Github
Demo codes
Developing Version

python™ HPC Workbench Toolbox [6]

HUMAN Connectome PROJECT

Yale Yale Psychiatry Cho Lab YALE CHILD STUDY CENTER Where discovery inspires care VANDERBILT UNIVERSITY

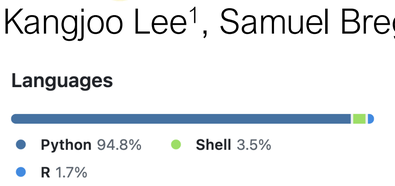
Correspondence to: kangjoo.lee@yale.edu

BrainCAP

An open-source neuroimaging toolkit to analyze brain co-activation patterns

Kangjoo Lee[1], Samuel Brege[1], Zailyn Tamayo[1], Catie Chang[2], Youngsun Cho[1,3]

1 Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA; 2 Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, USA; 3 Child Study Center, Yale University School of Medicine, New Haven, CT, USA



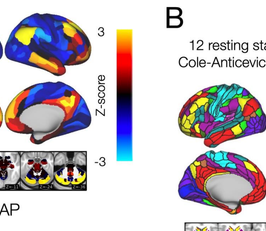
- The analysis of moment-to-moment changes in co-activation patterns (CAPs) in fMRI is useful to study dynamic properties of neural activity [1].
- It is based on Clustering fMRI timeframes to identify recurrent spatial patterns within and across subjects and to quantify their temporal profiles [2]
- We have developed an open-source platform for cross-sectional and longitudinal studies to allow a robust feature selection for reproducible brain-behavior mapping and benchmarking analytic choices. [3,4]
- Validation data: Human Connectome Project S1200 Young Adults, resting-state fMRI (M=337) [5]

Workflow Configuration

Global	Schedular	I/O	Modular	Performance	Version Control
•**Centralized parameter management in a single YAML file** •Consistent, reproducible setup across all workflow stages •Supports default and overridden values for flexibility	•Integrated scheduler settings (e.g., SLURM, local) •Control over resource allocation, job dependencies, and execution mode •Options to run steps **sequentially** or **independently**	•Define input datasets and file formats •Output directory structure with customizable prefixes •Automatic **logging** and directory creation	•**Configure step-specific parameters** •Enable or disable steps dynamically •Modular design supports easy extension and maintenance	Influenced by: •Sample size and data complexity •Input data format (e.g., voxels vs parcels) •Parallel processing capabilities •I/O throughput and compute infrastructure	•Tagged in **Github** •Singularity / Apptainer containers are created based on Github tagged versions. •Portable across different computing environments

Pipeline

Split-half cross-validation for reproducibility



- Spatial topography of CAPs
- Temporal metrics of CAPs
- Individual/group preference toward a specific number of CAPs
- Prediction of behavioral phenotypes and symptoms using CAP-derived neural features

Temporal sampling

- Seed-free vs seed-based
- Specific vs random temporal sampling
- Motion removal – scrubbing, interpolation
- Seed regions – parcel index, volumetric mask
- Seed signal generation – averaging, normalization
- Seed signal threshold – percentage

Spatial characterization

Spatial similarity of CAPs to canonical resting state networks

Adapted from Lee et al. PLoS Biology, 2024

Preprocessed data in BIDS format

Study Type
- Cross-sectional
- Longitudinal

Format
- Grayordinate CIFTI
- Volumetric NIFTI

Level
- Dense (voxel, vertex)
- Parcel

Input Data

- Imaging Data in BIDS format

```
Data/
└─ sub-id/
   ├─ anat/
   │  └─ T1w.mnii
   └─ func/
      ├─ bold1.nii
      ├─ bold2.nii
      ├─ bold3.nii
      ├─ bold4.nii
      ├─ bold1_motion.cvs
      ├─ bold2_motion.cvs
      ├─ bold3_motion.cvs
      └─ bold4_motion.cvs
```

Imaging Data

Motion Data (e.g. Frame Displacement)

