Data Article

MALINI (Machine Learning in NeuroImaging): A MATLAB toolbox for aiding clinical diagnostics using resting-state fMRI data

Pradyumna Lanka [a, b], D. Rangaprakash [a, c, d],
Sai Sheshan Roy Gotoor [a], Michael N. Dretsch [e, f, g],
Jeffrey S. Katz [a, g, h, j], Thomas S. Denney Jr. [a, g, h, j],
Gopikrishna Deshpande [a, g, h, i, j, k, l, m, *]

[a] AU MRI Research Center, Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA
[b] Department of Psychological Sciences, University of California Merced, Merced, CA, USA
[c] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[d] Division of Health Science and Technology, Massachusetts Institute of Technology, Boston, MA, USA
[e] U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, USA
[f] US Army Medical Research Directorate-West, Walter Reed Army Institute for Research, Joint Base Lewis-McChord, WA, USA
[g] Department of Psychology, Auburn University, Auburn, AL, USA
[h] Alabama Advanced Imaging Consortium, Birmingham, AL, USA
[i] Center for Health Ecology and Equity Research, Auburn University, Auburn, AL, USA
[j] Center for Neuroscience, Auburn University, Auburn, AL, USA
[k] Department of Psychiatry, National Institute of Mental and Neurosciences, Bangalore, India
[l] School of Psychology, Capital Normal University, Beijing, China
[m] Key Laboratory for Learning and Cognition, Capital Normal University, Beijing, China

ARTICLE INFO

ABSTRACT

Resting-state functional Magnetic Resonance Imaging (rs-fMRI) has been extensively used for diagnostic classification because it does not require task compliance and is easier to pool data from multiple imaging sites, thereby increasing the sample size. A MATLAB-based toolbox called Machine Learning in NeuroImaging (MALINI) for feature extraction and disease classification is presented. The MALINI toolbox extracts functional and effective connectivity features from preprocessed rs-fMRI data and performs classification between healthy and disease groups using any of 18

* Corresponding author. AU MRI Research Center, Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, 36849, USA.
E-mail address: gopi@auburn.edu (G. Deshpande).

Functional connectivity
Autism
ADHD
Alzheimer's disease
PTSD

popular and widely used machine learning algorithms that are based on diverse principles. A consensus classifier combining the power of multiple classifiers is also presented. The utility of the toolbox is illustrated by accompanying data consisting of resting-state functional connectivity features from healthy controls and subjects with various brain-based disorders: autism spectrum disorder from autism brain imaging data exchange (ABIDE), Alzheimer's disease and mild cognitive impairment from Alzheimer's disease neuroimaging initiative (ADNI), attention deficit hyperactivity disorder from ADHD-200, and post-traumatic stress disorder and post-concussion syndrome acquired in-house. Results of classification performed on the above datasets can be obtained from the main article titled "Supervised machine learning for diagnostic classification from large-scale neuroimaging datasets" [1]. The data was divided into homogeneous and heterogeneous splits, such that 80% could be used for training, model building and cross-validation, while the remaining 20% of the data could be used as a hold-out independent test data for replication of the classification performance, to ensure the robustness of the classifiers to population variance in image acquisition site and age of the sample.

Specifications Table

Subject area	*Brain imaging*
More specific subject area	*Functional magnetic resonance imaging, Diagnostic classification, Resting state functional connectivity (RSFC), Supervised machine learning classifiers*
Type of data	*MATLAB Toolbox, tables*
How data was acquired	*3T MRI scanners, resting-state functional MRI*
Data format	*Raw, Analyzed*
Experimental factors	*Resting-state functional connectivity-based features from healthy controls and autism spectrum disorder (ASD) from autism brain imaging data exchange (ABIDE), Alzheimer's disease (AD) and mild cognitive impairment (MCI) from Alzheimer's disease neuroimaging initiative (ADNI), attention deficit hyperactivity disorder (ADHD) from ADHD-200 and post-traumatic stress disorder (PTSD) and post-concussion syndrome (PCS) acquired in-house.*
Experimental features	*Resting-state: participants were requested to have their eyes open and were asked to not think of anything specific. Functional connectivity features calculated from resting-state functional Magnetic Resonance Imaging (Rs-fMRI) volumes. The data was further divided into two samples for training and testing. Various splits were performed resulting in both homogeneous and heterogeneous samples.*
Data source location	*USA, Germany, Netherlands, Belgium, Ireland, China*
Data accessibility	*The raw resting-state functional connectivity data is available with this article. The toolbox and the data can also be found at the following URL:* https://github.com/pradlanka/malini
Related research article	P. Lanka, D. Rangaprakash, M.N. Dretsch, J.S. Katz, T.S. Denney Jr., G. Deshpande, Supervised machine learning for diagnostic classification from large-scale neuroimaging datasets, Brain Imaging and Behavior (2019) in press, https://doi.org/10.1007/s11682-019-00191-8 [1].

