

# NeuroCAPs: A Python Package for Performing Co-Activation Patterns Analyses on Resting-State and Task-Based fMRI Data

Donisha Smith<sup>1</sup>

<sup>1</sup> Department of Psychology, Florida International University

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

Co-Activation Patterns (CAPs) is a dynamic functional connectivity technique that clusters similar spatial distributions of brain activity. To make this analytical technique more accessible to neuroimaging researchers, NeuroCAPs, an open source Python package, was developed. This package performs end-to-end CAPs analyses on preprocessed resting-state or task-based functional magnetic resonance imaging (fMRI) data, and is most optimized for data preprocessed with fMRIPrep, a robust preprocessing pipeline designed to minimize manual user input and enhance reproducibility ([Esteban et al., 2019](#)).

## Background

Numerous fMRI studies employ static functional connectivity (sFC) techniques to analyze correlative activity within and between brain regions. However, these approaches operate under the assumption that functional connectivity patterns, which change within seconds ([Jiang et al., 2022](#)), remain stationary throughout the entire data acquisition period ([Hutchison et al., 2013](#)).

Unlike sFC approaches, dynamic functional connectivity (dFC) methods enable the analysis of dynamic functional states, which are characterized by consistent, replicable, and distinct periods of time-varying brain connectivity patterns ([Rabany et al., 2019](#)). Among these techniques, CAPs analysis aggregates similar spatial distributions of brain activity using clustering techniques, typically the k-means algorithm, to capture the dynamic nature of brain activity ([Liu et al., 2013, 2018](#)).

## Statement of Need

The typical CAPs workflow can be programmatically time-consuming to manually orchestrate as it generally entails several steps:

1. implement spatial dimensionality reduction of timeseries data
2. perform nuisance regression and scrub high-motion volumes (excessive head motion)
3. concatenate the timeseries data from multiple subjects into a single matrix
4. apply k-means clustering to the concatenated data and select the optimal number of clusters (CAPs) using heuristics such as the elbow or silhouette methods
5. generate different visualizations to enhance the interpretability of the CAP

While other excellent CAPs toolboxes exist, they are often implemented in proprietary languages such as MATLAB (which is the case for TbCAPs ([Bolton et al., 2020](#))), lack comprehensive end-to-end analytical pipelines for both resting-state and task-based fMRI data with temporal

dynamic metrics and visualization capabilities (such as `capcalc` (Frederick & Drucker, 2022)), or are comprehensive, but generalized toolboxes for evaluating and comparing different dFC methods (such as `pydFC` (Torabi et al., 2024)).

NeuroCAPs addresses these limitations by providing an accessible Python package specifically for performing end-to-end CAPs analyses, from post-processing of fMRI data to creation of temporal metrics for downstream statistical analyses and visualizations to facilitate interpretations. However, many of NeuroCAPs' post-processing functionalities assumes that fMRI data is organized in a Brain Imaging Data Structure (BIDS) compliant directory and is most optimized for data preprocessed with fMRIPrep (Esteban et al., 2019) or preprocessing pipelines that generate similar outputs (e.g. NiBabies (Goncalves et al., 2025)). Furthermore, NeuroCAPs only supports the k-means algorithm for clustering, which is the clustering algorithm that was originally used and is often employed when performing the CAPs analysis (Liu et al., 2013).

## Modules

The core functionalities of NeuroCAPs are concentrated in three modules:

1. `neurocaps.extraction` Contains the `TimeseriesExtractor` class, which:
  - collects preprocessed BOLD data from an BIDS-compliant dataset (Yarkoni et al., 2019)
  - leverages Nilearn's ([contributors, n.d.](#)) `NiftiLabelsMasker` to perform nuisance regression and spatial dimensionality reduction using deterministic parcellations (e.g., Schaefer (Schaefer et al., 2018), AAL (Tzourio-Mazoyer et al., 2002))
  - scrubs high-motion volumes using fMRIPrep-derived framewise displacement values
  - reports quality control information related to high-motion or non-steady state volumes
2. `neurocaps.analysis` Contains the `CAP` class for performing the main analysis, as well as several standalone utility functions.
  - The `CAP` class:
    - performs k-means clustering (Pedregosa et al., 2011) to identify CAPs, supporting both single and optimized cluster selection with heuristics such as the silhouette and elbow method (Arvai, 2023)
    - computes subject-level temporal dynamics metrics (e.g., fractional occupancy, transition probabilities) for statistical analysis
    - converts identified CAPs back into NIFTI statistical maps for spatial interpretation
    - integrates multiple plotting libraries (Gale et al., 2021; Hunter, 2007; Inc., n.d.; Waskom, 2021) to provide a diverse range of visualization options
  - Standalone functions: Provide tools for data standardization (Harris et al., 2020), merging timeseries data across sessions or tasks, and creating group-averaged transition matrices.
3. `neurocaps.utils`
  - Contains a utility functions:
    - `fetch_preset_parcel_approach`: fetches a preset parcel approach (either "4S" or "HCPex" (Huang et al., 2022))
    - `generate_custom_parcel_approach`: automatically creates the necessary data structures from a parcellation's metadata file

## Workflow Example

The following code demonstrates a simple workflow example using NeuroCAPs to perform the CAPs analysis. An interactive variant of this workflow example is available on the [readthedocs](#).

