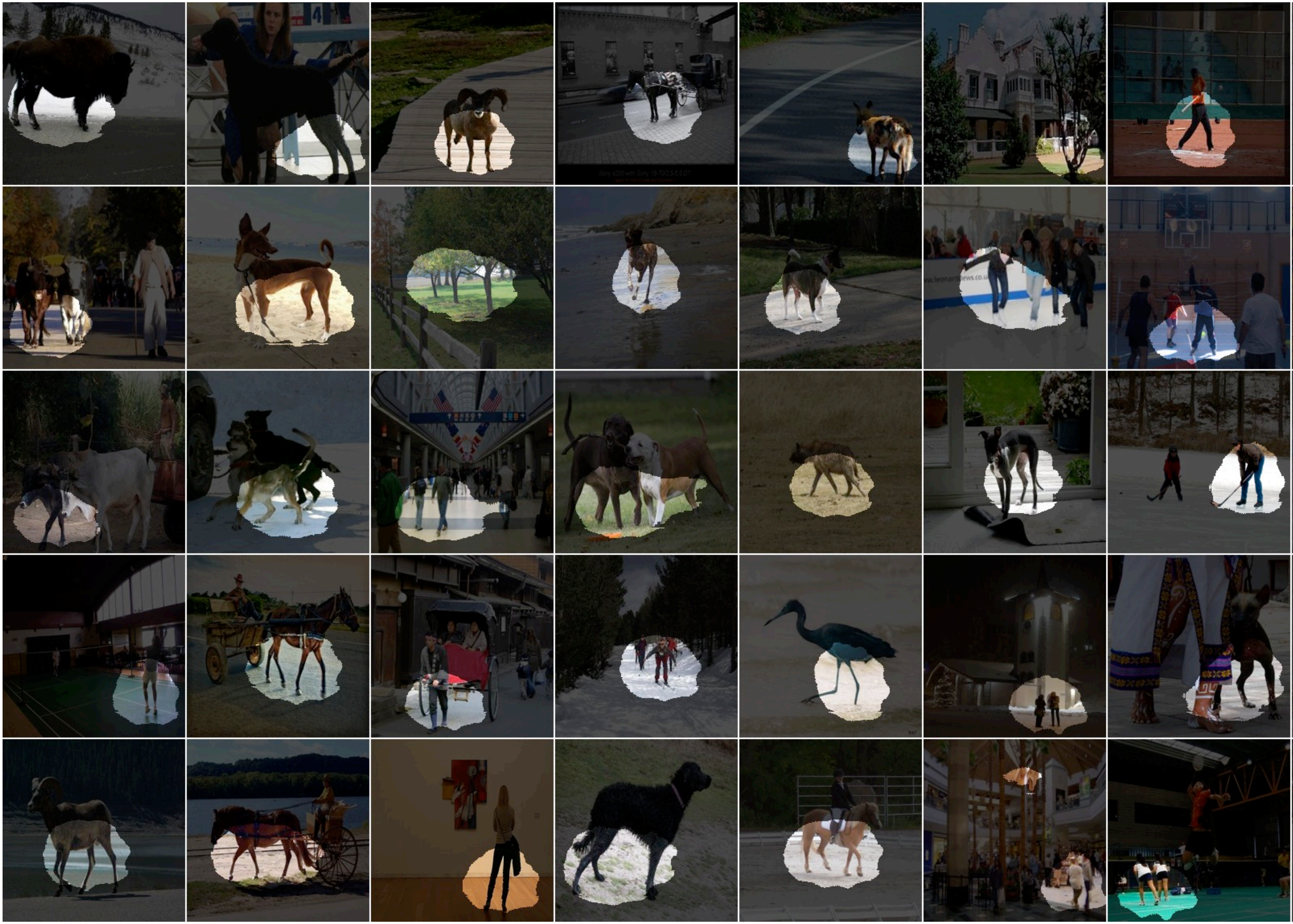
Zhou, Lapedriza, Xiao, Torralba & Oliva (2014), NIPS

