Characterizing functional discriminability across the brain using large-scale classification

Alejandro de la Vega[1], Marie Banich[1], & Tal Yarkoni[2]

delavega@colorado.edu

[1] University of Colorado Boulder

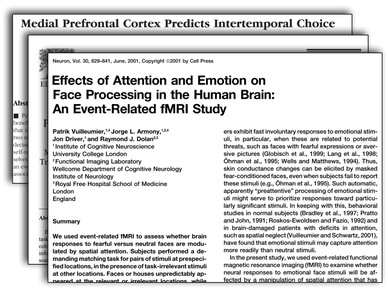
[2] The University of Texas at Austin

Introduction

- Large-scale fMRI databases are useful for robustly characterizing brain regions associated with cognitive functions
  - E.g. meta-analysis of 'reward' in Neurosynth
- However, it is not clear how useful these functions are for discriminating brain regions from each other
- Knowing what a study was about, how well can you predict a region was activated?
- **How well can we discriminate which regions of the brain a study activates, based on the semantics?**
  - How do regions differ?

Methods

- Neurosynth database (neurosynth.org): 8000+ fMRI studies[1]
  - Topic model of word frequencies in papers - semantics[2]
- Parcellated brain using k-means on coactivation of voxels (11, 20 and 60 regions)



Side	Area	X	Y	Z
Effects of Spatial Attention				
Faces > houses				
Right	fusiform gyrus	44	−54	−20
Left	fusiform gyrus	−44	−54	−20
Left	inferior temporal gyrus	−40	−78	−18
Houses > faces				
Right	parahippocampal gyrus	32	−44	−10
Right	parahippocampal gyrus	28	−40	−16
Left	parahippocampal gyrus	−26	−42	−14
Right	retrosplenial cortex	16	−56	16
Left	retrosplenial cortex	−12	−56	18
Right	lateral occipital gyrus	36	−82	24
Left	lateral occipital gyrus	−32	−86	34

social empathy mental

reward decision choice

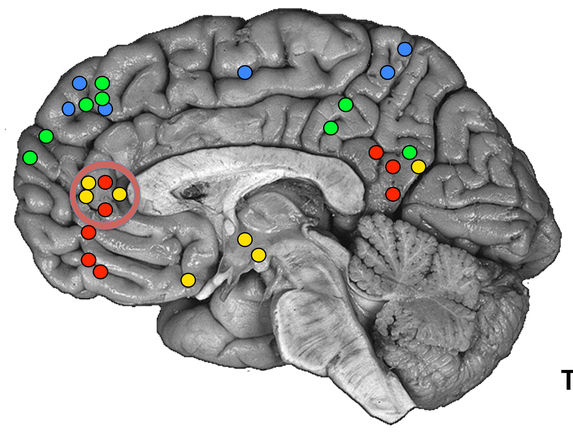
inhibition conflict control

body visual images

category adaptation representation

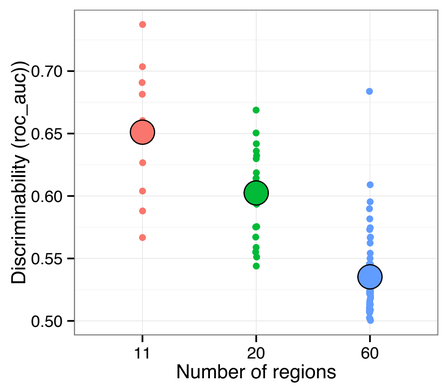
Classification analysis

1. Selected studies that activate and do not activate regions
   - Active - 5% of voxels in ROI active
   - Not active - zero voxels in ROI active
2. For each region individually, trained Ride classifier (python sklearn) to discriminate 'active' vs 'not active' studies using paper semantics (topics) as features
   - Ridge - regularized regression classifier - produces linearly interpretable weights for each topic
3. Evaluated using 4-fold cross-validation (tested models on unseen data) and scored using ROC area under curve

