

Harnessing cloud computing for high capacity analysis of neuroimaging data from NDAR



Daniel Clark¹, Christian Haselgrove², David Kennedy², Zhizhong Liu³,
Michael Milham¹, Petros Petrosyan⁴, Carinna Torgerson³, John Van Horn³, Cameron Craddock¹

¹Child Mind Institute, New York, NY, ² University of Massachussetts Medical School, Worcester, MA, ³University of Souther California, Los Angeles, CA, ⁴UCLA, Los Angeles, CA, ⁵Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY

Introduction

- ▶ The National Database for Autism Research (NDAR) hosts a vast collection of neuroimaging datasets that can be processed and utilized to yield significant scientific discoveries.
- ▶ This amount of resources necessitates a high-performance computing (HPC) infrastructure, which is not always readily available for researchers in-house.
- ▶ Amazon Web Services (AWS) Elastic Compute Cloud (EC2) computing service offers a “pay as you go” model that allows researchers to utilize HPC performance without the up-front captial costs and maintenance of an in-house solution.
- ▶ The developers of the Laboratory of Neuro Imaging (LONI) Pipeline, the Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC) Computational Environment (CE) and the Configurable Pipeline for the Analysis of Connectomes (C-PAC) have implemented pipelines in EC2 that interface with NDAR

Methods

- ▶ Launched an AWS-hosted miNDAR database by querying NDAR website for the data of interest (e.g. from a particular study)
- ▶ Built a subject list by querying the database for subjects of interest to pass to our pipeline
- ▶ Launch an AWS EC2 HPC cluster using Starcluster
- ▶ Log into the cluster and submit a Sun Grid Engine job using our pipeline software and the subject list
- ▶ The pipeline software will process the data, store raw outputs in an AWS S3 bucket and insert S3 filepaths and output measures into miNDAR database

