Gigantic mouses fMRI project

# Papers

## Cichy et al., 2016, Sci. Rep.

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific Reports*, *6*(1), 27755.<https://doi.org/10/f8q3v2> **- Manshan**

* Model RDMs from different layers of AlexNet (pretrained on ImageNet) → RSA with fMRI and MEG data (see Cichy et al., 2014)

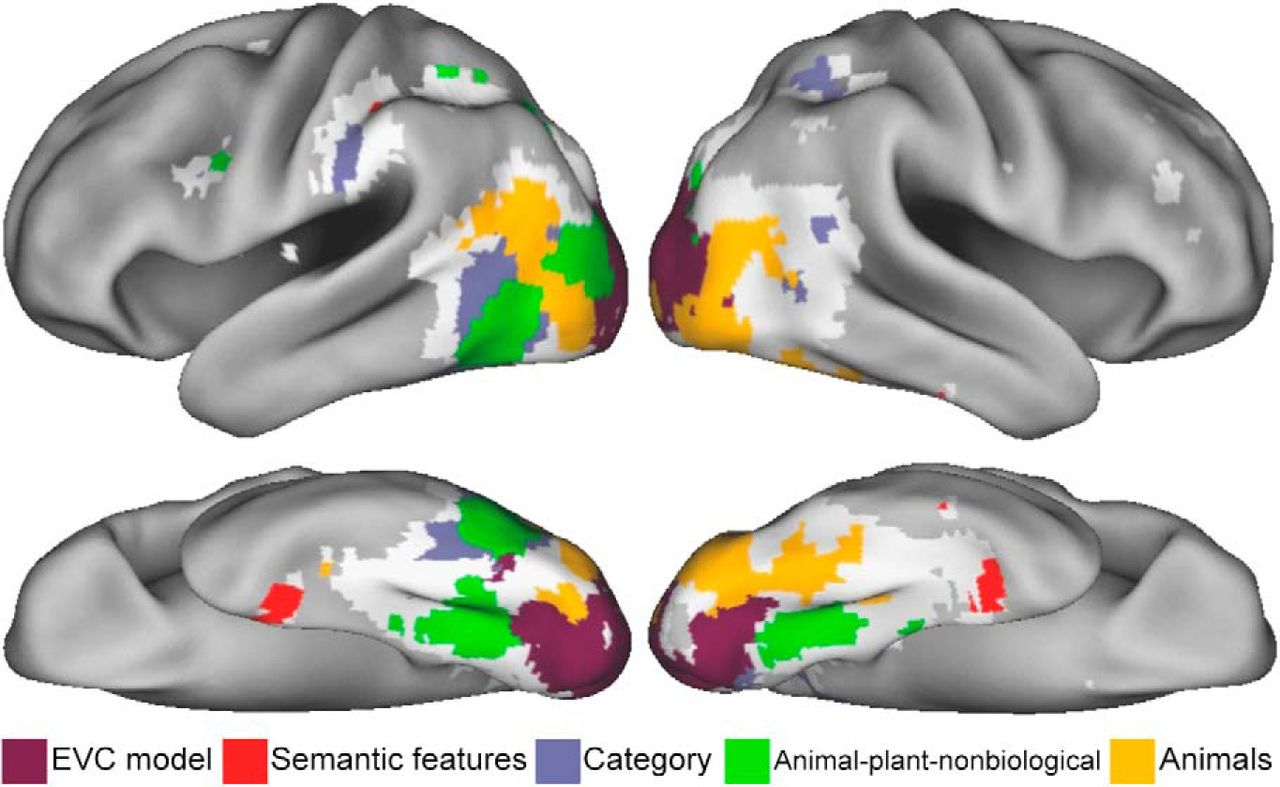
## Cichy et al., 2014, Nat. Neurosci.

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*(3), 455–462.<https://doi.org/10/gbfrr3> **- Gina**

## Clarke & Tyler, 2014, J. Neurosci.

Clarke, A., & Tyler, L. K. (2014). Object-specific semantic coding in human perirhinal cortex. *Journal of Neuroscience*, *34*(14), 4766–4775.<https://doi.org/10/f5wz94> **- Alex**