Value of the Data

- Our data can be used for replication, benchmarking and testing the performance of various machine learning algorithms to classify neurological diseases.
- These datasets would provide machine learning enthusiasts, who are not from the field of neuroimaging, an opportunity to explore disease classification without any prior knowledge or experience in either neuroscience or neuroimaging. However, researchers from outside the field of neuroimaging must familiarize themselves with the nuances, caveats, and limitations of the application of machine learning to questions of clinical diagnosis using neuroimaging-derived features before using these methods [1−3].
- The MALINI toolbox provides a one-stop solution for extracting the BOLD time series from regions in the CC200 template [4] and other brain parcellation templates, calculation of connectivity-based features from pre-processed fMRI data, and disease classification from the extracted features.
- The toolbox we include has 18 different machine learning algorithms embedded within it which can be used for disease classification. Further, consensus classification, by combining inferences from multiple classifiers, is also available. The code in the toolbox can be modified to include other machine learning classifiers as well.

1. Data

We utilized resting-state functional magnetic resonance imaging (rs-fMRI) data from several open-source neuroimaging databases such as autism brain imaging data exchange (ABIDE), Alzheimer's disease neuroimaging initiative (ADNI) and attention deficit hyperactivity disorder (ADHD-200) sample. The data also includes post-traumatic stress disorder (PTSD) data acquired at the Auburn University MRI Research Center. Along with this paper, we present tables with functional connectivity features from the 4 disorders/diseases for homogeneous and heterogeneous splits of the data for training/validation and testing as excel files. More information about the data and the splits can be obtained from the main paper corresponding to this article: Lanka et al. (2019) [1]. Tables with the labels corresponding to brain structures for the functional connectivity paths are included as well as excel files. The toolbox contains MATLAB script files for implementing machine learning algorithms.

2. Experimental design, materials and methods

Resting-state functional connectivity was estimated using data obtained from the ABIDE database which includes 556 healthy controls, 93 with Asperger's syndrome and 339 with Autism [5]. More information about the data can be obtained from http://fcon_1000.projects.nitrc.org/indi/abide/index. html. The data for ADHD was released by the ADHD-200 consortium [6]. The ADHD-200 dataset includes 573 healthy controls, 208 subjects with ADHD-C (Combined), 13 subjects with subtype ADHD-H (hyperactive), and 136 subjects with ADHD-I (inattentive), a total of 930 subjects [8]. Further information can be obtained from http://fcon_1000.projects.nitrc.org/indi/adhd200. Data for the mild cognitive impairment (MCI) and Alzheimer's disease was obtained from ADNI database. The data contains 132 subjects with 34 subjects diagnosed with early mild cognitive impairment (EMCI), 34 with late mild cognitive impairment (LMCI), 29 with Alzheimer's disease (AD) and 35 matched healthy controls [7]. More information about this database can be obtained from http://adni.loni.usc. edu.

