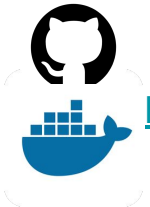
DeepPrep: an accelerated, scalable and robust pipeline for neuroimaging preprocessing empowered by deep learning

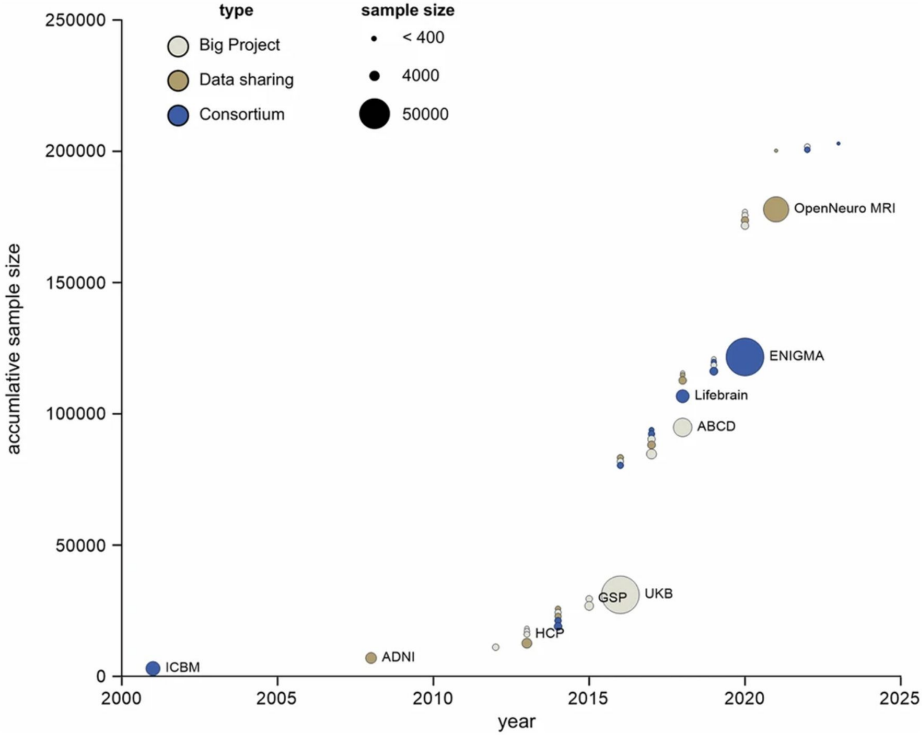
Jianxun Ren, Ning An, Cong Lin, Youjia Zhang, Zhenyu Sun, Wei Zhang, Shiyi Li, Ning Guo, Weigang Cui, Qingyu Hu, Weiwei Wang, Xuehai Wu, Yinyan Wang, Tao Jiang, Theodore D. Satterthwaite, Danhong Wang & Hesheng Liu
(Changping Laboratory, Beijing, China | University of Pennsylvania, Philadelphia, PA, USA)

DOI: https://doi.org/10.1038/s41592-025-02599-1
Nature Methods
Published: February 06 2025



https://github.com/pBFSLab/DeepPrep
https://hub.docker.com/r/pbfslab/deepprep
https://deepprep.readthedocs.io