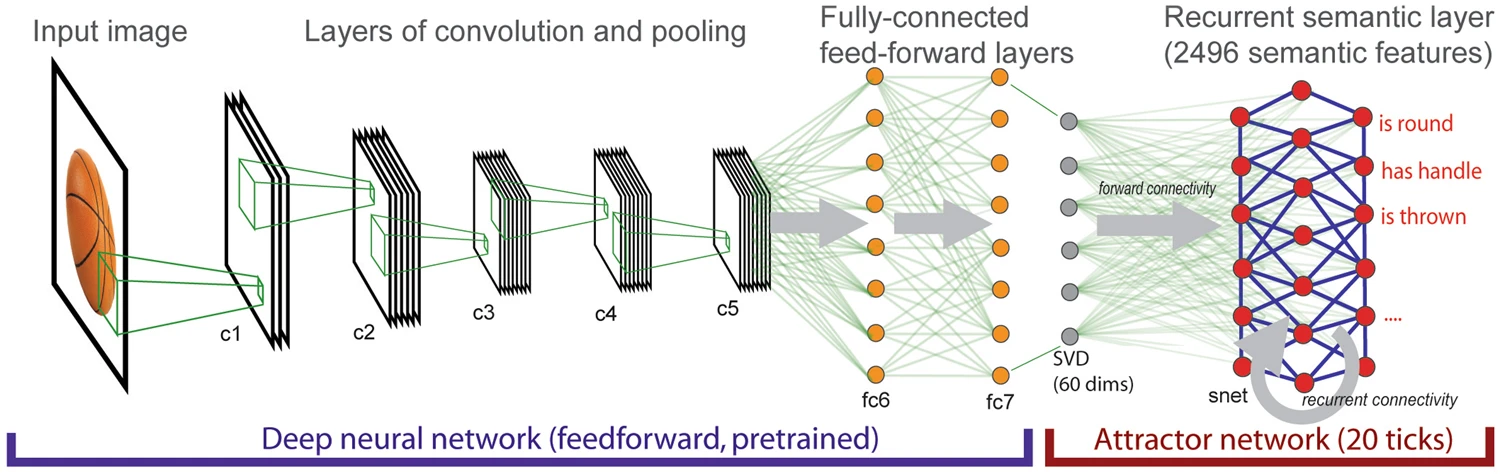
* *Research question*: What type of information (visual or semantic) about specific objects is represented where in the ventral visual stream?
* *General approach*: Different visual and semantic models <> RSA with fMRI data
* *fMRI data*: N = 16 human adults, naming 131 common objects from 6 categories
* *Comp. models*: early visual cortex model (HMax C1), object categories, individual object semantic features
* *Findings*: visual model fits EVC, object categories fit pVTC, object-specific semantic features fit PRC



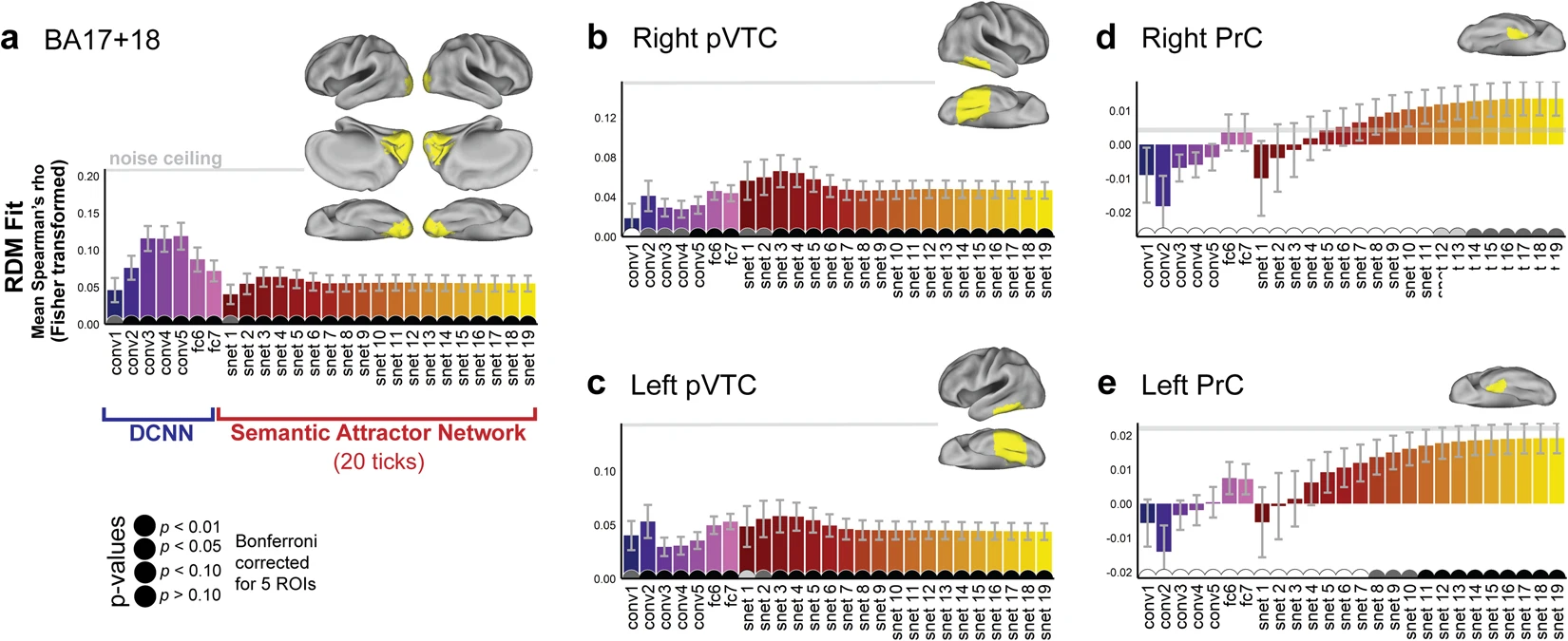
## Devereux et al., 2018, Sci. Rep.

