OPEN ACCESS
International Journal of Neural Systems, Vol. 31, No. 4 (2021) 2150009 (20 pages)

DOI: 10.1142/S012906572150009X

World Scientific
www.worldscientific.com

Impact of Machine Learning Pipeline Choices in Autism Prediction From Functional Connectivity Data

Manuel Graña*
Computational Intelligence Group
University of the Basque Country (UPV/EHU)
San Sebastian, Spain
manuel.grana@ehu.eus

Moises Silva
Universidad Mayor de San Andres
La Paz, Bolivia

Accepted 30 November 2020
Published Online 20 January 2021

Autism Spectrum Disorder (ASD) is a largely prevalent neurodevelopmental condition with a big social and economical impact affecting the entire life of families. There is an intense search for biomarkers that can be assessed as early as possible in order to initiate treatment and preparation of the family to deal with the challenges imposed by the condition. Brain imaging biomarkers have special interest. Specifically, functional connectivity data extracted from resting state functional magnetic resonance imaging (rs-fMRI) should allow to detect brain connectivity alterations. Machine learning pipelines encompass the estimation of the functional connectivity matrix from brain parcellations, feature extraction, and building classification models for ASD prediction. The works reported in the literature are very heterogeneous from the computational and methodological point of view. In this paper, we carry out a comprehensive computational exploration of the impact of the choices involved while building these machine learning pipelines. Specifically, we consider six brain parcellation definitions, five methods for functional connectivity matrix construction, six feature extraction/selection approaches, and nine classifier building algorithms. We report the prediction performance sensitivity to each of these choices, as well as the best results that are comparable with the state of the art.

Keywords: Autism; brain functional connectivity; feature extraction; brain parcellation; machine learning.

1. Introduction

Autism Spectrum Disorder (ASD)[1,2] is a highly prevalent, heritable, and heterogeneous neurodevelopmental disorder that has distinctive cognitive features often cooccurring with other psychiatric or neurological disorders. Across the Autism and Developmental Disabilities Monitoring (ADDM) Network sites, estimated ASD prevalence among children aged 8 years was 23.4 per 1000 (one in 43) boys and 5.2 per 1000 (one in 193) girls.[3] Thus, ASD appears to require specific research programs contemplating explicitly the impact of sex/gender-related issues.[4,5] Some references[6] state that the overall worldwide prevalence of ASD is as high as 1% of the population, while others[7] argue that nonspecific diagnostic criteria recently included in the diagnostic protocol produce as an artifact the 20-fold increase in its prevalence. Therefore, the heterogeneity in ASD features that hinders the identification of biomarkers could be a side effect of excessively wide

‡Corresponding author.

diagnostic criteria. ASD is currently diagnosed on the basis of qualitative information obtained from parent interviews and clinical observation, which leads to disturbing differences between sites.[8] Given its great prevalence, automated approaches to assist diagnosis[9,10] are highly desirable. Increasingly, clinical neuroscience focus is shifting to find brain imaging[11] derived metrics that may be useful to predict diagnostic category, disease progression, or response to intervention, e.g. looking for endophenotype using multivariate analysis approaches.[12] These metrics are raised from machine learning approaches to the study of brain structure and function. Some of them can be considered as neuroimage-based biomarkers that would be helpful to guide early interventions. Currently, the research community has not yet identified reliable and reproducible biomarkers for ASD. Issues related to clinical heterogeneity, methodological standardization, and cross-site validation must be addressed before further progress can be achieved.[13] A central role in the effort to obtain robust and reproducible biomarkers is played by the availability of public datasets, such as the Autism Brain Imaging Data Exchange (ABIDE) dataset[14,15] that includes demographic, clinical information and several magnetic resonance imaging (MRI) modalities allowing for a variety of studies such as brain maturity estimation as a biomarker of brain abnormality.[16] A wide variety of machine learning and pattern recognition studies on the computer-aided diagnostic of ASD have been reported[17] exploiting a variety of information sources, including structural, diffusion, and functional MRI, as well as electroencephalography (EEG),[18–20] additional demographic and clinical data,[21] and behavioral measurements captured by computer vision or other body measurement approaches.[22] There are meta-analysis confirmations of ASD imaging biomarkers from anatomical MRI,[23] and diffusion MRI imaging.[24,25] The latter showing white matter integrity disruption. Connectivity-based brain parcellation[26] provided additional evidence of altered white matter connectivity in ASD.