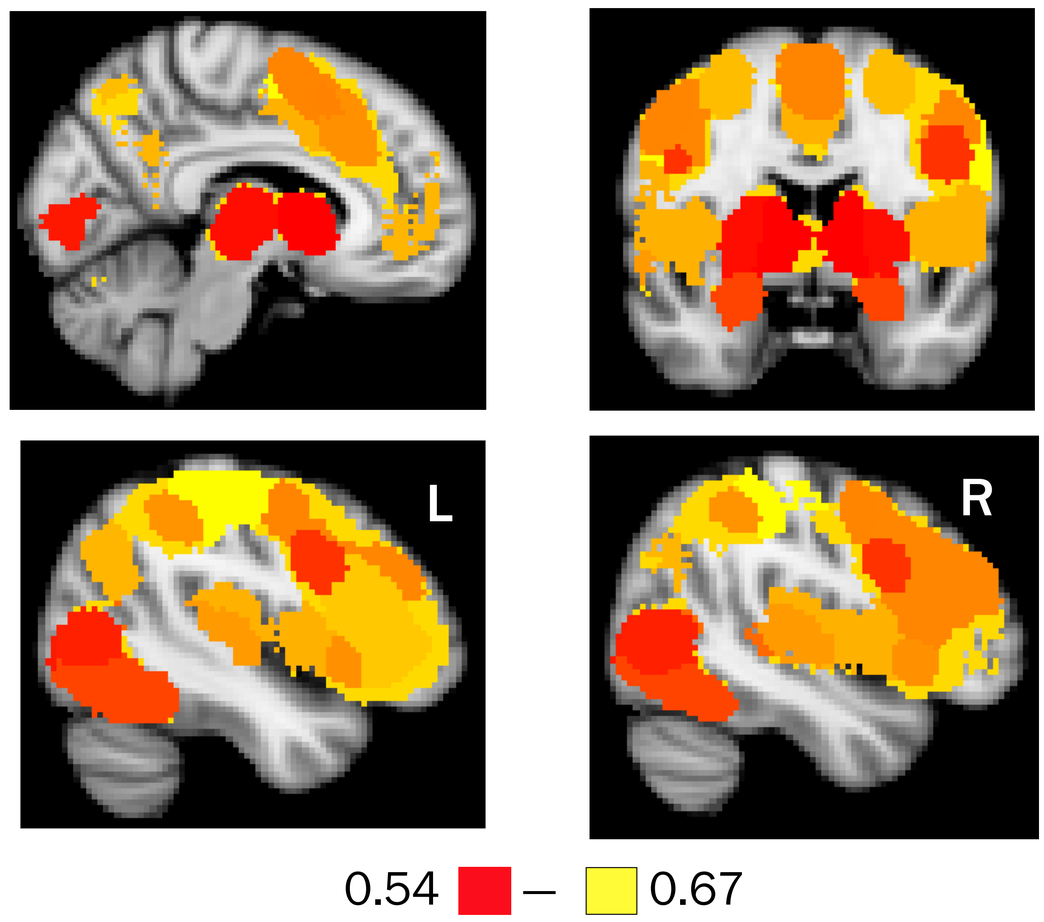


Active?	study	vision	social	motor
Yes	10.1016/j.neuroscie.2004	6	2	3
Yes	10.1012/science.2005.07.0	8	3	3
No	10.1037/neuron.2004.07	2	3	3
No	10.1016/j.neuropsychologia	1	1	4
Test: ???	12.1316/j.neurosci.2002.07	**7**	2	3

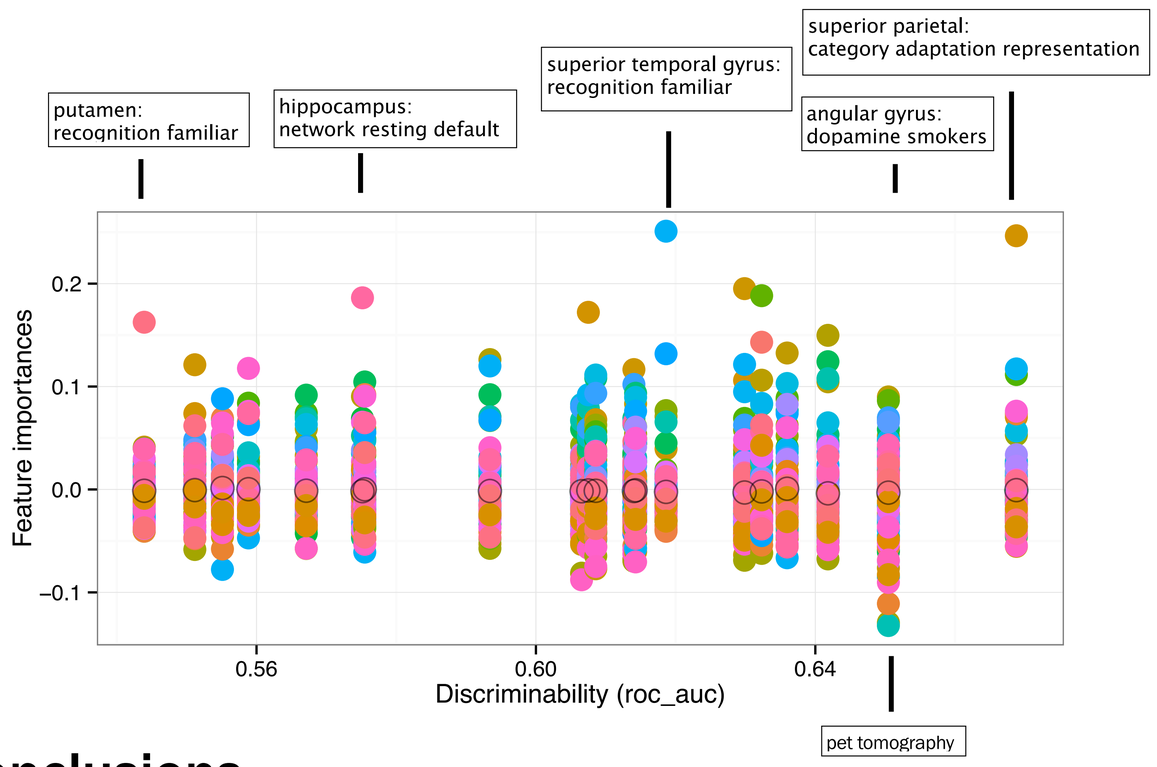
Classification results



- Above chance classification
- Chance is 0.5
- Decreased accuracy with increased resolution
- Wide variability across brain



Which topics are important for classification?



Conclusions

- We developed a measure of how discriminable brain regions are using cognitive function in Neurosynth
- Regions vary on discriminability
  - Less sparse combination of features yields better discriminability
- Features that support classification mostly agree with univariate approaches, but yield some surprising results

References

1. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature Methods
2. Poldrack, R. A., Mumford, J. A., Schonberg, T., Kalar, D., Barman, B., & Yarkoni, T. (2012). Discovering Relations Between Mind, Brain, and Mental Disorders Using Topic Mapping. PLoS Computational Biology,