Devereux, B. J., Clarke, A., & Tyler, L. K. (2018). Integrated deep visual and semantic attractor neural networks predict fMRI pattern-information along the ventral object processing pathway. *Scientific Reports*, *8*(1), 10636.<https://doi.org/10/gdzb3f> **- Alex**

* *Research question*: How can we model visual processing, taking into account not only visual but also semantic information about objects?
* *General approach*: Combined visual + semantic DNN <> RSA with fMRI data
* *fMRI data*: same as Clarke & Tyler (2014)
* *Comp. model*: AlexNet (Krievheksy et al., 2012) trained on ImageNet → 5 convolutional layers, 2 fully connected layers, 1 output layer → Output layer replaced with a distributed semantic feature network (2,469 feature nodes)



* *RSA*: 26 model RDMs for different layers in the network → Spearman’s correlation with fMRI data RDMs for 5 ROIs (MATLAB MRC-CBU RSA toolbox) + whole-brain searchlight
* *Findings*: Visual layers provide best fit in EVC, early semantic layers provide best fit in pVTC, late semantic layers provide best fit in PrC



* *Conclusion*: Posterior-to-anterior shift of representational content, from vision to semantics

## Kriegeskorte et al., 2008, Front. in Sys. Neurosci.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*.<https://doi.org/10/fh5dft> **- Katherine**

Representational Dissimilarity Matrices (RDMs): Quantitatively relate brain activity, behavioral measurement and computational modelling. Correspond different measurements of the brain, with different type of stimulus and various representations with neuronal methods with only RDMs.

Potential approaches if using Representational Similarity Analysis (RSA):

(1) Integration of computational modeling into the analysis of brain-activity data.

(2) Relating regions, subjects, species, and modalities of brain activity measurement.

(3) Relating brain and behavior.

(4) Addressing a broader array of neuroscientifi c questions with each experiment by means of condition-rich design

REPRESENTATIONAL SIMILARITY ANALYSIS – STEP-BY-STEP

STEP 1: ESTIMATING THE ACTIVITY PATTERNS: Univariate linear modeling of the voxels ROIs fMRI and the condition.

STEP 2: MEASURING ACTIVITY-PATTERN DISSIMILARITY: Correlation distance (1-correlation), euclidean distance, mahalanobis distance, etc

STEP 3: PREDICTING REPRESENTATIONAL SIMILARITY WITH A RANGE OF MODELS.

# Questions

**What exact aspect of data needs modeling?**

* Human fMRI responses to objects have a different (dis)similarity structure in different brain areas (e.g., EVC, IT)
* This indicates different types of representations in that regions that need to be modelled
* Research question: How do the representations in lower- vs. higher-level visual brain areas correspond to the representations in state-of-the-art computer vision models?
* Hypotheses:

1. Representations in an early visual brain area (EVC) are more similar to the representations of initial (as compared to deeper) DNN layers
2. Representations in a higher-level visual brain area (IT) are more similar to the representations of deeper (as compared to initial) DNN layers

* This is already done in the time domain (MEG <> DNN), let’s do it in the visual hierarchy (EVC vs. IT) → fMRI <> DNN
* We don’t (yet) focus on using another, more interesting network (e.g., including semantics)
* Our data are RDMs from:
  + MEG: 16 subjects × 2 repetitions × 1301 time points × 92 images × images
  + fMRI-EVC and fMRI-IT: 15 subjects × 92 images × 92 images

**· Define an evaluation method!**

o How will you know your modeling is good?