COLUMN_NAME	DATA_TYPE	NULLABLE	DATA_DEFAULT	COLUMN_ID	COMMENTS
1 IMAGE_ID	NUMBER(38, 0)	Yes	(null)	1	(null)
2 DATASET_ID	NUMBER(38, 0)	Yes	(null)	2	(null)
3 SUBJECT_KEY	VARCHAR2(255 BYTE)	Yes	(null)	3	(null)
4 SRC_SUBJECT_ID	VARCHAR2(255 BYTE)	Yes	(null)	4	(null)
5 INTERVIEW_DATE	TIMESTAMP(4) WITH TIME ZONE	Yes	(null)	5	(null)
6 INTERVIEW_AGE	NUMBER(38, 0)	Yes	(null)	6	(null)
7 GENDER	VARCHAR2(50 BYTE)	Yes	(null)	7	(null)
8 COMMENTS_NESC	VARCHAR2(1000 BYTE)	Yes	(null)	8	(null)
9 IMAGE_FILE	VARCHAR2(255 BYTE)	Yes	(null)	9	(null)
10 IMAGE_METADATA_FILE	VARCHAR2(255 BYTE)	Yes	(null)	10	(null)
11 IMAGE_DESCRIPTION	VARCHAR2(512 BYTE)	Yes	(null)	11	(null)
12 IMAGE_FILE_FORMAT	VARCHAR2(50 BYTE)	Yes	(null)	12	(null)
13 IMAGE_MODALITY	VARCHAR2(50 BYTE)	Yes	(null)	13	(null)
14 SCANNER_MANUFACTURER_PO	VARCHAR2(50 BYTE)	Yes	(null)	14	(null)
15 SCANNER_TYPE_PO	VARCHAR2(50 BYTE)	Yes	(null)	15	(null)
16 SCANNER_SOFTWARE_VERSION_PO	VARCHAR2(50 BYTE)	Yes	(null)	16	(null)
17 MAGNETIC_FIELD_STRENGTH	VARCHAR2(50 BYTE)	Yes	(null)	17	(null)
18 MRI_REFERENCE_TIME_PO	PLAT	Yes	(null)	18	(null)
19 MRI_SCAN_TIME_PO	PLAT	Yes	(null)	19	(null)
20 FLIP_ANGLE	VARCHAR2(50 BYTE)	Yes	(null)	20	(null)
21 ACQUISITION_MATRIX	VARCHAR2(50 BYTE)	Yes	(null)	21	(null)
22 MRI_FIELD_OF_VIEW_PO	VARCHAR2(50 BYTE)	Yes	(null)	22	(null)
23 PATIENT_POSITION	VARCHAR2(50 BYTE)	Yes	(null)	23	(null)
24 PATIENT_ORIENTATION	VARCHAR2(50 BYTE)	Yes	(null)	24	(null)
25 RECEIVE_COIL	VARCHAR2(50 BYTE)	Yes	(null)	25	(null)
26 TRANSFORM_CODE	VARCHAR2(50 BYTE)	Yes	(null)	26	(null)
27 TRANSFORMATION_PERFORMED	VARCHAR2(50 BYTE)	Yes	(null)	27	(null)
28 TRANSFORMATION_TYPE	VARCHAR2(50 BYTE)	Yes	(null)	28	(null)
29 IMAGE_HISTORY	VARCHAR2(1024 BYTE)	Yes	(null)	29	(null)
30 IMAGE_RAW_CONSISTENCY	NUMBER(38, 0)	Yes	(null)	30	(null)
31 IMAGE_EXTENT1	NUMBER(38, 0)	Yes	(null)	31	(null)
32 IMAGE_EXTENT2	NUMBER(38, 0)	Yes	(null)	32	(null)
33 IMAGE_EXTENT3	NUMBER(38, 0)	Yes	(null)	33	(null)
34 IMAGE_EXTENT4	NUMBER(38, 0)	Yes	(null)	34	(null)
35 EXTENT1_TYPE	VARCHAR2(50 BYTE)	Yes	(null)	35	(null)
36 EXTENT2_TYPE	VARCHAR2(50 BYTE)	Yes	(null)	36	(null)
37 EXTENT3_TYPE	VARCHAR2(50 BYTE)	Yes	(null)	37	(null)
38 IMAGE_UNIT1	VARCHAR2(20 BYTE)	Yes	(null)	38	(null)
39 IMAGE_UNIT2	VARCHAR2(20 BYTE)	Yes	(null)	39	(null)
40 IMAGE_UNIT3	VARCHAR2(20 BYTE)	Yes	(null)	40	(null)
41 IMAGE_UNIT4	VARCHAR2(20 BYTE)	Yes	(null)	41	(null)
42 IMAGE_UNITS	VARCHAR2(20 BYTE)	Yes	(null)	42	(null)
43 IMAGE_RESOLUTION	PLAT	Yes	(null)	43	(null)
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