Unlike the other three datasets, the data for subjects with PTSD and post-concussion syndrome (PCS) were acquired in-house at the Auburn University MRI Research Center. It consists of rs-fMRI data acquired from 87 active duty male US Army soldiers (who served in the wars in Iraq and/or Afghanistan) with 28 combat controls, 17 diagnosed with only PTSD, and 42 diagnosed with both PCS and PTSD. The subject groups were matched for age, race, education and deployment history. More information about the diagnosis of the subjects for PTSD or for both PCS and PTSD can be found in Lanka et al. [1] as well as in our previous publications [9−12]. The procedure and the protocols in this study were approved by the Auburn University Institutional Review Board (IRB) and the Headquarters U.S. Army Medical Research and Material Command, IRB (HQ USAMRMC IRB). The participants were scanned in a Siemens 3T MAGNETOM Verio Scanner (Siemens, Erlangen, Germany) with a 32-channel head coil. Two runs of resting-state data were acquired from all subjects using a T2* weighted multiband echo-planar imaging (EPI) sequence with the following acquisition parameters:

TR = 600 ms, TE = 30 ms, FA = 55°, multiband factor = 2, Voxel size = 3×3×5 mm^3 and 1000 time points. It must be noted that the brain coverage for the volumes was limited to just the cortical and subcortical structures, with the cerebellum excluded.

2.1. Pre-processing of resting-state fMRI data and calculation of functional connectivity

Data Processing Assistant for Resting-State fMRI Toolbox (DPARSF) [13] was used for preprocessing the rs-fMRI data. Pre-processing steps included slice timing correction, 3D volume realignment, co-registration of the T1-weighed structural image to the mean functional image, nuisance variable regression which included linear detrending, mean global signal, white matter and cerebrospinal fluid signals and 6 motion parameters. The data were normalized to Montreal Neurological Institute (MNI) template. The blood-oxygen-level-dependent (BOLD) time series from every voxel in the brain was deconvolved by estimating the voxel-specific hemodynamic response function (HRF) using a blind deconvolution procedure to obtain the latent neural signals [14]. The data were then temporally filtered with a band pass filter of bandwidth 0.01−0.1 Hz. Mean time series were extracted from 200 functionally homogeneous brain regions as defined by the CC200 template [4]. After extracting the time series, pair-wise functional connectivity (FC) between the 200 regions was calculated as the Pearson's correlation coefficient between pairs of time series, giving us a total of 19,900 FC values. These connectivity values were then used as features for the classification procedure. For ADHD and PTSD datasets, we did not have whole brain coverage. Therefore, time series were obtained from only 190 regions and 125 regions respectively, lowering the number of functional connectivity paths in the data. The connectivity measures for two of the datasets we release with this paper (ABIDE and ADHD) are from publicly available databases containing resting-state fMRI data obtained from different sites with different TRs. However, the PTSD dataset was acquired in-house using the same scanner and scanning parameters while ADNI data that we have employed was also acquired on the same scanner with identical parameters. The classifier accuracy on PTSD and ADNI datasets would represent performance in a more homogeneous dataset.

2.2. Data splits for training/validation and hold-out test data

For each of the four neuroimaging datasets, we split the entire data into two sets: 80% of the data can be used for training/validation, and the remaining 20% can be used as an independent hold-out test data for replicating the classification performance. In a few data splits, the training/validation and test data came from homogeneous populations, i.e. they were matched for age and acquisition site. In some other splits, the training/validation and hold-out test data were not matched, i.e. they had different age range or acquisition site. Therefore, training/validation and the test data for the heterogeneous splits were matched in race, gender and education but differed in either age range or acquisition site (but not both). For the homogeneous split, the training/validation data and the test data were matched, in all the factors described above along with age and acquisition site. There are two heterogeneous splits (differing in age range or acquisition site) for ABIDE data, and one heterogonous split (differing in age range) for PTSD and ADNI. For ADHD, unlike the other three datasets, the data was split into training/validation and test data as released by the ADHD-200 consortium [6]. The split information is summarized in Fig. 1. More information on the splits can be obtained from Lanka et al. [1].