- List of subjects

	A	B
1	subject_id	group
2	sub_01	control
3	sub_02	control
4	sub_03	control
5	sub_04	SCZ
6	sub_05	SCZ
7	sub_06	SCZ
8	sub_07	Bipolar
9	sub_08	Bipolar

Multi-session, multi-group fMRI concatenation

Temporal sampling

Clustering & CAP identification

Basis CAP generation

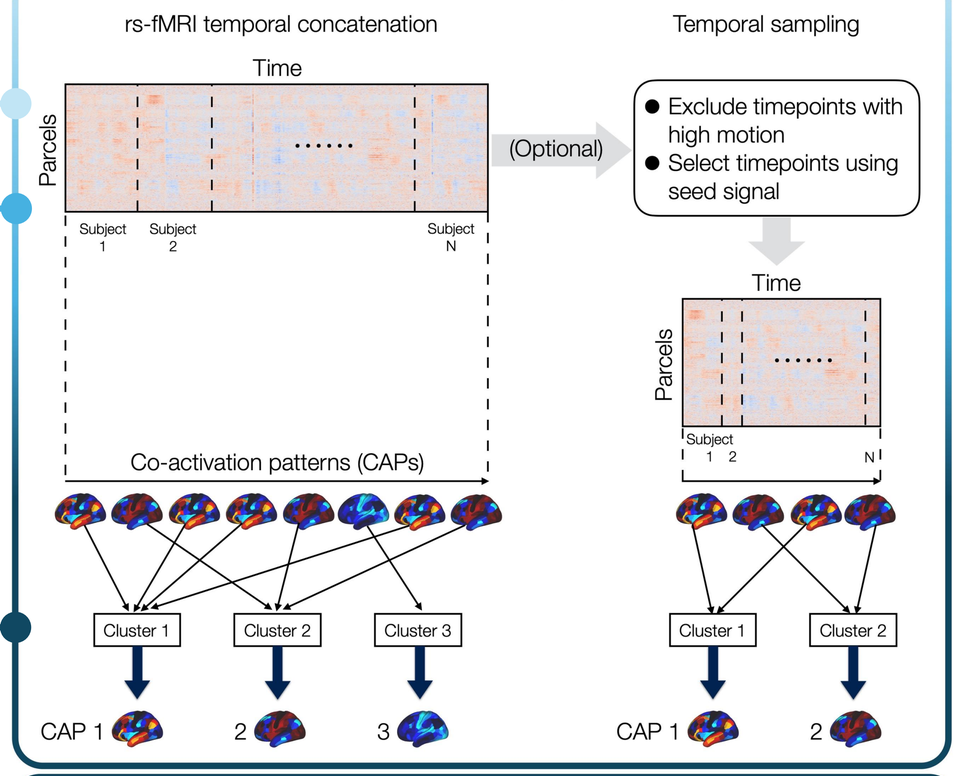
Identification of spatial characteristics

Quantification of temporal metrics

2nd level analyses

Quality control

Estimation of co-activation patterns from fMRI



Analytic Choices

Clustering Algorithms
- All algorithms supported by *scikit-learn.clustering* library

Number of Clusters
- Elbow approach
- Silhouette score

Quantification of temporal metrics

- Fractional Occupancy
- Dwell time
- Transition Probability
- Fractional Entry

Adapted from Lee et al. PLoS Biology, 2024

2nd level analyses

Feature selection

Dimension Reduction

Adapted from Lee et al. PLoS Biology, 2024

Behavioral Prediction

Cross-sectional / Longitudinal Modelling

(Under Development)

References

1. Liu, X. (2013). Proc Natl Acad Sci U S A, 110(11), 4392-7.
2. Liu, X. (2018). Neuroimage, 180(Pt B), 485-494.
3. Lee, K. (2024). PLoS Biol, 22(9), e3002808.
4. https://github.com/Kangjoo/BrainCAP/tree/develop
5. Van Essen, DC. (2012). Neuroimage, 62(4), 2222-31.
6. Gorgolewski, KJ. (2016). Sci Data, 3:160044.

Fundings

NIH National Institute of Mental Health