Figure 1 : miNDAR database

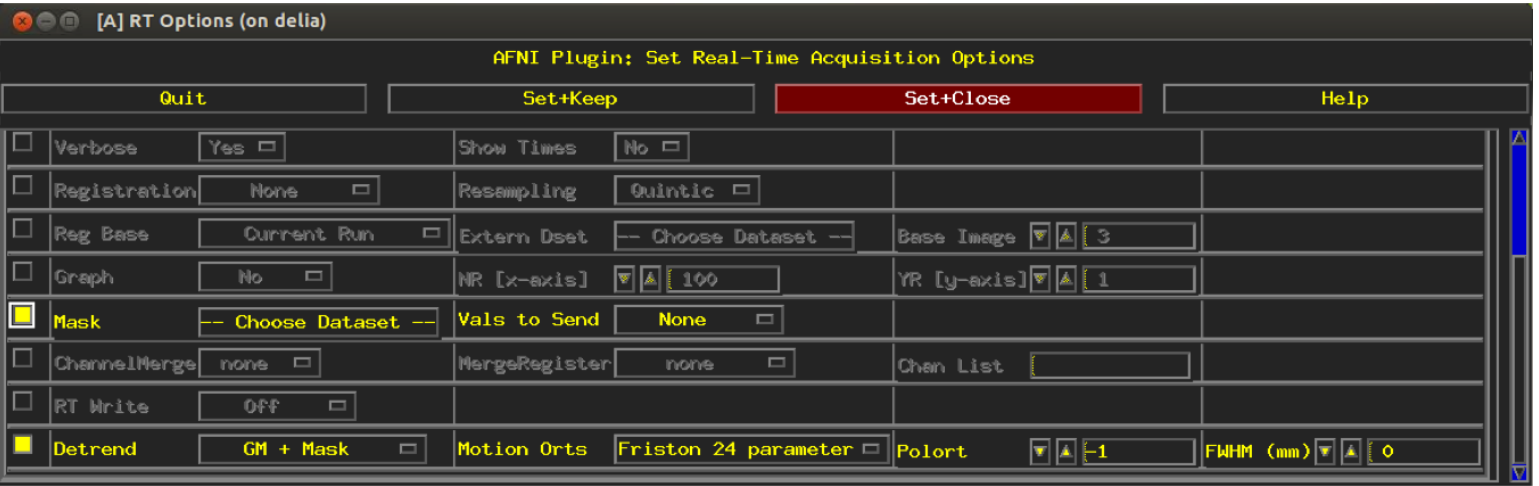
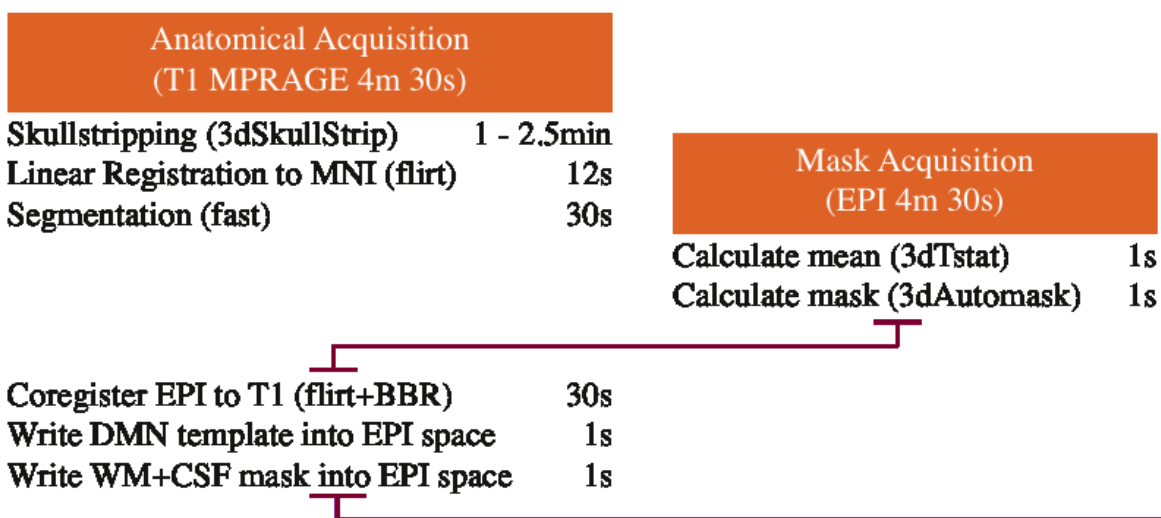


Figure 2 : AFNI interface for online denoising.

Online Preprocessing

- ▶ RT denoising implemented in AFNI¹⁰ to remove contributions of confounds (intensity modulations induced by head motion, physiological noise, scanner drift, ...)
 - ▷ N^{th} order polynomial
 - ▷ Global mean
 - ▷ Mask average time series (i.e. WM, CSF)
 - ▷ Motion parameters (6 or 24 regressor models)
 - ▷ Spatial smoothing
- ▶ Adds < 5 ms of delay