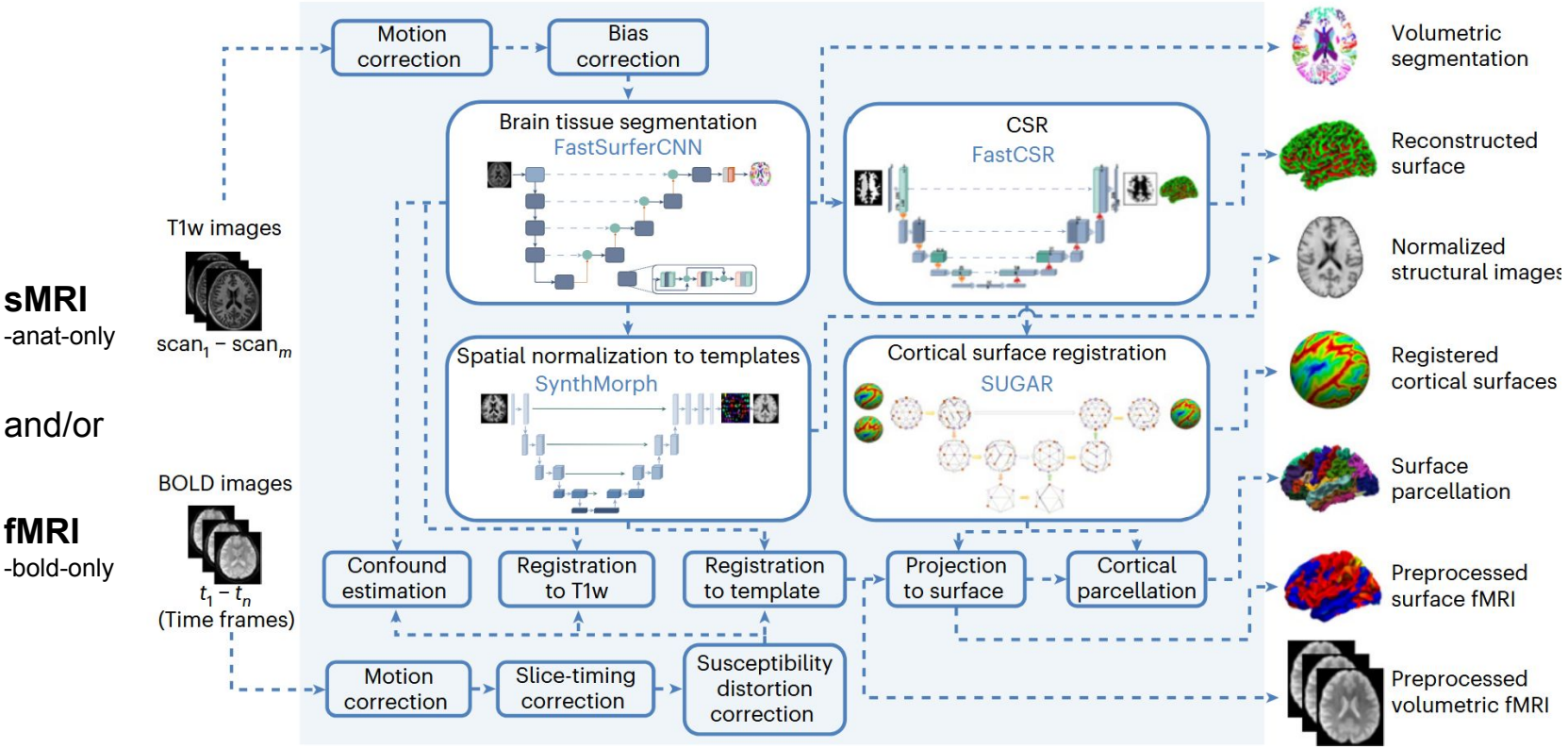
Context



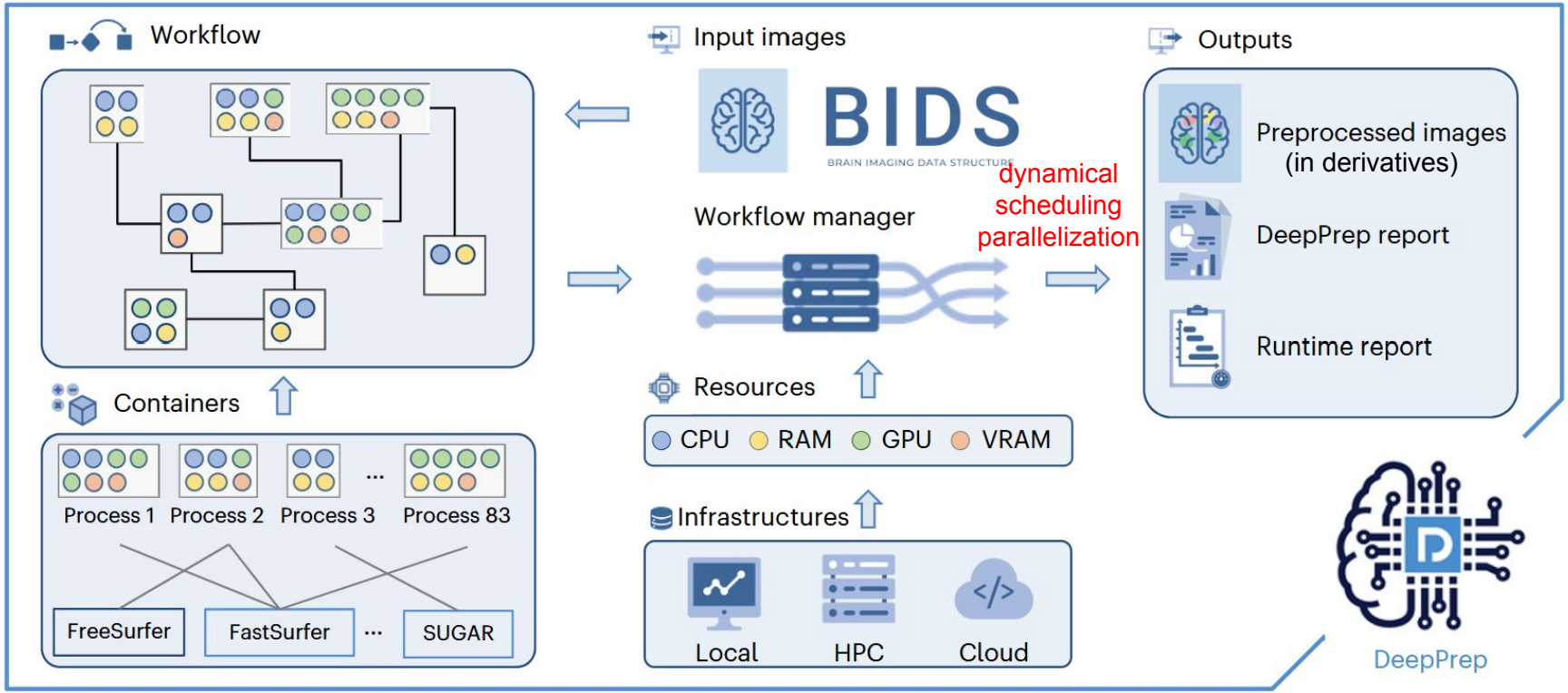
- Growing amounts of publicly available data from large-scale projects (like UKBB, ENIGMA, OpenNeuro)
- Prevailing preprocessing pipelines too slow (FreeSurfer, fMRIPrep, QSIPrep and ASLPrep)
- Segmentation issues for distorted brains (due to traumas, gliomas or strokes)