Brain functional connectivity analysis based on resting state functional MRI (rs-fMRI) can be done by seed analysis, where the specific connectivity relative to a selected brain region is compared across subjects and populations,[26] or on the basis of brain parcellations into a set of regions of interest (ROIs) which can be defined either by anatomical guidelines or by data-driven unsupervised segmentation.[27] In any case, rs-fMRI connectivity analysis has been accepted as a source of information for the discovery of biomarkers of psychiatric disorders[28] such as schizophrenia[29] and ASD.[13]

Neuroimage biomarker discovery may be guided by statistically significant differences between ASD and typically developing (TD) subjects. For instance, t-test on the dynamical network strength of ASD versus TD was reported to confirm identification of aberrant connectivity in ASD subjects,[30] and significant differences between ASD and TD in the level of activation of thalamic connectivity have been identified by independent component analysis (ICA).[31] In predictive analysis approaches to biomarker identification, the subject's condition (ASD versus TD) prediction performance achieved is the measure of the biomarker significance. Predictor models are built by machine learning techniques, often consisting of two steps: a dimensionally reduction (aka feature extraction or feature selection) followed by a classification step for class prediction. Although discrimination between ASD and TD is the most common paradigm, some works[32,33] compare ASD with Schizophrenia over a small cohort, while others consider patterns for discrimination among low-functioning and high-functioning ASD subjects.[34]

The general pipeline of predictive analysis for brain connectivity-based biomarkers is illustrated in Fig. 1. Functional connectivity matrices are extracted from preprocessed rs-fMRI data as follows: (1) a parcellation of the brain is defined; (2) the time series corresponding to the voxels in each region of the parcellation are aggregated into one representative time series often by averaging; (3) the connectivity matrix is built computing the similarity among the representatives of each pair of regions in the parcellation. Hence, the connectivity matrix is always a symmetric matrix that can be interpreted as the adjacency matrix of a graph representing the relations among brain regions. (4) The connectivity matrix is then used as the raw data for machine learning processes which may involve feature extraction/selection. (5) Predictive performance is estimated by the training/testing of classification models often in a cross-validation scheme. Predictive performance may be heavily influenced by the decisions made at each step, namely by the cohort selection, the brain volume parcellation, the

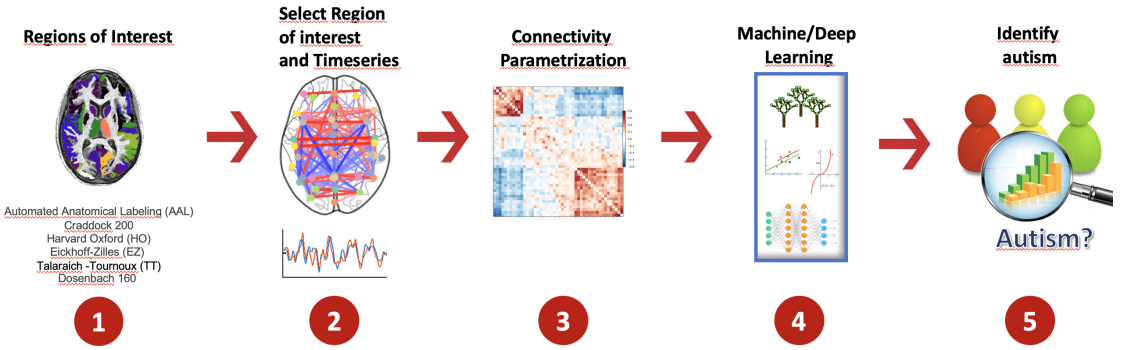

Fig. 1. Functional connectome predictive analysis pipeline steps after rs-fMRI data preprocessing (not shown): (1) given a parcellation of the brain, (2) obtain the representative time series of each region by averaging the time series of voxels within the region, (3) build the connectivity matrix by computing a similarity measure between each pair of representative time series, (4) carry out cross-validation experiments, using Machine Learning algorithms for feature extraction and classifier training; (5) report test results on the prediction of the ASD versus TD.

