Encoding and decoding brain activity

Bertrand Thirion,
Parietal team, INRIA Saclay-Île-de-France
DSV, I2BM, CEA, Neurospin
bertrand.thirion@inria.fr

Syllabus

- Sketch the goals of cognitive neuroscience
- Human brain mapping through Magnetic Resonance Imaging (fMRI)
- Encoding models: representations of stimuli comparable to brain activity
- Decoding or stimulus identification from brain activity
- Example of vision

Outline

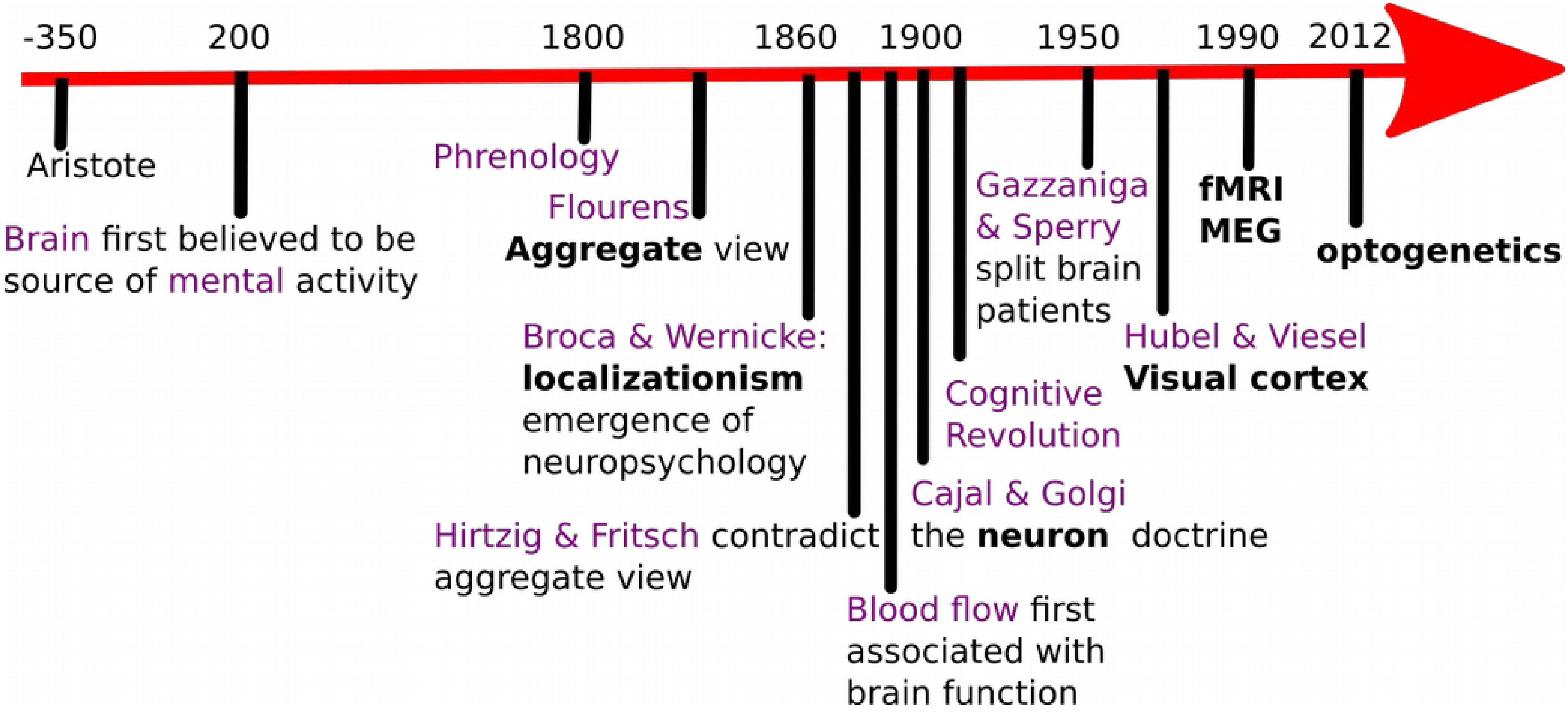
- A brief history of cognitive neuroscience
- functional Magnetic Resonance Imaging: the BOLD contrast
- Encoding vision models for fMRI
- Decoding visual percepts

Cognitive neuroscience

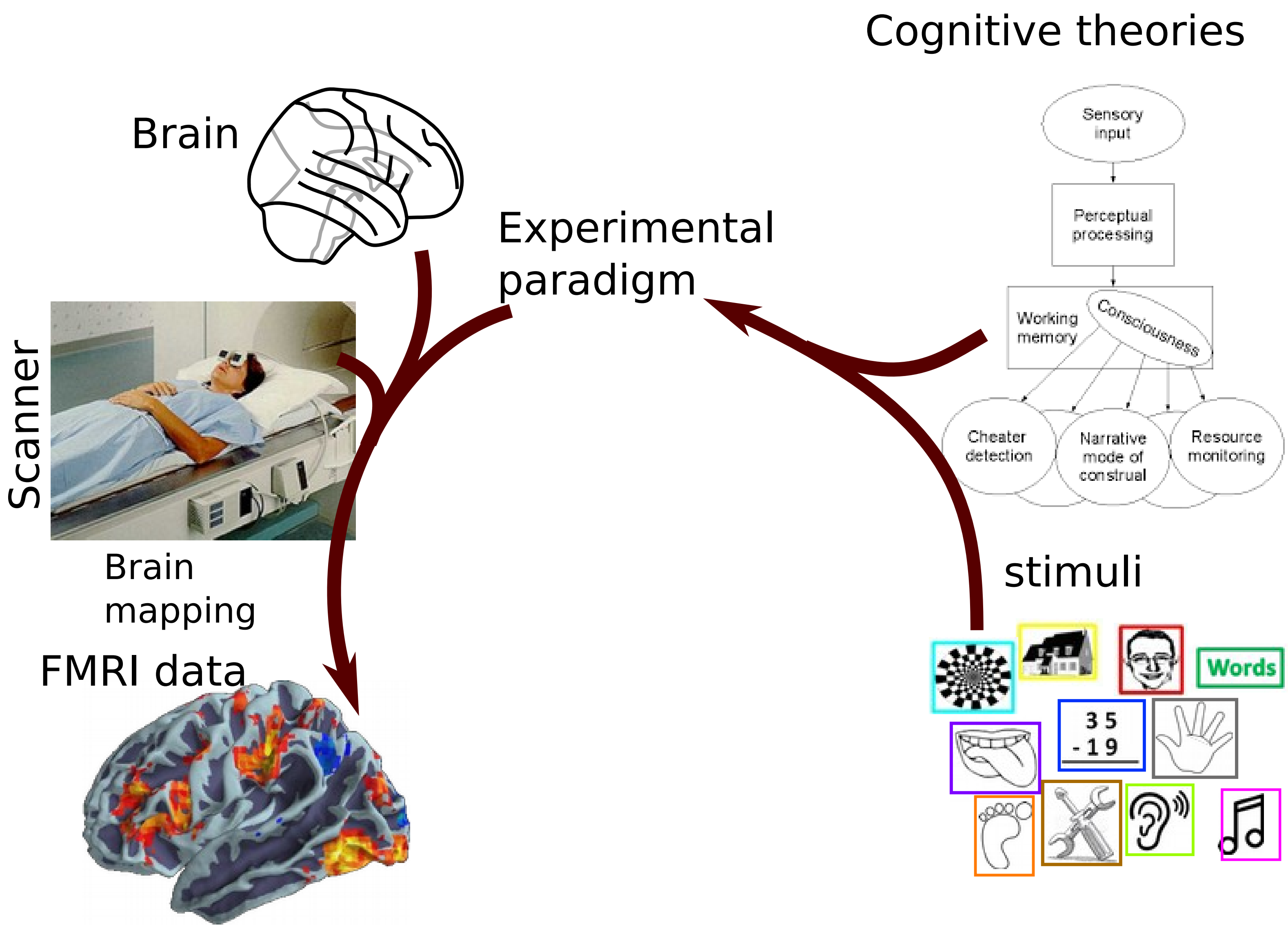
How are cognitive activities affected or controlled by neural circuits in the brain ?

Cognitive neuroscience

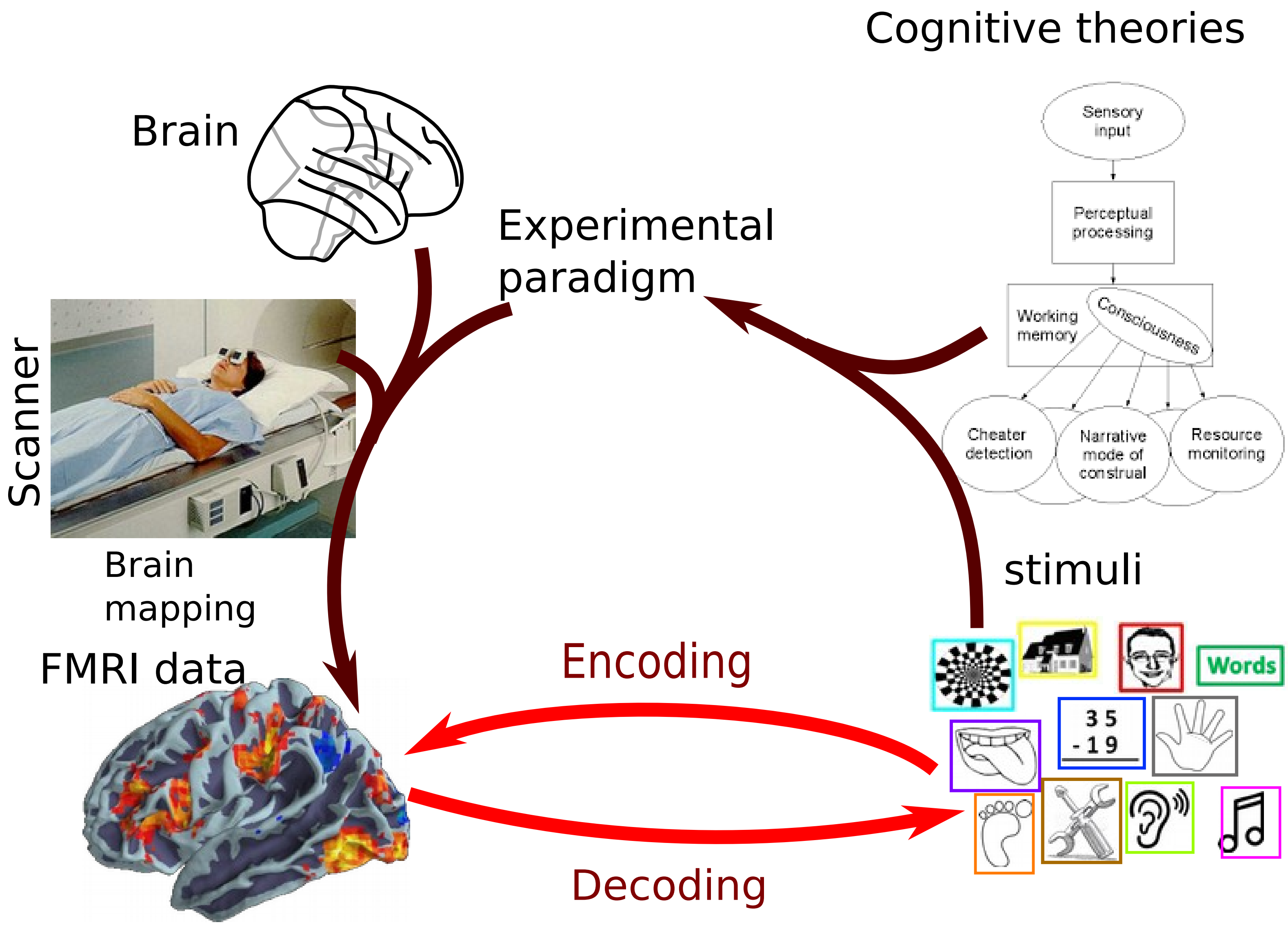
How are cognitive activities affected or controlled by neural circuits in the brain ?



The brain, the mind and the scanner



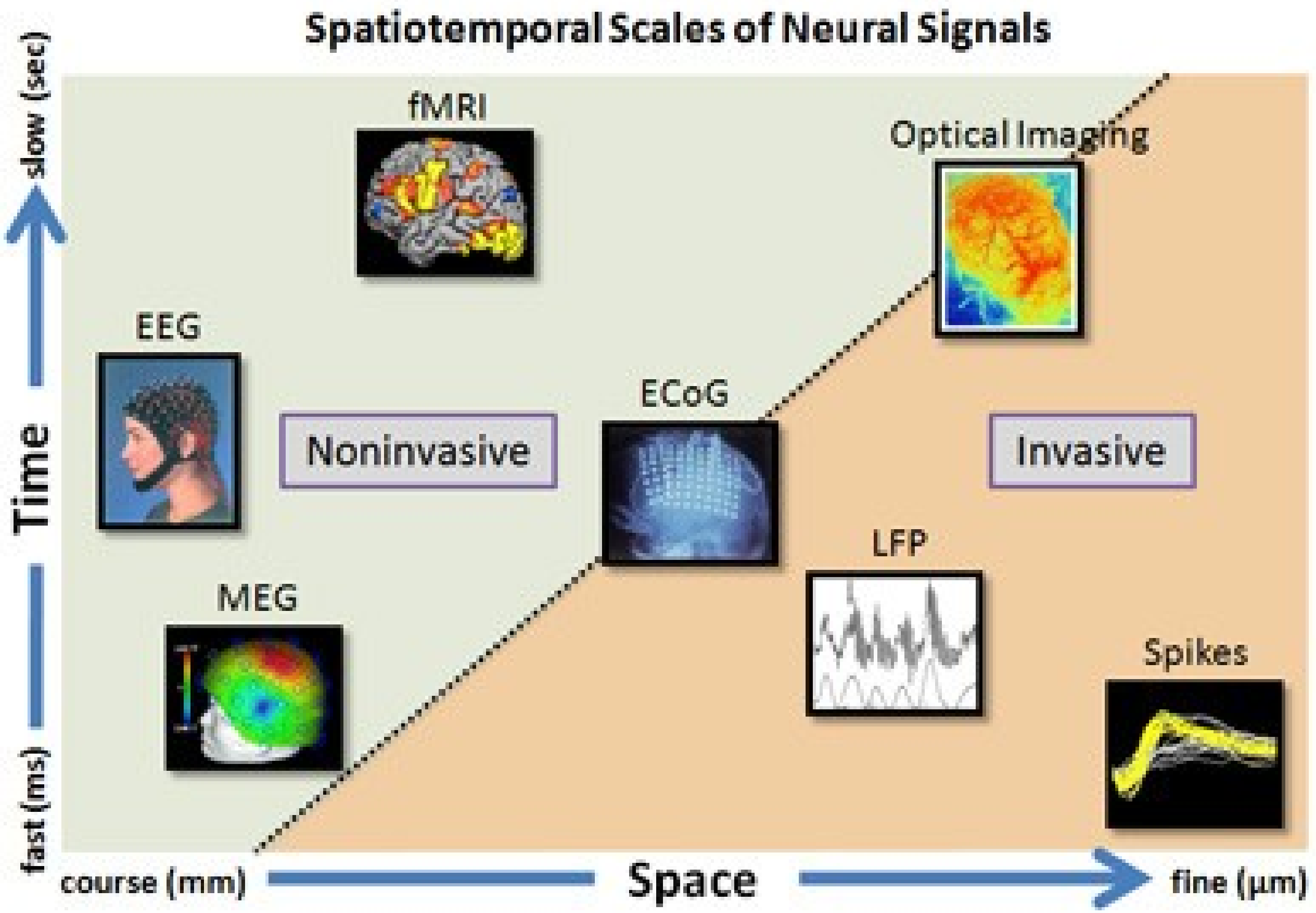
The brain, the mind and the scanner



Outline

- A brief history of cognitive neuroscience
- functional Magnetic Resonance Imaging: the BOLD contrast
- Encoding vision models for fMRI
- Decoding visual percepts

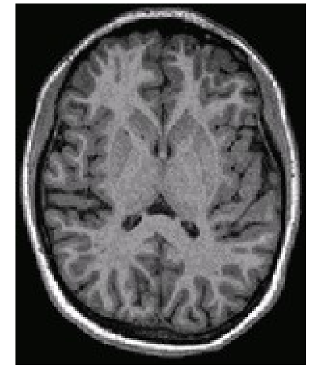
The resolution of different functional brain imaging modalities



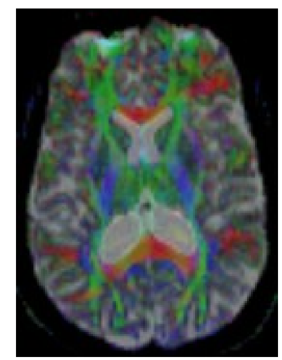
http://lifesciences.ieee.org/publications/newsletter/april-2012/96-building-brain-machine-interfaces-neuroprosthetic-control-with-electrocorticographic-signals

Magnetic Resonance Imaging

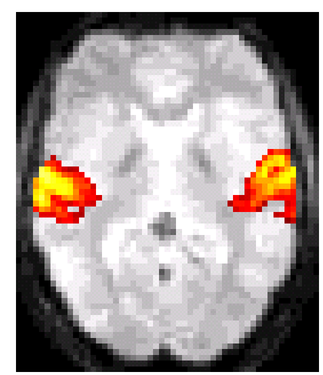
Magnetic Resonance Imaging



(T1)
Anatomical MRI



Diffusion-weighted MRI

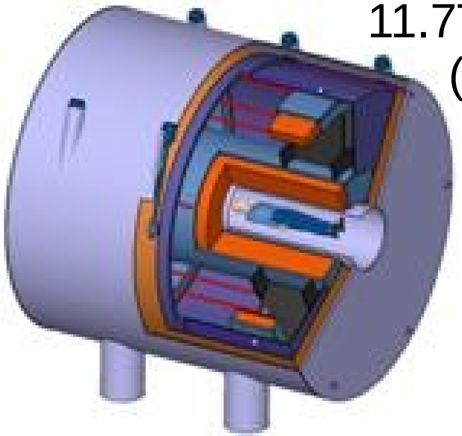


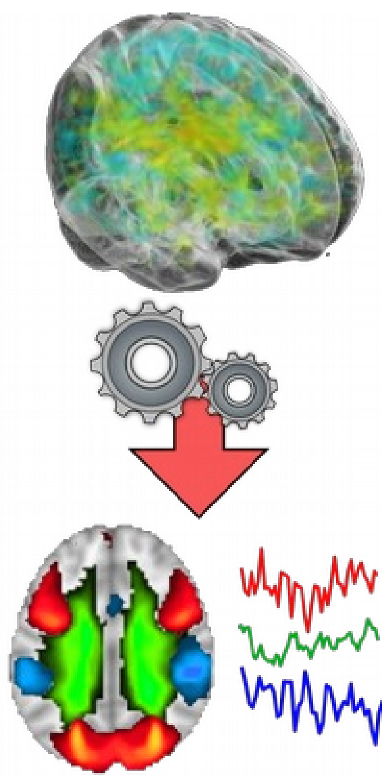
functional MRI

Neuroimaging technologies

7T MRI (2007)

11.7T MRI (2022)



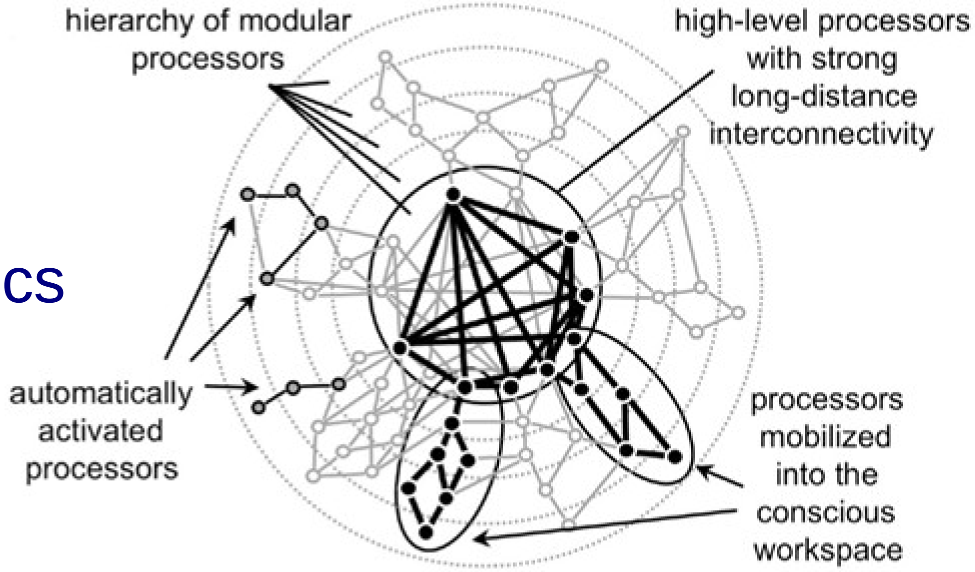


Physics (RMN, EM)

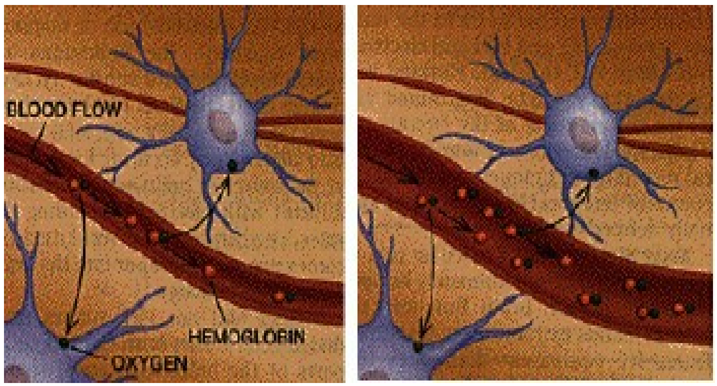
medical diagnosis

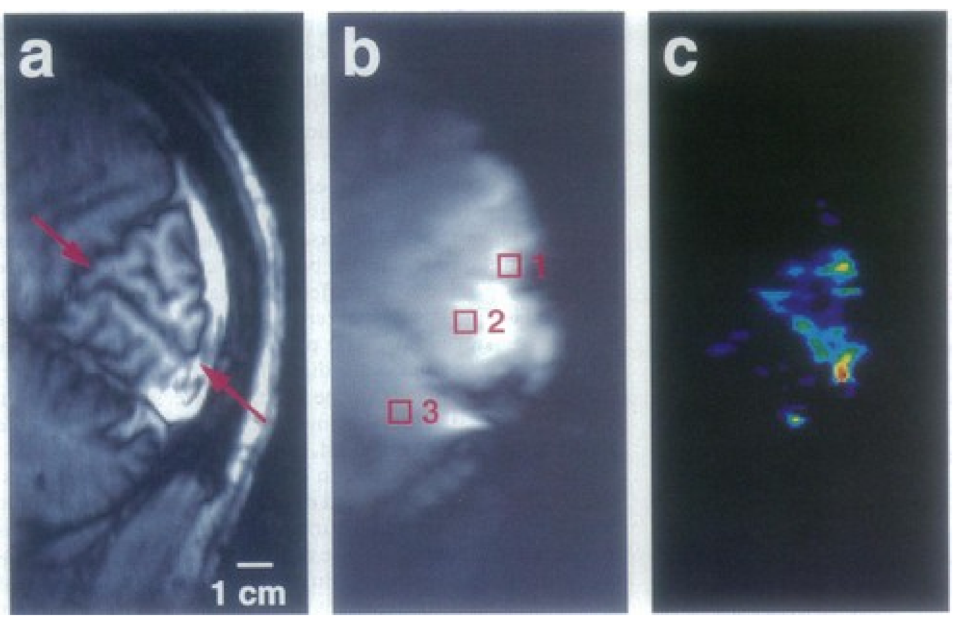
cognitive sciences

modeling and informatics



BOLD imaging





Experimental paradigm

↓

Neural activity

↓

Massive local oxygenated blood inflow (5s)

↓

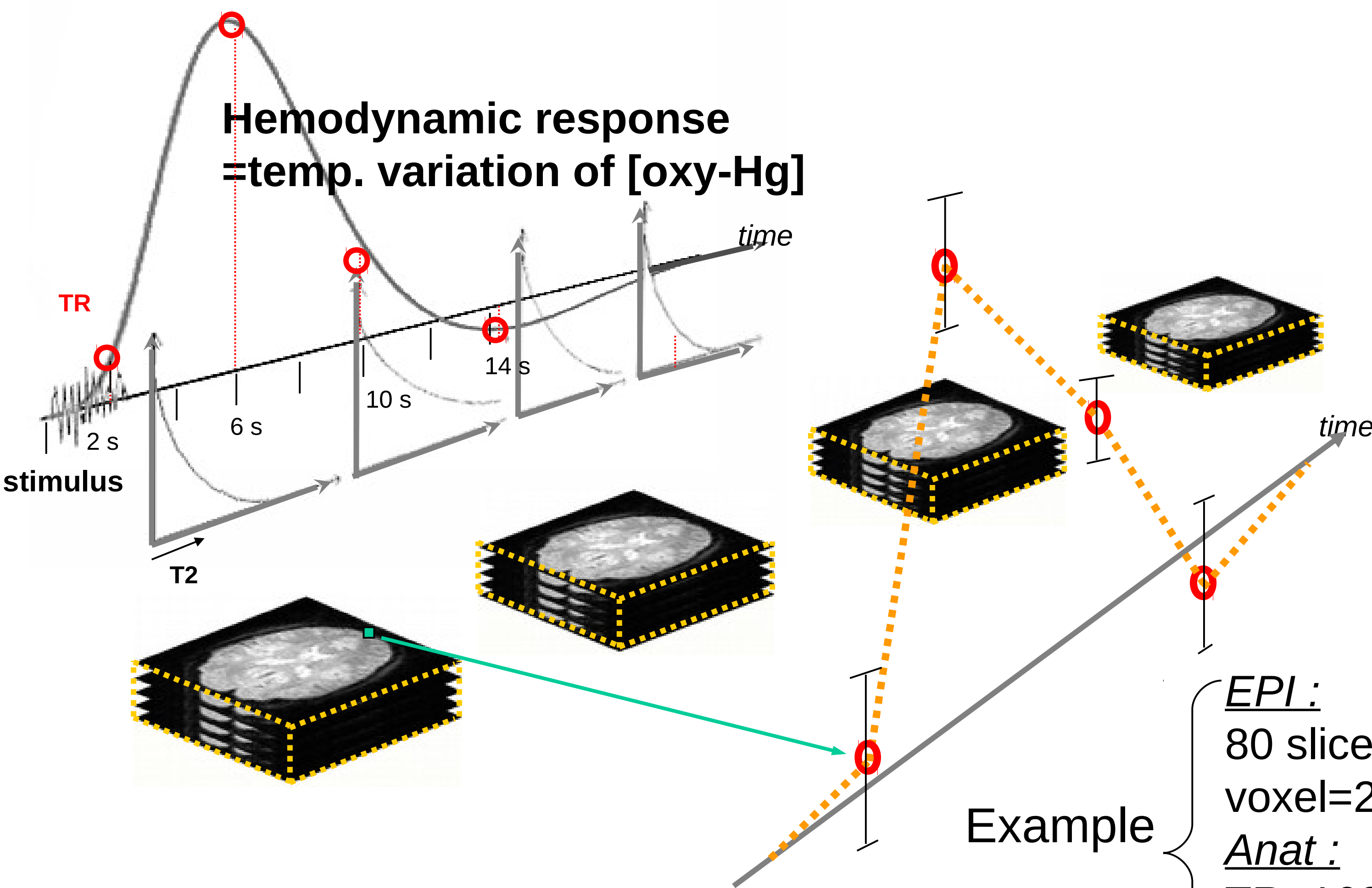
MR signal increase

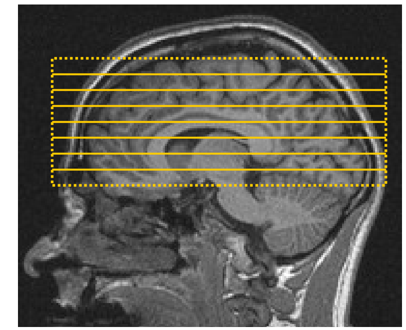
BOLD: blood oxygen-level dependent

Oxyhemoglobin diamagnetic
≠
deoxy-hemoglobine

[Ogawa et al, PNAS, 1992]