1. We want to modulate the correspondence of the early layers DNNs w.r.t. V1 in fMRI for different image categories.
2. We want to modulate the correspondence of the higher layers on DNNs w.r.t. IT in fMRI for different image categories.

o E.g. comparison to specific data (quantitative method of comparison?)

· For computational models: think of an experiment that could test your model

1. We will use Representative Similarity Analysis (RSA) metrics to make comparisons between different layers of DNNs and the fMRI representation within each category.

o You essentially want your model to interface with this experiment, i.e. you want to simulate this experiment

· Survey the literature

o What’s known?/What has already been done? → Correspondence between different areas in the ventral visual stream and DNNs

o Previous models as a starting point? AlexNet

o What hypotheses have been emitted in the field?

o Are there any alternative / complementary modeling approaches?

· What skill sets are required?

o Do I need learn something before I can start?

RSA analysis, DNNs architecture and activation, how to measure the activity patterns of the DNNs for different layers.

o Ensure that no important aspect is missed

· Potentially provides specific data sets / alternative modeling approaches for comparison

· What parameters / variables are needed?]

o Constants?

o Do they change over space, time, conditions…?

o What details can be omitted?

o Constraints, initial conditions?

o Model inputs / outputs?

· Variables needed to describe the process to be modelled?

o Brainstorming!

o What can be observed / measured? latent variables?

o Where do these variables come from?

o Do any abstract concepts need to be instantiated as variables?

§ E.g. value, utility, uncertainty, cost, salience, goals, strategy, plant, dynamics

§ Instantiate them so that they relate to potential measurements!

· You think about the hypotheses in words by relating ingredients identified in Step 3

o What is the model mechanism expected to do?

o How are different parameters expected to influence model results?

· You then express these hypotheses in mathematical language by giving the ingredients identified in Step 3 specific variable names.

o Be explicit, e.g. y(t)=f(x(t),k)y(t)=f(x(t),k) but z(t)z(t) doesn’t influence y

# Draft proposal (200–300 words)

Brain areas along the human ventral visual stream contain increasingly abstract representations of visual objects. To characterize these, Cichy et al. (2014) collected fMRI (N = 15) and MEG (N = 16) data. The fMRI data were collected from two visual regions (early visual cortex [EVC] and inferotemporal cortex [IT]) while participants observed 92 different images (two repetitions each). To model the representations in EVC and IT, we apply Representational Similarity Analysis (RSA). We compare the representational dissimilarities in these regions to those of different layers in a pretrained deep neural network (AlexNet; Krizhevsky et al., 2012). We expect that representations in EVC are more similar to the representations of initial (as compared to deeper) DNN layers as indicated by their activation patterns, whereas representations in IT are more similar to the representations of deeper (as compared to initial) DNN layers. After testing these initial hypotheses, other routes could be pursued in an exploratory fashion. These include: (a) using alternative DNN architectures, (b) extending this methodology to the other fMRI datasets available that are applied to other stimulation inputs, (c) using different types of candidate computational models instead of visual DNNs (e.g., language models).

Keywords:

fMRI, MEG, Representational Similarity Analysis, vision, perception, Deep Neural Networks

Group members:

[mzaghi@umh.es](mailto:mzaghi@umh.es); Regina Zaghi

[enge@cbs.mpg.de](mailto:enge@cbs.mpg.de); Alexander Enge

[renzfabian@googlemail.com](mailto:renzfabian@googlemail.com); Fabian Renz

[kabaquerod@unal.edu.co](mailto:kabaquerod@unal.edu.co): Katherine Baquero

[ms0501@mail.ustc.edu.cn](mailto:ms0501@mail.ustc.edu.cn): Manshan Guo

# RSA method

* Tutorial notebook (not from NMA, but looks cool): [https://colab.research.google.com/  
  github/ljchang/dartbrains/blob/master/content/RSA.ipynb](https://colab.research.google.com/github/ljchang/dartbrains/blob/master/content/RSA.ipynb)