⇒ Need of a computationally efficient, scalable and robust preprocessing pipeline

DeepPrep: Deep learning-based neuroimaging preprocessing pipeline



Nextflow: Workflow manager of DeepPrep



Evaluation datasets

accuracy and reliability assessment

evaluating computational efficiency and scalability

Manual brain segmentation

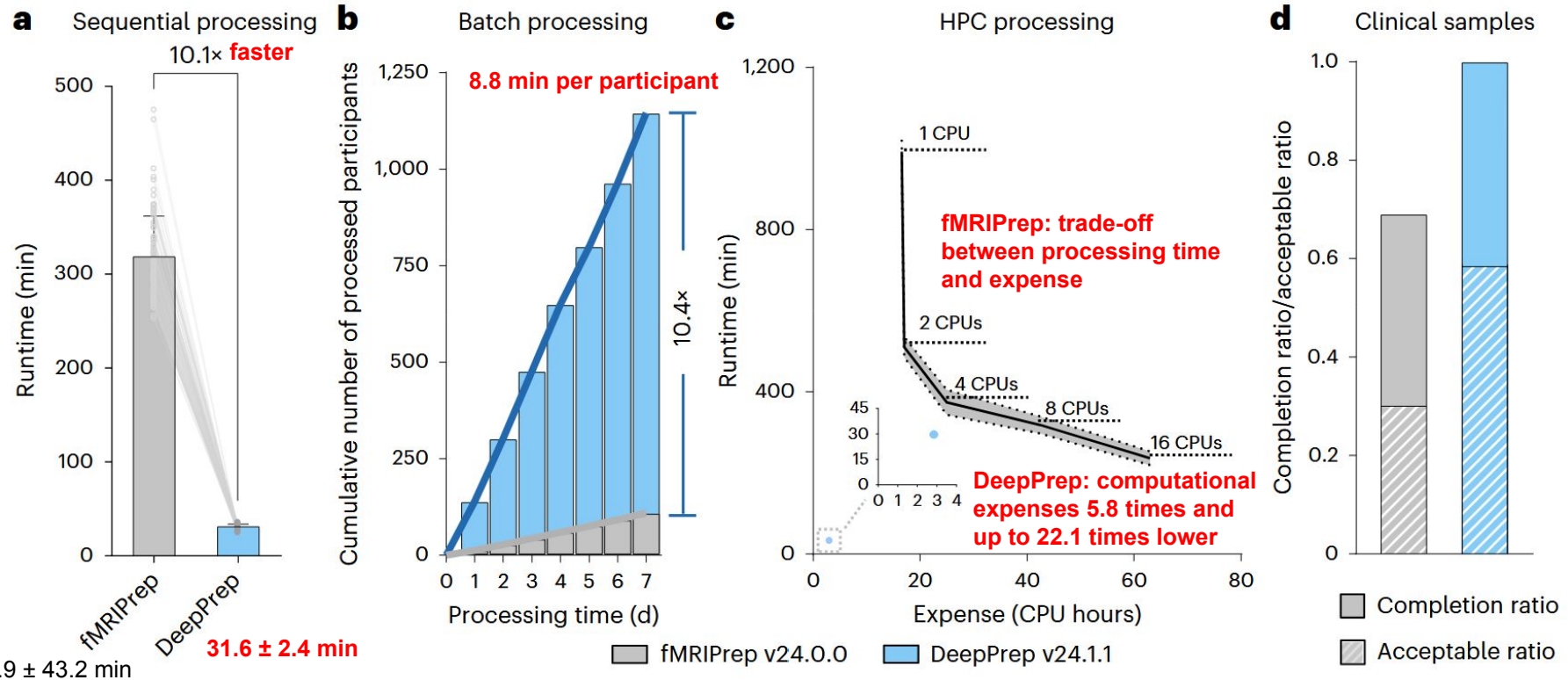
Three clinical datasets for robustness assessment