Sampling the BOLD response with fMRI





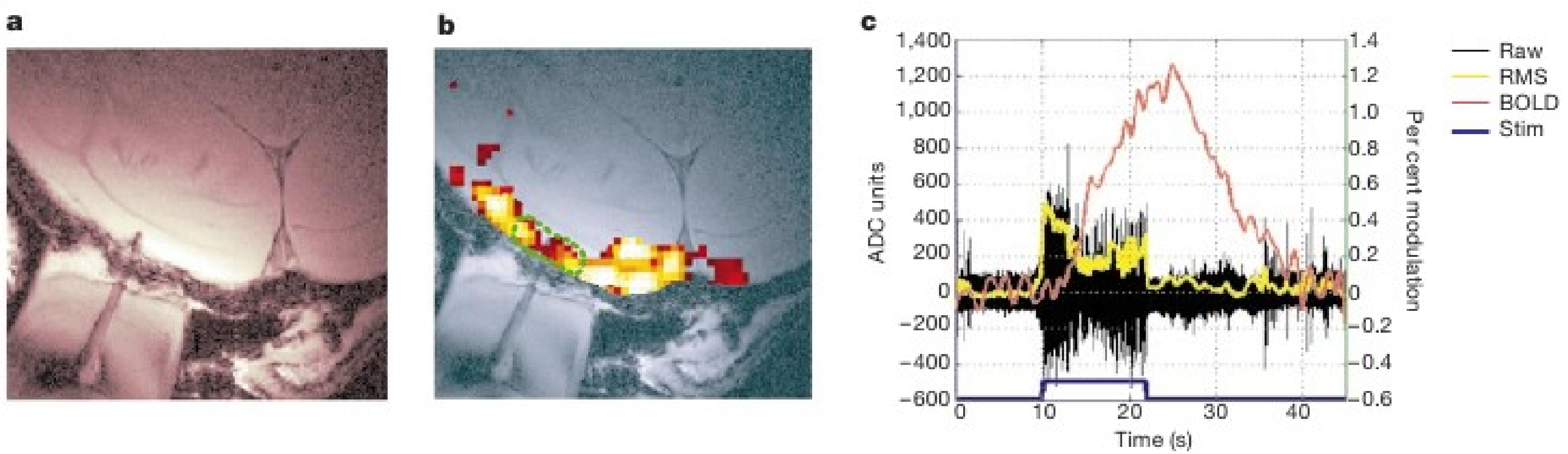
Example

EPI :
80 slices, TR=2000 ms, TE=30ms,
voxel=2x2x2mm, 64x64 matrix

Anat :
TR=1600ms, FOV=256x256mm^2,
256X256 matrix, 192 slices
Slice thickness 1mm

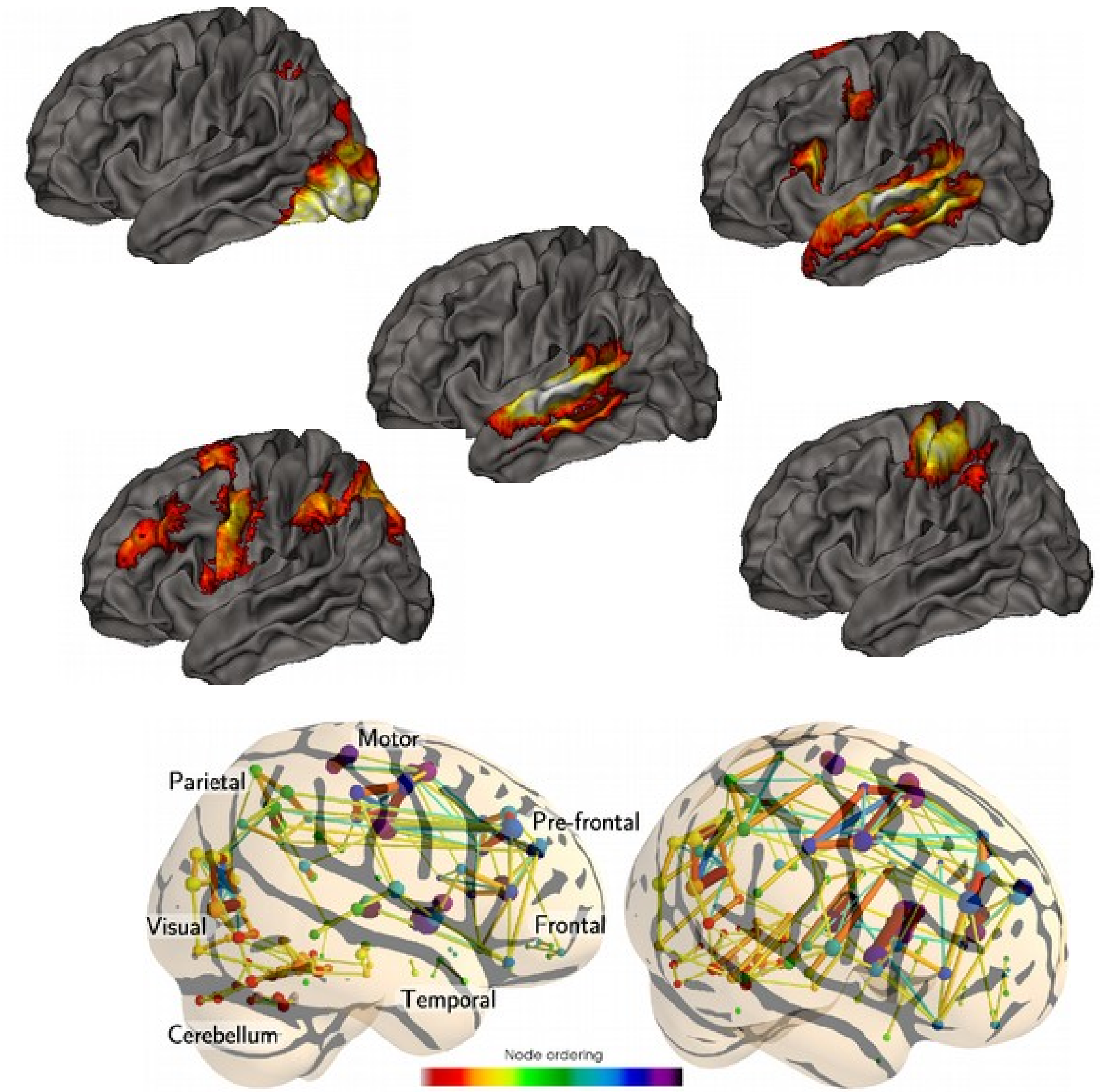
27 years of human BOLD imaging

- BOLD response approximately linear in the stimulus function
  - Application of simple linear model for data analysis [Friston et al. 1995]
- BOLD signal highly correlated with LFPs [Logothetis et al. Nature 2001]
- High spatial accuracy (~2mm) [Ugurbil et al. NeuroImage 2007]
- Poor temporal resolution, no consensual model on the signal

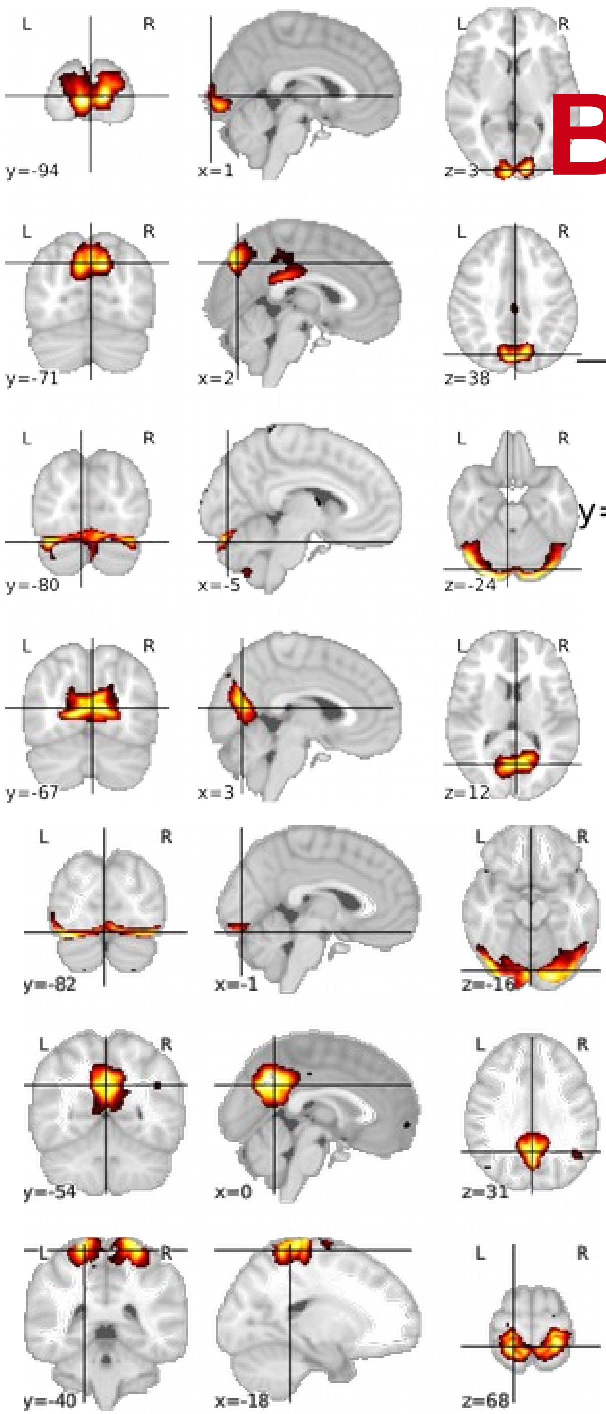


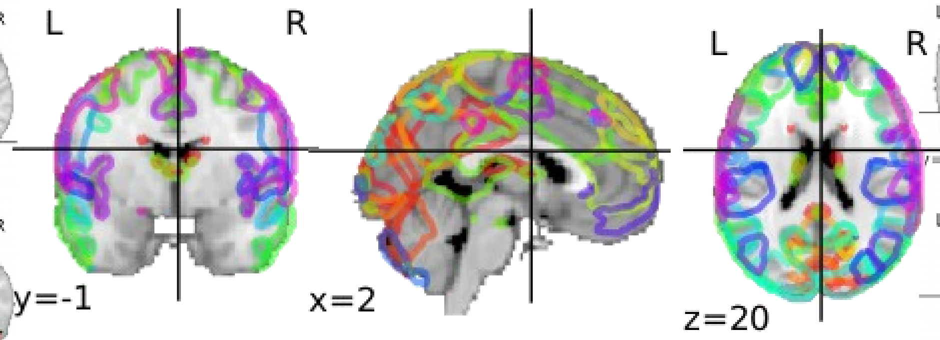
27 years of Human BOLD imaging

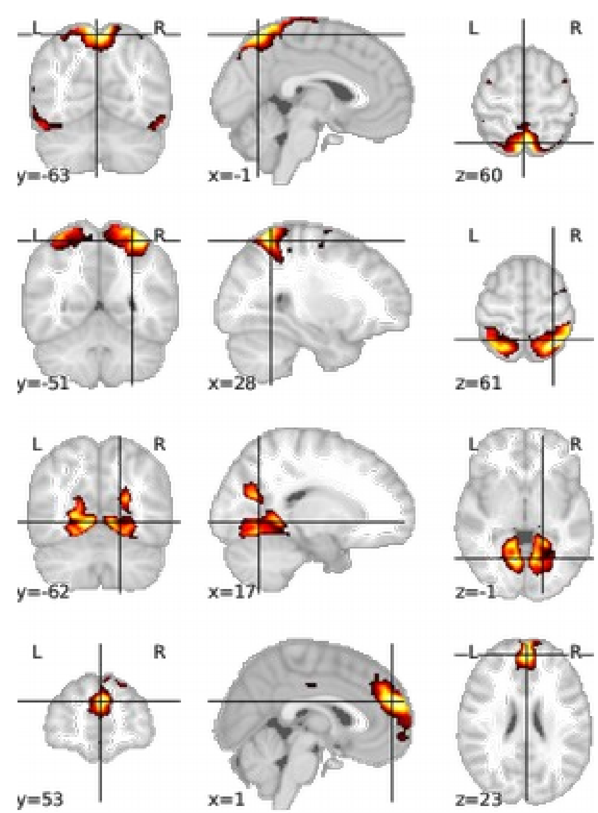
- map cognitive functions in the human brain & identify brain regions
  - Segregation principle
- Network models / connectivity analysis
  - Integration principle
- Prediction of behavior / psychological tasks
  - “brain activity decoding” or diagnosis

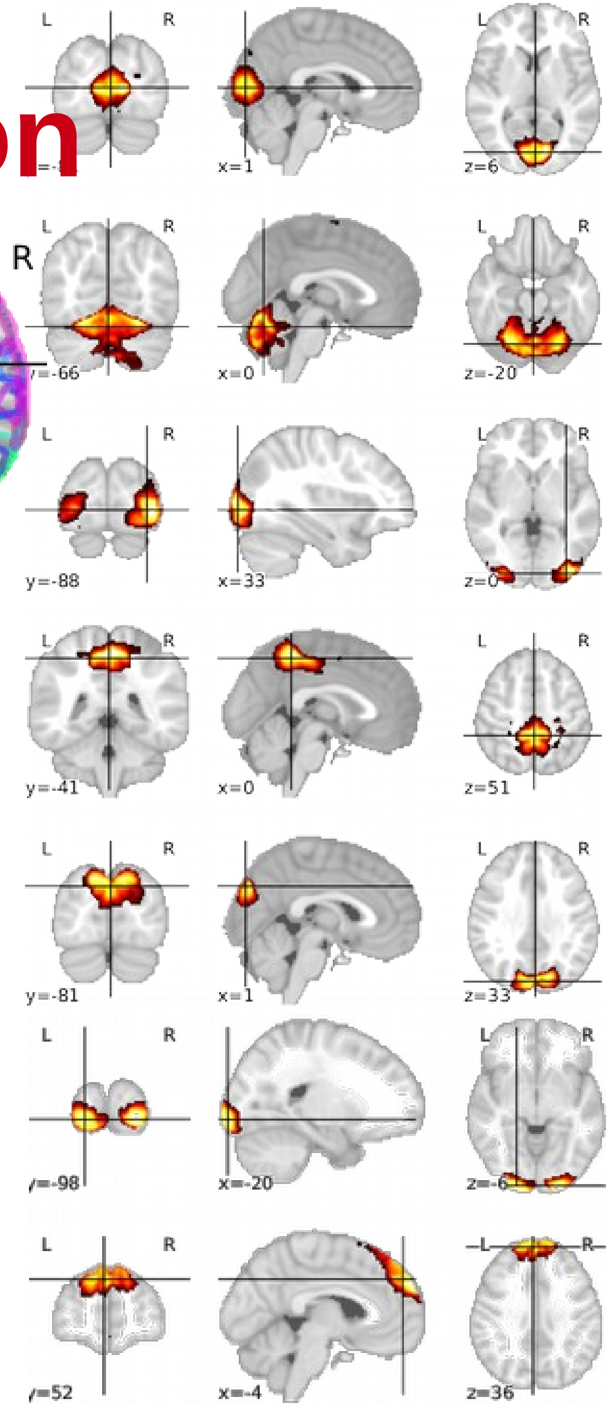


Brain segregation

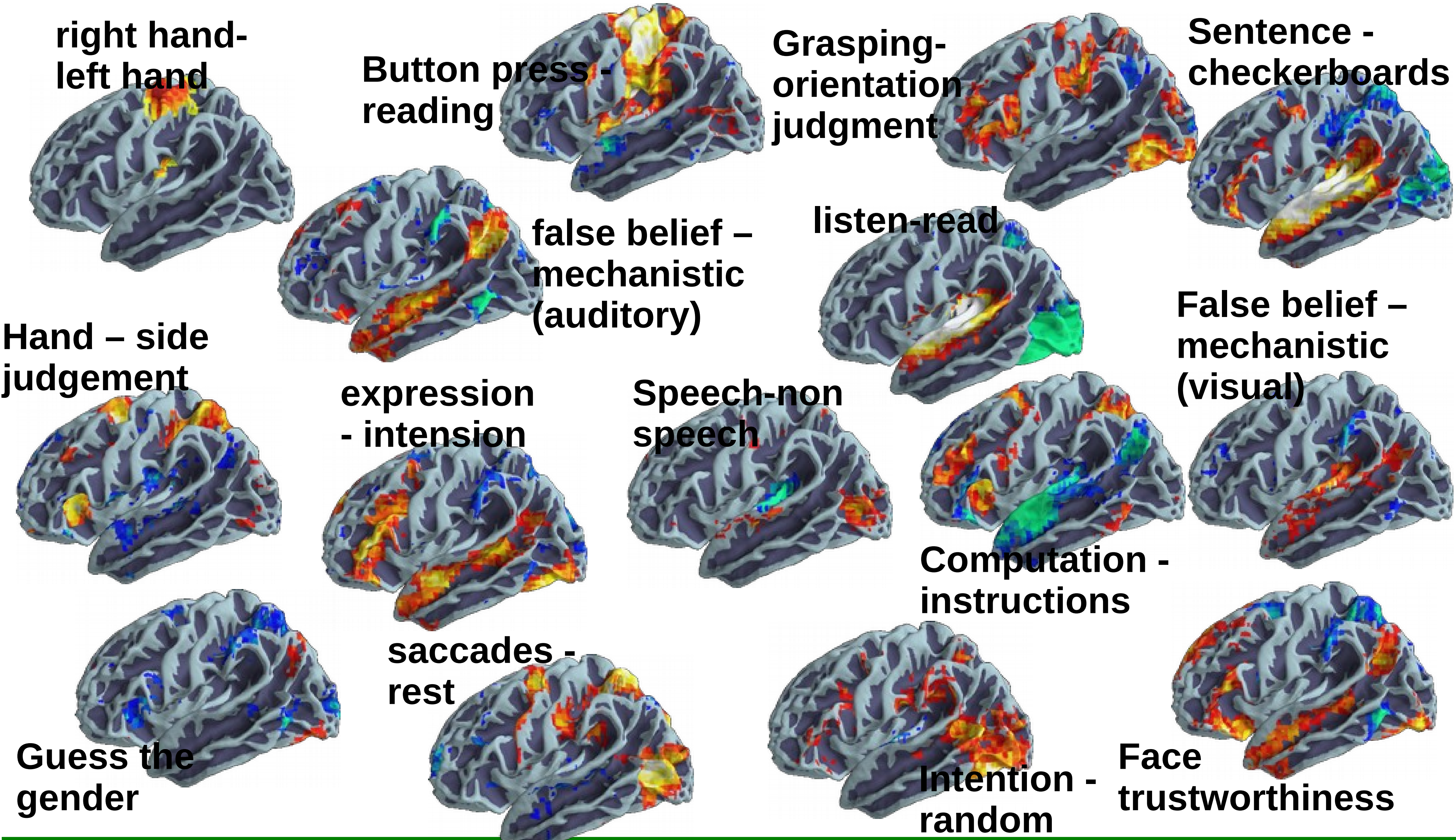




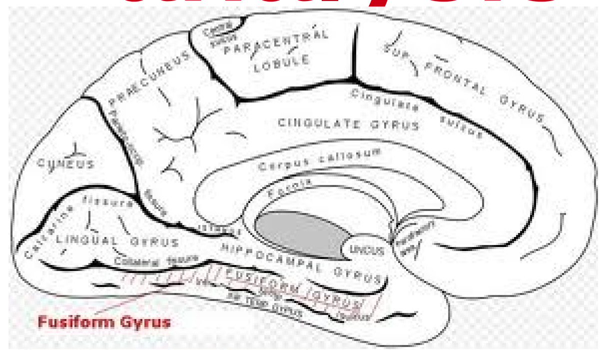


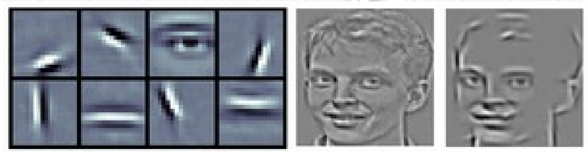


Brain mapping...

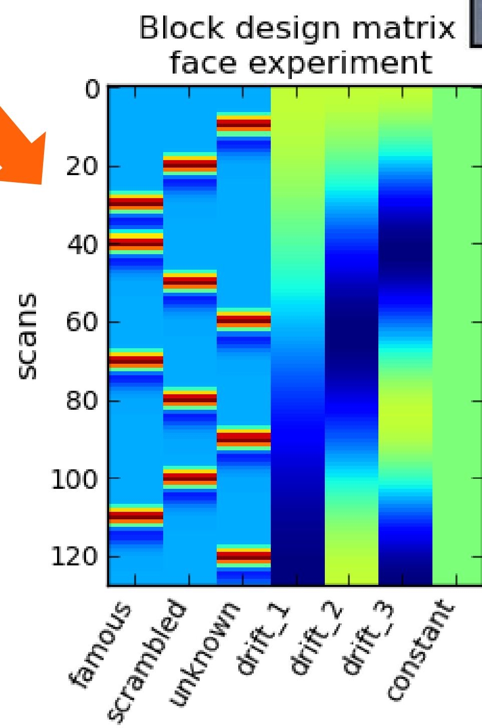


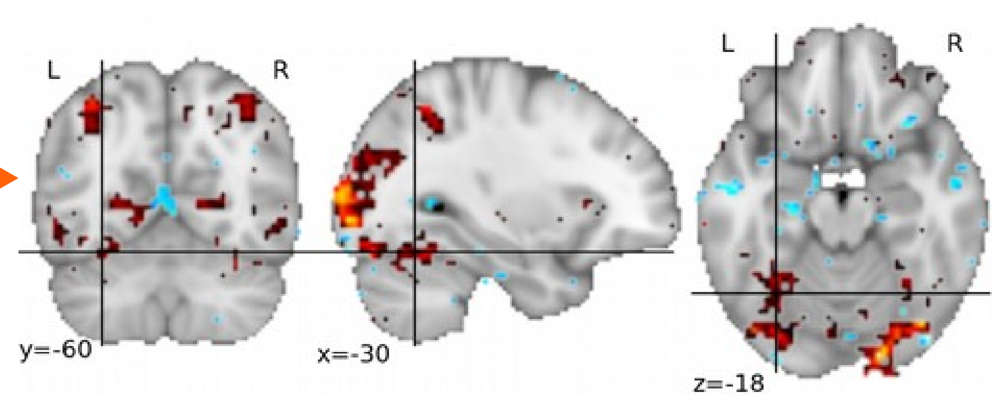
fMRI data from acquisition to analysis





Complex metabolic pathway

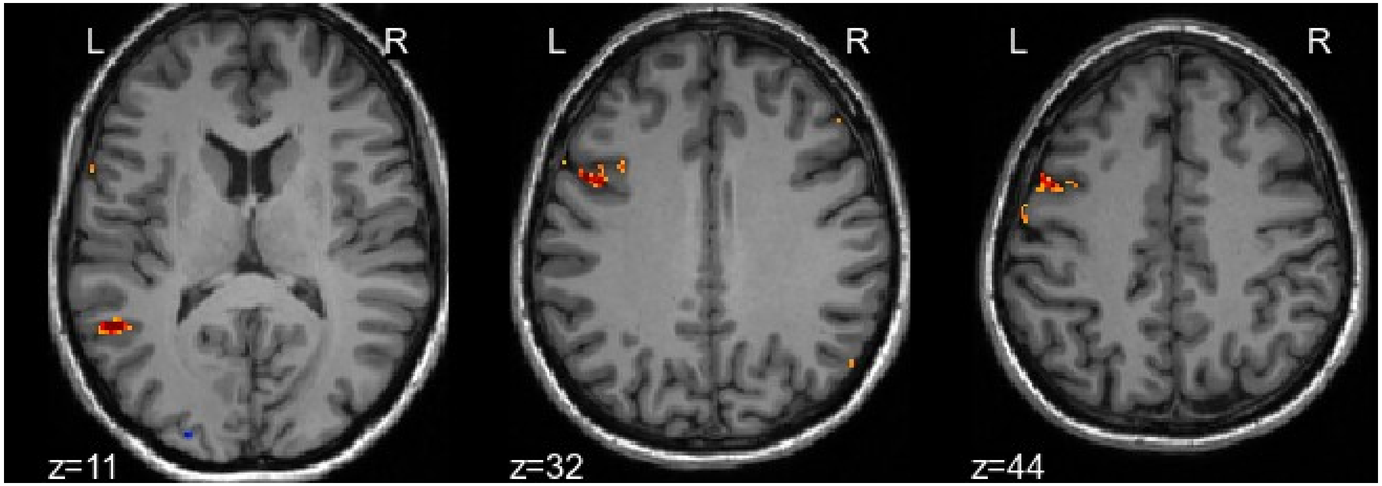




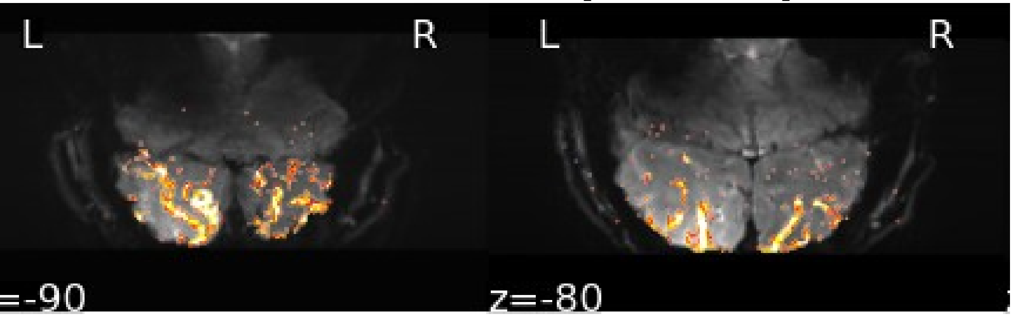
Raw fMRI data are hard to interpret

Towards high-resolution MRI

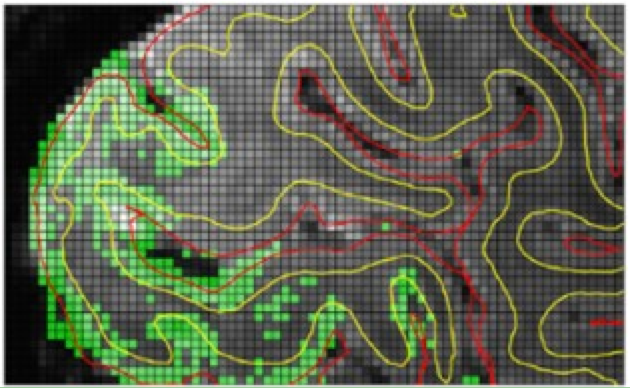
- Current resolution: moving to 1.5mm isotropic (3T)



- 1mm at 7T (2010)



Current: sub-millimeter (7T) [Lawrence et al., nimg 2017]



Is BOLD fMRI sufficient to study human vision ?

