CNNP Cortical Folding User Interface (UI)

Unser Manual
This manual explains functionalities and requirement for proper use of *CNNP Cortical Folding UI*. For further assistance or to report any issues, visit our Github repo or Lab website.

The current version of the UI requires MATLAB to be executed. Go to the official website for further information. For optimal performance and appearance, use MATLAB 2023a or later version. The oldest version tested was 2021a, showing complete functionality but having some issues with appearance. Additional software will be required for certain functionalities. System and hardware requirements are:

- *Processor*: Intel® (or MATLAB for Intel® processor)
- *RAM*: 16GB or more recommended.
- *Operative System*: Linux, Mac, and Windows all support main functionalities. **Windows has some restricted functionalities**; further details in "*Processing pial-outer-smoothed files*".

Complete Usage Workflow

The UI is designed for easy and fluent use, by dynamically changing appearance to indicate the next steps. To start the UI, double click on the file "*Main.mlapp*". On Linux device follow these steps: open MATLAB, open "*Main.mlapp*" file, press 'Run'. For a complete workflow the following steps should be taken.

1. Dataset/s selection and management

The UI is designed to allow for the analysis of several dataset simultaneously. After launching the app, the initial interface should be as shown in Figure 1. Press "*Add Datasets*" and folder selection user interface will be initiated. Select the **root folder of the dataset(s)**. The program requires one of the specific file structures as follows:

- *Dataset > freesurfer >* Subject > Session >
- *Dataset > freesurfer >* Subject >

The **root folder** in these examples is called *"Dataset"*. The following folder must be spelled as shown with flexibility on the casing: "*freesurfer*", "*FreeSurfer*" and "*FREESURFER*" are allowed. FreeSurfer output files (including the smoothed surfaces) should be placed inside the "*subject/session*" folder. The UI will analyse file structure and content identifying the number of subjects/sessions available in the root folder, based on the presence of FreeSurfer folder structures. This information will be internally stored for next steps and presented to the user. Avoid the use of spaces on files and folders naming.
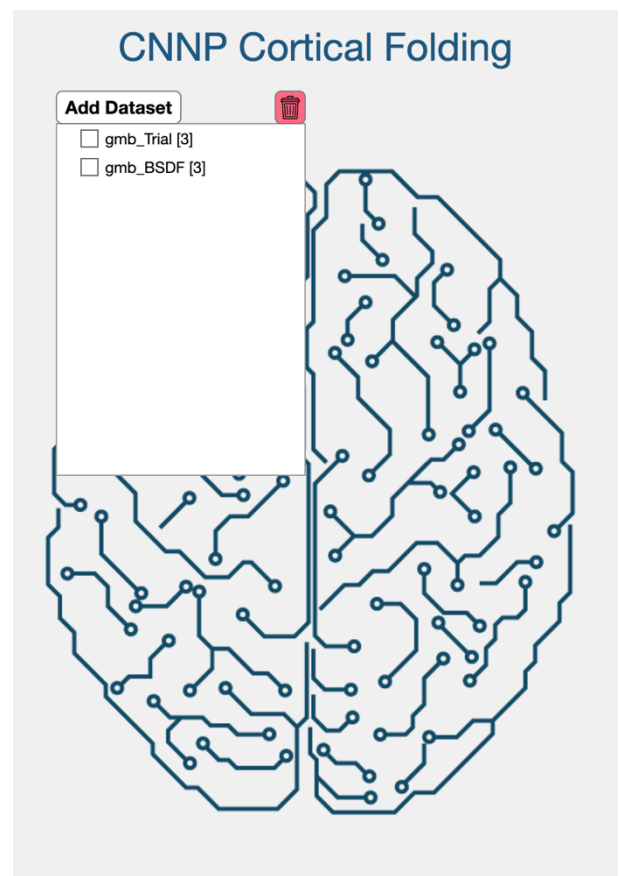

Figure 1 Initial UI appearance

Multiple datasets can be introduced to the UI: press "*Add Datasets*", select the new dataset root folder. The user can then choose which datasets to process by toggling the selection buttons to the left of each dataset. The red button with a red trash bin symbol (to the right of the "*Add Datasets*" button) **resets the UI** removing all stored information relating datasets and configurations.