Deep architectures: Place and Object Recognition

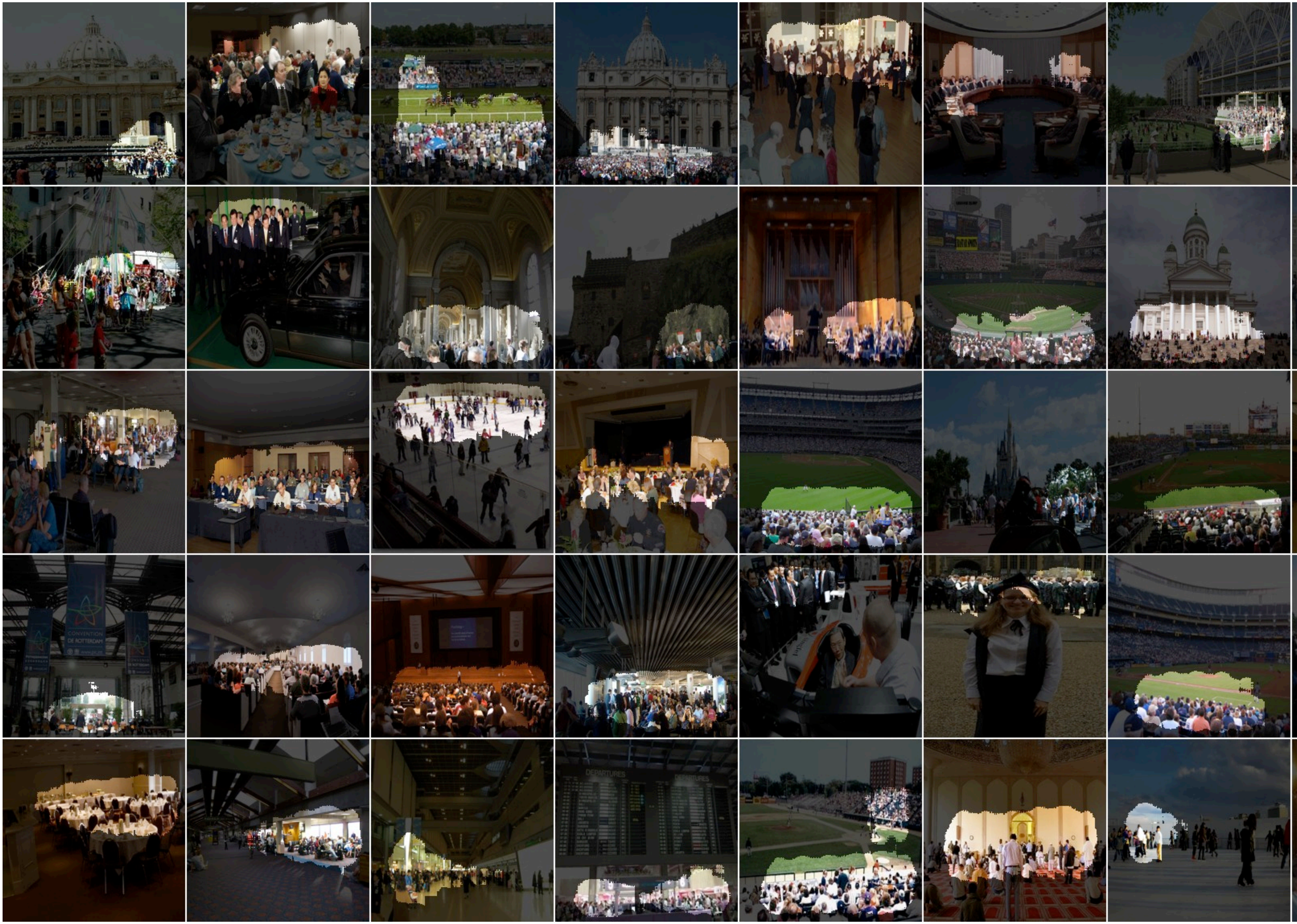


Object detectors emerge within CNN trained to classify scenes, without any object supervision

NYPD
PEARL
POPCORNERS
POLAROID LAND CAMERA




SPORT


CONVENTION
DE ROTTERDAM
DEPARTURES
DEPARTURES