

LOCALLY LINEAR EMBEDDING AND fMRI FEATURE SELECTION IN PSYCHIATRIC CLASSIFICATION DATASET NOMENCLATURE

CHILDMIND DATA RELEASE

- For each dataset, the name of each subdirectory is a unique identifier for the respective subject:
 - [CATEGORY] / [DATASETNAME] / [SUBJECTID]
- Note #1: The square brackets denote a string that can be substituted for the respective dataset and subject ID.
- Note #2: This nomenclature generalises for every dataset used in the evaluation.
- where CATEGORY crudely describes the patient group or project:
 - abide1** healthy controls and patients diagnosed with Autism Spectrum Disorder from ABIDE I.
 - abide2** healthy controls and patients diagnosed with Autism Spectrum Disorder from ABIDE II.
 - psychosis** healthy controls and patients with psychosis.
 - ADHD200** healthy controls and patients with Attention-Deficit Hyperactivity Disorder (ADHD) from the ADHD200 competition.
 - pknsn** healthy controls and patients with psychosis from the INDI Retrospective Parkinson's data release.
- Every SUBDIRECTORY [SUBJECTID]/ contains:
 - session_1/¹
 - * {rest_{1,2}²/func³}
 - rest_{1,2}w.nii.gz (task{1,2,3}.nii.gz for MRN/MCIC block-design fMRI dataset).
 - fMRIdata.mat
 - rp_rest_0001.txt
 - rest.mat
 - LLE/

where:

rest{w}.nii.gz – the preprocessed functional image

rp_rest_0001.txt⁵ the estimation parameters produced by SPM12 after registering the functional volumes to the first volume in the scan.

rest.mat⁵ The realignment parameters produced by SPM12 if the functional volume's origin was adjusted.

LLE/ Folder containing the reconstructed functional image for various dimensionality parameters d .

fMRIdata.mat MATLAB data file that contains datastructure fMRIdata, where:

fMRIdata.subID unique subject ID

fMRIdata.orig = original four-dimensional ($L \times W \times H \times T$) functional image

fMRIdata.K = number of neighbouring voxels used to reconstruct each voxel.

fMRIdata.dx = label for respective subject. 0 is control, 1 is patient.

fMRIdata.mask – binary-valued mask used to remove voxels outside of the brain. The mask is fixed for all subjects within a dataset.

fMRIdata.recon = struct containing reconstructed fMRI ($L \times W \times H \times d$) for 18 different reconstruction parameters.

fMRIdata.recon.d[value] is the four-dimensional reconstructed fMRI using reconstruction parameter $d=value$.

¹The psychosis/MRN dataset uses subdirectories K{26,124}, each of which maintain the data structure mentioned below.

²The session_1/rest_{1,2} subdirectory is only present for the collection sites comprising ABIDE II, where only the IP_1 and OHSU_1 sites contain rest_2 subdirectory.

³The session_1/rest_{1,2} subdirectory is only present for the collection sites comprising ABIDE II, where only the IP_1 and OHSU_1 sites contain rest_2 subdirectory.

⁴The session_1/func subdirectory is only present for the TaoWu collection site that is part of the INDI Retrospective Parkinson's data release.

⁵The rest_{1,2}w.nii.gz images are only applicable to the neurocon dataset that is part of the INDI Retrospective Parkinson's data release.

⁶These files are not present for earlier datasets that were preprocessed with C-PAC (Stanford, UCLA_1,UM_1,UM_2,USM, Yale)