- Advantages
  - Full brain coverage
  - Non invasive, cheap
  - Relatively good spatial resolution ...
- Limitations
  - … Not enough for cortical column mapping
  - Temporal resolution too coarse for activity chronometry
  - No access to neural mechanisms

Outline

- A brief history of cognitive neuroscience
- functional Magnetic Resonance Imaging: the BOLD contrast
- Encoding vision models for fMRI
- Decoding visual percepts

Encoding visual stimuli

- Brain activation encodes some information about the stimuli
- Population receptive field
  - Location-specific (depends on where you are on the cortex)
  - Feature-related (position, contrast, speed, color...)

$$\mathbf{y} = \phi(\mathbf{X})\beta + \varepsilon$$

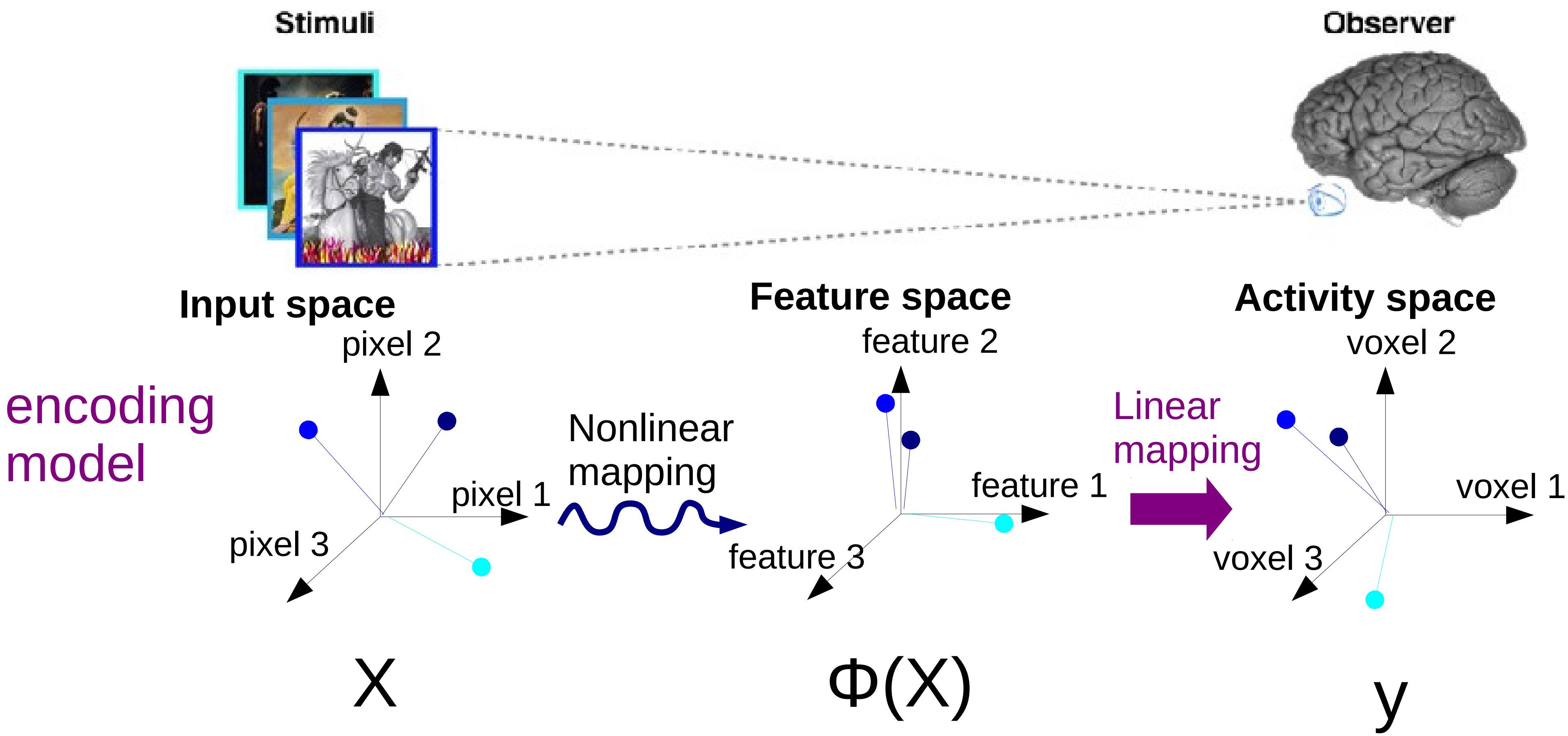
e.g. ϕ = wavelet transform

Signal in a voxel

Input image

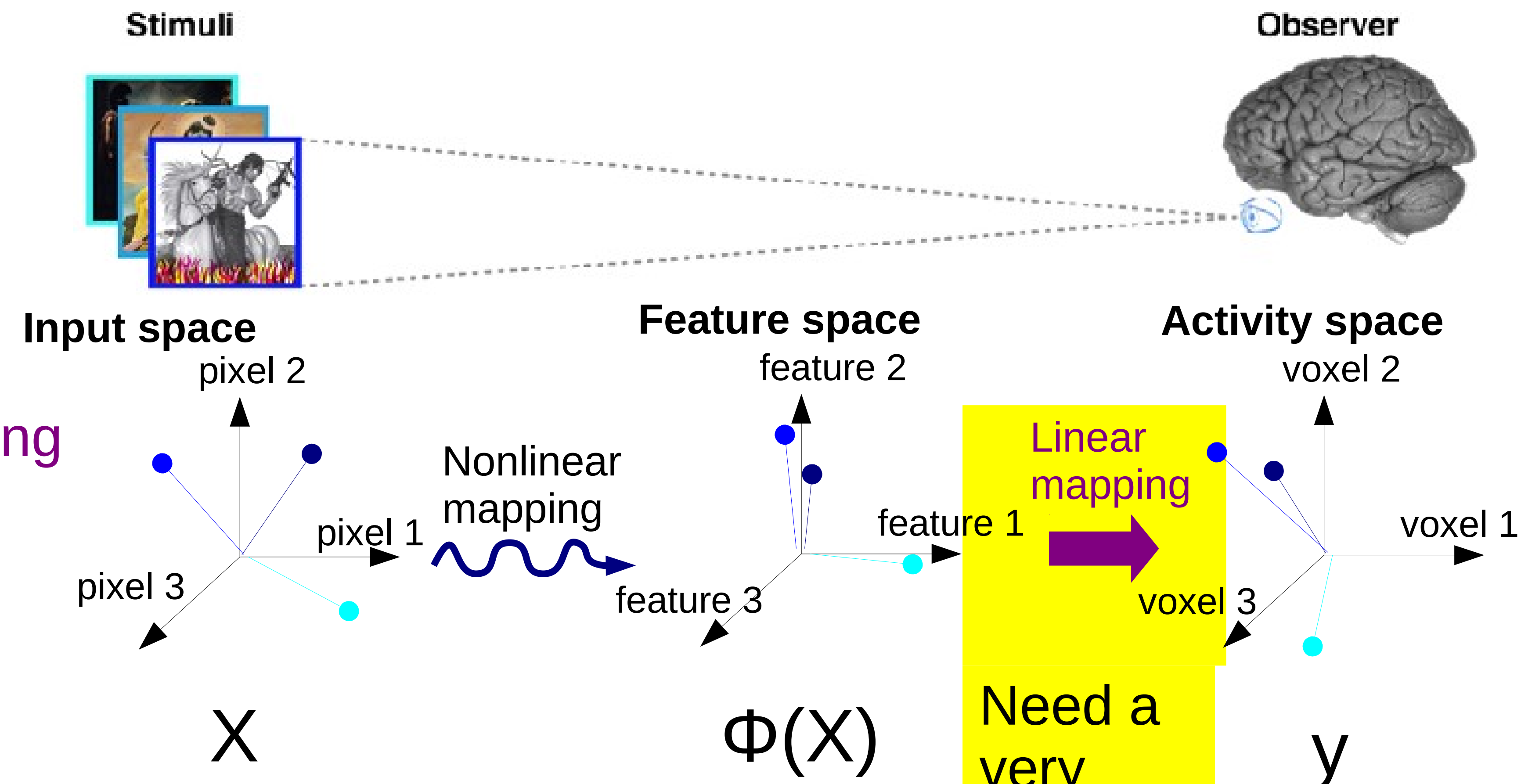
Voxel-specific response

Encoding: modeling challenge



Encoding: modeling challenge

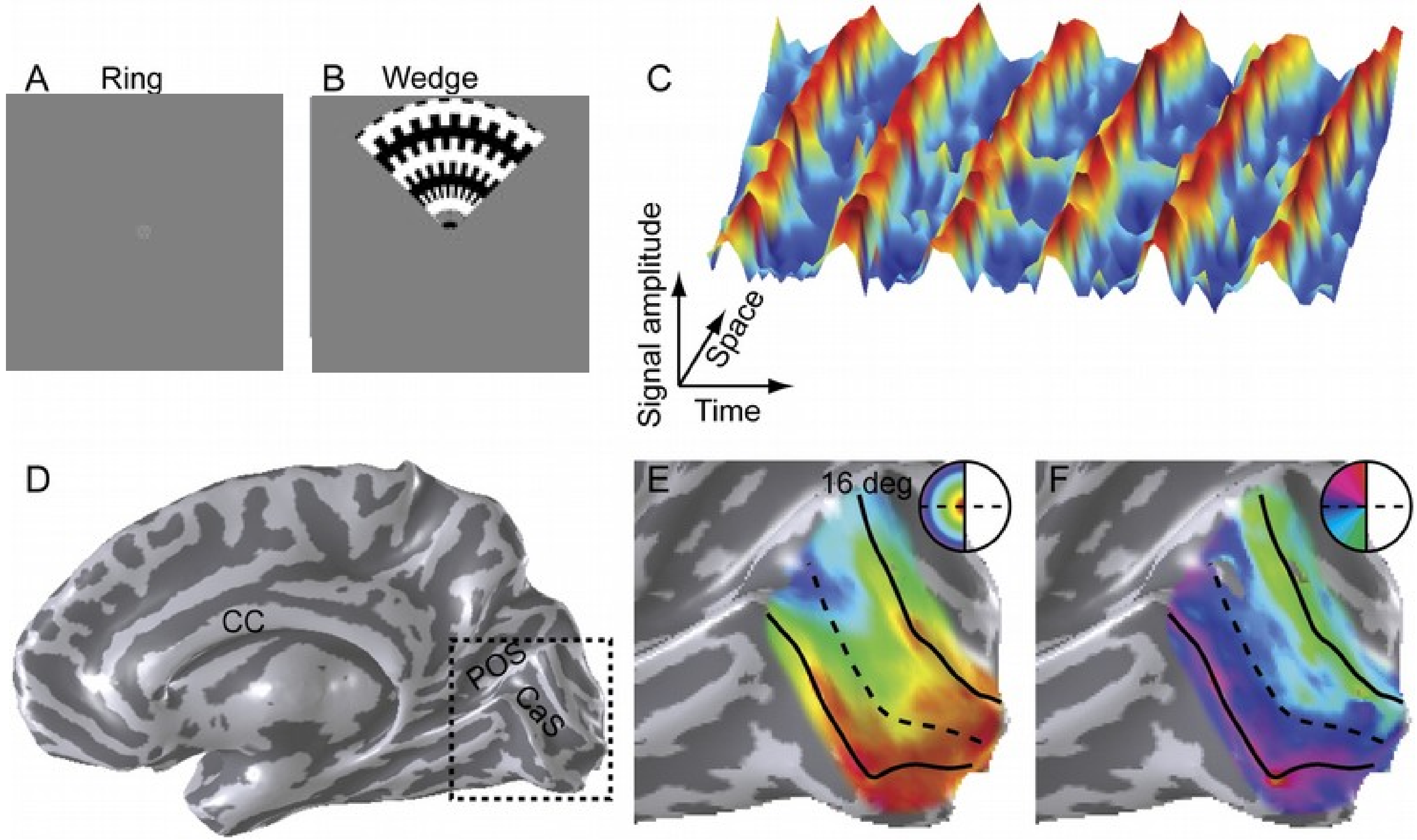
encoding model



Need a very efficient learner

Visual encoding w/ fMRI: retinotopy

- In early visual areas brain voxels respond to location-specific contrasts
- Traveling-wave stimuli map the visual field to visual cortex



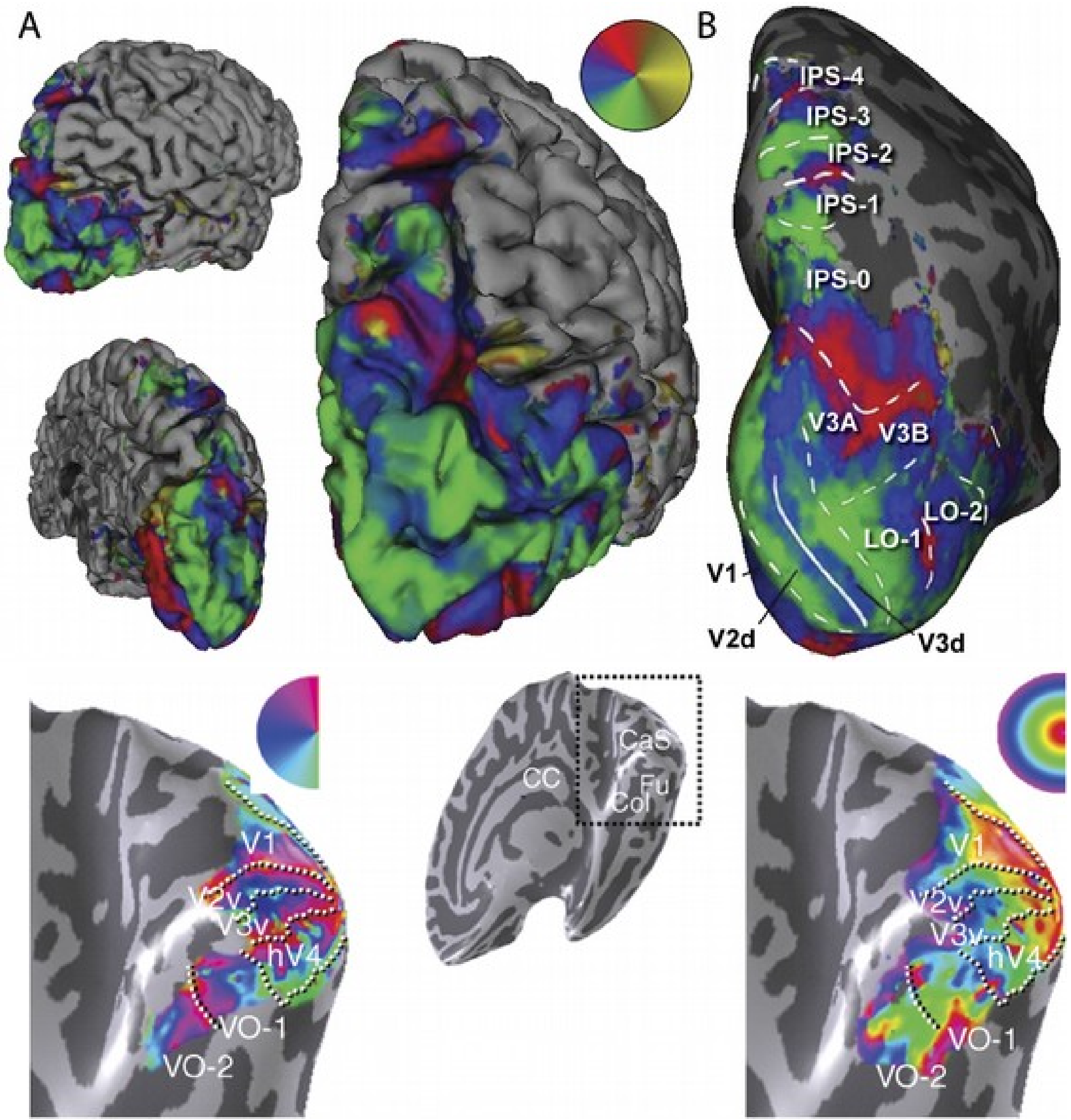
- The phase of the BOLD response at the reference frequency varies smoothly across the cortical surface (space)

[Sereno et al. 1994,...,Wandell et al. 2007]

Accessing visual encoding w/ fMRI: retinotopy

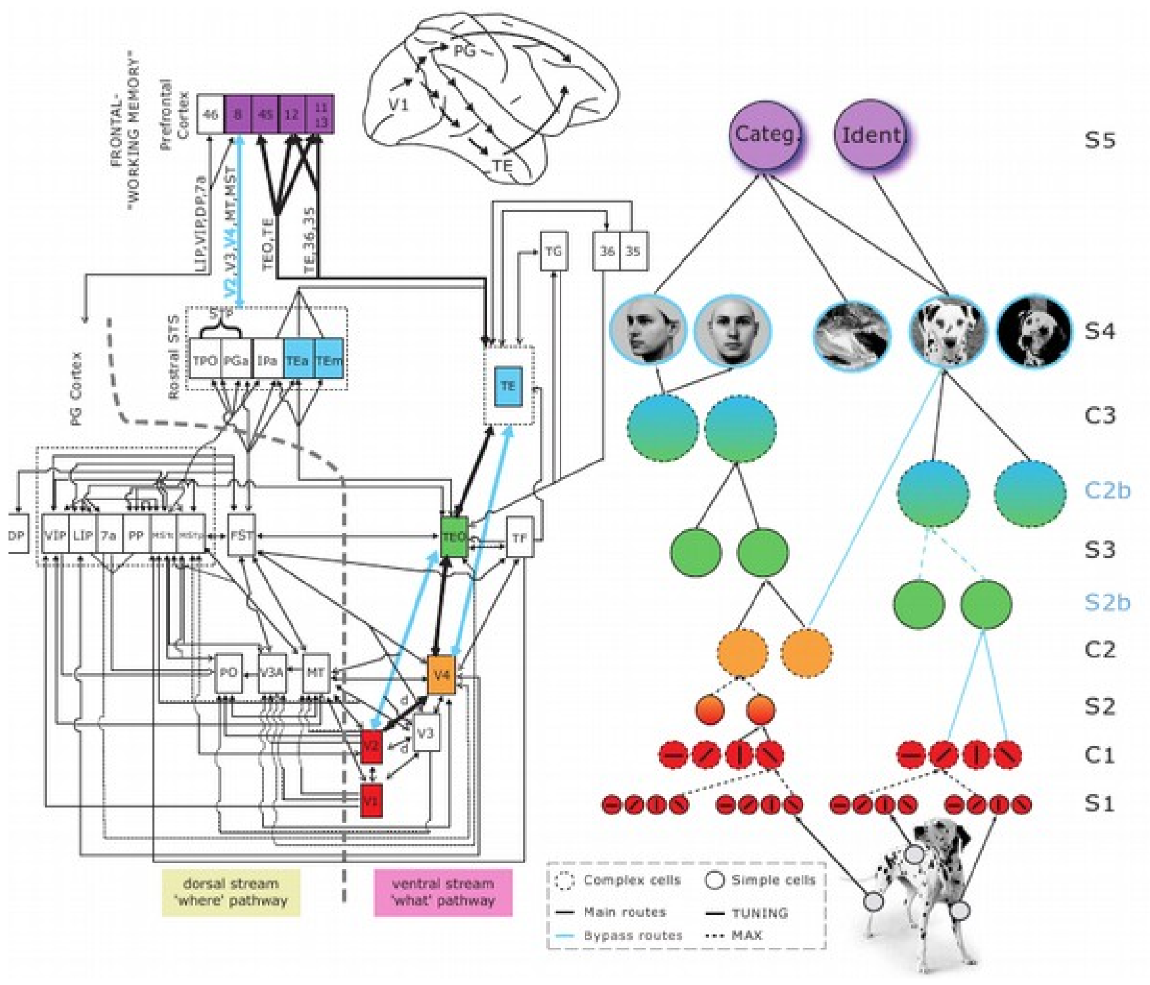
- Retinotopy provides the main access to the definition of visual cortical areas

  [Sereno et al. 1994, Wandell et al. 2007...]

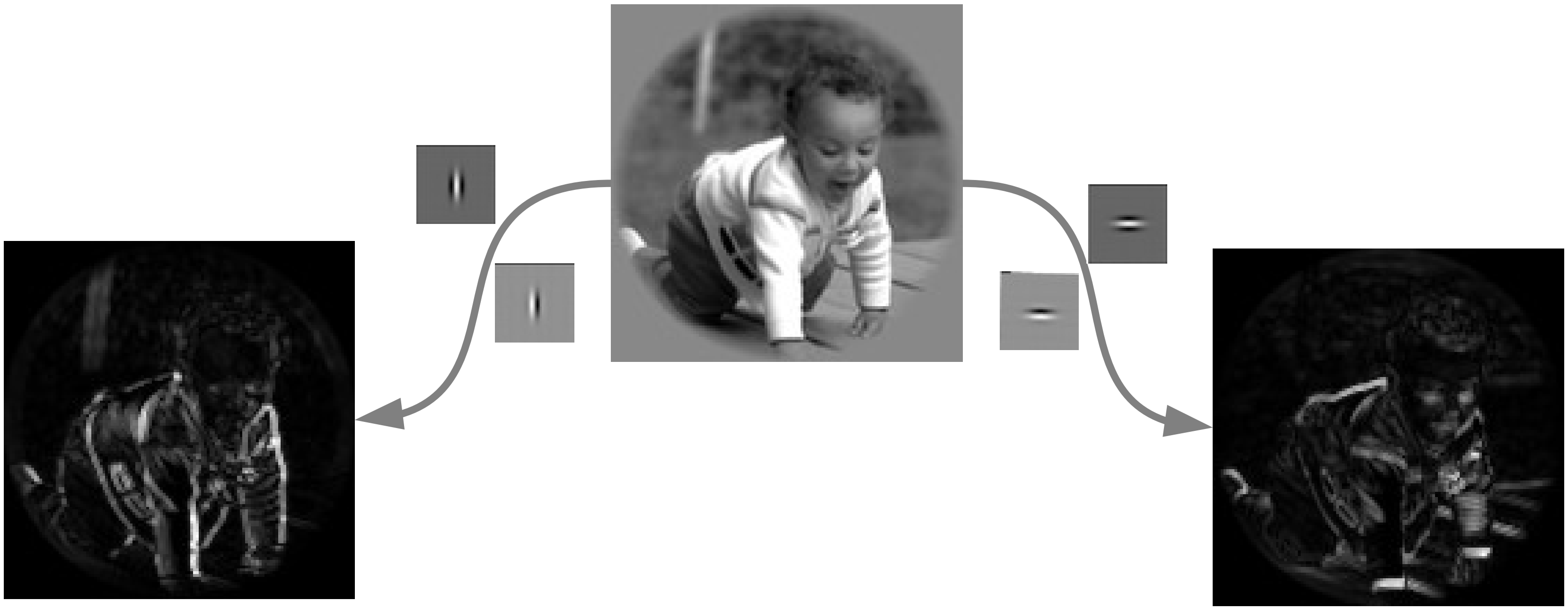


Neurocomputational models of vision

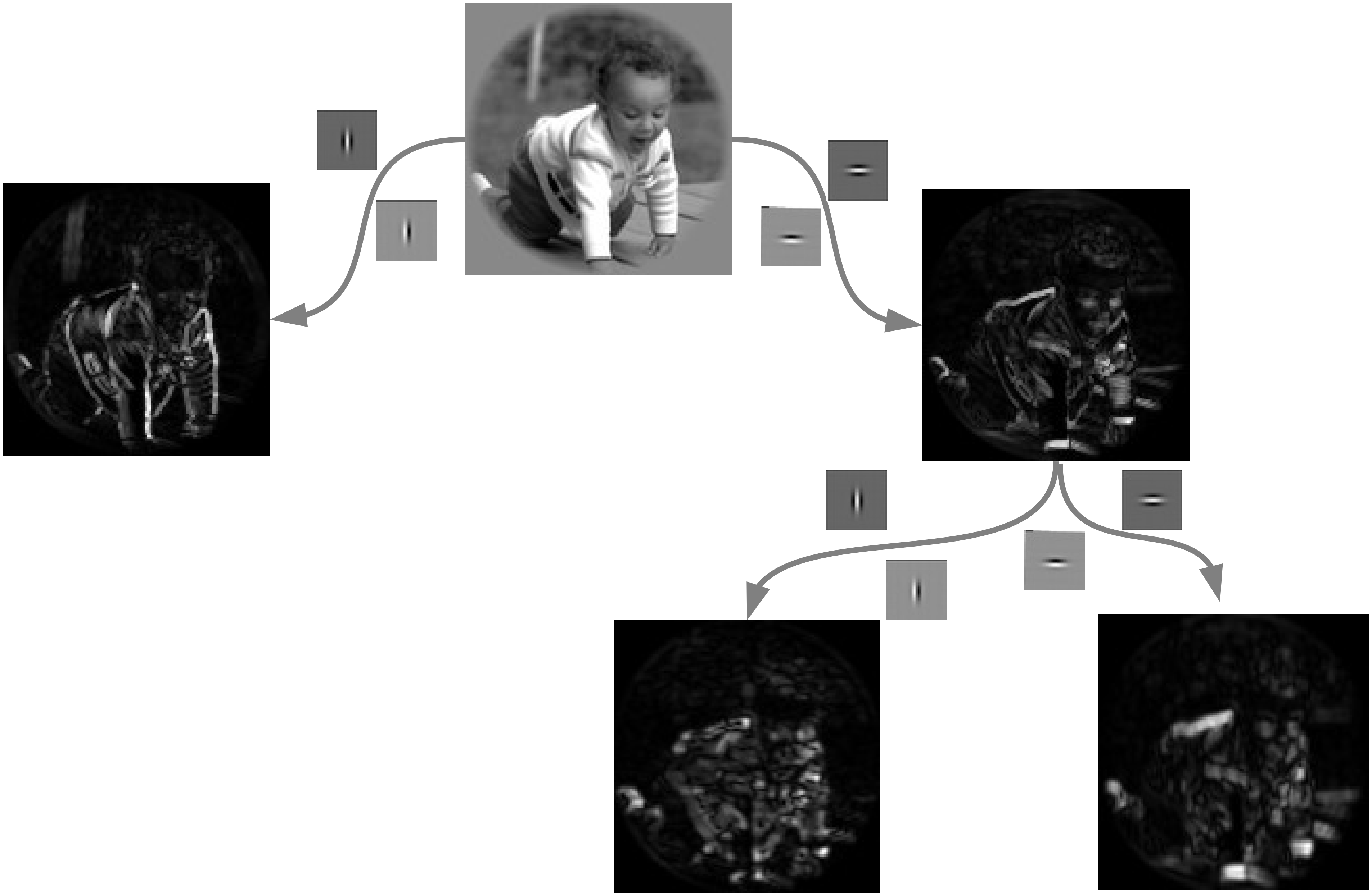
- “feedfoward architectures”:
  - Local operations (convolutions)
  - Non-linear operation
- **Hmax** [Riesenhuber and Poggio, 1999, Serres et al. 2007]
- scattering transform,
- convolutional networks



