Neuroscout: a cloud-based platform for flexible re-analysis of naturalistic fMRI datasets

Alejandro de la Vega[1]*, Quinten McNamara[1], Michael Hanke[2], & Tal Yarkoni[1]**
[1]Department of Psychology, University of Texas at Austin. [2]Institute of Psychology, University of Magdeburg
*delavega@utexas.edu; **tyarkoni@utexas.edu

Introduction

• The utility of fMRI is constrained by its resource-intensive nature, as only a small fraction of research hypotheses are ever tested, even among open datasets.
• We present a platform for rapid and flexible fMRI analysis, enabling researchers to test novel theoretical hypotheses in existing datasets.
• We focus on extracting maximum utility from experiments that use intrinsically high dimensional naturalistic stimuli such as movies and audio narratives.

Neuroscout Architecture

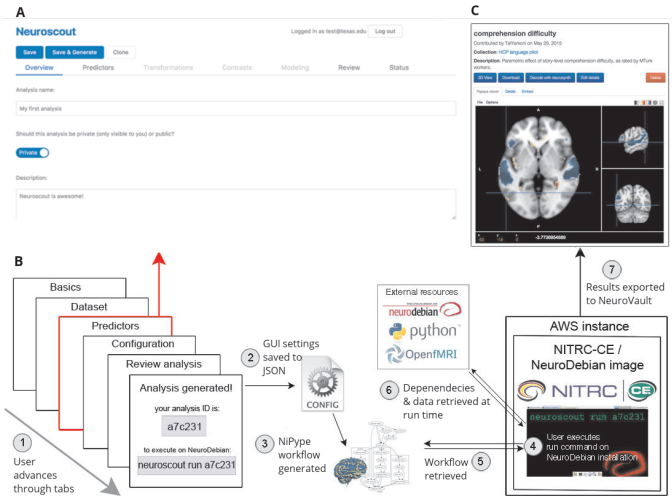

Figure 1. Neuroscout end-to-end analysis.
A) Online analysis builder used to define analyses.
B) Back-end API validates analysis, and generates a NiPype workflow which can then be deployed locally or in the cloud with ease
C) Results are made available as interactive, editable, and sharable NeuroVault reports.

Backend RESTFul API

• Built using Python Flask web microframework.
• Analysis history is linked to prevent selective reporting.
• Containerized using Docker

Frontend

• Singe page app (SPA) built using React.js

Analysis engine command line interface

• NeuroDebian image manages software dependencies
• DataLad manages data dependencies, and only necessary data is downloaded as necessary
• BIDS-compatible NiPype workflow executes fMRI model

Pliers - Automated feature extraction framework[1]

• Naturalstic datasets are rich and ecologically valid
• Manually coding of features is time-consuming, limiting
• Uniform, flexible interface for extractors, including feature extraction APIs and manual annotations
• Extracts up to thousands of features for each dataset

1. McNamara, Q., De la Vega, A., & Yarkoni, T. (2017). Developing a comprehensive framework for multimodal feature extraction. Proc. of the 22nd ACM SIGKDD

fMRI Analysis Demonstration

• Human Connectome Project dataset, movie watching task
60 minutes; N=35 subjects
• Standard univariate two-level FSL pipeline

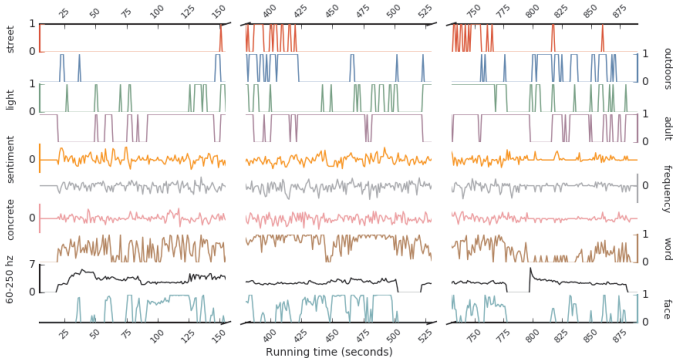

Figure 2. Timeline of automatically extracted features.
• Clarifai image labels ('street', 'outdoors', 'light' and 'adult')
• Google Vision Face Detection ('face')
• Short-time fourier transform ('60-250 Hz')
• IBMWatson was used to transcribe movie audio. We modeled onsets ('word') and extracted lexical norms ('frequency', 'concreteness') and 'sentiment' using Indico.

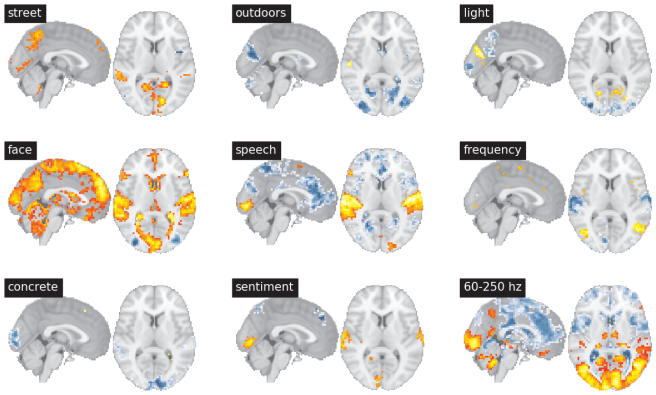

Figure 3. Brain activity associated with extracted features, thresholded at p<0.001.

• Brain activity patterns resembled neural correlates previously identified in conventional factorial experiments.
• Speech was associated with language processing regions, such as the superior frontal gyrus.
• Image tags were associated with differential activation in brain regions important for natural scene recognition (e.g. visual and retrosplenial cortices)

Conclusions

• Neuroscout will provide a turnkey solution for extremely rapid analysis and visualization of existing fMRI data at a marginal cost very close to zero.
• We aim to incentivize its use by seamlessly integrating results with broader ecosystem of data sharing, visualization and interpretation.

www.github.com/PsychoinformaticsLab/neuroscout
www.github.com/tyarkoni/pliers