The more training data we use, the more the classifier learns and performs better, but the flipside is that we have less data for testing which makes it difficult to estimate the generalizability of the performance of our classifiers. If we used an even split (instead of an 80:20 split), then the classifiers will not have enough data to learn from. Unfortunately, we cannot use cross-validation accuracy as our primary metric of performance given that it is known to inflate performance. Therefore, we split the data into training/validation data (which uses cross-validation) and an independent hold-out test data to essentially illustrate the difficulty in the generalization of machine learning classifiers to differences in data acquisition site and age groups. Given these factors, we chose 80:20 split since it is a commonly used rule-of thumb. This allowed the classifier to have enough training data to learn and still be left with enough independent test data to get an unbiased estimate of the generalization accuracy of the classifier.

a

ABIDE
Autism Brain Imaging Data Exchange

Different age split

Diagnosis	Training/Validation Data Age 7-21	Independent Test Data Age 22-58
Controls	429	127
Asperger's	63	30
Autism	279	60
Total Subjects	761	217

Matched split

Diagnosis	Training/Validation Data Matched	Independent Test Data Matched
Controls	447	111
Asperger's	74	19
Autism	271	68
Total Subjects	791	198

Different Acquisition site split

Diagnosis	Training/Validation Data 12 Institutions	Independent Test Data 3 Institutions
Controls	451	106
Asperger's	74	19
Autism	268	71
Total Subjects	792	196

b

ADHD-200
Global Competition

Diagnosis	Training/Validation Data 7 imaging sites	Hold-out Test Data 6 imaging sites
Controls	479	94
ADHD-C	159	49
ADHD-H	11	2
ADHD-I	110	26
Total Subjects	759	171

c

PTSD
Post-Traumatic Stress Disorder

Different age split

Diagnosis	Training/Validation Data Age 23-37	Independent Test Data Age 38-53
Controls	46	10
PCS & PTSD	66	18
PTSD	26	8
Total Subjects	138	36

Age-matched split

Diagnosis	Training/Validation Data Age 23-53	Independent Test Data Age 23-53
Controls	45	11
PCS & PTSD	67	17
PTSD	27	7
Total Subjects	139	35

d

ADNI
Alzheimer's Disease Neuroimaging Initiative

Different age split

Diagnosis	Training/Validation Data Age 56-76	Independent Test Data Age 77-88
Controls	23	12
Early Mild Cognitive Impairment (EMCI)	28	6
Late Mild Cognitive Impairment (LMCI)	27	7
Alzheimer's Disease (AD)	22	7
Total Subjects	100	32

Age-matched split

Diagnosis	Training/Validation Data Age 56-88	Independent Test Data Age 56-88
Controls	28	7
Early Mild Cognitive Impairment (EMCI)	27	7
Late Mild Cognitive Impairment (LMCI)	27	7
Alzheimer's Disease (AD)	23	6
Total Subjects	105	27

Fig. 1. Number of subjects in each subgroup after splitting the entire data for the following datasets — (A) autism brain imaging data exchange (ABIDE), (B) attention deficit hyperactivity disorder-200 (ADHD-200), (C) post-traumatic stress disorder (PTSD), and (D) Alzheimer's disease neuroimaging initiative (ADNI), into training/validation and hold-out test data. PCS: post-concussion syndrome; ADHD−I (inattentive); ADHD−H (hyperactive/impulsive); ADHD−C (combined).

A *t*-test/ANOVA with p<0.05 (FDR corrected), was performed on the features in the training/validation data to reduce the number of features to around 1000, after controlling for differences in the head motion, age, race, and education between the groups. This was performed to reduce the computational time and ensure that the classifiers do not overfit the data. The hold-out test datasets were not used for feature selection by a *t*-test/ANOVA. Hence, it can be used to provide an unbiased estimate of the performance of the classifiers.