Supplementary Table.2 | Characteristics of evaluation datasets

No.	Dataset	# used subs	# total subs	age	source	state	scanner
1	CoRR_HNU	30	30	20-30	http://fcon_1000.projects.nitrc.org/indi/CoRR/html/hnu_1.html	Healthy	GE MR750 3T
2	MSC	10	10	24-36	https://openneuro.org/datasets/ds000224/versions/1.0.3	Healthy	Siemens TRIO 3T
3	UKB_20252_2.0	49300	49300	44-83	https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=20252	Healthy	Siemens Skyra 3T
4	UKB_20252_3.0	5215	5215	44-83	https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=20252	Healthy	Siemens Skyra
5	Mindboggle-101	97[a]	101	19-61	http://mindboggle.info/data	Healthy	Siemens, Philip
6	CRRC-Stroke	19[b]	218	37-71	Collected in Beijing Bo'ai Hospital	Stroke	Philip Ingenia 3T
7	BTH-Glioma	19[b]	168	21-62	Collected in Beijing Tiantan Hospital	Glioma	Siemens TrimTrio 3T
8	SHH-DoC	15[b]	38	31-53	Collected in Shanghai Huashan Hospital	DoC	Siemens MAGNETOM 3T

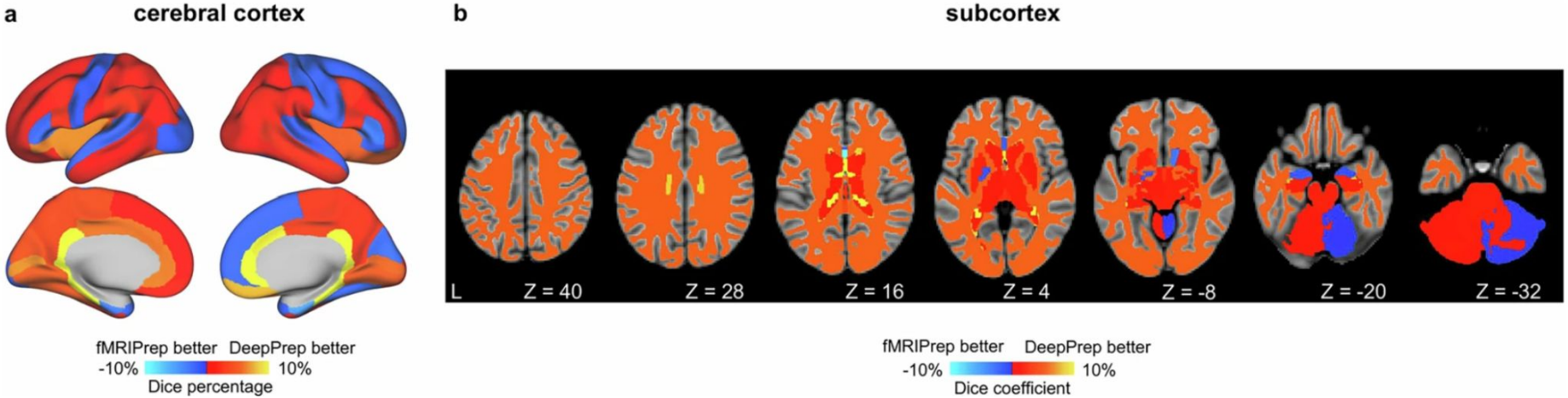
[a] Four cases have been excluded due to either poor data quality or the absence of necessary anatomical annotations.
[b] Cases were selected from the clinical datasets if FreeSurfer v6.0 failed to complete processing within 48 CPU hours.