Results

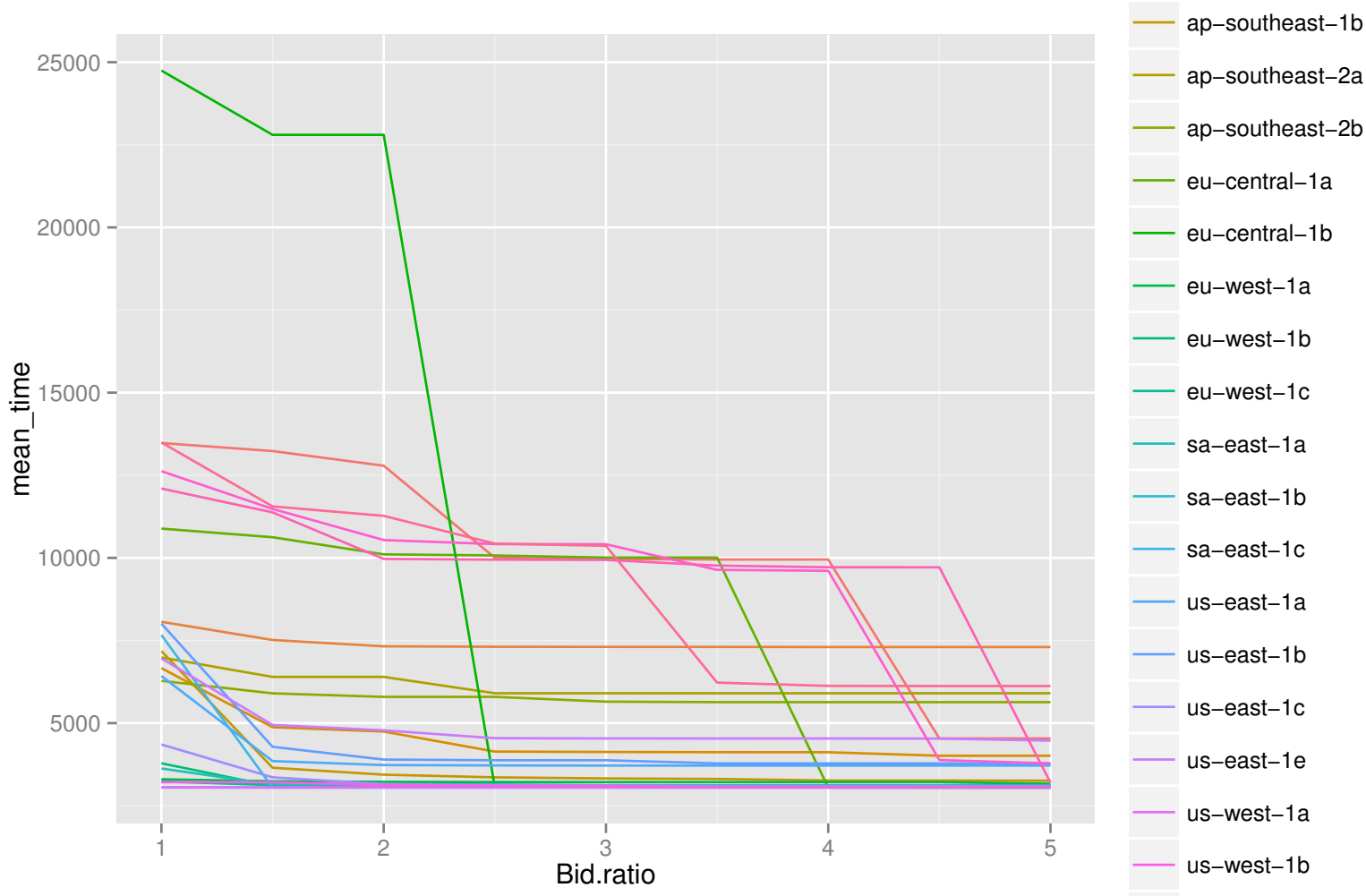


Figure 4 : Bid ratio vs Computation time in minutes for the Freesurfer pipeline

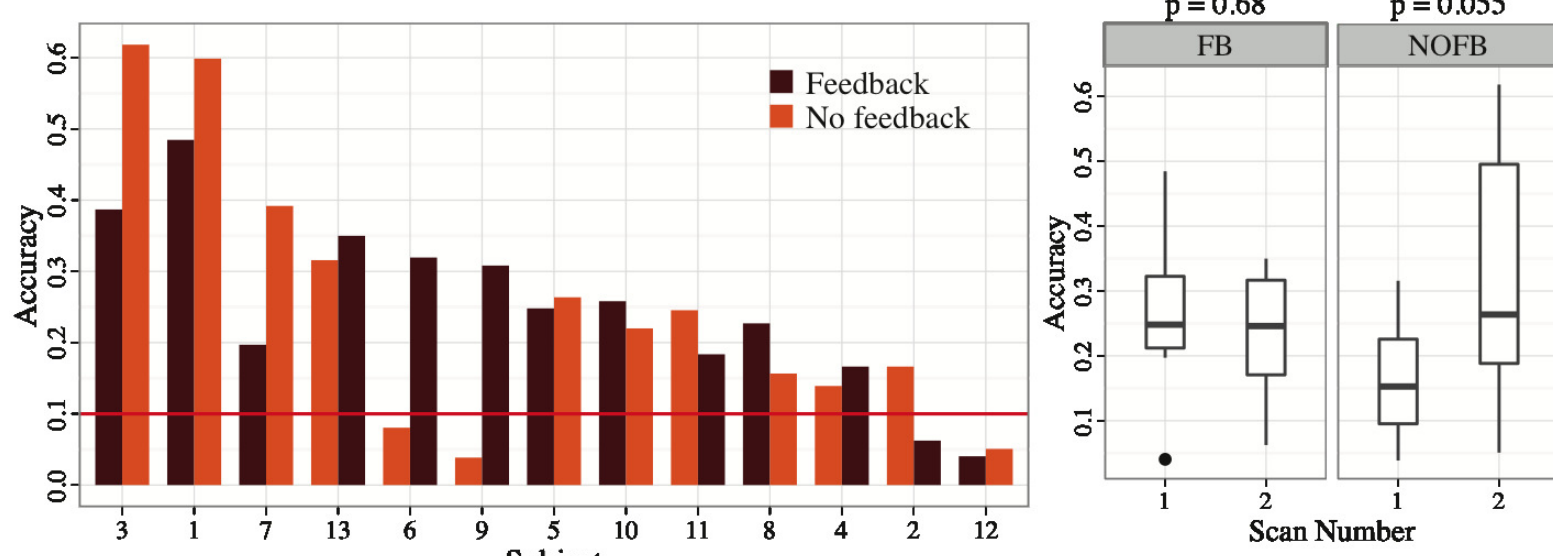


Figure 5 : Performance across participants (A) differs between feedback and neurofeedback scans as determined by their order (B).

Conclusion

- ▶ We developed a system for measuring DN regulation using realtime neurofeedback.
- ▶ Participants were able to modulate their DN activity and their ability to do so was correlated with phenotype.
- ▶ This system provides a new experimental paradigm for understanding network dysregulation and how it maps to disease states and phenotype.

References and Acknowledgements

1. Sonuga-Barke, E. et al. (2007), Neuroscience and Biobehavioral Reviews 31:977-986.
2. Broyd, S. J. et al. (2009), Neuroscience and Biobehavioral Reviews 33: 279-296.
3. Sheline, Y.I. et al. (2009), PNAS 106: 1942-1947.
4. Whitfield-Gabrieli, S. et al. (2009), PNAS 106: 1279-1284.
5. Hamilton, J.P. et al. (2011), Biol. Psychiatry 70: 327-333.
6. Sylvester, C.M. et al. (2012), Trends Neurosci. 35, 527-535.
7. Castellanos, F.X. et al. (2012), Trends Cogn. Sci. 16, 17-26.
8. LaConte, S.M. et al. (2004). Human Brain Mapping, 11: 2551.
9. LaConte, S.M. et al. (2011). Neurolmage, 56(2), 440-454.
10. Cox, R.W. (1996) Comput. Biomed. Res. 29: 162-173.

Data collection and salary support was provided by a NARSAD Young Investigator Award to RCC.