functional connectivity matrix estimation procedure, the feature selection/extraction algorithm, and the classification model building algorithm. After the cross-validation assessment, the cross-validation classification performance results may be used to identify biomarkers in the connectivity matrix.

Contributions and contents of the paper

We explore the impact of choices made in the implementation of the machine learning pipeline of Fig. 1 for the prediction of ASD versus TD. In this study, the functional connectivity matrix computed from rs-fMRI data is the sole source of information for classification. We have carried out extensive cross-validation experiments over the algorithmic choices at each step of the classification model building pipeline. We report the impact of all combinations of five feature extraction/selection approaches, six brain parcellations, five functional connectivity matrix computation methods, and ten classification model building techniques. The comprehensive comparison encompasses more than 11,000 (11,500) cross-validation experiments. We report statistically significant differences in performance found as well as direct comparison to state-of-the-art published results. We found that specific combinations of pipeline choices can boost the performance of ASD classification based on brain functional connectivity data. The software needed to replicate the experiments reported in this paper has been published in github.[a]

Section 2 provides a quick revision of the state of the art in the classification of ASD versus TD using functional connectivity data, Sec. 3 provides the details of the data and the computational methods used. Section 4 provides the computational results and a discussion on their relevance to the field. Finally, Sec. 5 provides our conclusions and ideas for future work.

2. Related Works

Artificial intelligence tools and problem-solving approaches[57] are contributing to the understanding and predictive analysis of ASD. Overall, there is increasing evidence that specific features extracted from MRI neuroimaging can be used to discriminate ASD from TD. However, there is a wide variety of methodological and computational approaches tested on widely different cohorts.[13,58] Regarding anatomical brain imaging, predictive models can be built based on anatomical differences computed by voxel-based morphometry (VBM) over gray and white matter segmentations of T1-weighted MRI data. Experiments over of a small cohort ($N = 120$) have reported average accuracies cross-validation experiments below 70% for a series of stratified computational experiments.[59] However, a similar study[60] using VBM results of diverse stratifications of female and male subjects of the ABIDE dataset reported much higher accuracies, above 90% in several subgroups. On the other hand, the use of 3D convolutional neural networks (3D-CNN)[61] on the complete

[a]https://github.com/mmscnet/Impact-feature-extraction-in-Autism

ABIDE T1-weighted dataset reported accuracy of 70%. Hence, there is a strong selection effect that may bias significantly the reported results. Significance weighted principal component analysis allows to remove the effects of site data acquisition improving discrimination based on anatomical imaging.[62]

Works based on functional connectivity information extracted from rs-fMRI data have been predominant in the latter times. They have been carried out over a wide variety of cohorts, testing many computational approaches. Relevant brain connection selection using logistic regression[63] achieved an accuracy of 85% by a linear classifier in a leave one out validation over a small cohort (74 high-functioning adult ASDs and 107 adult TDs). Further validation on an independent subset of the ABIDE dataset ($N = 88$) achieved a remarkable accuracy of 75%. Another work[64] reports accuracies over 80% on a small cohort of paired 20 ASD and 20 TD children using hyperconnectivity networks as features for an SVM classifier, while other authors[65] reported on the results of multilinear regression over the functional connectivity matrices after PCA dimensionality reduction of a cohort of 85 ASD and 163 TD children finding specific imbalances in brain connectivity for ASD children.