3. MALINI toolbox

We developed Machine Learning in NeuroImaging (MALINI), a MATLAB (MathWorks, Natick, MA) based toolbox which automates the entire process. It allows one to automatically extract the functional connectivity (undirected measure based on Pearson's correlation between time series) and effective connectivity (directed measure based on Granger causality inferred from multivariate autoregressive models of time series [15,16]) features from the preprocessed 4D fMRI volumes. A picture of the GUI for the toolbox is shown in Fig. 2. Firstly, the toolbox divides the entire dataset into training data and test data based on a value entered in a data split ratio field. Using the preprocessed 4D fMRI volumes, the toolbox extracts the time-series from the CC200 template [4], CC400 [4] or Dosenbach 160 ROIs [17]. The toolbox also gives the user a chance to use their own brain parcellation template. BOLD time series is a convolution of the hemodynamic response function (HRF) with latent neuronal activity. Since the HRF varies across brain regions and individuals, it affects both functional [9] and effective connectivity metrics [18]. Therefore, blind HRF deconvolution becomes especially important to obtain reliable connectivity metrics in the latent neural space. Considering the advantages of HRF deconvolution [19−21], the toolbox gives the user an option to extract connectivity-based features either from the raw BOLD-time series or from the deconvolved data after performing a blind hemodynamic deconvolution of the time-series as detailed in Wu et al. [14]. The toolbox calculates pair-wise functional connectivity metrics from Pearson's correlation coefficient and pair-wise effective connectivity metrics utilizing Granger causality, either from the original time series or the HRF deconvolved latent neural time series. Using the time series, the toolbox then calculates (i) Static functional connectivity (SFC), (ii) Dynamic functional connectivity (DFC) [22], and (iii) Static effective connectivity (SEC) [9]. Consequently, the user has the option of using static functional connectivity (SFC), the variance of dynamic functional connectivity (varDFC) [23], or static effective connectivity (SEC) as features from the time series for disease classification. The toolbox also gives the user the opportunity to directly classify any type of feature (i.e. need not necessarily be connectivity) arranged as rows and subjects as columns in an excel file.
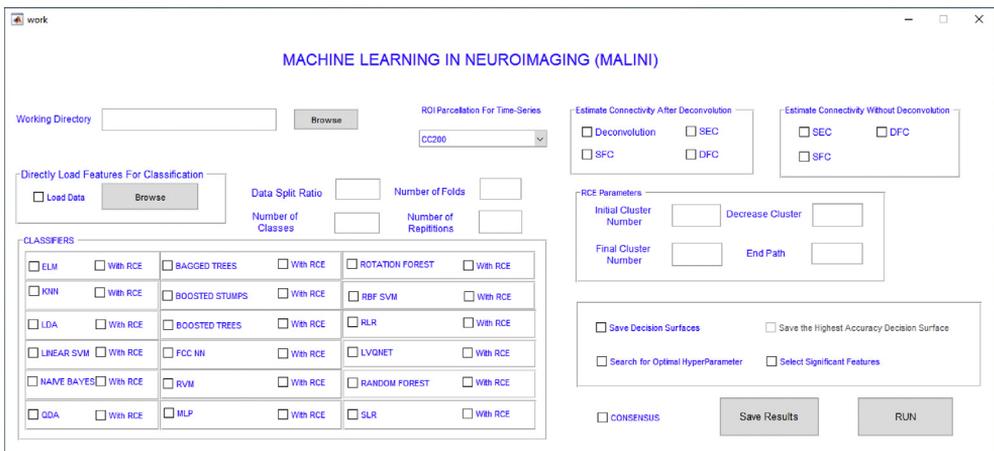
Fig. 2. A snapshot of the GUI for the proposed MALINI toolbox.

The classification procedure in the toolbox implements the classifiers in a nested cross-validation procedure with the inner cross-validation loop used for feature selection and the outer cross-validation loop used for performance estimation. Since the number of features obtained are $\binom{N}{2}$ for static and dynamic functional connectivity measures and twice as many for effective connectivity, feature reduction can help avoid overfitting and give better generalization. Reducing the features quickly is accomplished by *t*-test filtering within the inner cross-validation loop to include only the features whose means are significantly different between the groups. Consequently, the cross-validation estimate of the classifier performance is unbiased and ensures clear separation between feature selection and performance estimation. The users also have the option of implementing the classifiers in a recursive cluster elimination (RCE) framework [24] for further feature elimination, though it requires exponentially more time, as it is a wrapper method. Given the computational and time constraints, RCE framework is recommended if the number of features is still large compared to the number of class instances in the sample, even after *t*-test/ANOVA filtering. The RCE classification procedure requires additional parameters such as the initial number of feature clusters, the final number of feature clusters and the number of feature clusters to be eliminated at every step in the RCE procedure. These parameters can be set in the toolbox.