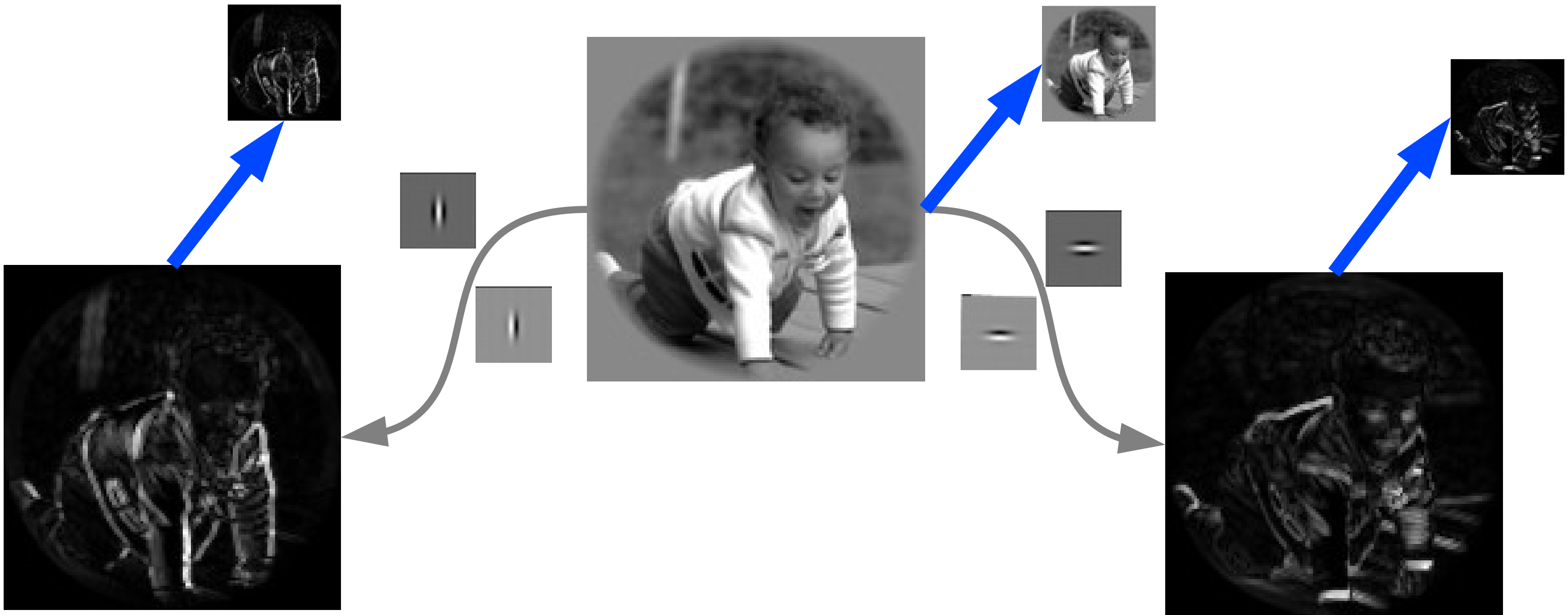
The Scattering Transform

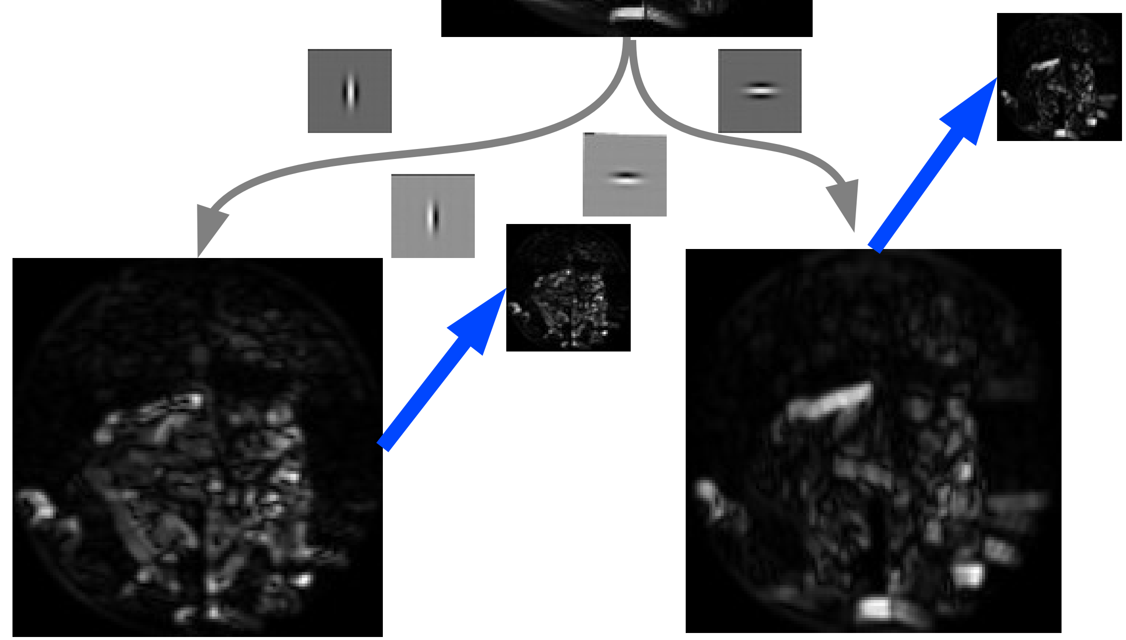


The Scattering Transform



The Scattering Transform

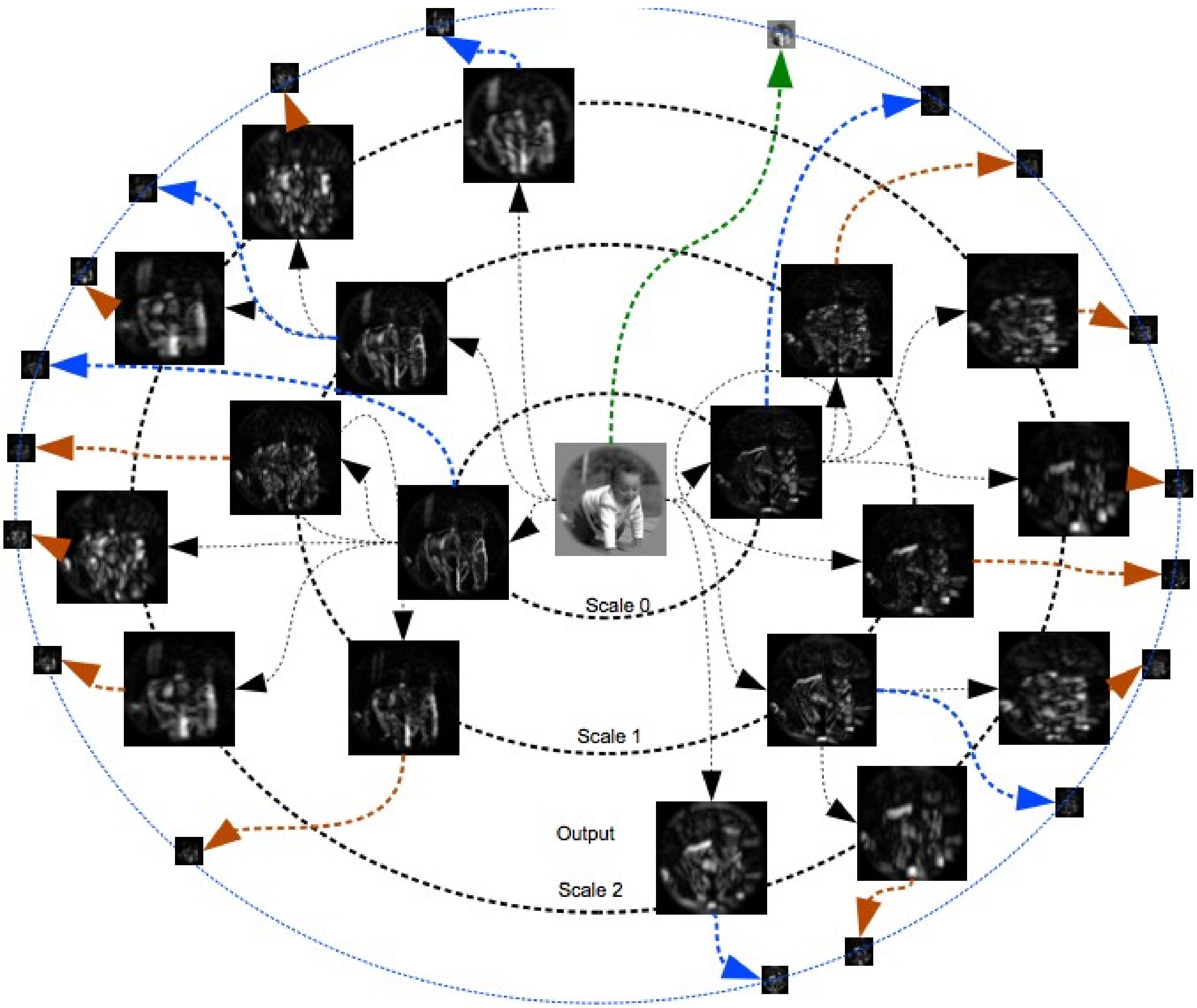




Joan Bruna and Stéphane Mallat. Invariant Scattering Convolution Networks, PAMI 2012

Laurent Sifre and Stéphane Mallat. Rotation, Translation and Deformation Invariant Scattering, CVPR 2013

The Scattering Transform



Joan Bruna and Stéphane Mallat. Invariant Scattering Convolution Networks, PAMI 2012

Laurent Sifre and Stéphane Mallat. Rotation, Translation and Deformation Invariant Scattering, CVPR 2013

Example: visual field seen in a voxel

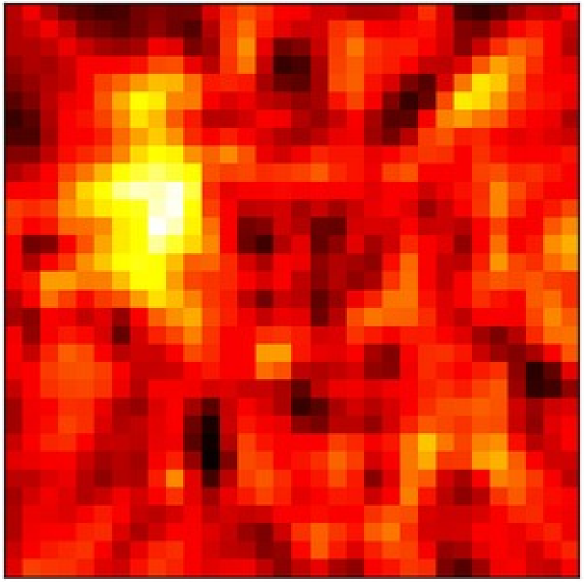
Receptive field of one voxel (obtained by ridge regression over a large set of images)

$$\mathrm{argmin}_\beta \|\mathbf{y} - \phi(\mathbf{X})\beta\|^2 + \lambda\|\beta\|^2$$

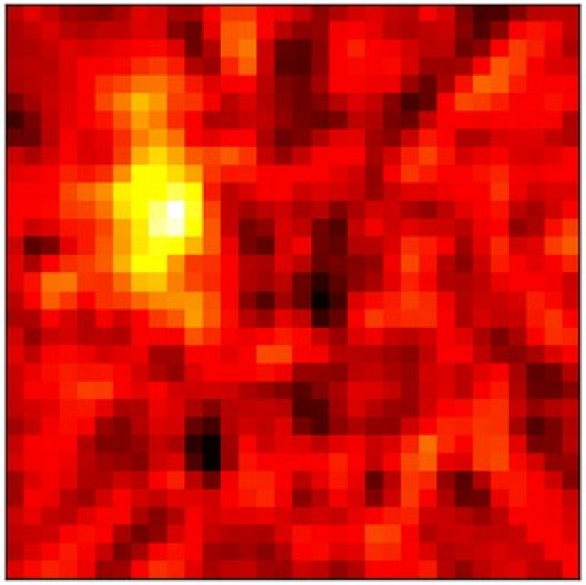
$$\mathbf{y} = \phi(\mathbf{X})\beta + \varepsilon$$

$\phi(\mathbf{X})$= scattering transform

Scattering/layers
0 and 1



Scattering/layers
0, 1 and 2

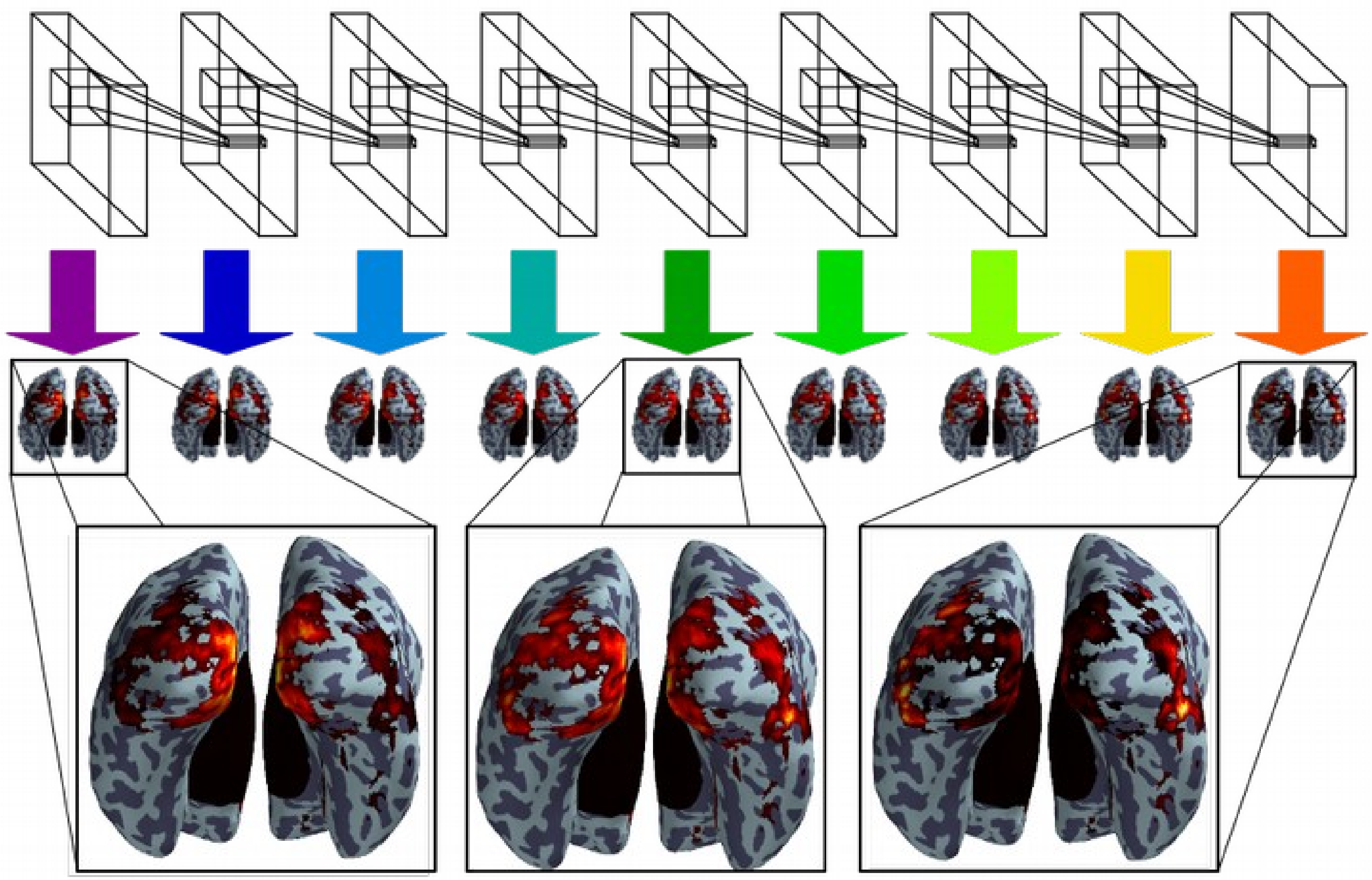


[Eickenberg et al. PRNI 2012]
[Eickenberg et al. ICML 2012 neuroimaging workshop]

Encoding visual stimuli

Create Features

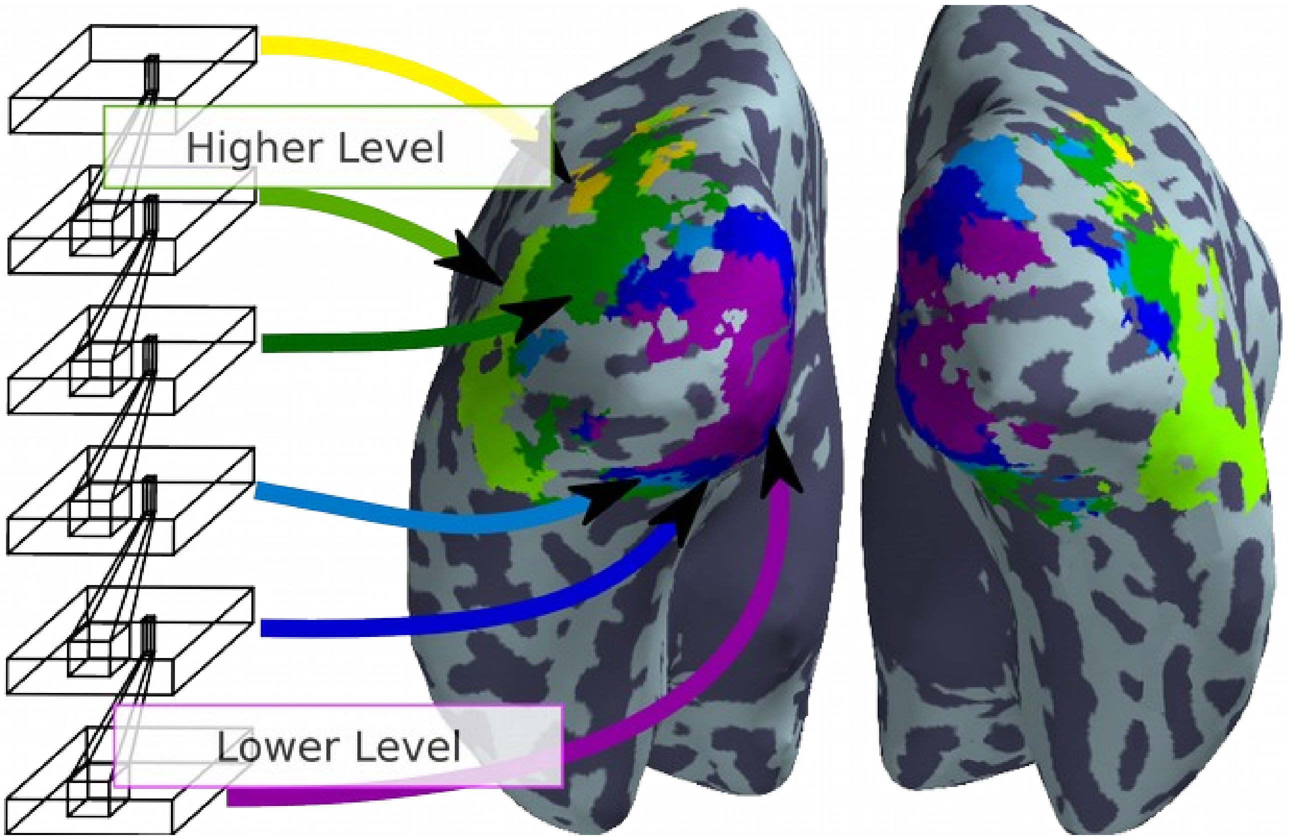
Convolution model:
Feedfoward model of vision for object recognition

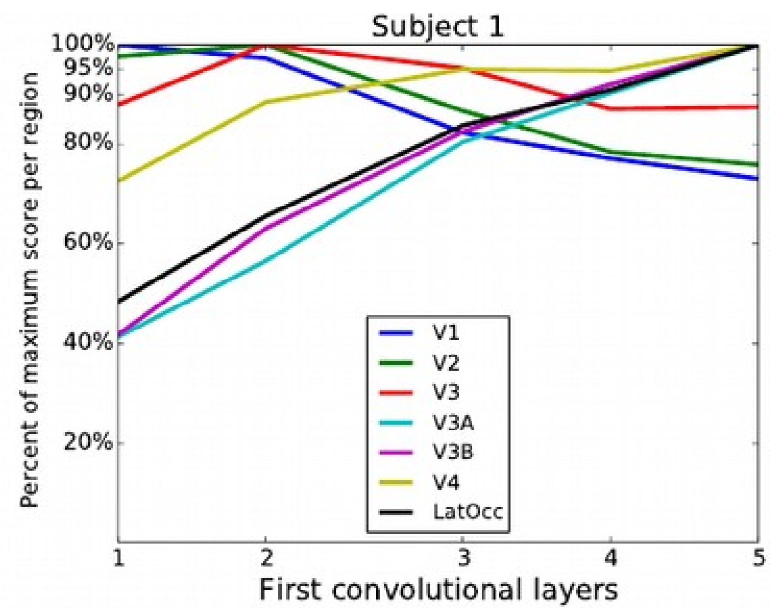


Comparison with fMRI
How well are brain regions explained by the features ?

[Eickenberg et al. NeuroImage 2017]

Mapping convnet layers to the cortex

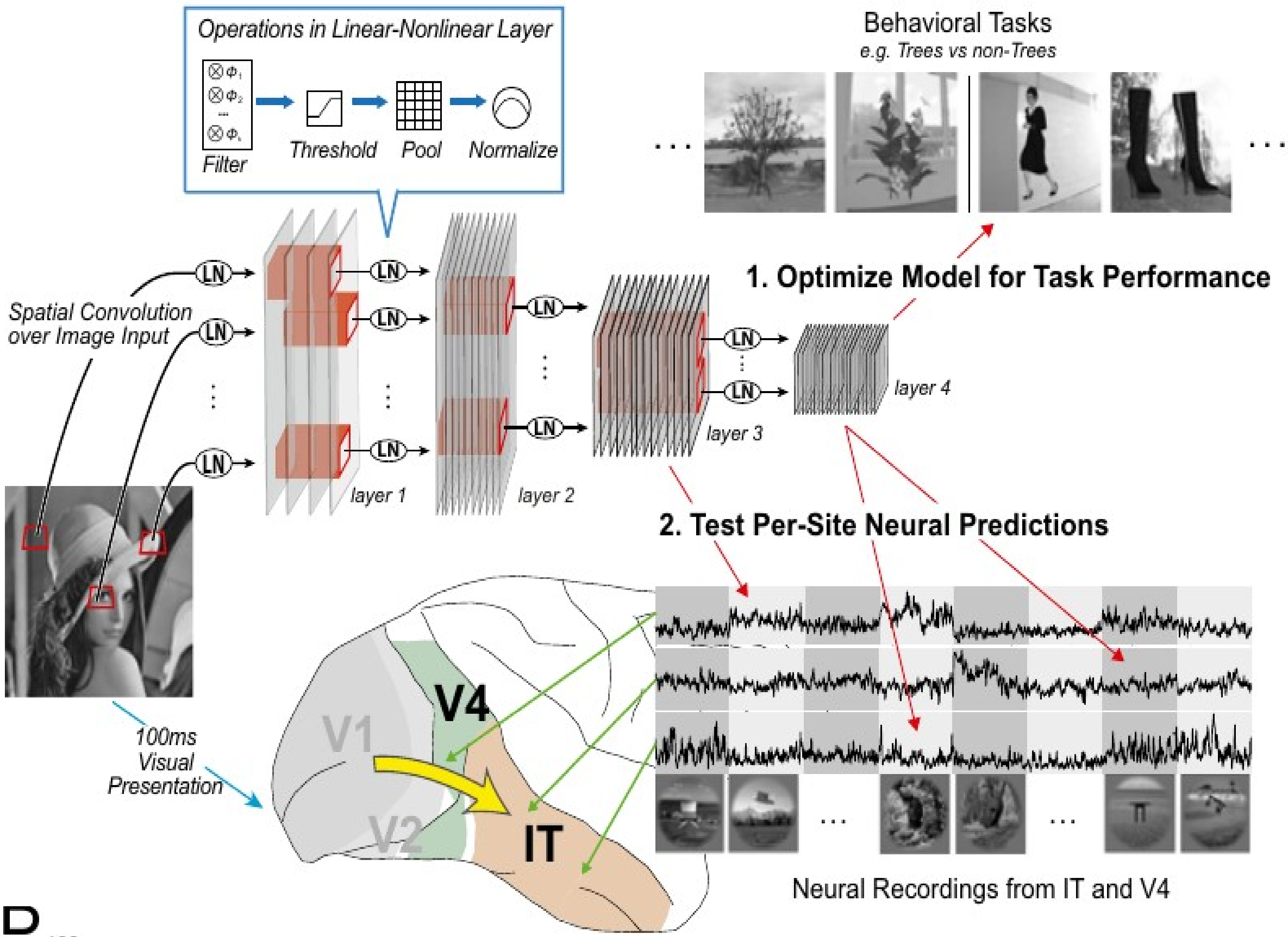




The convolutional network reproduces the cortex structure !

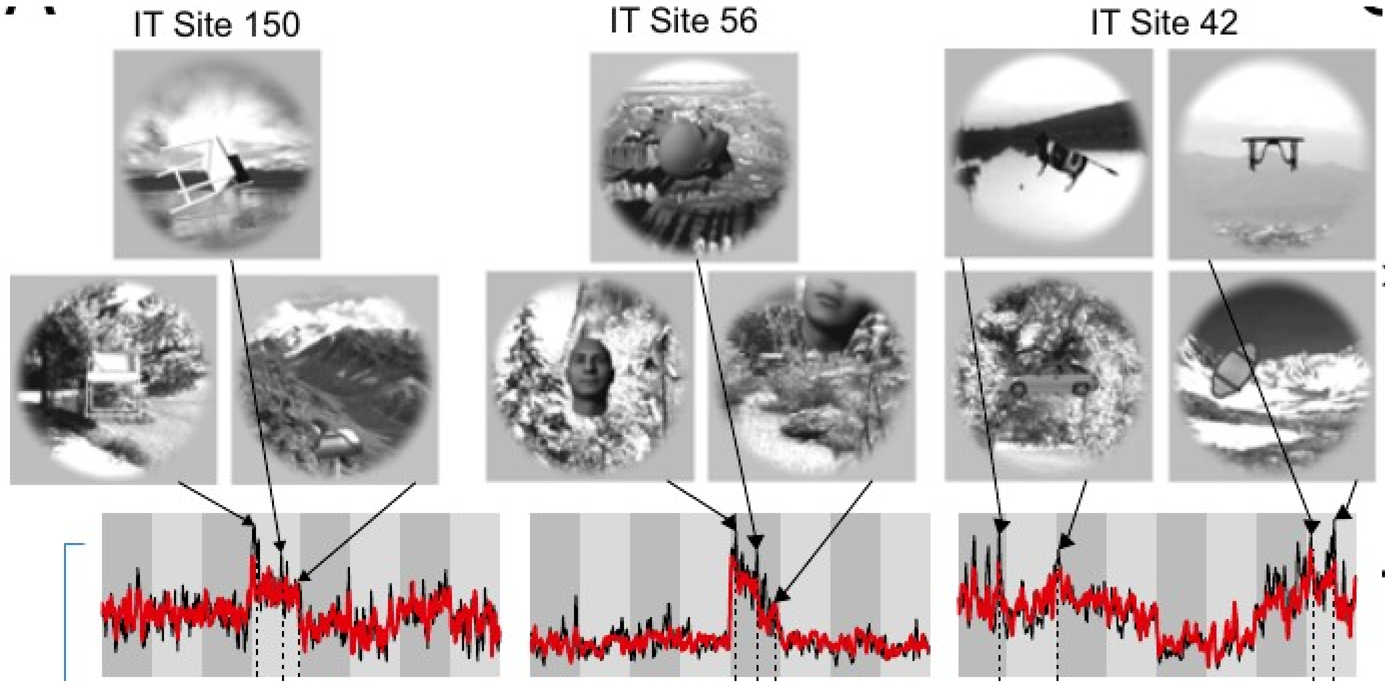
[Yamins et al. 2014, diCarlo et al. 2014, Güçlü et al. 2015, Eickenberg et al.nimg 2017]

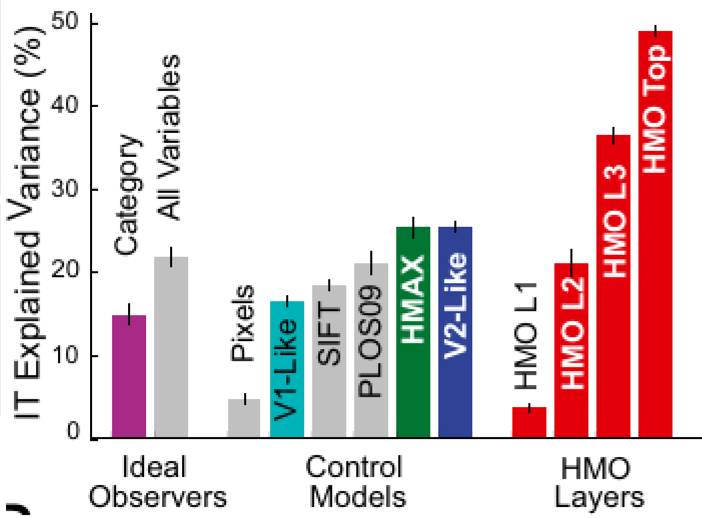
// studying IT selectivity in monkeys



[Yamins et al. PNAS 2014]

Studying IT selectivity in monkeys



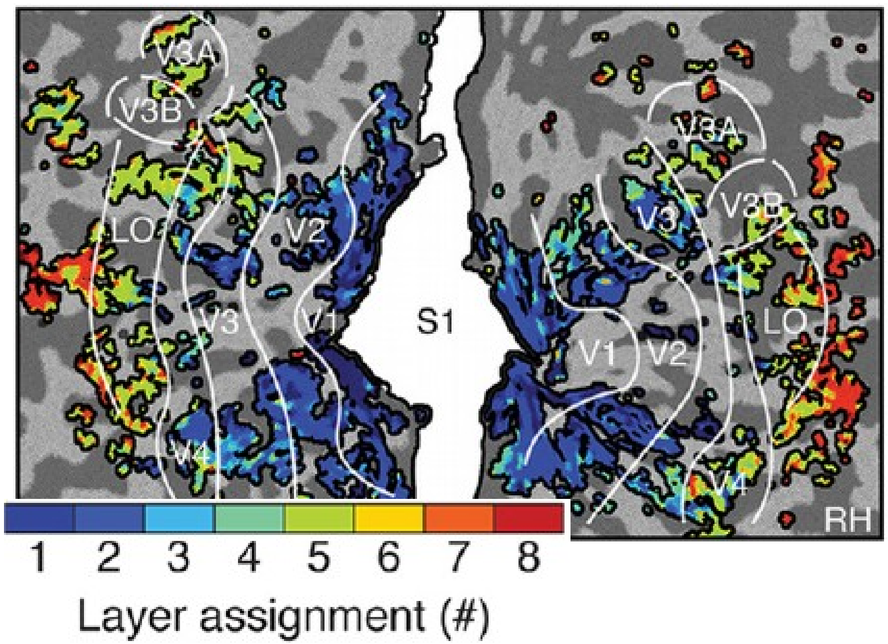


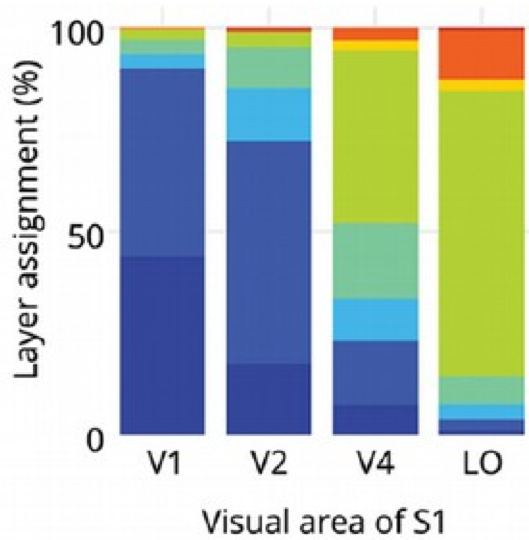
Different IT sites responsive to different stimuli categories (chairs, faces)