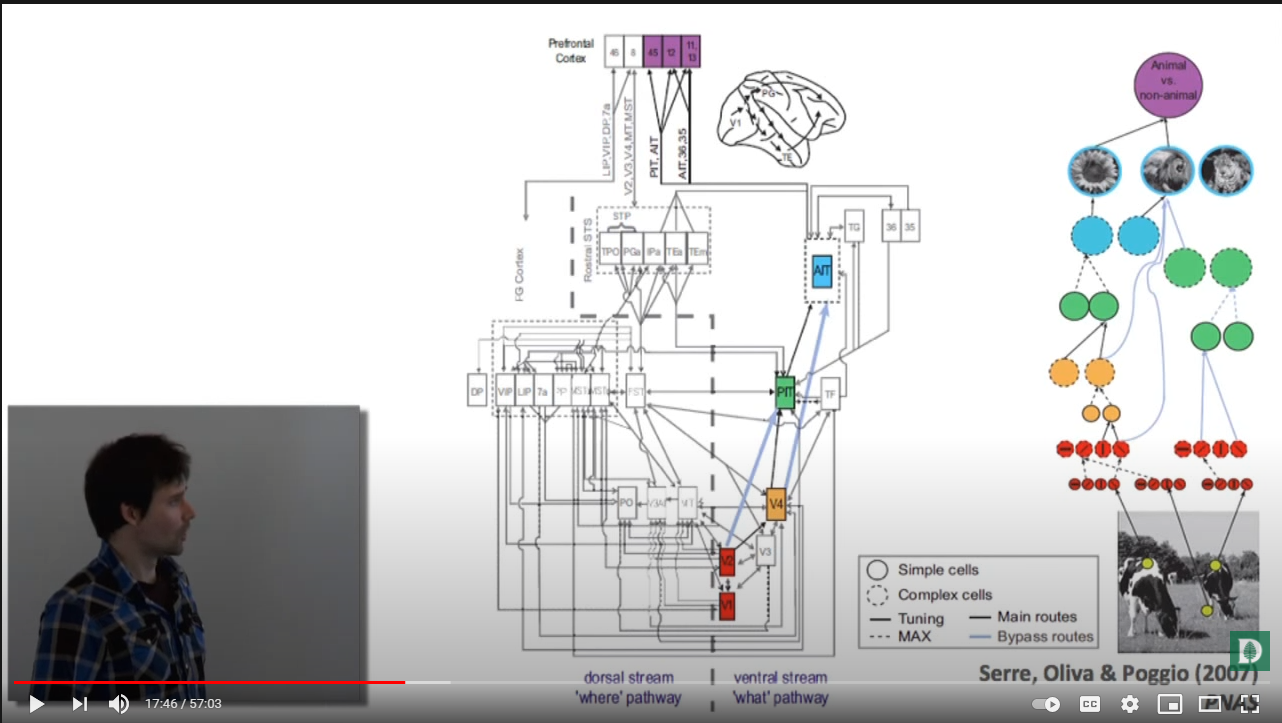
Questions to Tejas Savalia 08/July

1. What are your thoughts about our hypothesis on testing which models (mainly DNNs) and on what layers have a higher similarity with the RDMs of the fMRI data for the different categories?
2. How did they come up with the RSD matrix for the fMRI data for these two ROIs?
   1. From the time series BOLD data
3. How did they come up with the RSD matrix for the MEG? Why does the matrix contain decoding accuracies instead of correlations?
4. Would it be possible to test our model with the other datasets that are available?
5. What kind of other DNNs could be used to compare it? Does it make sense?
6. What would he suggest how we should drive on these three weeks working together and such?
7. Who could help us at the moment of implementing the other DNN models for our project? What packages should we use? Tensorflow, Pytorch. Suggested Pytorch.
   1. check other NNs for this dataset
      1. H1 max model: <https://github.com/wmvanvliet/pytorch_hmax>
      2. VGG 16 <https://github.com/pytorch/vision/blob/master/torchvision/models/vgg.py>
      3. Resnet https://pytorch.org/hub/pytorch\_vision\_resnet/
8. Can we use simpler models (e.g., pixel values, category structure, semantic features) that would rival the performance of AlexNet (or other DNNs)?

RSA = Representational similarity analysis → Correlation between two RDMs

RDM = Representational dissimilarity matrix





Papers to consider:

<https://www.biorxiv.org/content/biorxiv/early/2020/05/08/2020.05.07.082743.full.pdf>

Diverse deep neural networks all predict human IT well, after training and fitting

Here they show the error of different NNs applied to the Imagenet dataset, but focused on Disimilarity matrices and 62 categorical images and compared to fMRI data of the IT region.