The number of folds for cross-validation and the number of repetitions of the cross-validation procedure can be set in the toolbox. The classifiers implemented in the toolbox include Probabilistic/Bayesian methods: Gaussian naïve Bayes (GNB), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), sparse logistic regression (SLR), regularized logistic regression (RLR), (ii) Kernel methods: linear and radial -basis function (RBF) kernel support vector machine (SVM) [25], relevance vector machine (RVM), (iii) Artificial neural networks: multilayer perceptron neural net (MLP-Net), fully connected neural net (FC-Net), extreme learning machine (ELM), linear vector quantization net (LVQNET), (iv) Instance-based learning: k-nearest neighbors (KNN), (v) Decision tree based ensemble methods: bagged trees [26], boosted trees, boosted stumps, random forest, rotation forest. More information about the classifiers can be obtained from Lanka et al. [1]. The toolbox also gives the user the opportunity to modify the default parameters for optimizing hyperparameters in the included classifiers. The optimal hyperparameters are chosen by a simple grid search via a linear or an exponential search procedure. When more than one classifier is selected, the user can select the consensus classifier, which uses the all the selected classifiers to predict the disease class of the observations in the test data (For details regarding the consensus classifier, the readers are referred to Lanka et al. [1]).

The toolbox provides the overall accuracy (balanced and unbalanced accuracies), individual class accuracies, and the confusion matrix. We did not add specificity, sensitivity and positive predictive value (PPV) measures because it would be difficult to interpret such measures for multi-class classification scenarios with controls and subcategories of clinical populations. But since, the toolbox does output the confusion matrix, the users can use the confusion matrix to derive measures that they might be interested in. The toolbox also gives the user an option to save the classification models (outer k-fold x no of resamples), which can then be used to predict the disease status on an independent hold-out test dataset for replication to ensure the reliability and robustness of the classifier model obtained from training data.

Acknowledgements

the MATLAB code available at https://www.mathworks.com/matlabcentral/fileexchange/38792-rotation-forest-algorithm. The Matlab code for Extreme learning machine (ELM) [32,33] was obtained from http://www.ntu.edu.sg/home/egbhuang/elm_kernel.html. For the remaining classifiers, we used the inbuilt classifier functions available in MATLAB.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2020.105213.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P. Lanka, D. Rangaprakash, M.N. Dretsch, J.S. Katz, T.S. Denney Jr., G. Deshpande, Supervised Machine Learning for Diagnostic Classification from Large-Scale Neuroimaging Datasets, Brain Imaging and Behavior, 2019 in press, https://doi.org/10.1007/s11682-019-00191-8.

[2] M.R. Arbabshirani, S. Plis, J. Sui, V.D. Calhoun, Single subject prediction of brain disorders in neuroimaging: promises and pitfalls, Neuroimage 145 (2017) 137—165, https://doi.org/10.1016/j.neuroimage.2016.02.079.

[3] L.Q. Uddin, D.R. Dajani, W. Voorhies, H. Bednarz, R.K. Kana, Progress and roadblocks in the search for brain-based biomarkers of autism and attention-deficit/hyperactivity disorder, Transl. Psychiatry 7 (2017) e1218, https://doi.org/10.1038/tp.2017.164.

[4] R.C. Craddock, G. James, P.E. Holtzheimer, X.P. Hu, H.S. Mayberg, A whole brain fMRI atlas generated via spatially constrained spectral clustering, Hum. Brain Mapp. 33 (2012) 1914—1928, https://doi.org/10.1002/hbm.21333.

[5] M.A. Syed, Z. Yang, X.P. Hu, G. Deshpande, Investigating brain connectomic alterations in autism using the reproducibility of independent components derived from resting state functional MRI data, Front. Neurosci. 11 (2017) 459, https://doi.org/10.3389/fnins.2017.00459.