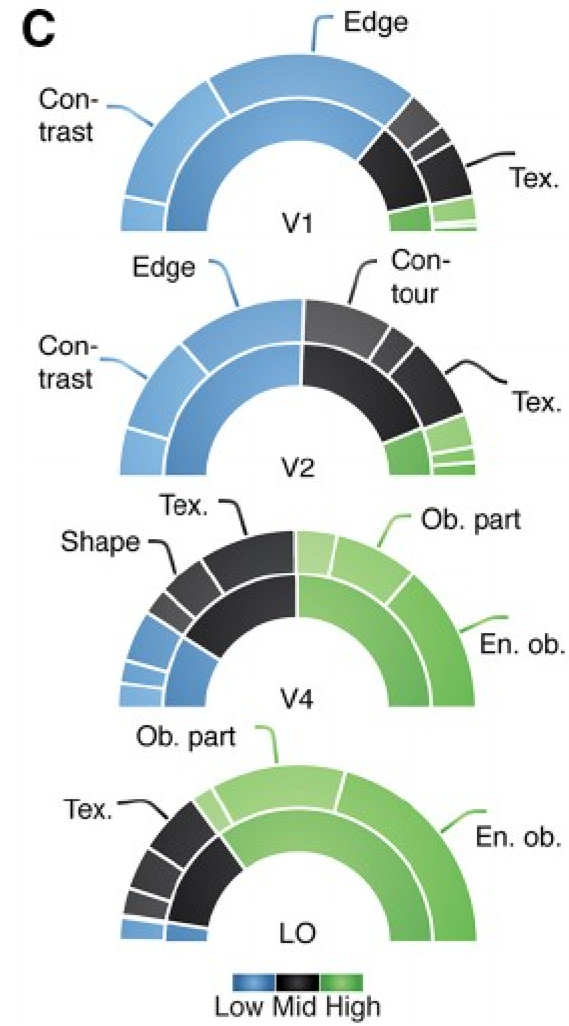
Only end layers of deep nets explain IT activity well

[Yamins et al. PNAS 2014]

Deep encoding: different levels of complexity

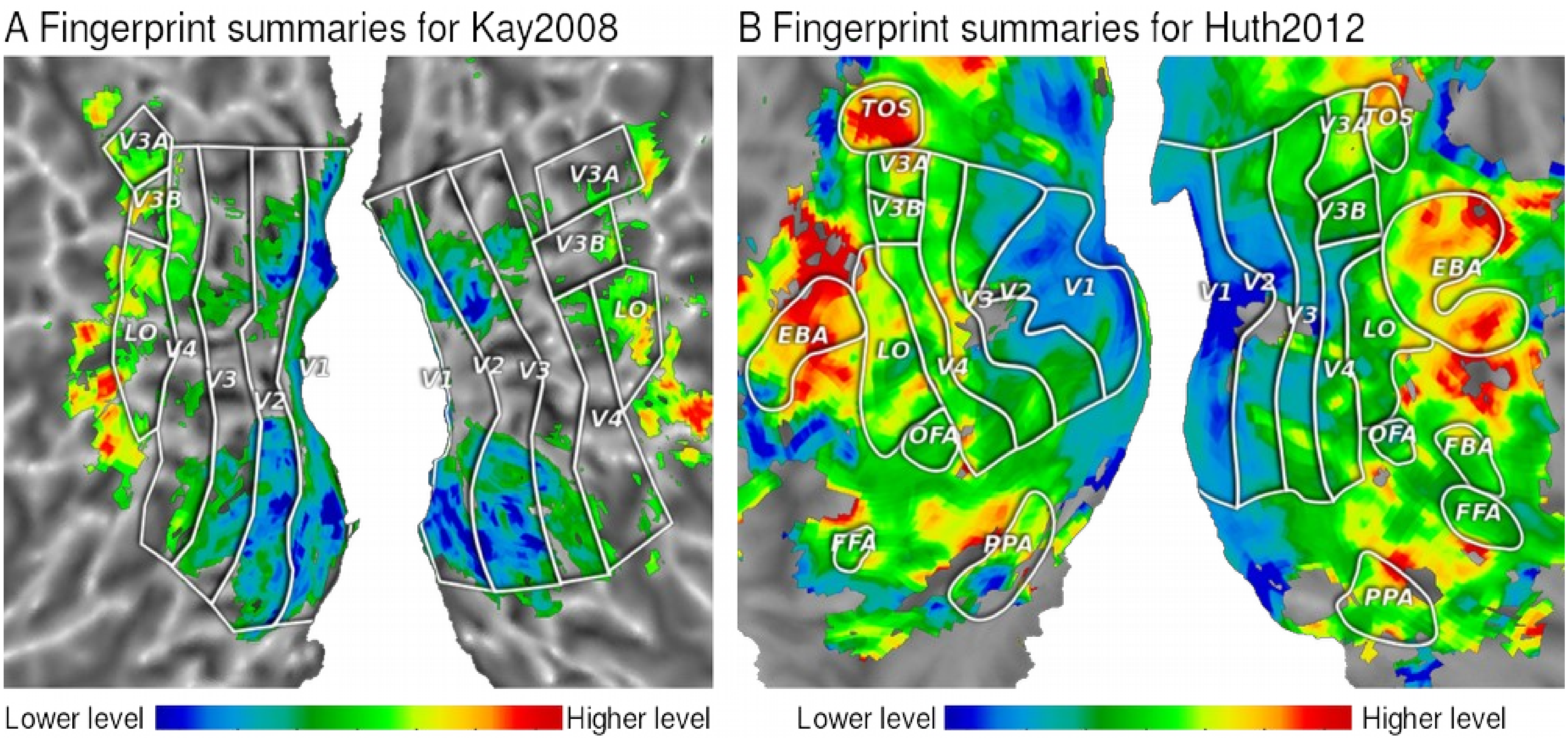






[Güçlü and van Gerven, J.Neuroscience, 2015]

Fingerprint summary statistic



Photos
[Kay et al., 2008]

Videos
[Huth et al., 2012]

More validation

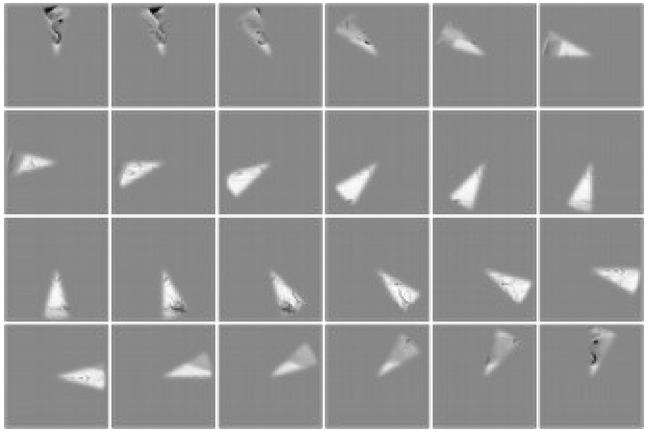
Synthesizing Brain activation maps

- Our model is generative: we can use it to reproduce known experiments
- Generate BOLD response, do GLM analysis

New Stimuli → Convolutional Net Forward Model → Activation Maps

- Low level properties: Retinotopy
- High level properties: Faces/Places contrast

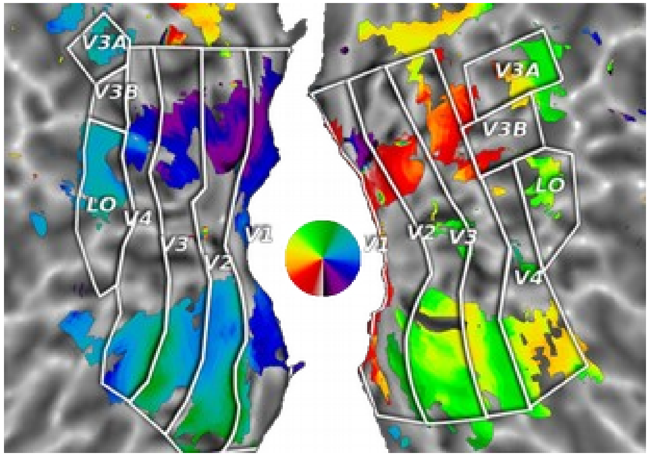
Low-level Validation: Retinotopy



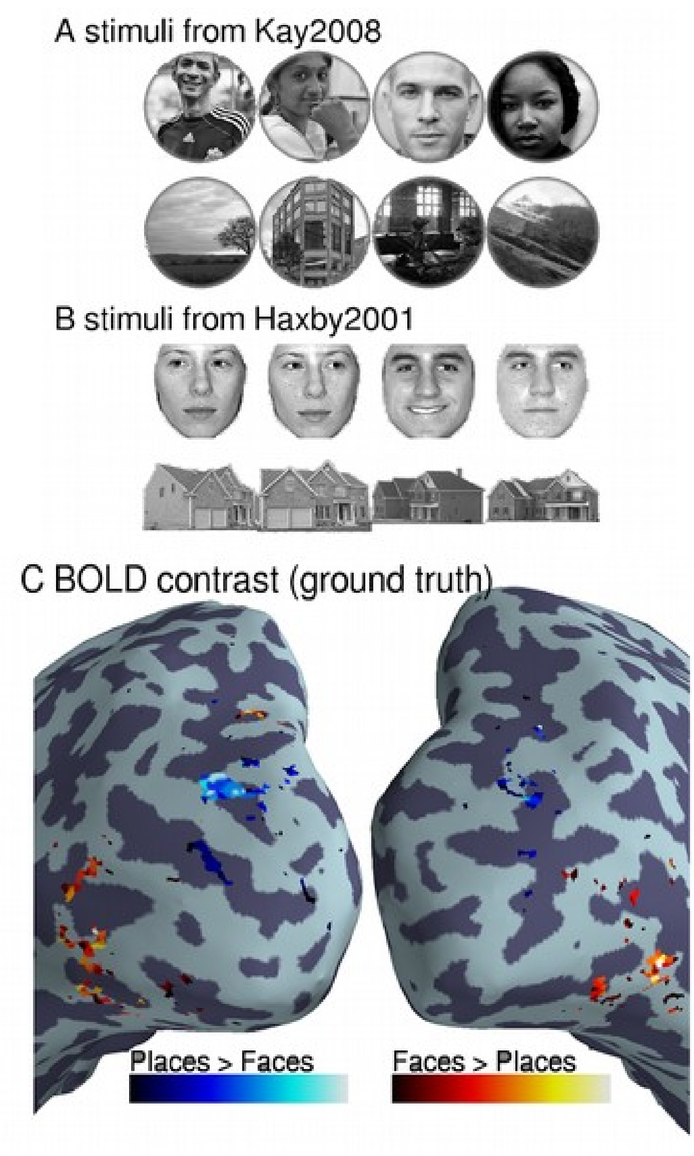
Convolutional Net Forward Model

Activation Maps

Phase Coding GLM



High-level Validation: Faces vs Places

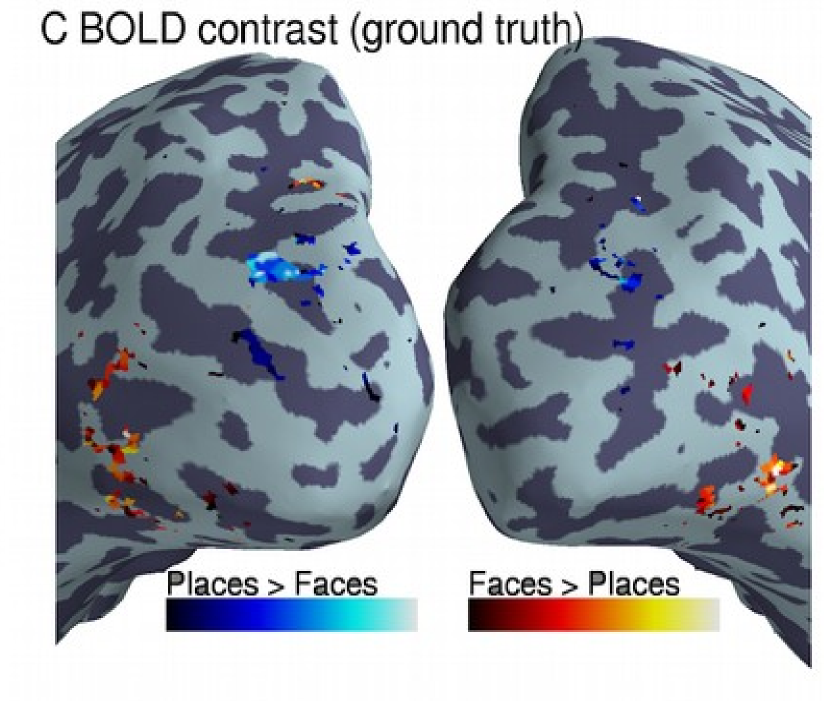


Convolutional Net Forward Model

Activation Maps

GLM Contrast Maps

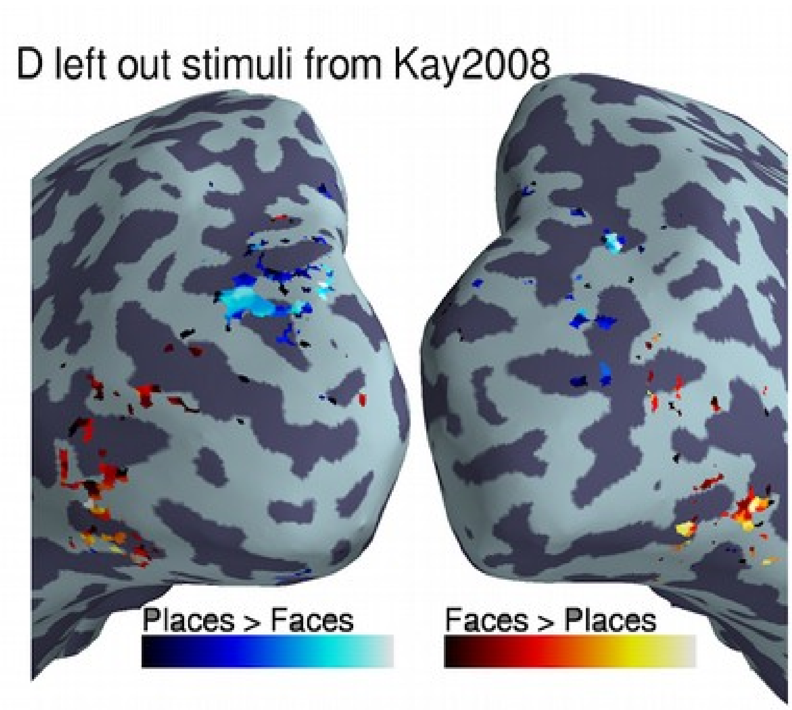
Faces vs Places: Ground Truth

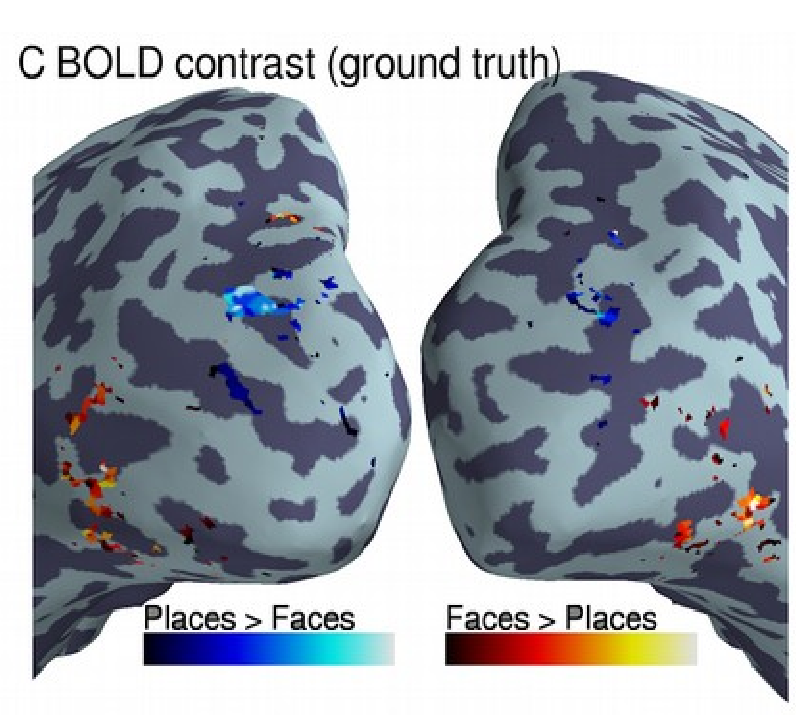


Stimuli from [Kay et al., 2008]: Close-up Faces and Scenes

Contrast of activation maps from [Kay et al., 2008]

Faces vs Places

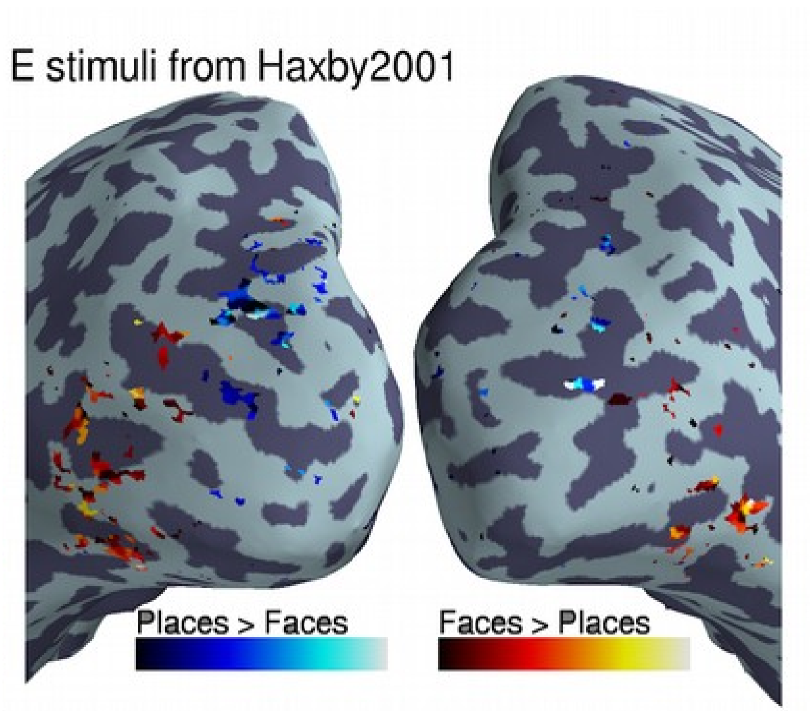


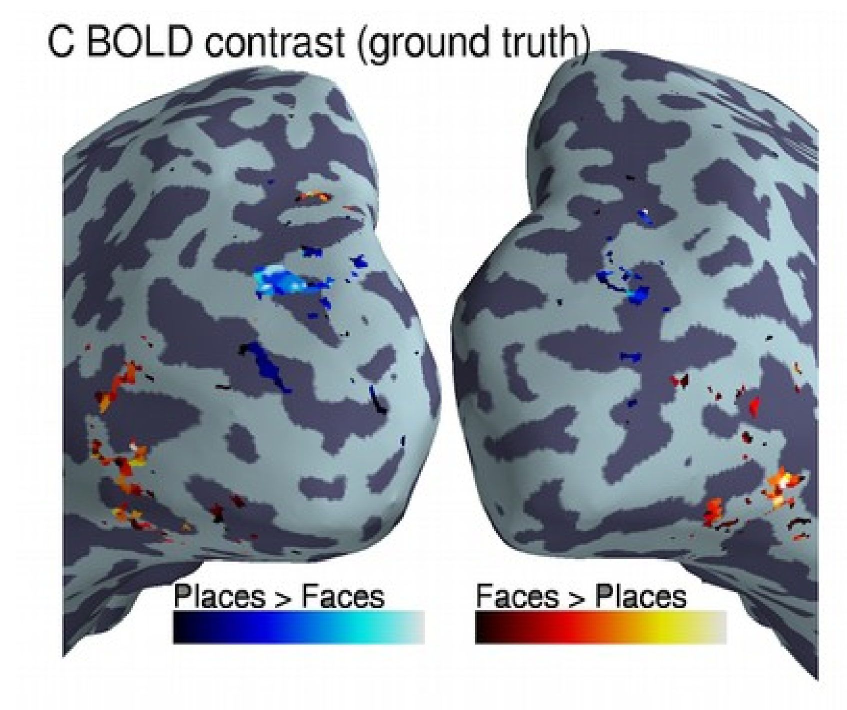


Simulation on [Kay2008] left-out stimuli

BOLD ground truth

Faces vs Places

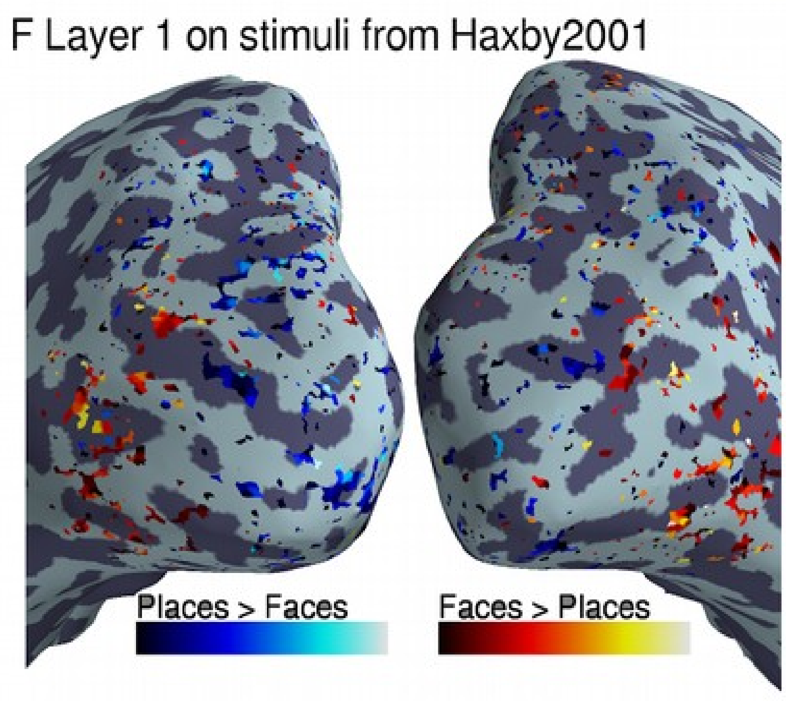


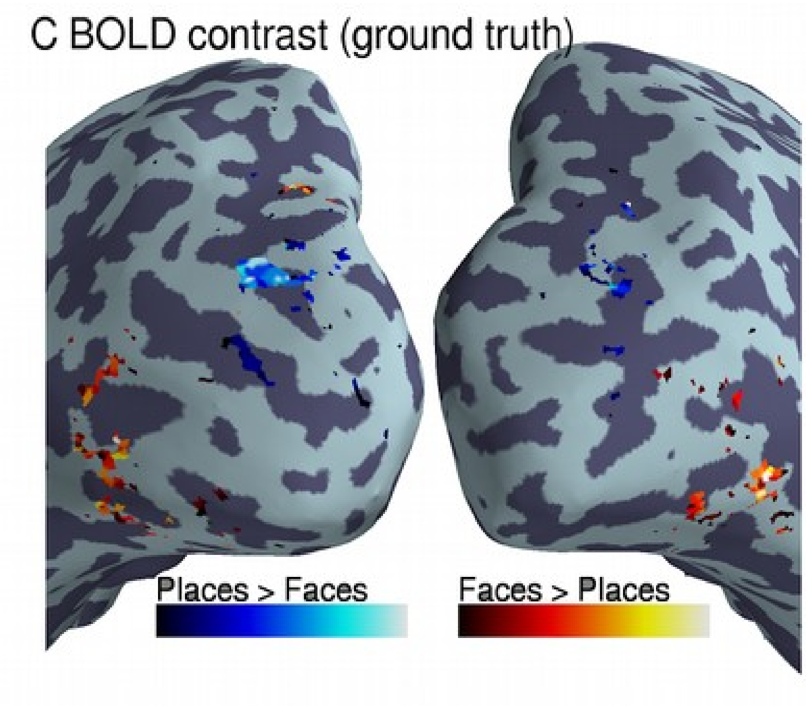


Simulation on [Haxby2001] stimuli

BOLD ground truth

Faces vs Places



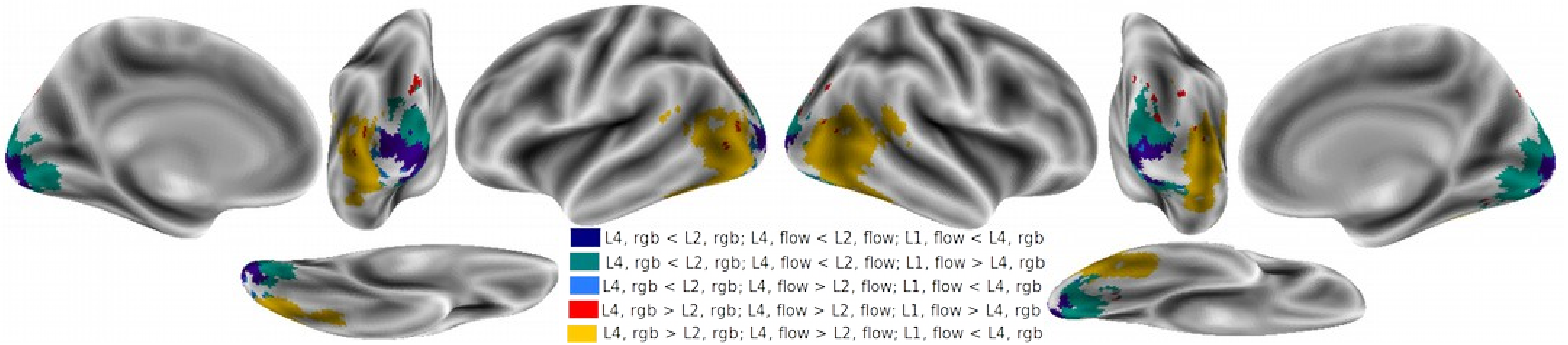


Simulation on
[Haxby et al., 2001]
Layer 1 only
Identification Impossible

BOLD ground truth

Encoding videos

Video clips (4hrs, 3.6kframes)



Brain regions that are significantly explained
by the deep video analysis network.