DeepPrep (v24.1.1) VS fMRIPrep (v24.0.0)



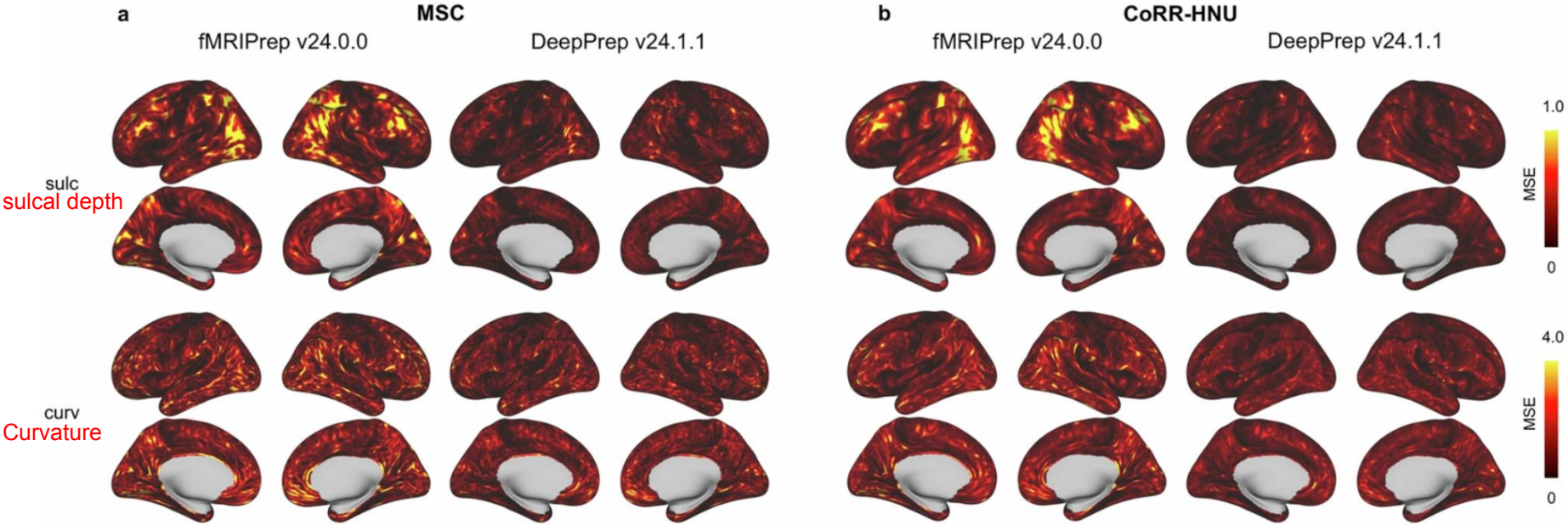
two-tailed paired t-test, $t(99) = 67.0$, $P = 2.6 \times 10^{-84}$

DeepPrep (v24.1.1) VS fMRIPrep (v24.0.0)



Mindboggle-101 dataset (manual segmentation)
Comparison of **dice coefficient** for each brain region
fMRI: 52/62 cortical regions better segmented with DeepPrep
sMRI: 31/39 subcortical regions