[6] The ADHD-200 Consortium, the ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience, Front. Syst. Neurosci. 6 (2012) 62, https://doi.org/10.3389/fnsys.2012.00062.

[7] G. Deshpande, P. Wang, D. Rangaprakash, B. Wilamowski, Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data, IEEE Trans. Cybern. 45 (2015) 2668—2679, https://doi.org/10.1109/TCYB.2014.2379621.

[8] S. Zhao, D. Rangaprakash, A. Venkataraman, P. Liang, G. Deshpande, Investigating focal connectivity deficits in Alzheimer's disease using directional brain networks derived from resting-state fMRI, Front. Aging Neurosci. 9 (2017) 211, https://doi.org/10.3389/fnagi.2017.00211.

[9] D. Rangaprakash, G. Deshpande, T.A. Daniel, A.M. Goodman, J.L. Robinson, N. Salibi, J.S. Katz, T.S. Denney, M.N. Dretsch, Compromised hippocampus-striatum pathway as a potential imaging biomarker of mild traumatic brain injury and post-traumatic stress disorder, Hum. Brain Mapp. 38 (2017) 2843–2864, https://doi.org/10.1002/hbm.23551.

[10] X. Zhao, D. Rangaprakash, T.S. Denney Jr., J.S. Katz, M.N. Dretsch, G. Deshpande, Identifying neuropsychiatric disorders using unsupervised clustering methods: data and code, Data in Brief 22 (2019) 570–573, https://doi.org/10.1016/j.dib.2018.01.080.

[11] D. Rangaprakash, M.N. Dretsch, A. Venkataraman, J.S. Katz, T.S. Denney Jr., G. Deshpande, Identifying disease foci from static and dynamic effective connectivity networks: illustration in soldiers with trauma, Hum. Brain Mapp. 39 (2018) 264–287, https://doi.org/10.1002/hbm.23841.

[12] D. Rangaprakash, M.N. Dretsch, W. Yan, J.S. Katz, T.S. Denney Jr., G. Deshpande, Hemodynamic variability in soldiers with trauma: implications for functional MRI connectivity studies, Neuroimage: Clinic 16 (2017) 409–417, https://doi.org/10.1016/j.nicl.2017.07.016.

[13] C.G. Yan, Y.F. Zang, DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI, Front. Syst. Neurosci. 4 (2010) 13, https://doi.org/10.3389/fnsys.2010.00013.

[14] G.R. Wu, W. Liao, S. Stramaglia, J.R. Ding, H. Chen, D. Marinazzo, A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data, Med. Image Anal. 17 (2013) 365–374, https://doi.org/10.1016/j.media.2013.01.003.

[15] G. Deshpande, S. LaConte, G.A. James, S. Peltier, X. Hu, Multivariate Granger causality analysis of brain networks, Hum. Brain Mapp. 30 (2009) 1361–1373, https://doi.org/10.1002/hbm.20606.

[16] G. Deshpande, P. Santhanam, X. Hu, Instantaneous and causal connectivity in resting state brain networks derived from functional MRI data, Neuroimage 54 (2011) 1043–1052, https://doi.org/10.1016/j.neuroimage.2010.09.024.

[17] N.U.F. Dosenbach, B. Nardos, A.L. Cohen, D.A. Fair, J.D. Power, J.A. Church, S.M. Nelson, G.S. Wig, A.C. Vogel, C.N. Lessov-Schlagga, K.A. Barnes, J.W. Dubis, E. Feczko, R.S. Coalson, J.R. Pruett Jr., D.M. Barch, S.E. Petersen, B.L. Schlaggar, Prediction of individual brain maturity using fMRI, Science 329 (2010) 1358–1361, https://doi.org/10.1126/science.1194144.

[18] G. Deshpande, X. Hu, Investigating effective brain connectivity from fMRI data: past findings and current issues with reference to Granger causality analysis, Brain Connect. 2 (2012) 235–245, https://doi.org/10.1089/brain.2012.0091.

[19] D. Rangaprakash, G.R. Wu, D. Marinazzo, X. Hu, G. Deshpande, Hemodynamic response function (HRF) variability confounds resting state fMRI connectivity, Magn. Reson. Med. 80 (2018) 1697–1713, https://doi.org/10.1002/mrm.27146.