Table 1 summarizes the state of the art regarding the classification of subjects into ASD or TD on the basis of functional connectivity matrices extracted from the rs-fMRI data published in the ABIDE I dataset.[14,15] The criteria for inclusion in this table are (1) that the references report results on the (almost) complete ABIDE I dataset in order to be comparable to our own results reported below, and (2) that they report results using only features extracted from the functional connectivity matrices. We have excluded results obtained adding other kinds of information, such as the graph convolutional networks (GCN) enriched with demographic information,[35] and the features extracted from structural and MRI data.[55] The heterogeneity of the data in ABIDE as illustrated by the demographic information shown in Table 2 is a source of frustration for the machine learning approaches, so results concerning single sites report overly optimistic results that cannot be achieved with the entire dataset or a large subsample.[66] Another example of this practice carries out the cross-validation experiments intra-site reporting the average of the separate intrasite results.[50,67]

Columns of Table 1 reflect the choices made in the steps of the process of Fig. 1, namely on brain parcellation, feature extraction, and classification method. Most works do not report on the specific functional connectivity matrix estimation procedure. The most popular brain parcellations are the Harvard–Oxford (HO),[68] and the Automated Anatomical Labeling (AAL),[69,70] which are guided by anatomical criteria. However, data-driven parcellations have also been assessed in the literature applying dictionary learning, independent component analysis (ICA), clustering approaches, stochastic parcellations according to a random selection of sites (SP), and the selections of sites following biomarkers reported in the literature, such as Ref. 71. The feature extraction processes applied are widely varying among references. Some works report graph measures, other PCA and recursive feature selection (RFE), sequential feature selection (SFS), or the use of ANOVA to select the most relevant connections.[41] Unlike conventional machine learning classifier model building approaches, studies using deep learning[52] do not have a separate feature extraction process. Convolutional Neural Networks (CNN)[72–78] learn from the data to carry out a hierarchy of feature extraction processes.

The selected references of Table 1 apply the conventional machine learning validation methodology uniformly. Works report the average results of repetitions of k-fold cross-validation results where the training and any feature extraction is restricted to the training dataset avoiding the double dipping issues,[79,80] with training and testing datasets selected across original sites contributing to ABIDE listed in Table 2. The performance reports in the references of Table 1 are usually in terms of the average Accuracy. Some works report the AUC as a more robust performance measure,[35,41,51] and some report the median and 5% and 95% percentiles of the AUC.[51] Maximal accuracy and AUC results found in the literature are 77% and .75, respectively. Regarding reproducibility of the results, one key issue is the availability of the actual data used in the experiments, which is heavily dependent on rs-fMRI preprocessing, brain parcellation, and functional connectivity matrix estimation. In many instances, obtaining the same dataset is not possible, so we prefer to work on the publicly available preprocessed

Table 1. State-of-the-art results of the classification ASD versus TD on the functional connectivity data of ABIDE I dataset. 24ISC=24 inter-hemispheric selected connections, Acc = Accuracy, ADM = additional demographic features, AUC = Area under the ROC, BASC = Bootstrap Analysis of Stable Clusters, CC200 = Craddock200, DNN = Deep neural network, E3DCNN=ensemble 3D convolutional networks, ECM = eigenvalues of connectivity matrix; EGC = Ensemble of GCNs, GNB = Gaussian naive Bayes, GNG = G naive graph, HO = Harvard Oxford, kNN = k-Nearest Neighbors, LDA = linear discriminant analysis, LR = logistic regression, LSTM = Long Short-Term Memory networks, maLRR = multisite adaptation low rank decomposition, MODL = massive online dictionary learning, MSDL = multi-subject dictionary learning, MTGM = multi-task graphical model, NC = network centralities; RF = random forest, RRC = ridge regression classification, SFM = spatial feature detection method, SFS = sequential feature selection, SGC = spectral graph convolution, SP = stochastic parcellation, TT=Teilarach & Tournoux, WC=Ward's clustering.