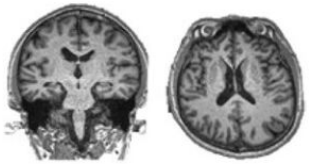
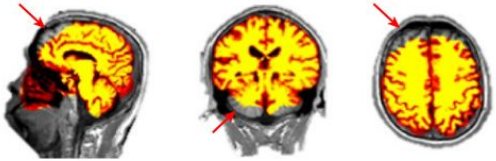
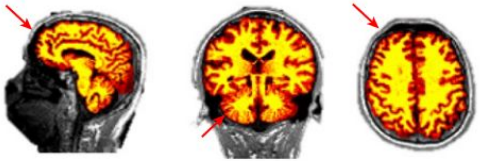
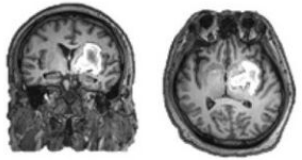
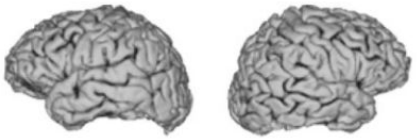
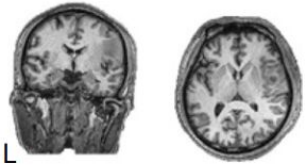
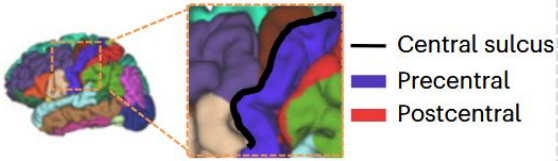
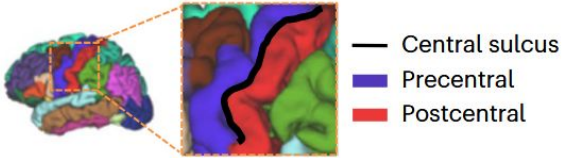
DeepPrep (v24.1.1) VS fMRIPrep (v24.0.0)



Comparison of both cortical surfaces features with target atlases

Performance of DeepPrep in speed and robustness

e Typical examples of three preprocessing error categories

Error categories	Original images	fMRIPrep v24.0.0	DeepPrep v24.1.1
Segmentation (skull stripping)			
CSR		Failed to reconstruct surface	
Cortical surface registration	 L	 Central sulcus Precentral Postcentral	 Central sulcus Precentral Postcentral

Summary and future work

- DeepPrep = High efficiency and robustness in processing large-scale neuroimaging datasets and complex clinical samples

- Future comprehensive platform for processing multimodal neuroimaging
- Will integrate additional modalities, such as arterial spin labeling and diffusion imaging

Dependencies

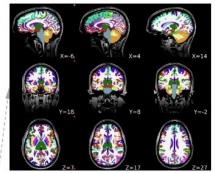
Software used by DeepPrep		
Processing step	*Implementation*	*Software Version*
Input data		
Data in NIfTI format	BIDS format	
Anatomical processing		
brain extraction brain tissue segmentation	FastSufer	v1.1.0
bias correction	SimpleITK's N4	v2.3.0
reconstruct cortical surfaces from T1w structural images	FastCSR	v1.0.0
cortical surface registration	SUGAR	1.0.0
cortical surface parcellation	FreeSurfer	7.2.0
BOLD processing		
BOLD reference image	niworkflows	1.10.0
head motion correction	FSL MCFLIRT	6.0.5.1
slice time correction	AFNI 3dTshift	24.0.00
susceptibility distortion correction	SDCFlow	2.8.1
coregistration	FreeSurfer bbregister	7.2.0
resampling EPI to standard space	SynthMorph	2
confounds estimation	In-house codes	
Workflow		
workflow manager	Nextflow	23.10.4
in-memory database	Redis	6.0.16
lock context manager	python-redis-lock	4.0.0

Python packages used by DeepPrep		
Packages		*Version*
fastcsr		1.0.0
sugar		1.0.0
nighres		1.5.0
torch		2.0.1+cu118
torchvision		0.15.2+cu118
templateflow		23.2.0
lapy		1.0.1
voxelmorph		0.2
tensorflow		2.11.1
torch-geometric		2.2.0
fmriprep		23.2.0
mriqc		23.2.0
nnunet		1.7.1
dynaconf		3.2.3
wand		0.6.11
bids		0.0
nipype		1.8.6
niworkflows		1.10.0
SimpleITK		2.3.0
sdcflows		2.8.1
open3d		0.17.0
onnxruntime		1.16.3
pytorch3d		0.7.4
python-redis-lock		4.0.0
neurite		0.2

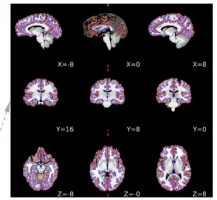
Reports

Summary

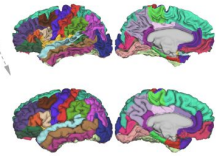
Anatomical processing



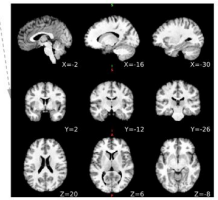
Brain mask and brain tissue segmentation of the T1w.
This panel shows skull-stripped brain and cortical/subcortical segmentation o the T1w image.