1. Extract timeseries data

```
from neurocaps.extraction import TimeseriesExtractor

# Using Schaefer, one of the default parcellation approaches
parcel_approach = {"Schaefer": {"n_rois": 100, "yeo_networks": 7}}

# List of fMRIPrep-derived confounds for nuisance regression
confound_names = [
    "cosine*",
    "trans_x",
    "trans_x_derivative1",
    "trans_y",
    "trans_y_derivative1",
    "trans_z",
    "trans_z_derivative1",
    "rot_x",
    "rot_x_derivative1",
    "rot_y",
    "rot_y_derivative1",
    "rot_z",
    "rot_z_derivative1",
    "a_comp_cor_00",
    "a_comp_cor_01",
    "a_comp_cor_02",
    "a_comp_cor_03",
    "a_comp_cor_04",
]

# Initialize extractor with signal cleaning parameters
extractor = TimeseriesExtractor(
    space="MNI152NLin2009cAsym",
    parcel_approach=parcel_approach,
    confound_names=confound_names,
    standardize=False,
    fd_threshold={
        "threshold": 0.50,
        "outlier_percentage": 0.30,
    },
)

# Extract BOLD data from preprocessed fMRIPrep data
# which should be located in the "derivatives" folder
# within the BIDS root directory
# The extracted timeseries data is automatically stored
extractor.get_bold(
    bids_dir="path/to/bids/root",
    pipeline_name="fmriprep",
    session="1",
    task="rest",
    tr=2,
    verbose=False,
)

# Retrieve the dataframe containing QC information for each subject
# to use for downstream statistical analyses
qc_df = extractor.report_qc()
```

```
print(qc_df)
```

- 82     2. Use k-means clustering to identify the optimal number of CAPs from the data using a  
83         heuristic

```
from neurocaps.analysis import CAP
```

```
# Initialize CAP class
```

```
cap_analysis = CAP(parcel_approach=extractor.parcel_approach)
```

```
# Identify the optimal number of CAPs (clusters)
```

```
# using the elbow method to test 2-20
```

```
# The optimal number of CAPs is automatically stored
```

```
cap_analysis.get_caps(  
    subject_timeseries=extractor.subject_timeseries,  
    n_clusters=range(2, 21),  
    standardize=True,  
    cluster_selection_method="elbow",  
    max_iter=500,  
    n_init=10,  
    random_state=0,  
)
```

- 84     3. Compute temporal dynamic metrics for downstream statistical analyses

```
# Calculate temporal fraction of each CAP for all subjects
```

```
output = cap_analysis.calculate_metrics(extractor.subject_timeseries, metrics=["temporal
```

```
print(output["temporal_fraction"])
```

- 85     4. Visualize CAPs

```
# Project CAPs onto surface plots
```

```
# and generate cosine similarity network alignment of CAPs
```

```
cap_analysis.caps2surf().caps2radar()
```

## 86     Documentation

87     Comprehensive documentations and interactive tutorials of NeuroCAPS can be found at  
88     <https://neurocaps.readthedocs.io/> and on its [repository](#).

## 89     Research Utility

90     NeuroCAPs was originally developed (and later expanded and refined for broader use) to  
91     facilitate the analysis in Smith et al. (2025), which has been submitted for peer review by the  
92     same author.

## 93     Acknowledgements

94     Funding provided by the Dissertation Year Fellowship (DYF) Program at Florida International  
95     University (FIU) assisted in further refinement and expansion of NeuroCAPs.