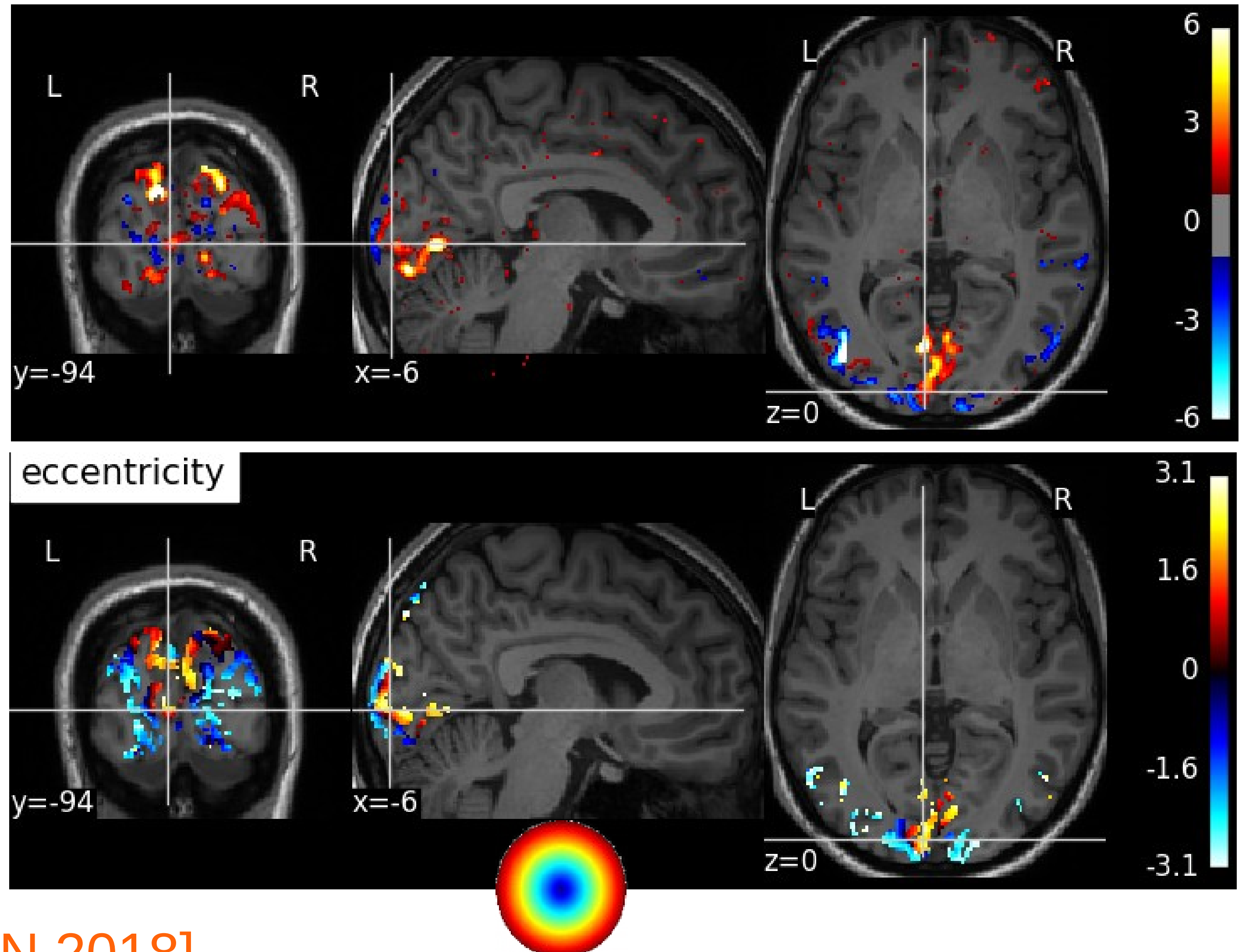
Regions characterized by the layers that best explain voxel activity

- Overall consistency with retinotopic results
- Agreement with functional properties of the voxels, e.g. eccentricity

[Richard et al CCN 2018]

Response explained not by Visual areas but eccentricity

Example: response to flow layer 1 vs ultimate RGB layer (top) very highly correlated with eccentricity contrast on the same subject (bottom)



[Richard et al CCN 2018]

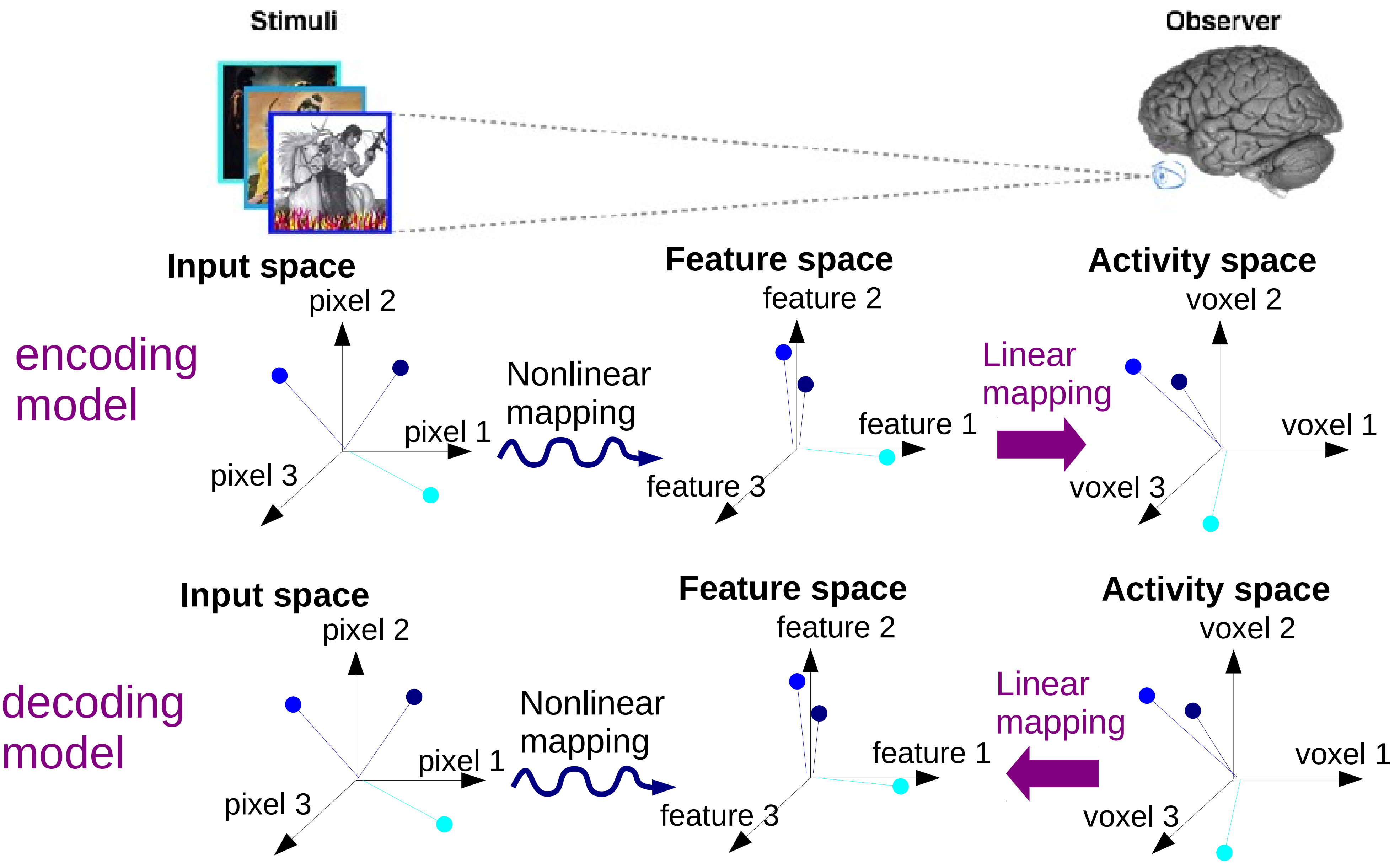
Decoding visual stimuli

- Infer stimulus information from brain maps: inverse problem of the encoding.
  - Related to brain computer interfaces
- Detect distributed effects. Not aimed at localizing feature representations
- Requires high-dimensional generalized linear models

$$\phi(\mathbf{X}) = \mathcal{L}(\mathbf{Y}\mathbf{w}) + \eta$$

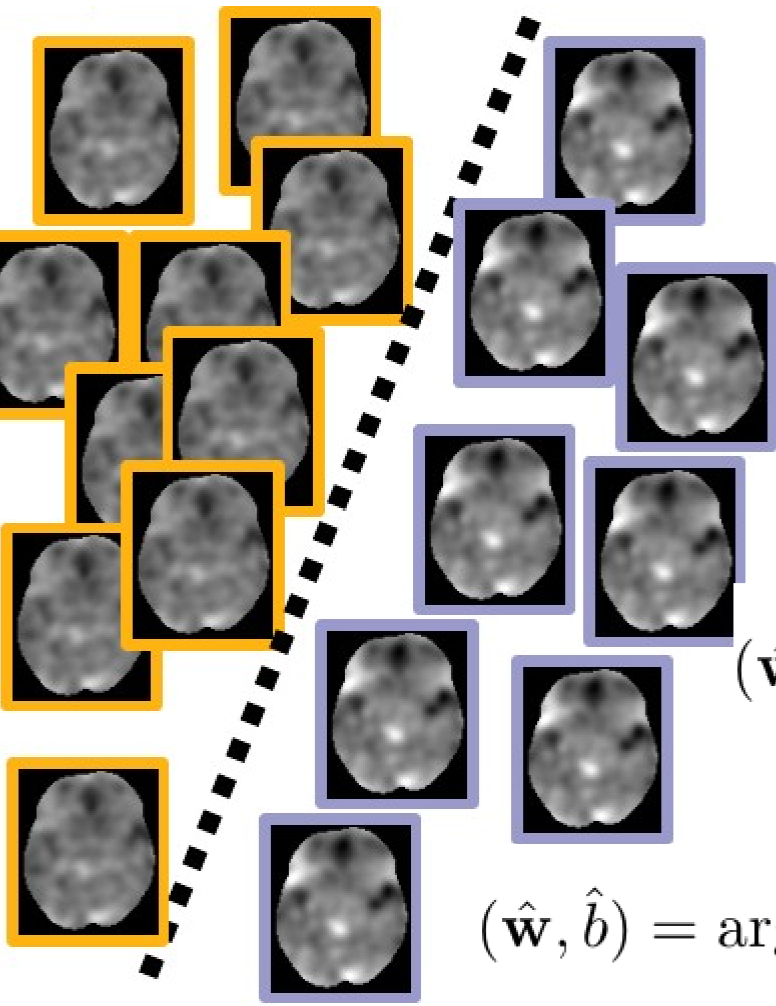
$\mathcal{L}$ Possible nonlinearity (e.g. sign function)

  - Suffers from curse of dimensionality



[Naselaris et al. NeuroImage 2009]

Training classifiers for FMRI data



- Given x in R^p, (fMRI volume with p voxels), predict a label y in {-1, 1} i.e. or

  or better the class probability Proba(x = 1|y)

- Use of logistic regression: learn the weight w and bias b such that

$$(\hat{\mathbf{w}}, \hat{b}) = \operatorname{argmin}_{\mathbf{w},b} \sum_{i=1}^{n} \log\left(1 + \exp\left(-x_i(y_i \mathbf{w} + b)\right)\right)$$

- With regularization

$$(\hat{\mathbf{w}}, \hat{b}) = \operatorname{argmin}_{\mathbf{w},b} \sum_{i=1}^{n} \log\left(1 + \exp\left(-x_i(y_i \mathbf{w} + b)\right)\right) + \lambda \|\mathbf{w}\|_2^2$$

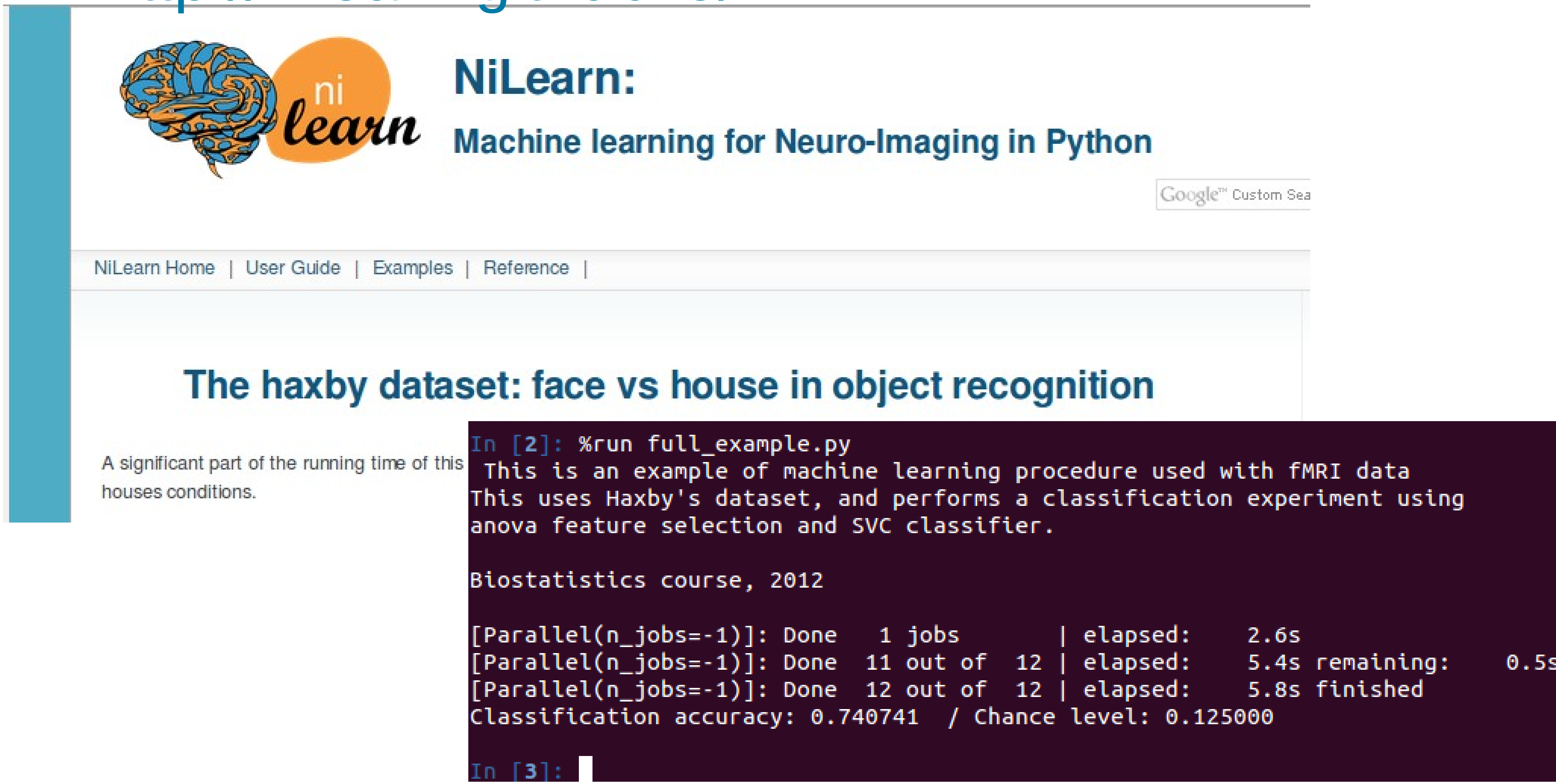
Decoding visual categories

Visual categories very well discriminated indivudally

[Haxby et al. Science 2001]

Do it yourself !

- http://nilearn.github.io/



Reconstructing a perceived stimulus

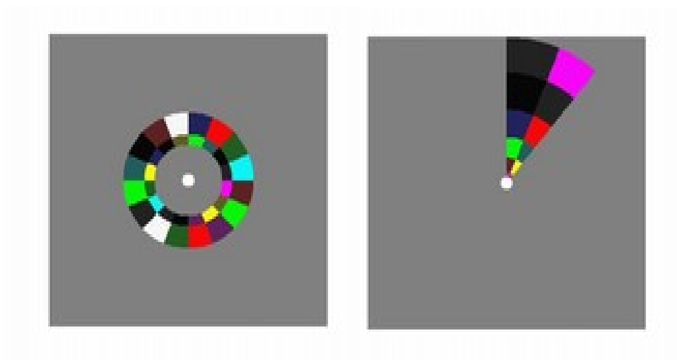
$$\phi(\mathbf{X}) = \mathcal{L}(\mathbf{Yw}) + \eta$$

- Predicting **X** given **Y**
- Two possible approaches:
  - Inverse problem: assume that ϕ is the identity / invertible $\mathbf{X} = \mathcal{L}(\mathbf{Yw}) + \eta$
  - Identification among a finite set of samples

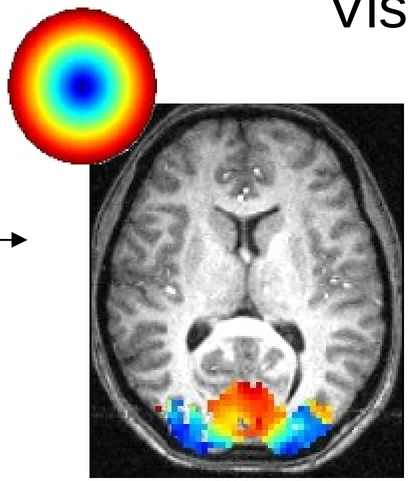
$$\hat{i}(\mathbf{y}) = \mathrm{argmin}_{i \in [1..n]} \|\phi(\mathbf{x_i}) - \mathcal{L}(\mathbf{y^T w})\|^2$$

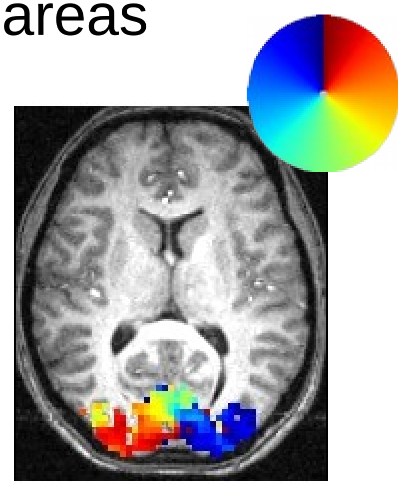
Testing vision models with fMRI: inverse retinotopy

Retina



Visual areas

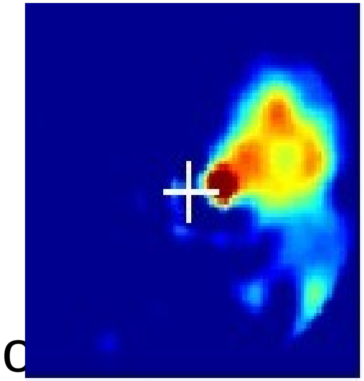




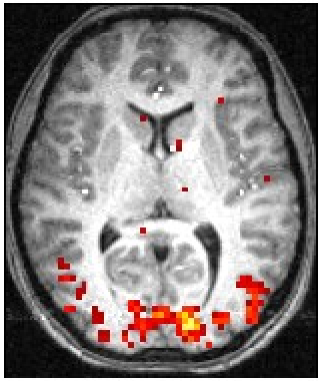
learning

Retinotopy experiment

Retinal representatio of the map



inverse inference: 'brain reading'

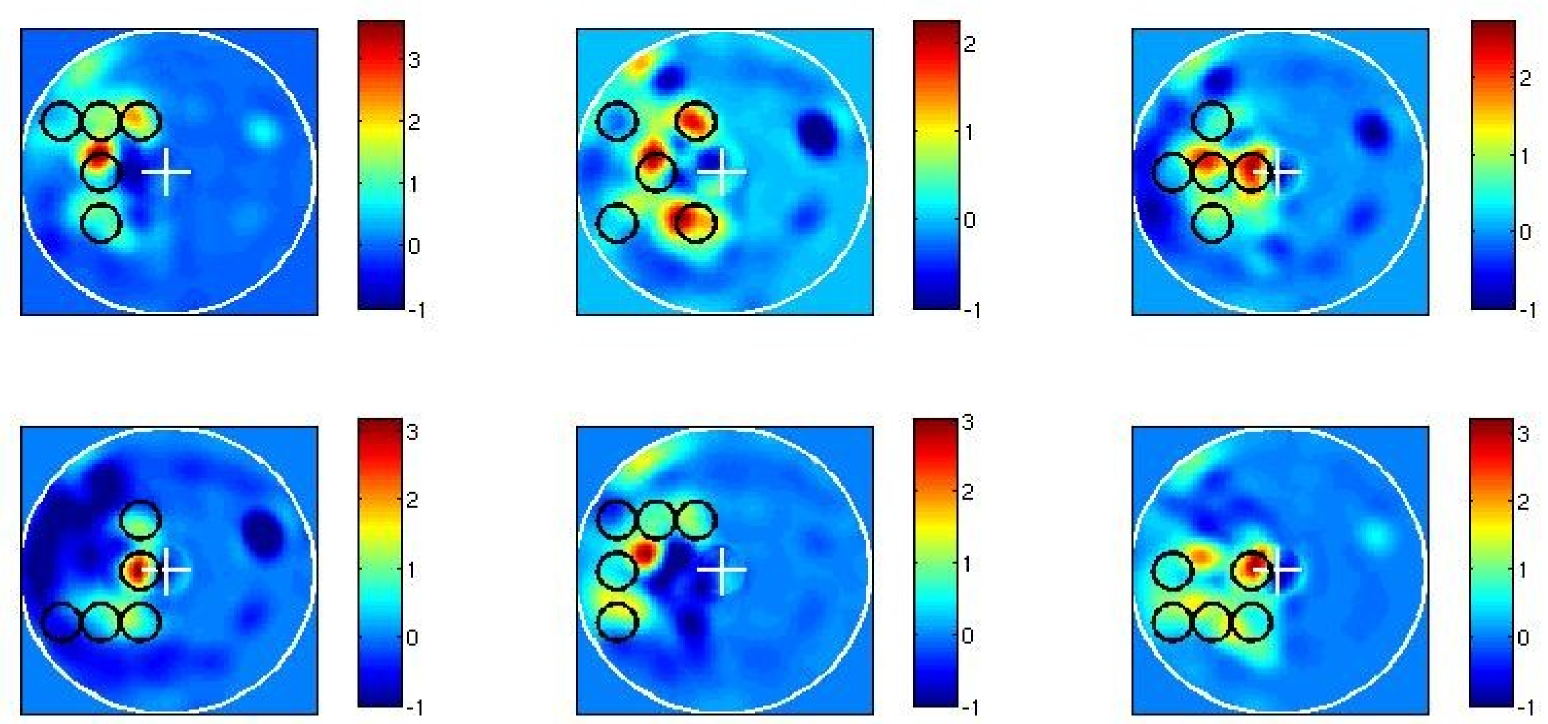


Any candidate activation map

[Thirion et al. NeuroImage 2006]

Inverse retinotopy : results

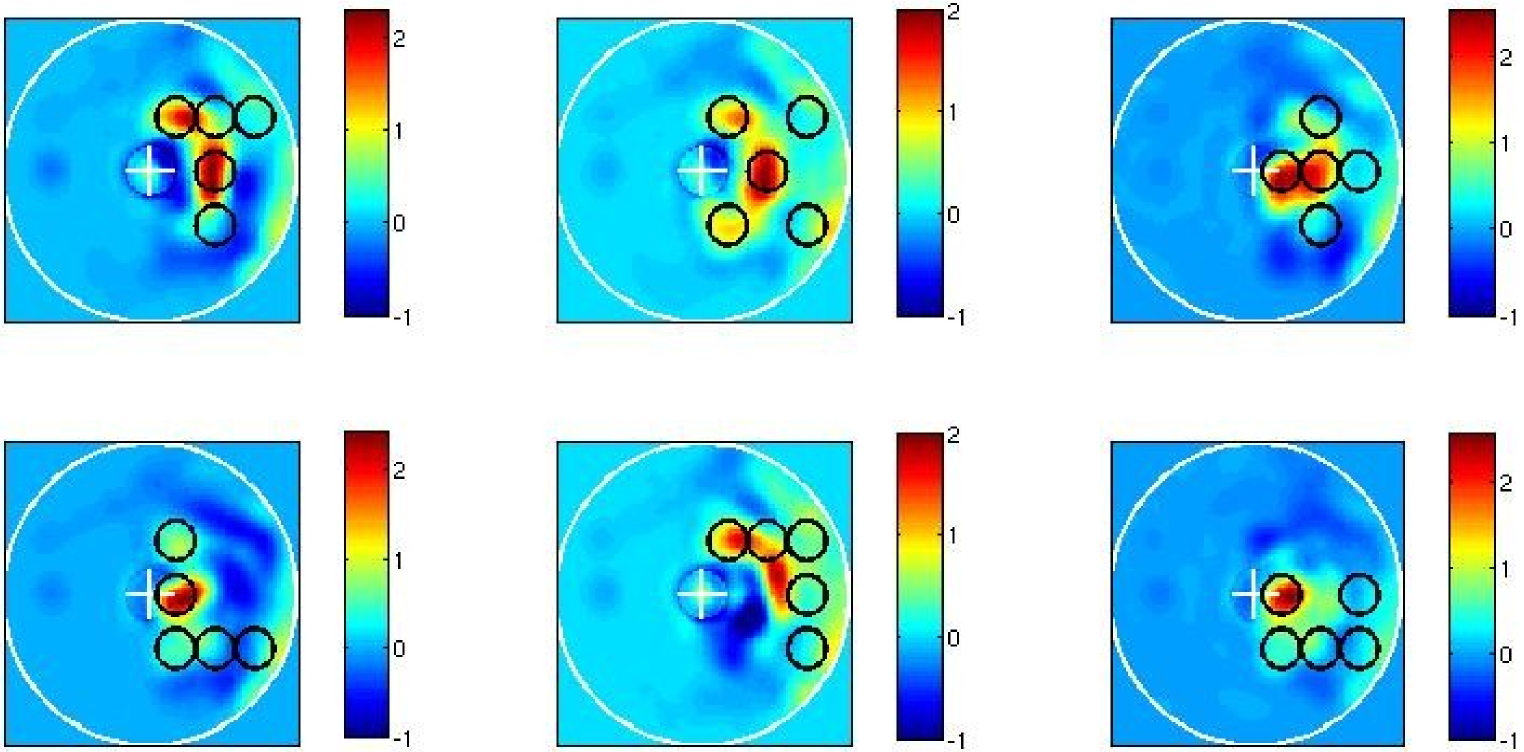
Average response for arbitrary visual patterns presented by their outline



The reconstruction ifs good enough to allow the identification of the stimuli on a trial-by-trial basis

Inverse retinotopy: results

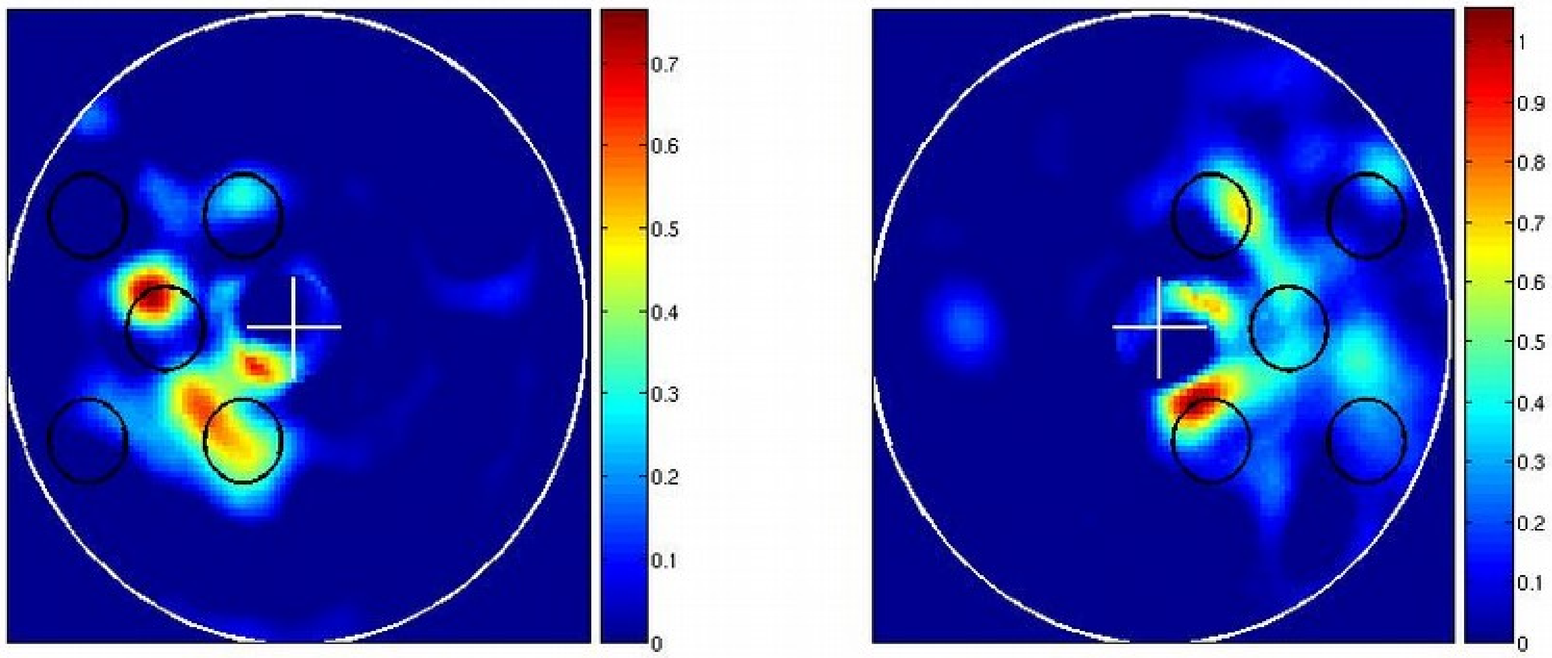
Average response for arbitrary visual patterns presented by their outline



The reconstruction ifs good enough to allow the identification of the stimuli on a trial-by-trial basis

Can we make inference about mental imagery ?