Surface reconstruction.
The white surface (blue contours) and pial surface (red contours) were reconstructed with FastCSR and are overlaied on the native T1w image.



Cortical surface anatomical parcellations.
The cortical parcellations were generated based on the cortical surface registration of SUGAR. Parcellations ar shown on the white (upper row) and pia surfaces (lower row).

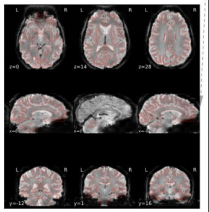


Spatial normalization of the anatomical T1w reference.
SynthMorph was used to perform nonlinear registration between the T1w reference and the template space. Hover on the panels with the mouse pointer to transition between both native T1w and template spaces.Spati normalization of the T1w image to the MNI152NLin6Asym template.

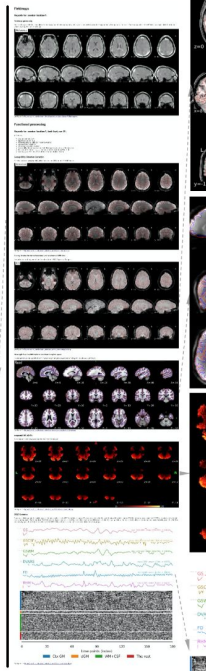
Fieldmaps processing

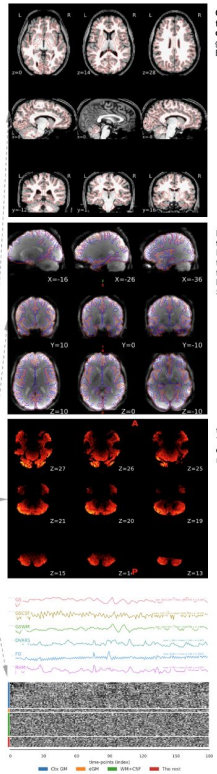
When fieldmaps were found, some mosaics will show the field inhomogeneity with the "magnitude map" as the reference

Functional processing



Susceptibility distortion correction.
If fieldmap information was found, the step is assessed with a dynamic mosaic that transitions between the unwarped ("after") and original ("before"). Contours of the white-matter boundaries are shown for reference.





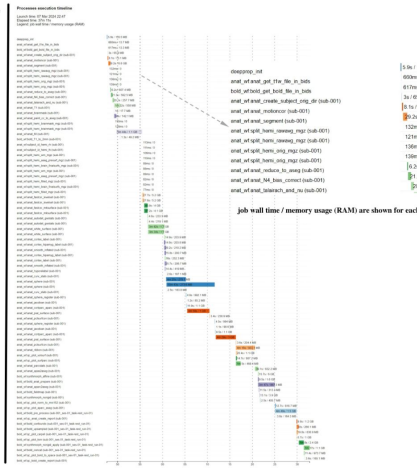
Co-registration between functional and anatomical MRI data. bbregister was used to generate transformations from EPI-BOLD space to T1w-space.

Resample functional MRI data to the template space.
Functional MRI was resampled to the MNI152NLin6Asym template space based on bbregister, SUGAR and SynthMorph.

temporal SNR (tSNR).
The temporal SNR map was estimated by the *TSNR* from nipype.

BOLD Summary.
Summary statistics are plotted, which may reveal trends or artifacts in the BOLD data. Global signals (GS) were calculated within the whole-brain, and the white-matter (GSWM) and the cerebro-spinal fluid (GSCSF) were calculated with their corresponding masks. The standardized DVARS, framewise-displacement measures (FD), and relative head motion (RHM) were calculated. A carpet plot shows time series for all voxels within the brain mask, including cortical gray matter (Cox GM), deep (subcortical) gray matter (dGM), white-matter and CSF (WM+CSF), and the rest of the brain (The rest).

Processer execution timeline



Thank you for your attention !