## 96     References

97     Arvai, K. (2023). *Kneed*. Zenodo. <https://doi.org/10.5281/ZENODO.8127224>

- 98 Bolton, T. A. W., Tuleasca, C., Wotruba, D., Rey, G., Dhanis, H., Gauthier, B., Delavari,  
99 F., Morgenroth, E., Gaviria, J., Blondiaux, E., Smigielski, L., & Van De Ville, D. (2020).  
100 TbCAPs: A toolbox for co-activation pattern analysis. *NeuroImage*, 211, 116621. <https://doi.org/10.1016/j.neuroimage.2020.116621>  
101
- 102 contributors, N. (n.d.). *nilearn*. <https://doi.org/10.5281/zenodo.8397156>
- 103 Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent,  
104 J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez,  
105 J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing  
106 pipeline for functional MRI. *Nature Methods*, 16(1), 111–116. <https://doi.org/10.1038/s41592-018-0235-4>  
107
- 108 Frederick, B. D., & Drucker, D. M. (2022). *Bbfrederick/capcalc: Version 1.2.2.2 - 8/30/22*  
109 *deployment bug fix*. Zenodo. <https://doi.org/10.5281/ZENODO.7035806>
- 110 Gale, D. J., Vos de Wael, R., Benkarim, O., & Bernhardt, B. (2021). *Surfplot: Publication-*  
111 *ready brain surface figures*. Zenodo. <https://doi.org/10.5281/ZENODO.5567926>
- 112 Goncalves, M., Markiewicz, C. J., Esteban, O., Feczko, E., Poldrack, R. A., & Fair, D. A.  
113 (2025). *NiBabies: A robust preprocessing pipeline for infant functional MRI*. Zenodo.  
114 <https://doi.org/10.5281/ZENODO.14811979>
- 115 Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D.,  
116 Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk,  
117 M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., ... Oliphant,  
118 T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>  
119
- 120 Huang, C.-C., Rolls, E. T., Feng, J., & Lin, C.-P. (2022). An extended human connectome  
121 project multimodal parcellation atlas of the human cortex and subcortical areas. *Brain*  
122 *Structure and Function*, 227(3), 763–778. <https://doi.org/10.1007/s00429-021-02421-6>
- 123 Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science &*  
124 *Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- 125 Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta,  
126 M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., Handwerker, D. A.,  
127 Keilholz, S., Kiviniemi, V., Leopold, D. A., De Pasquale, F., Sporns, O., Walter, M., &  
128 Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations.  
129 *NeuroImage*, 80, 360–378. <https://doi.org/10.1016/j.neuroimage.2013.05.079>
- 130 Inc., P. T. (n.d.). *Chart title*. <https://plotly.com/python/radar-chart/>; Plotly Technologies  
131 Inc.
- 132 Jiang, F., Jin, H., Gao, Y., Xie, X., Cummings, J., Raj, A., & Nagarajan, S. (2022). Time-  
133 varying dynamic network model for dynamic resting state functional connectivity in fMRI  
134 and MEG imaging. *NeuroImage*, 254, 119131. <https://doi.org/10.1016/j.neuroimage.2022.119131>  
135
- 136 Liu, X., Chang, C., & Duyn, J. H. (2013). Decomposition of spontaneous brain activity  
137 into distinct fMRI co-activation patterns. *Frontiers in Systems Neuroscience*, 7. <https://doi.org/10.3389/fnsys.2013.00101>  
138
- 139 Liu, X., Zhang, N., Chang, C., & Duyn, J. H. (2018). Co-activation patterns in resting-state  
140 fMRI signals. *NeuroImage*, 180, 485–494. <https://doi.org/10.1016/j.neuroimage.2018.01.041>  
141
- 142 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,  
143 Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,  
144 Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python.  
145 *Journal of Machine Learning Research*, 12, 2825–2830.

- 146 Rabany, L., Brocke, S., Calhoun, V. D., Pittman, B., Corbera, S., Wexler, B. E., Bell, M. D.,  
147 Pelphrey, K., Pearlson, G. D., & Assaf, M. (2019). Dynamic functional connectivity in  
148 schizophrenia and autism spectrum disorder: Convergence, divergence and classification.  
149 *NeuroImage: Clinical*, 24, 101966. <https://doi.org/10.1016/j.nicl.2019.101966>
- 150 Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff,  
151 S. B., & Yeo, B. T. T. (2018). Local-global parcellation of the human cerebral cortex  
152 from intrinsic functional connectivity MRI. *Cerebral Cortex*, 28(9), 3095–3114. <https://doi.org/10.1093/cercor/bhx179>  
153
- 154 Smith, D. D., Bartley, J. E., Peraza, J. A., Bottenhorn, K. L., Nomi, J. S., Uddin, L. Q., Riedel,  
155 M. C., Salo, T., Laird, R. W., Pruden, S. M., Sutherland, M. T., Brewe, E., & Laird, A. R.  
156 (2025). *Dynamic reconfiguration of brain coactivation states associated with active and*  
157 *lecture-based learning of university physics*. <https://doi.org/10.1101/2025.02.22.639361>
- 158 Torabi, M., Mitsis, G. D., & Poline, J.-B. (2024). On the variability of dynamic functional  
159 connectivity assessment methods. *GigaScience*, 13, giae009. <https://doi.org/10.1093/gigascience/giae009>  
160
- 161 Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix,  
162 N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in  
163 SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain.  
164 *NeuroImage*, 15(1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>
- 165 Waskom, M. L. (2021). Seaborn: Statistical data visualization. *Journal of Open Source*  
166 *Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- 167 Yarkoni, T., Markiewicz, C., De La Vega, A., Gorgolewski, K., Salo, T., Halchenko, Y.,  
168 McNamara, Q., DeStasio, K., Poline, J.-B., Petrov, D., Hayot-Sasson, V., Nielson, D.,  
169 Carlin, J., Kiar, G., Whitaker, K., DuPre, E., Wagner, A., Tirrell, L., Jas, M., ... Blair, R.  
170 (2019). PyBIDS: Python tools for BIDS datasets. *Journal of Open Source Software*, 4(40),  
171 1294. <https://doi.org/10.21105/joss.01294>