2. Configuring cortical folding feature estimation

When at least one valid dataset is introduced to the UI and selected as shown in Figure 2, several panels will become available. These are configuration options to set up the cortical folding feature estimation (section 5) and the grouping of outputs (section 6).

Any change on these configuration parameters will reset the progress on the following sections.

"Hemisphere" Panel: Indicates which hemisphere will be analysed: left, right or both hemispheres. Note that the options average and sum of hemisphere is not yet implemented, thus are not available yet.

"Mode" Panel: Indicates the level of parcellation (hemisphere or lobes) and the spatial scale options to estimate cortical folding features. Note standard parcellations (e.g. Desikan-Killiany) will be released later.

- *'Hemisphere'*: Estimation of novel cortical folding metrics based on entire hemispheres. Here are some example papers that used this type of analysis: *Y. Wang 2021*, *Y. Wang 2016*.
- *'Lobe'*: Divides each hemisphere in lobes, based on selected atlas. Currently, only FSDK atlas (see details here) is included: more atlases will be implemented in the future. Here are some example papers that this this type of analysis: *Y. Wang 2019*
- *'Multi-scale'*: Select the scale for feature estimation: *Higher Scales* and *Lower Scales* are explained below[(*)] and *No Scales* uses the native scale (i.e., the spatial scale that is normally used in FreeSurfer). Here are some example papers that this this type of analysis: *Y. Wang 2024*, *K. Leiberg 2023*.

3. Validity of available subjects/sessions (Screening)

After configuration, the screening section will appear as shown in Figure 3. Press the '*Screening*' button to start the validation process on subjects/sessions included. The UI will check the presence and validity of:

1. '*lh.pial-outer-smoothed*' & '*rh.pial-outer-smoothed*' files in the FreeSurfer folder structure, required to estimate cortical folding features.
2. Intermediate subject-level output files (go to "*Estimating the cortical folding parameters*" for more information) including cortical features matching the defined configuration.
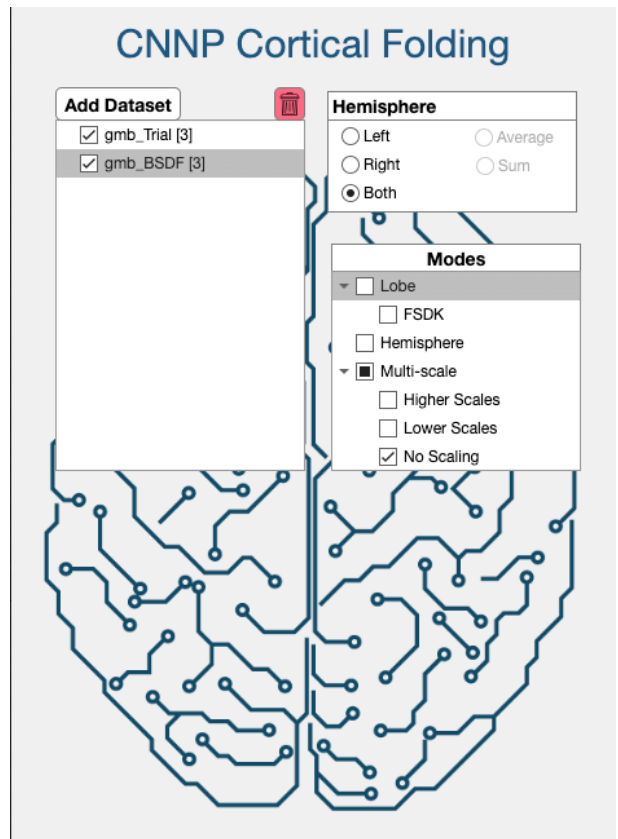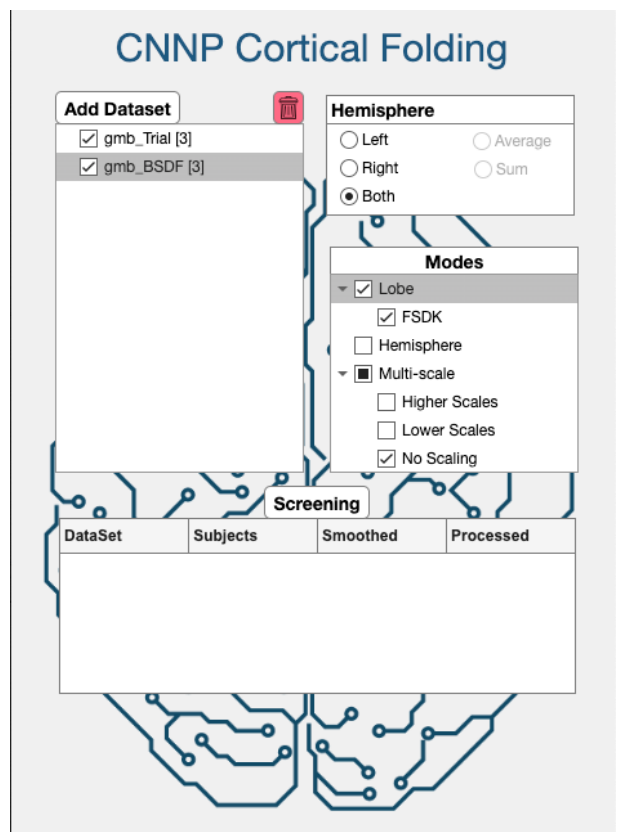