[20] D. Rangaprakash, G.R. Wu, D. Marinazzo, X. Hu, G. Deshpande, Parameterized hemodynamic response function data of healthy individuals obtained from resting-state functional MRI in a 7T MRI scanner, Data in Brief 17 (2018) 1175–1179, https://doi.org/10.1016/j.dib.2018.01.003.

[21] D. Rangaprakash, M.N. Dretsch, W. Yan, J.S. Katz, T.S. Denney Jr., G. Deshpande, Hemodynamic response function parameters obtained from resting-state functional MRI data in soldiers with trauma, Data in Brief. 14 (2017) 558–562, https://doi.org/10.1016/j.dib.2017.07.072.

[22] J. Hao, X. Hu, G. Deshpande, Behavioral relevance of the dynamics of the functional brain connectome, Brain Connect. 4 (2014) 741–759, https://doi.org/10.1089/brain.2014.0300.

[23] C. Jin, H. Jia, P. Lanka, D. Rangaprakash, L. Li, T. Liu, X. Hu, G. Deshpande, Dynamic brain connectivity is a better predictor of PTSD than static connectivity, Hum. Brain Mapp. 38 (2017) 4479–4496, https://doi.org/10.1002/hbm.23676.

[24] G. Deshpande, Z. Li, P. Santhanam, C.D. Coles, M.E. Lynch, S. Hamann, X. Hu, Recursive cluster elimination based support vector machine for disease state prediction using resting state functional and effective brain connectivity, PloS One 5 (2010) e14277, https://doi.org/10.1371/journal.pone.0014277.

[25] L. Palaniyappan, G. Deshpande, P. Lanka, D. Rangaprakash, S. Iwabuchi, S. Francis, P.F. Liddle, Effective connectivity within a triple network brain system discriminates schizophrenia spectrum disorders from psychotic bipolar disorder at the single-subject level, Schizophr. Res. 214 (2019) 24–33, https://doi.org/10.1016/j.schres.2018.01.006.

[26] L.E. Libero, T.P. DeRamus, A.C. Lahti, G. Deshpande, R.K. Kana, Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates, Cortex 66 (2015) 46–59, https://doi.org/10.1016/j.cortex.2015.02.008.

[27] M. Yousef, S. Jung, L.C. Showe, M.K. Showe, Recursive Cluster Elimination (RCE) for classification and feature selection from gene expression data, BMC Bioinf. 8 (2007) 144, https://doi.org/10.1186/1471-2105-8-144.

[28] O. Yamashita, M-a. Sato, T. Yoshioka, F. Tong, Y. Kamitani, Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns, Neuroimage 42 (2008) 1414–1429, https://doi.org/10.1016/j.neuroimage.2008.05.050.

[29] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (27) (2011) 1–27, https://doi.org/10.1145/1961189.1961199, 27.

[30] A. Thayananthan, R. Navaratnam, B. Stenger, P.H.S. Torr, R. Cipolla, Multivariate relevance vector machines for tracking, in: A. Leonardis, H. Bischof, A. Pinz (Eds.), Proceeding on Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, vol. 3953, Heidelberg Springer, Berlin, 2006, pp. 124–138, https://doi.org/10.1007/11744078_10.

[31] J.J. Rodriguez, L.I. Kuncheva, C.J. Alonso, Rotation forest: a new classifier ensemble method, IEEE Trans. Pattern Anal. Mach. Intell. 28 (2006) 1619–1630, https://doi.org/10.1109/TPAMI.2006.211.

[32] G. Huang, G.B. Huang, S. Song, K. You, Trends in extreme learning machines: a review, Neural Network. 61 (2015) 32–48, https://doi.org/10.1016/j.neunet.2014.10.001.

[33] G.B. Huang, D.H. Wang, Y. Lan, Extreme learning machines: a survey, Int. J. Mach. Learn. Cybern. 2 (2011) 107–122, https://doi.org/10.1007/s13042-011-0019-y.