Reconstruction of a pattern that was imagined by a subject
(disclosed by the subject after the scanning session)

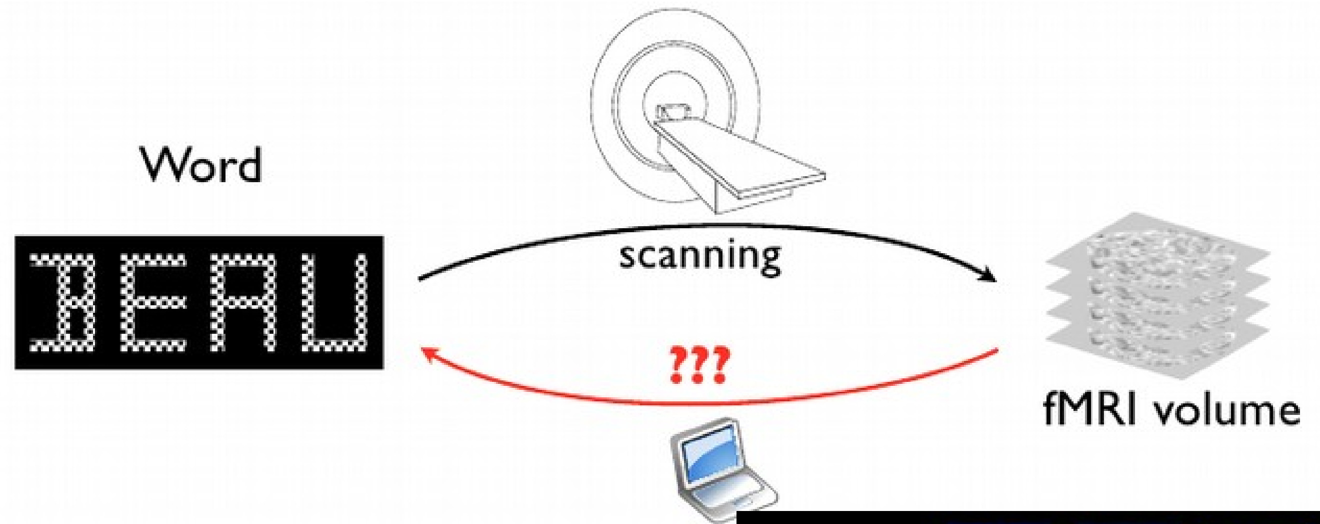


The true pattern was predicted in 5/16 hemispheres ($p<0.05$)

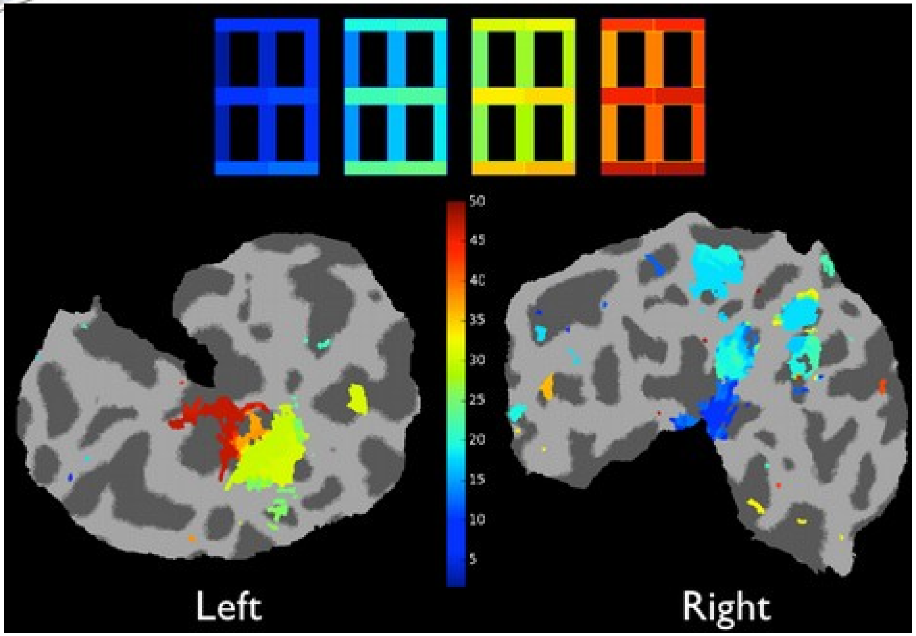
[Thirion et al. Nimg 2006]

[Senden et al. Brain Struct Func 2019]

Application: decoding 4-letters words from the visual cortex



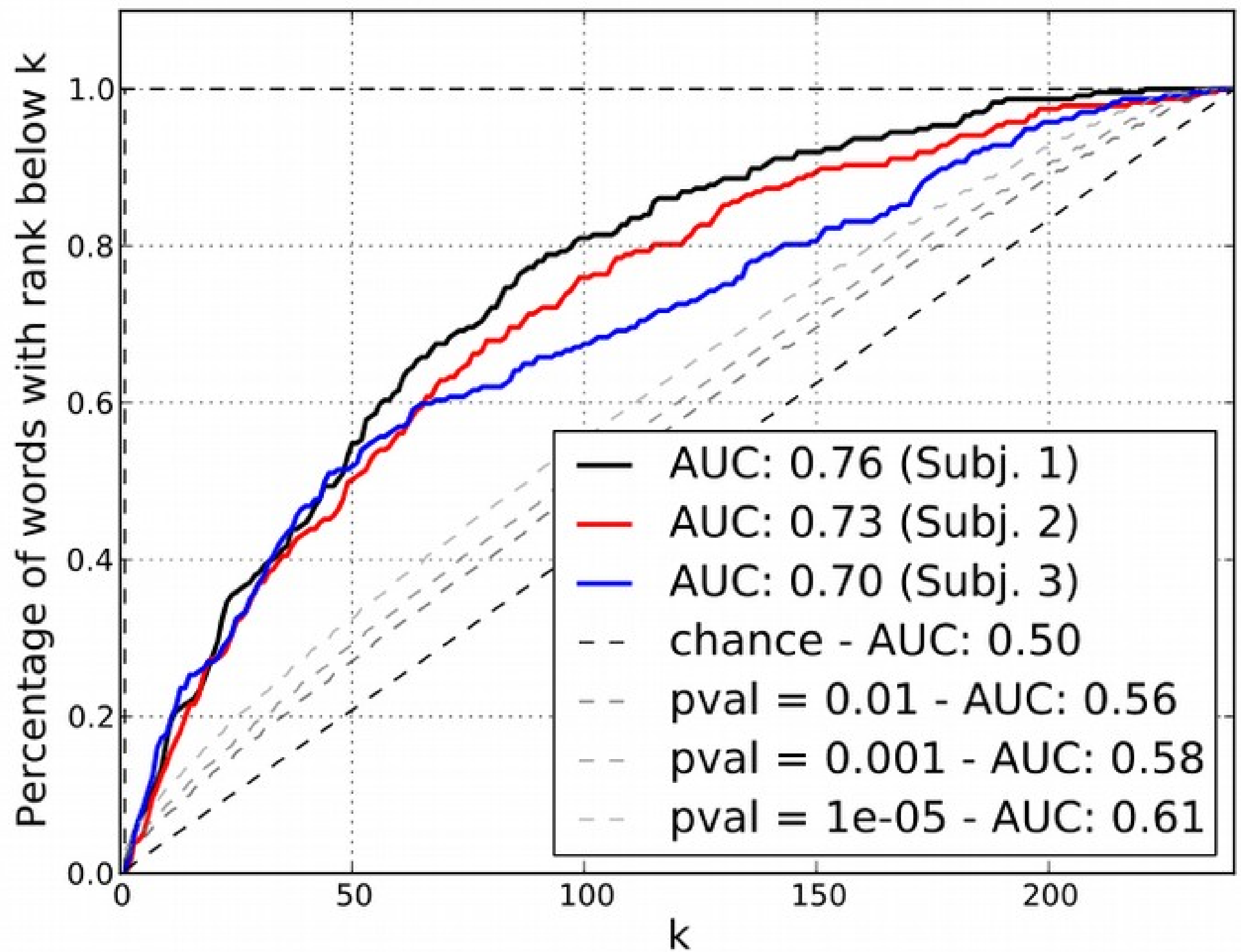
- Achieve the reconstruction of 4-letter words presented at fixed position



Application: decoding 4-letters words from the visual cortex

Validation:
Receiver-operating characteristic
1-session-out cross-validation.

Prediction significantly above chance in all subjects

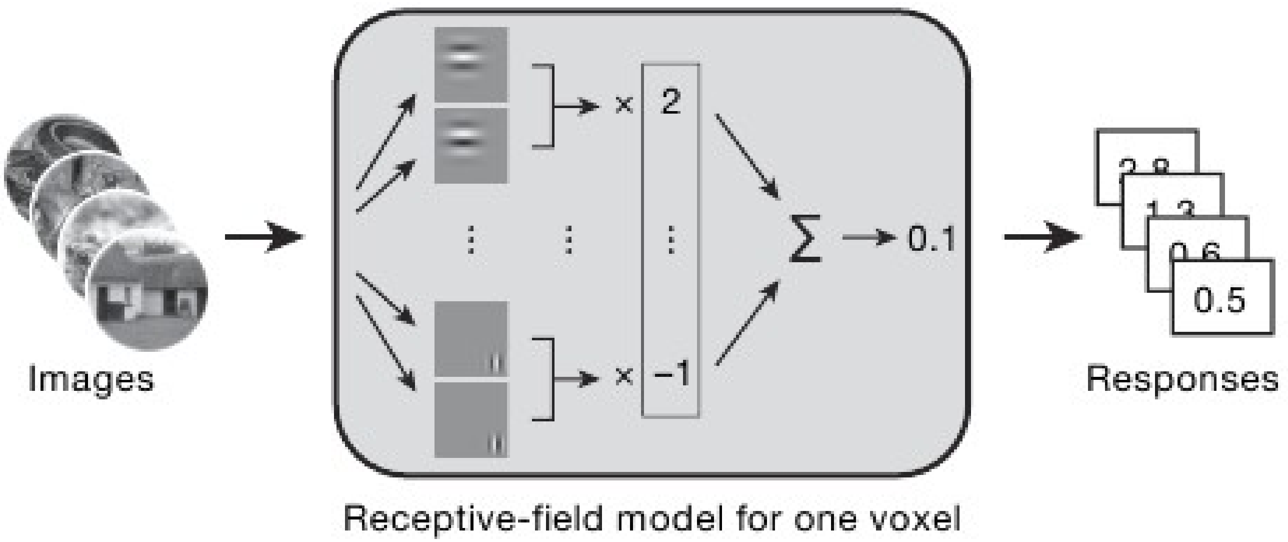


[Gramfort et al. PRNI, 2012, Miyawaki et al. Neuron 2008]

Decoding through identification

[Kay et al. Nature Neuroscience 2008]
predict which image has been observed by the subject

Step 1: model the response in each voxel based on a training set (1750 images)



$$\mathbf{y} = \phi(\mathbf{X})\beta + \varepsilon$$

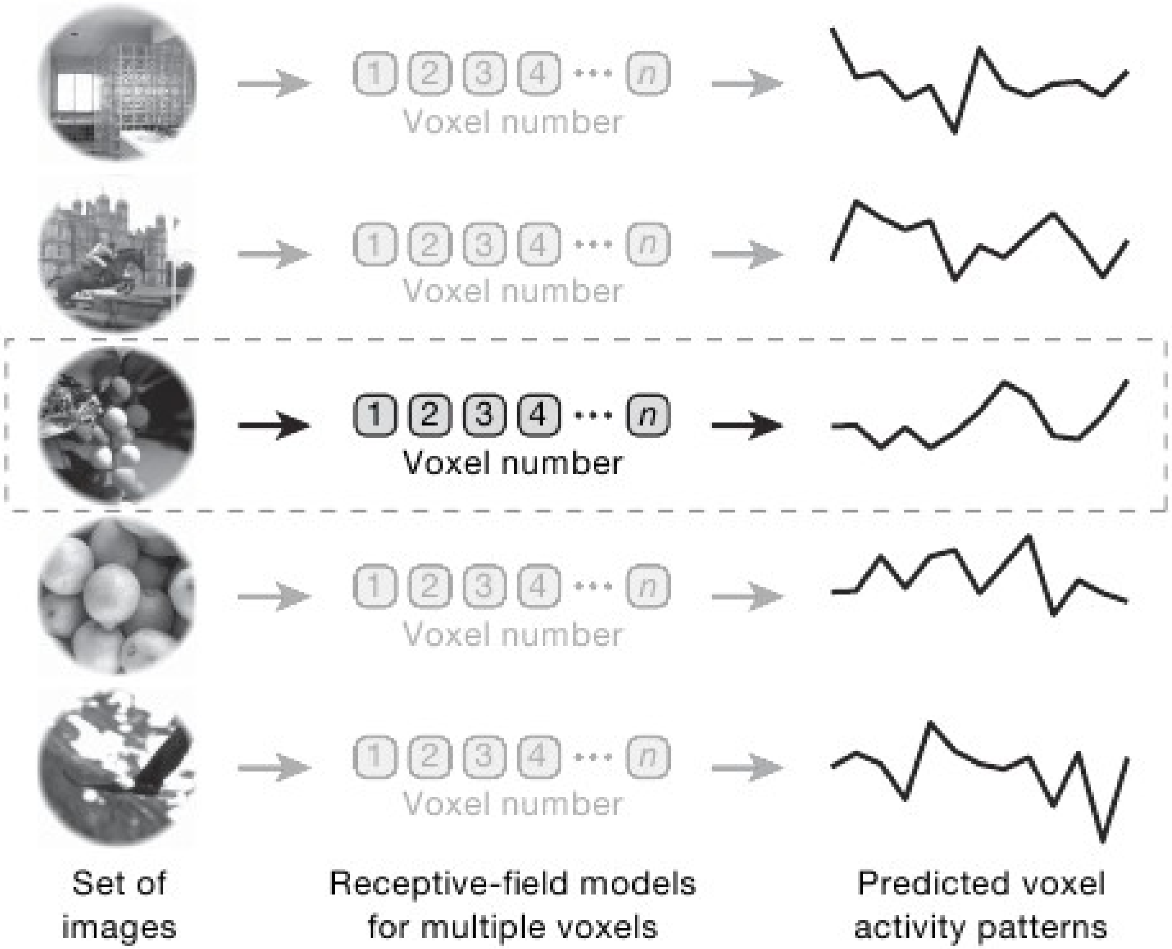
Encoding step

Decoding through identification

$$\hat{i}(\mathbf{y}) = \operatorname{argmin}_{i \in [1..n]} \|\phi(\mathbf{x_i})\hat{\beta} - \mathbf{y}_{\text{test}}\|^2$$

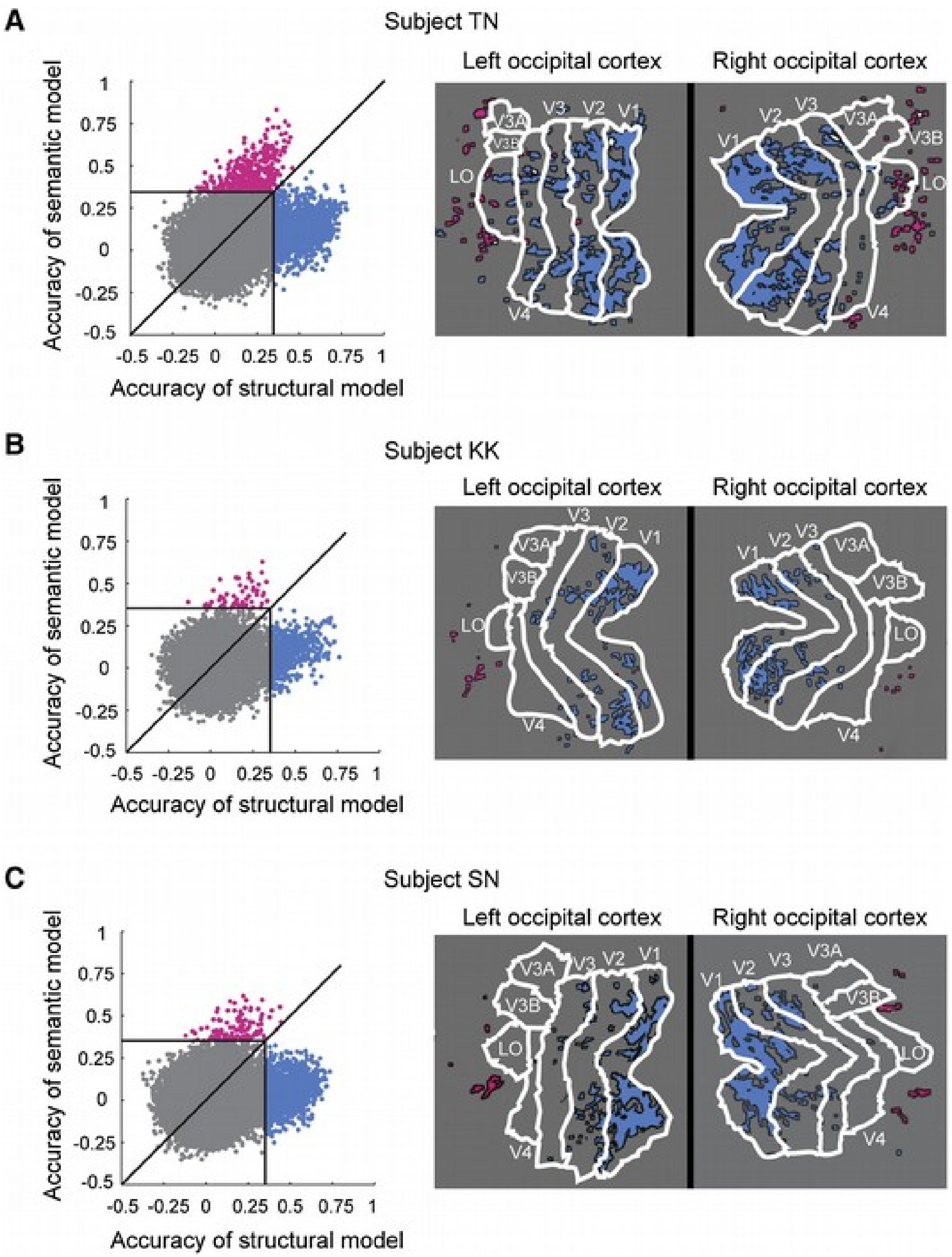
Step 2: among 100 possible images, select the image whose predicted brain activity is most similar to the measured brain activity

accuracy: 92%



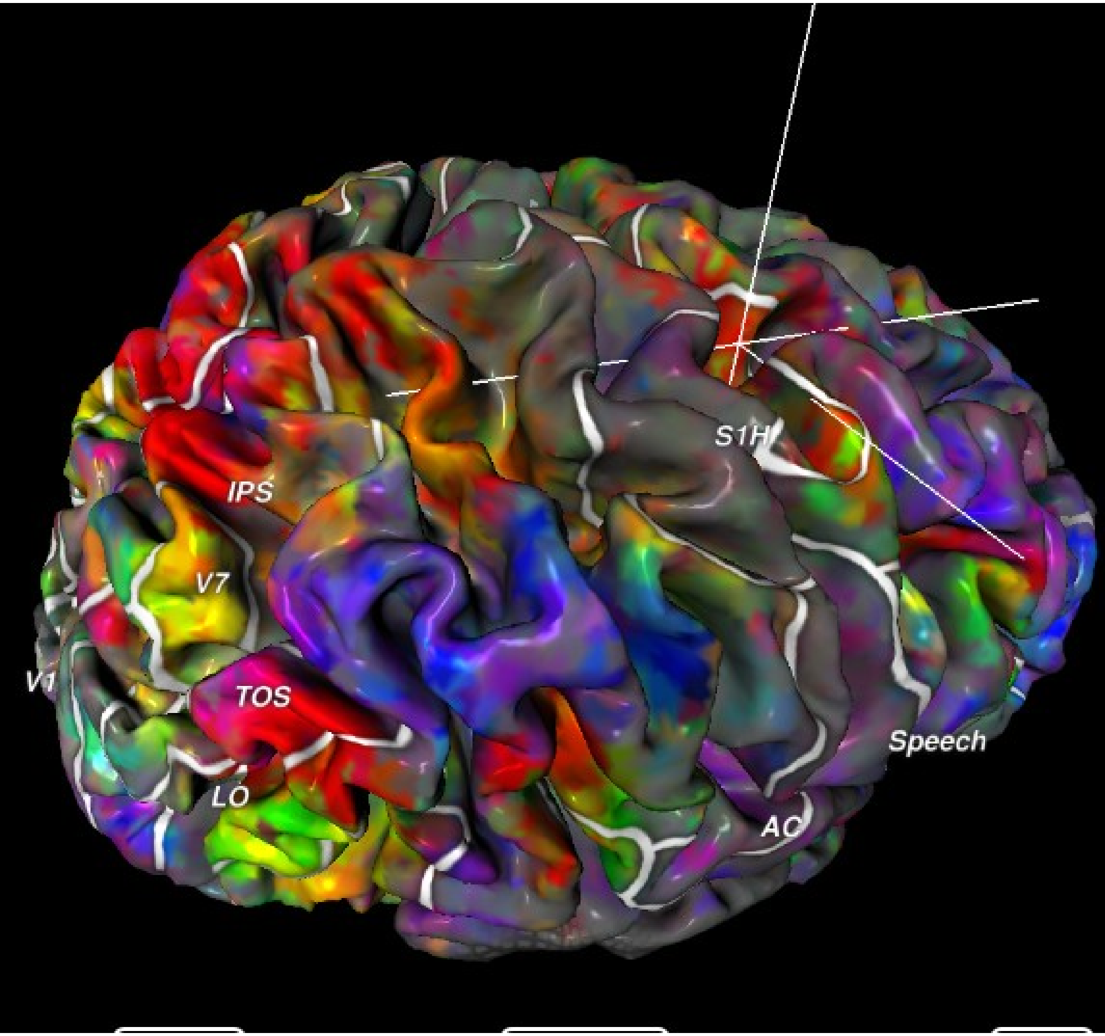
- Encoding semantic information in parallel increases prediction accuracy a lot
- *obvious* mistakes are avoided
- The relative weighting of low-level and high-level features requires careful design

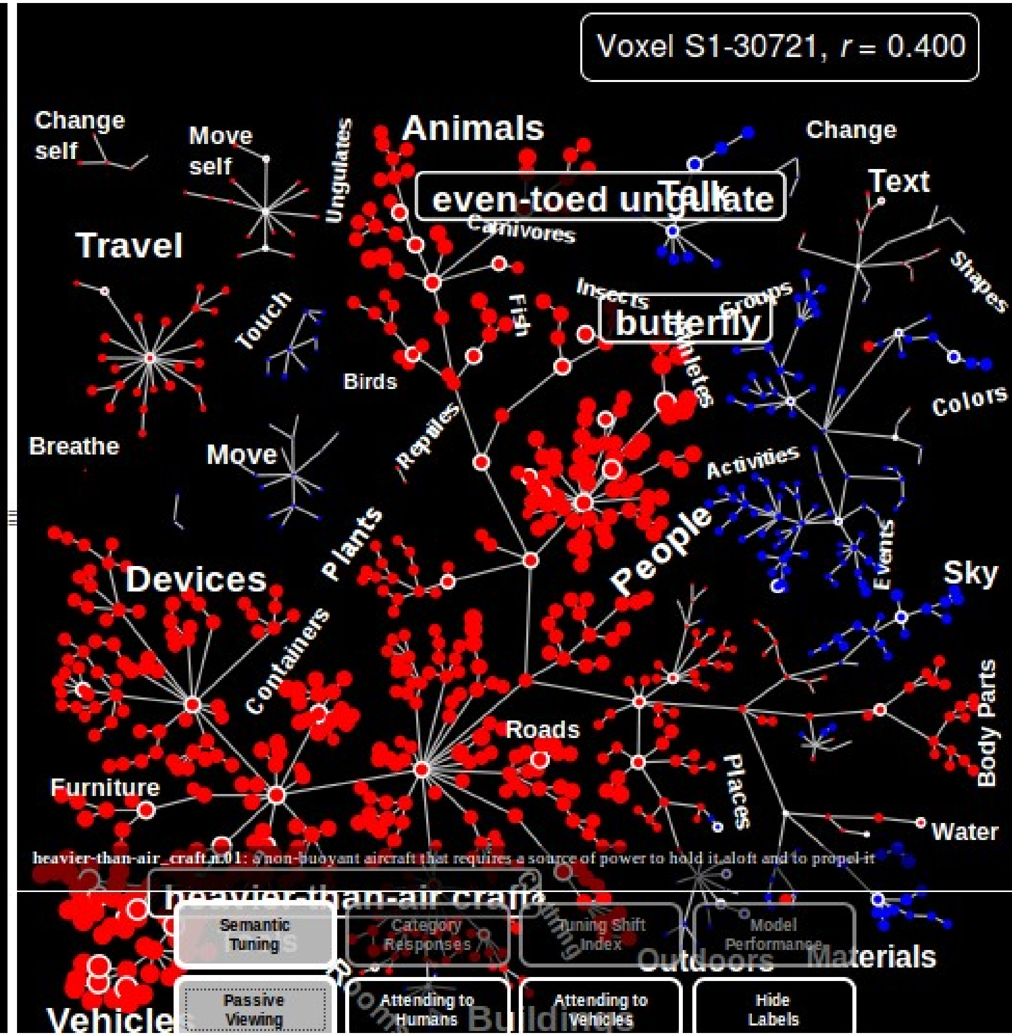
[Naselaris et al. 2009]



- Not surprisingly, the low level and high level features are encoded in non-overlapping regions of the visual pathway

Mapping the semantic space

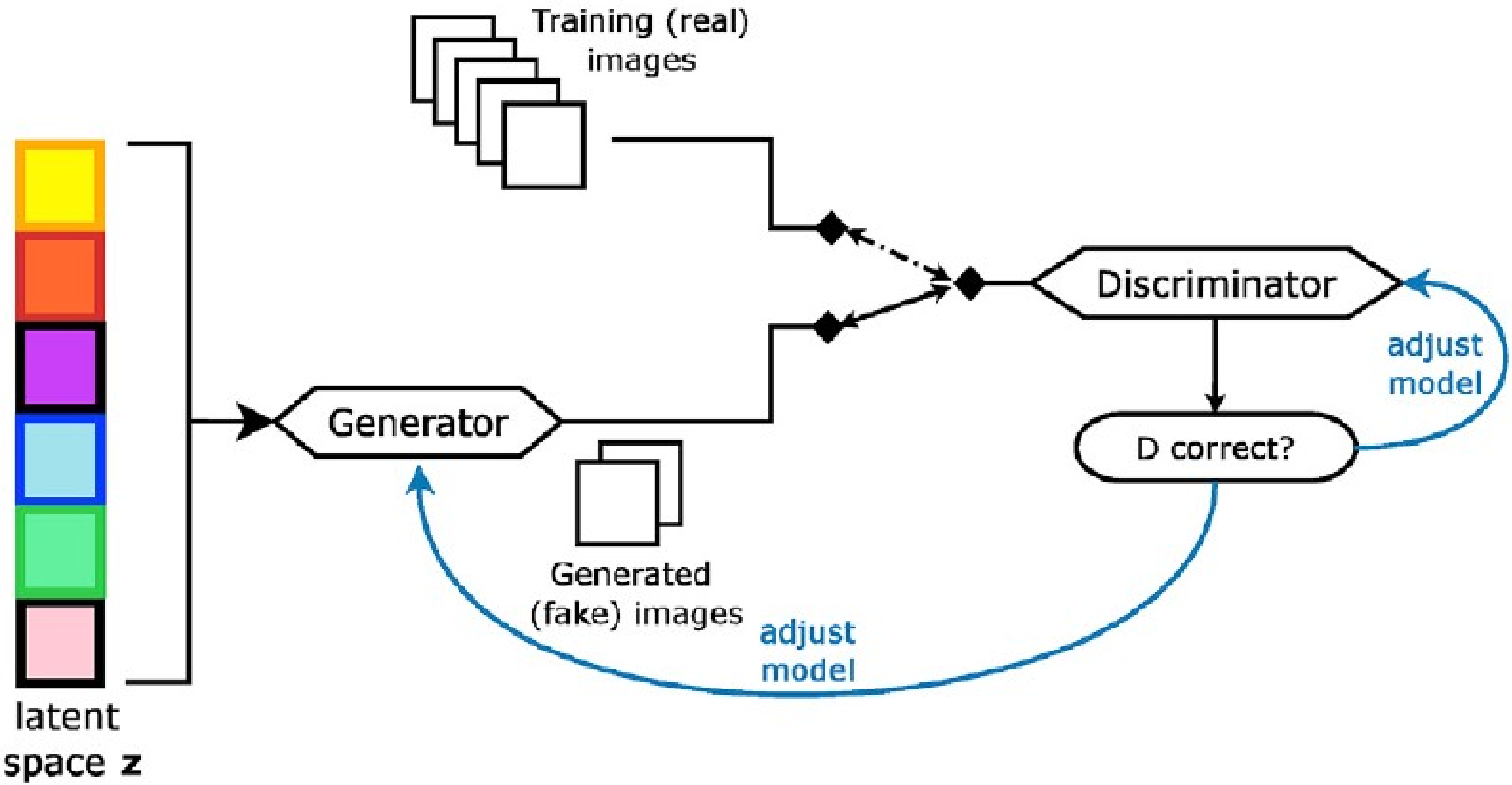




http://gallantlab.org/brainviewer/cukuretal2013/

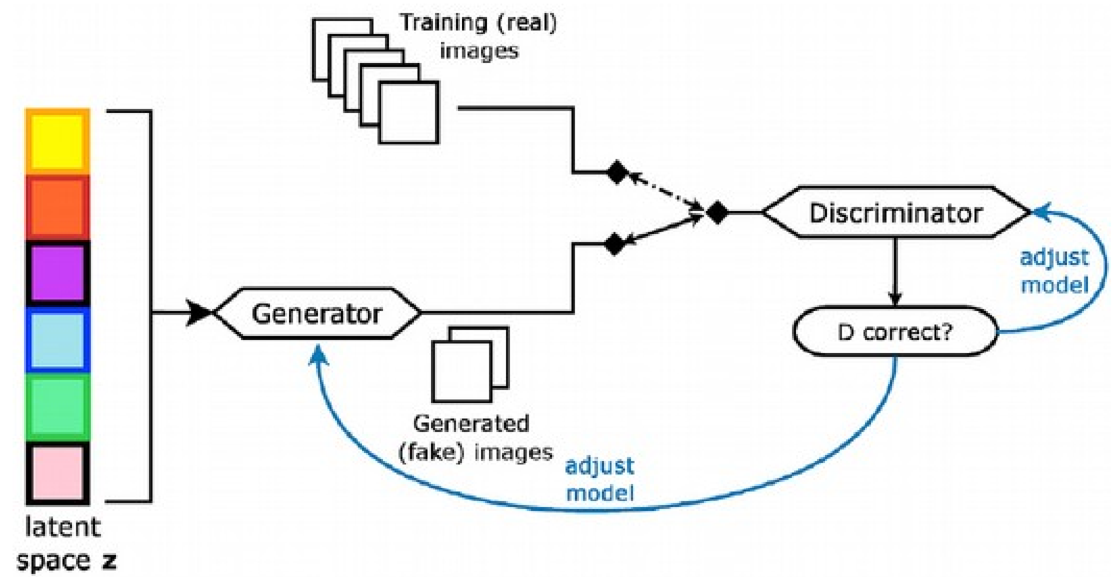
Associating brain regions with categories; modulation of attention