Figure 2 UI appearance during configuration mode

Figure 3 Screening step

[(*)] Scales are generated with an initial scale [s] and the number of iterations [3]. The initial scale is increased by a power of two every iteration [$s*2^0$, $s*2^1$, $s*2^2$, $s*2^3$]. *Higher Scales* are generated with initial scales [0.325, 0.35, 0.375] and 4 iterations. *Lower Scales* are generated with initial scales [0.225, 0.25, 0.275, 0.3] and [5, 5, 4, 4] iterations respectively.

The screening process may take several minutes, depending on the total number of subjects/sessions. Progress bars will indicate the estimated remaining time. The screening report table will show:

- '*DataSet*', the name of the tested dataset.
- '*Subjects*', the number of available subjects.
- '*Smoothed*', the number of subjects including both valid '*xh.pial-outer-smoothed*'
- '*Processed*', the number of subjects with extracted cortical parameters matching the set configuration.

Several new options will also become available after screening (Figure 4). These can be used to generate smooth surfaces files (*xh.pial-outer-smoothed*) if they were missing (section 4), estimate cortical folding parameters (section 5), and group outputs into a single file (section 6).

4. Processing *pial-outer-smoothed* files

WARNING
This process is not available for Windows devices as *FreeSurfer* toolbox is not available for this operative system
ADVICE
This process is optional and can be skipped if not needed

If there are fewer '*Smoothed*' subjects than '*Subjects*' in the screening report table (Figure 4), this indicates missing or corrupted *pial-outer-smoothed* files. We provide the option to generate these missing or corrupt files given some conditions:

a) *FreeSurfer* toolbox is installed on the device at a known location. Visit the *official webpage* to obtain more information about installing FreeSurfer.
b) If using Linux device, *dos2unix* is installed. This will be requested in case of need.
c) MATLAB must have sudo/administrator access (by execution of computer user).

Press "*Extract Smooth Surfaces*" to begin the *pial-outer-smoothed* generation. On Windows devices this command will be blocked. The UI will request the location of FreeSurfer toolbox. Please provide **the FreeSurfer version folder** to prevent execution issues.

If all conditions are meet, the process will begin, and a progress bar will indicate the estimated remaining time. Upon completion, a new screening process will be automatically run updating the screening report table.
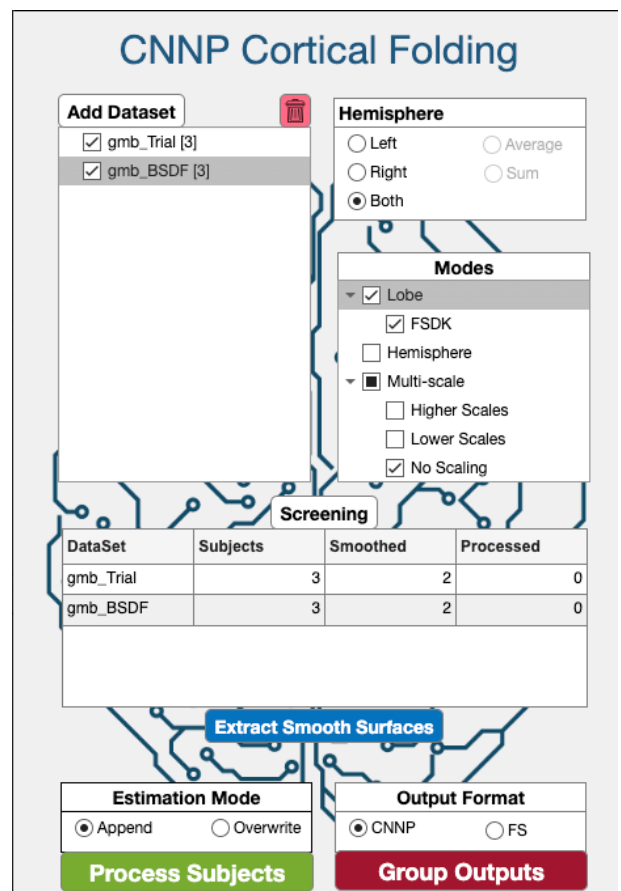

Figure 4 Latest stages of the UI

Please note that this is the most time-consuming process of the entire workflow, and large datasets could take several hours to complete. If possible for large datasets, please generate the *xh.pial-outer-smoothed* surfaces outside of this toolbox using parallel computing facilities. In case of proceeding, please remove any spaces on folder and file names, and execute the UI though MATLAB. The "*Extract Smooth Surfaces*" will report the process on MATLAB Command Window.

5. Estimating the cortical folding parameters

The estimation of cortical folding features can be started by pressing "*Process Subjects*". The initiated workflow estimates cortical folding features for all subjects/sessions following the configuration set by the user. Each subject's parameters are stored as individual .csv files in *CNNP Format* (go to "*Obtaining grouped outputs*" for more details) in the subject directory.

In the presence of pre-existing subject-level csv files (reported as '*Processed*' on the screening report table), matching configuration entries are looked for. If new unmatched configurations have been requested (for example, a new parcellation option), features are estimated and included on the same file. If matching configurations are found the settings in the "*Estimation Mode*" panel indicates how to manage them:

- '*Append*': the estimation process is skipped keeping the previously estimated values.
- '*Overwrite*': parameters are estimated, and the **preexisting entries are overwritten**.

A screening process will automatically run at the end to update the screening report table.

6. Obtaining grouped outputs

The last step is grouping the estimated cortical folding parameters into a single location. Specify on the output format on "*Output Format*" panel and press "*Group Outputs*". There are two available formats:

CNNP Format

Output format also used for subject-level storage of parameters. This consists of a single .csv file including cortical folding parameters, with several additional fields.

- *Dataset*, *Subject* and *Session*: indicate the origin of the data (_*_ is used in case of no *Session*)
- *Hemisphere*, *Atlas*, *Finess* and *Region*: represent the anatomical location of the brain.
- *Scale*: indicates the scaling[(*)], with 0 representing the native scale.