Improving decoding with GANs

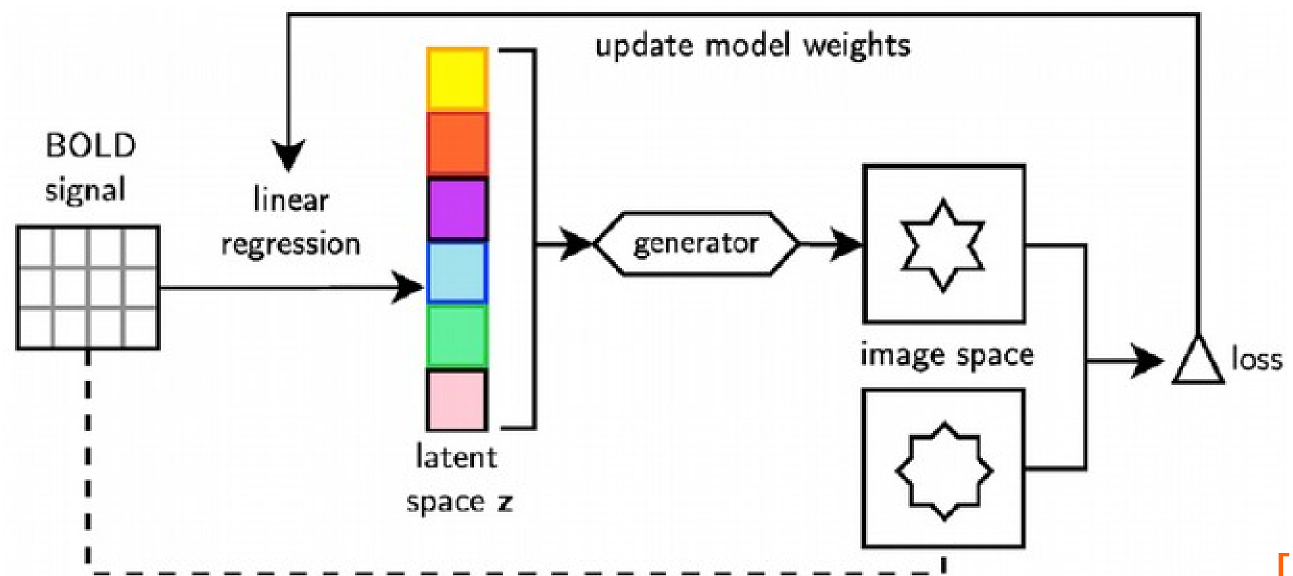


GAN training: generating from latent space [Seeliger et al. Neuroimage 2018]

Improving decoding with GANs



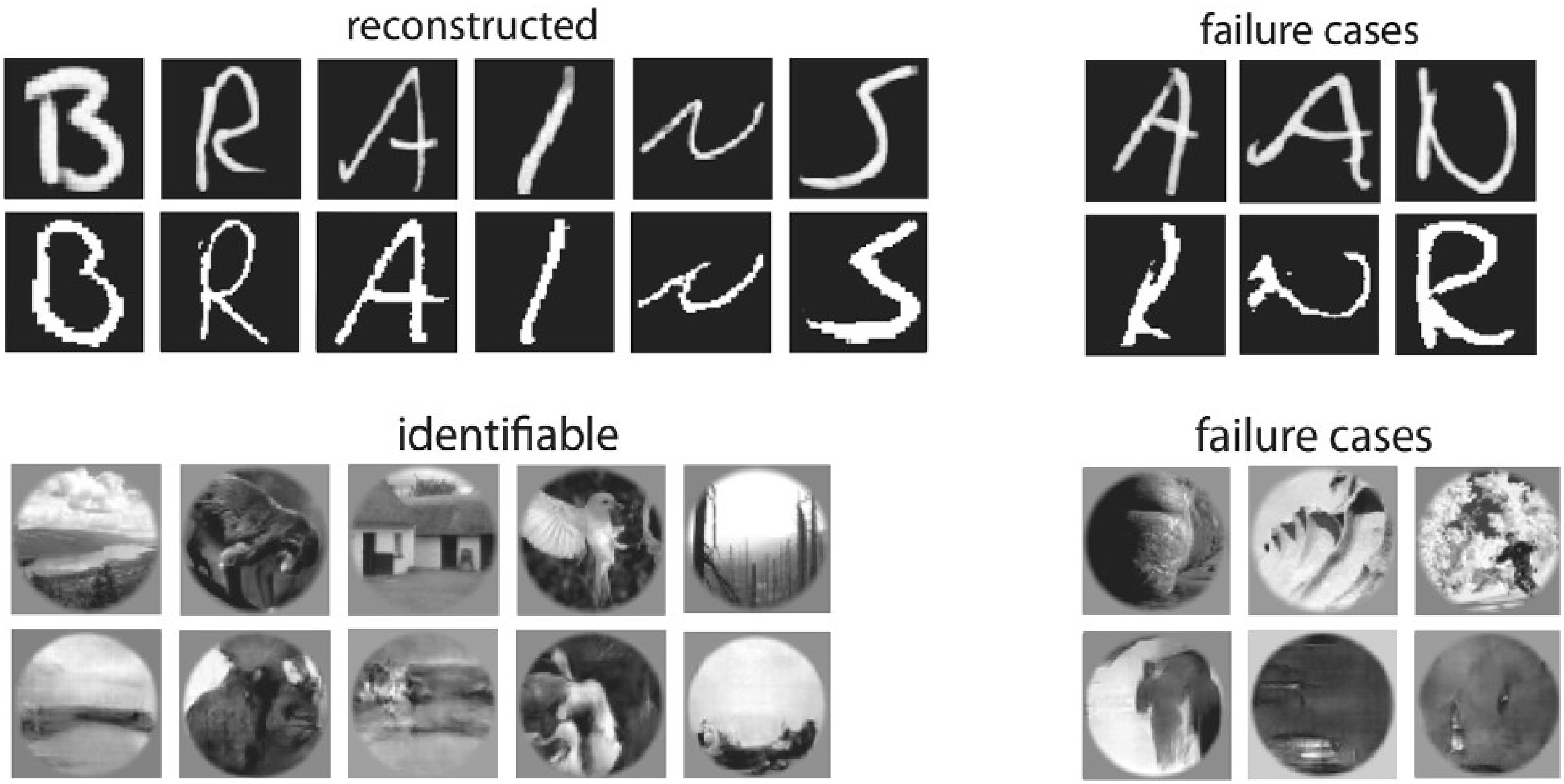
GAN training: generating from latent space



BOLD-based image reconstruction

[Seeliger et al. Neuroimage 2018]

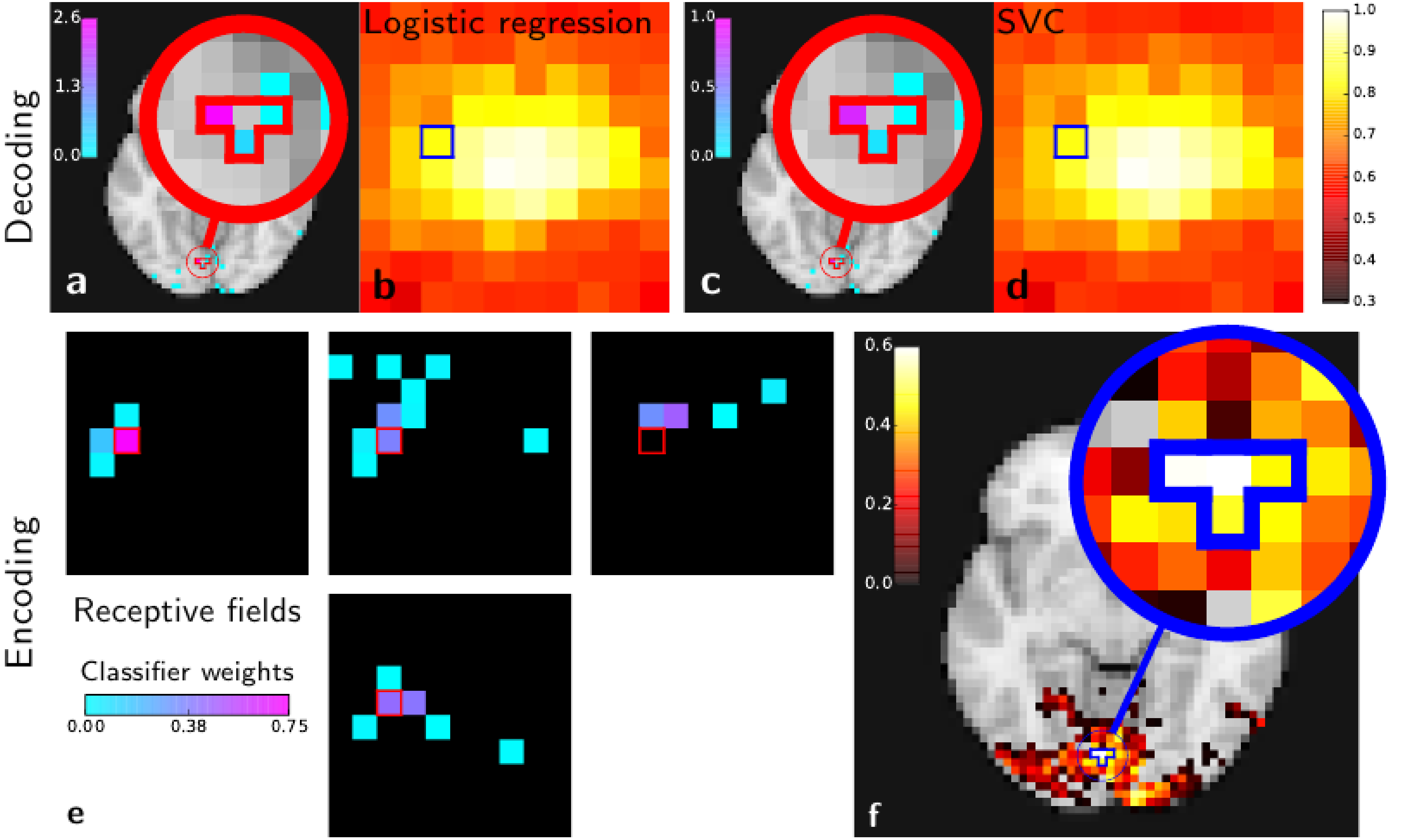
Improving decoding with GANs



[Seeliger et al. Neuroimage 2018]

Do it yourself !

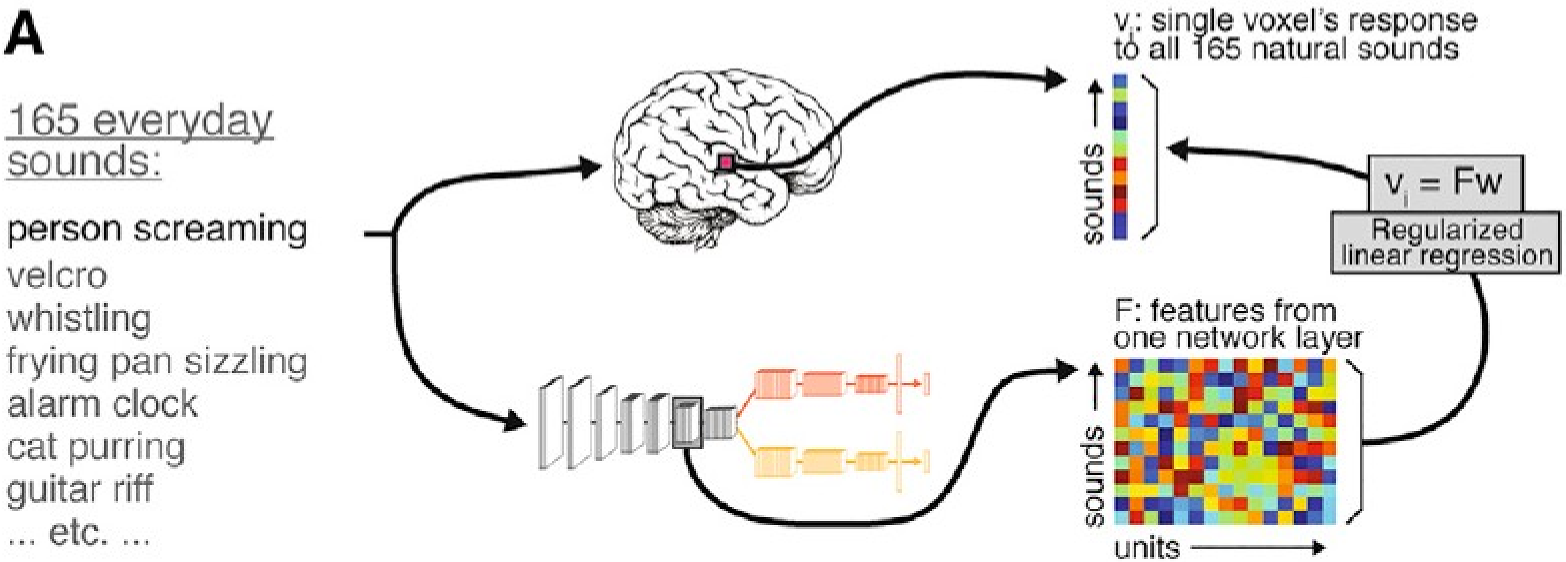
http://nilearn.github.io



Conclusion

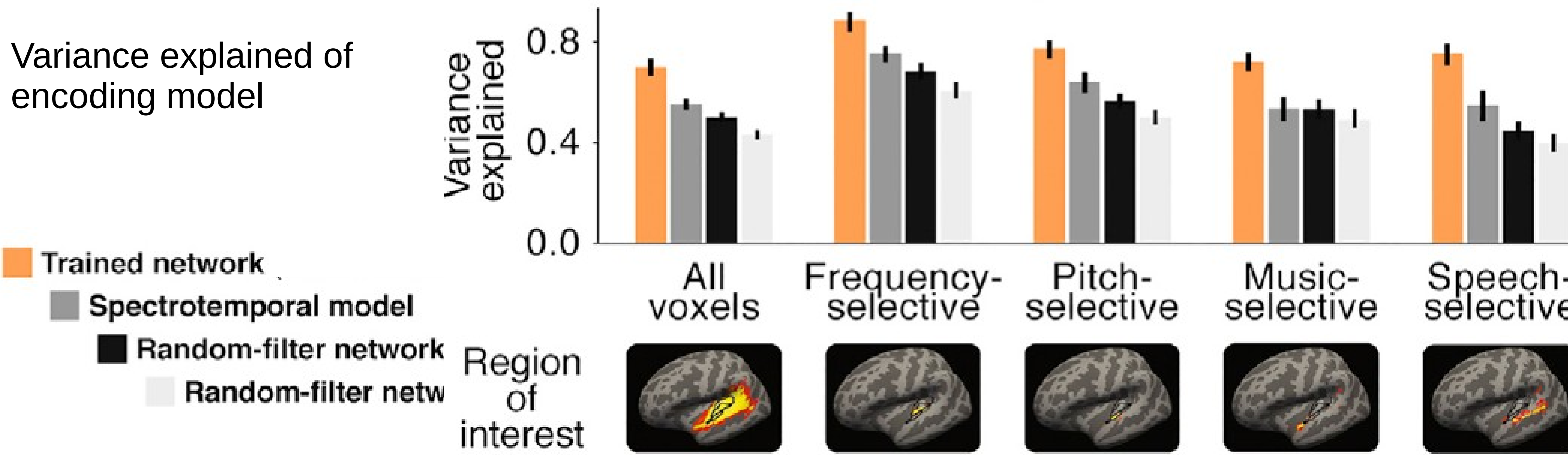
- Compare complex representations of cognitive content to brain activity
  - Recent deep learning methods → great candidate models
- Bottleneck: Limited SNR and number of samples
- Works for audition too, maybe NLP
- Future:
  - High(er) resolution fMRI
  - task-specific architectures

Audition [Kell et al. Neuron 2018]

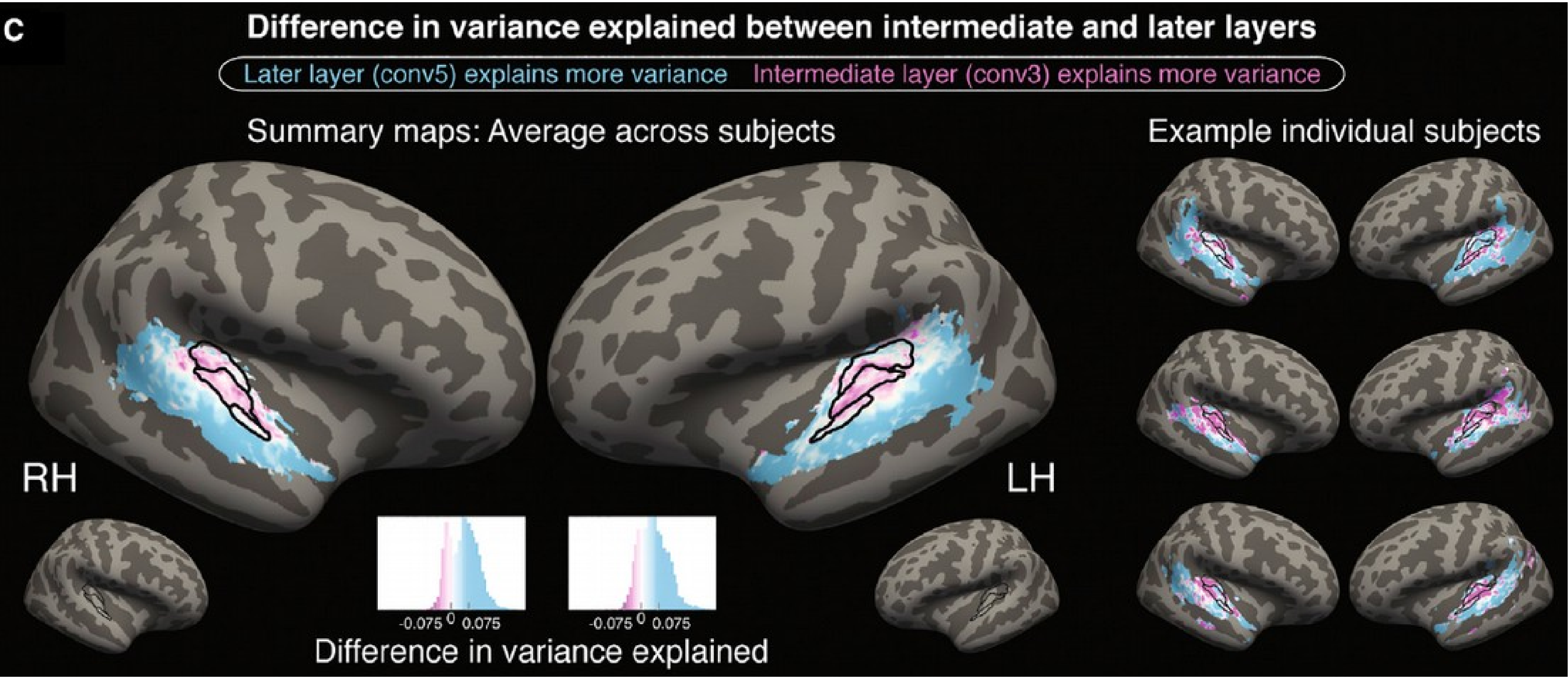


Encoding model based on a dual task network (speech and music)

Variance explained of encoding model



Audition [Kell et al. Neuron 2018]



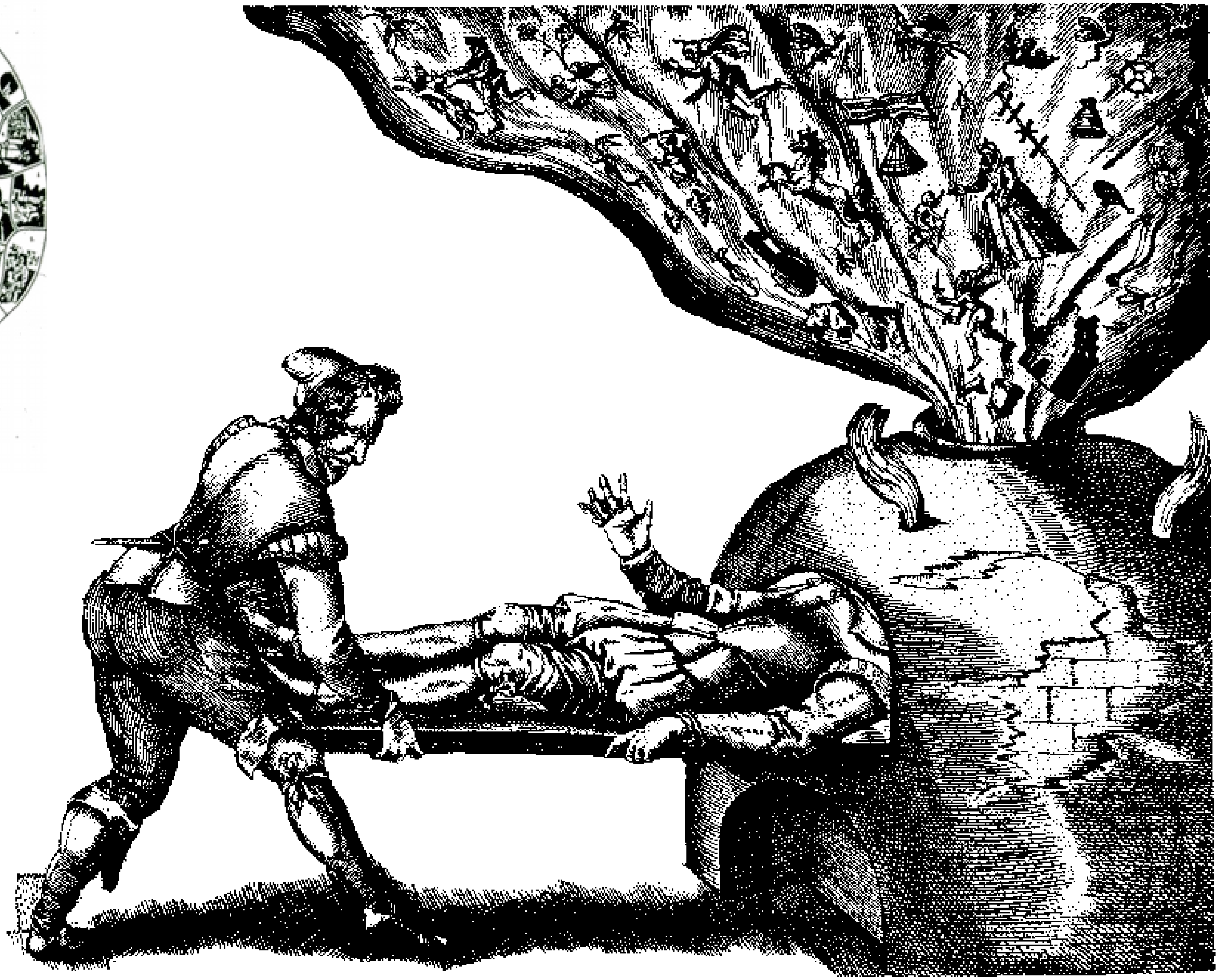
Gradient of complexity along auditory areas

Acknowledgements

Parietal
G. Varoquaux,
A. Gramfort
A.L. Grilo Pinho,
H. Richard
M.Eickenberg

Other collaborators
S. Dehaene
S.Mallat

Human Brain Project
université PARIS-SACLAY
ANR



Bilbiography: deep encoding of vision

- Eickenberg M, Gramfort A, Varoquaux G, Thirion B. Seeing it all: Convolutional network layers map the function of the human visual system. Neuroimage. 2017 May 15;152:184-194.
- Seeliger K, Güçlü U, Ambrogioni L, Güçlütürk Y, van Gerven MAJ. Generative adversarial networks for reconstructing natural images from brain activity. Neuroimage. 2018 Nov 1;181:775-785.
- Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc Natl Acad Sci U S A. 2014 Jun 10;111(23):8619-24.
- Güçlü U, van Gerven MA. Deep Neural Networks Reveal a Gradient in the Complexity of Neural Representations across the Ventral Stream. J Neurosci. 2015. Jul 8;35(27):10005-14.
- Hugo Richard, Ana Pinho, Bertrand Thirion, Guillaume Charpiat. Optimizing deep video representation to match brain activity CCN 2018 – Conference on Cognitive Computational Neuroscience, Sep 2018,

Some landmark experiments

- Kay KN, Naselaris T, Prenger RJ, Gallant JL. Identifying natural images from human brain activity. Nature. 2008 Mar 20;452(7185):352-5.
- Huth AG, Nishimoto S, Vu AT, Gallant JL. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron. 2012 Dec 20;76(6):1210-24.
- Naselaris T, Prenger RJ, Kay KN, Oliver M, Gallant JL. Bayesian reconstruction of natural images from human brain activity. Neuron. 2009 Sep 24;63(6):902-15.
- Miyawaki Y, Uchida H, Yamashita O, Sato MA, Morito Y, Tanabe HC, Sadato N, Kamitani Y. Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron. 2008 Dec 10;60(5):915-29.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 2001 Sep 28;293(5539):2425-30.

Hmax and scattering

- Serre T, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust object recognition with cortex-like mechanisms. IEEE Trans Pattern Anal Mach Intell. 2007 Mar;29(3):411-26.
- Riesenhuber M, Poggio T. Hierarchical models of object recognition in cortex. Nat Neurosci. 1999 Nov;2(11):1019-25.
- Bruna J, Mallat S. Invariant scattering convolution networks. IEEE Trans Pattern Anal Mach Intell. 2013 Aug;35(8):1872-86.
- Michael Eickenberg, Fabian Pedregosa, Senoussi Mehdi, Alexandre Gramfort, Bertrand Thirion Second order scattering descriptors predict fMRI activity due to visual textures PRNI 2013 – 3nd International Workshop on Pattern Recognition in NeuroImaging, Jun 2013

Bibliography: reconstruction of percepts & mental images

- Senden M, Emmerling TC, van Hoof R, Frost MA, Goebel R. Reconstructingimagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception. Brain Struct Funct. 2019 Apr;224(3):1167-1183
- Thirion B, Duchesnay E, Hubbard E, Dubois J, Poline JB, Lebihan D, Dehaene S. Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage. 2006 Dec;33(4):1104-16.

Bibliography: retinotopy with fMRI

- Sereno MI, McDonald CT, Allman JM. Analysis of retinotopic maps in extrastriate cortex. Cereb Cortex. 1994 Nov-Dec;4(6):601-20.
- Dumoulin SO, Wandell BA. Population receptive field estimates in human visual cortex. Neuroimage. 2008 Jan 15;39(2):647-60.