Data matching the specified configuration is pulled for each subject's file and stored in a new csv file. The UI will request a file name (the '*_CNNP*' suffix will be added by the UI) and location to store the grouped outputs.

FS Format

The FS format is an output format inspired by *FreeSurfer* output file structure. In this case a folder structure is formed as follows: <u>Output_Folder > Dataset > Atlas ></u>. Within the *Atlas* folder, a .csv file for each cortical folding parameter is generated. Each one includes the values for all *Subject/Sessions* (rows) and *Regions* (columns). *Scale* is defined on the name of the .csv file. *Output_Folder* will be requested to store the dataset. **It is recommended to generate a new folder** for this step, using the corresponding file manager of your device (i.e. Finder for Mac). The selected folder will be overwritten and renamed with the original name and the suffix '*_FS*'.

Notes on output

Empty columns: Some variables in the output files are specific to either hemisphere-wise or lobe-wise data. Therefore, columns for the lobe-specific variables will be empty for hemisphere data, and vice versa.

Lobe data: Please note that while lobe data is provided for the Insula and Corpus Collosum, it is not advised to use this data in analyses. Insula volumes may be integrated into other lobes; for more information, see our publication *Y. Wang 2019* and its associated supplementary materials.

7. Testing dataset

If a testing dataset is needed, access the public dataset *IDEAS* and download a couple of individuals. For proper functionality of the app, follow *Dataset/s selection and management* section to place the download content on the right file structure

8. Common Issues and how to approach them

First, we apologize for any inconveniences you may encounter while using this UI. Our team has toughly tested it to identify any issues and address them. Below, we presented several issues users may face and how to approach them. We are continuously improving the app. If this guide does not provide a solution to your issue, please contact us to the provided details.

Not found files

During dataset selection, the UI will estimate of the number of files with the proper structure found on the root folder. After the *Screening* process this number may changed due to file structure or name incompatibility. Make sure all subjects have the proper file structure as indicated in *Dataset/s selection and management* to be included on further analysis.

The screening report table may show different numbers on *Subjects* and *Smoothed*, even after going thought the *Extract Smooth Surfaces* process. The main reason for this discrepancy is that one of the *xh.pial* files (right and/or left depending defined configurations) is missing or corrupted. This file is needed to obtain the smooth surfaces files (*xh.pial-outer-smoothed*). Although this will not impair the use of the app, the best solution will be to identify the corrupted individual and substitute the subject with a new version of the file.

UI freezes during *Extract Smooth Surfaces*

If you choose to run the *Extract Smooth Surfaces*, the UI should be executed though MATLAB. If this is not the case, please follow these steps: open MATLAB, open "*Main.mlapp*" file, press 'Run'. If any issues occur during this process they will be reported on MATLAB Command Window.

A common issue is the presence of spaces on the dataset path, which will be reported as "*not found file*". Locate any spaces and remove or replace them.

Another Common issue is a not found function error. In MATALB Command Window a message may appear asking you to change the MATLAB current folder. There are 2 potential causes for this.
  1) The FreeSurfer directory has not been properly selected. Further details for proper FreeSurfer directory selection are provided on section Processing *pial-outer-smoothed*.
  2) A faulty FreeSurfer installation may be causing this issue. Remove the current FreeSurfer toolbox and reinstall it.

UI closes without warning

If the UI closes unexpectedly, it may be due to a shortage of RAM, especially if the number of subjects is large or the computer has limited memory. To resolve this, divide the dataset in two and process them independently. The processed datasets can be then merged into one keeping the intermediate files.

Brain MoNoCle

This cortical folding toolbox has been designed to flow into a further pipeline -*Brain MoNoCle*- that can provide estimates of brain abnormalities using normative modelling approaches based on a large normative reference sample (n>3,200, aged 5-95 years old). To use Brain MoNoCle, please export the grouped outputs in the CNNP format described above. For more information about our normative modelling pipeline Brain MoNoCle, please see our latest preprint: *B. Little 2024*

Figure 5 Screenshot of Brain MoNoCle