Ref.	Year	Classifier method	Feature extraction/ selection	Brain parcellation	Best result
35	2018	GCN	PCA, MLP, ADM	HO	Acc = 70.4 AUC = 0.75
36	2017	GCN	PCA, GCN	HO	Acc = 62.9
37	2018	DN, RF,SVC		AAL, CC200, D160, EZ, HO, TT	Acc=70.0
38	2016	GCN	RFE	HO	Acc = 69.5
39	2019	SVC, GCN, EGCN	RFE	HO	Acc = 70.86
40	2015	DT, RF, SVC, LR, kNN,MLP		PCD features	Acc = 62.0, AUC=0.65
41	2019	kNN, GNB, RF, SVC, RRC, LR	ANOVA-SVC	AAL, HO, BASC, MODL,	Median AUC = 0.71
42	2017	LSTM		CC200	Acc = 68.5
43	2017	SVC, RF, GNB, MLP		CC200	Acc = 61.8
44	2017	SCVC		AAL	Acc = 62.0
45	2017	LR	Structural MRI		Acc = 62.0
46	2017	SVC, RRC		HO, Yao, CC200, kM, WC, ICA, MSDL	Acc = 69.7
47	2017	MTGM		dosenbach 40	Acc = 58.6
48	2019	E3DCNN		SP, HO, TT, EZ, CC200, CC400	Acc = 73
49	2019	LR	Subnetwork extraction		Acc = 66.74
50	2019	SVM	24ISC		Acc = 88 (intrasite)
51	2018	SGC	PCA	HO	AUC = 0.58, Acc<70
52	2020	Ensemble CNN (300)		AAL	AUC = 0.67, Acc = 67
53	2020	SVM	Graph measures	Glassner	Acc = 60
54	2017	SVM	SFM	AAL (90)	Acc = 77
55	2020	ensemble MLP	Autoencoders	AAL,CC200	Acc=75
56	2020	kNN	maLRR	AAL	Acc = 73

Table 2. Demographics distribution per site of the ABIDE I dataset. Test = the subject underwent DSM IV TR test, A = Autism, C = Control.

				Male		Female		Test		Male		Female	
Site	N	A	C	A	C	A	C	Y	N	A	C	A	C
CALTECH	38	19	19	15	15	4	4	37	1	15	14	4	4
CMU	27	14	13	11	10	3	3	5	22	3	1		1
KKI	55	22	33	18	24	4	9	39	16	9	20	3	7
LEUVEN1	29	14	15	14	15			29		14	15		
LEUVEN2	35	15	20	12	15	3	5	32	3	11	14	2	5
MAX_MUN	57	24	33	21	29	3	4	42	15	15	23	3	1
NYU	184	79	105	68	79	11	26	171	13	64	72	9	26
OHSU	28	13	15	13	15			23	5	12	11		
OLIN	36	20	16	17	14	3	2	25	11	11	9	3	2
PITT	57	30	27	26	23	4	4	45	12	18	20	4	3
SBL	30	15	15	15	15			26	4	14	12		
SDSU	36	14	22	13	16	1	6	33	3	12	15		6
STANFORD	40	20	20	16	16	4	4	36	4	13	15	4	4
TRINITY	49	24	25	24	25			44	5	21	23		
UCLA_1	82	49	33	42	29	7	4	55	27	26	23	2	4
UCLA_2	27	13	14	13	12		2	20	7	8	10		2
UM_1	110	55	55	46	38	9	17	82	28	28	31	8	15
UM_2	35	13	22	12	21	1	1	31	4	11	18	1	1
USM	101	58	43	58	43			61	40	38	23		
YALE	56	28	28	20	20	8	8	48	8	15	19	7	7
	1112	539	573	474	474	65	99	884	228	358	388	50	88

connectomes. We feel that results reported over this dataset are fairly comparable.

The selection of the experimental cohort among the ABIDE subjects varies among studies, often for unexplained reasons. For instance, the benchmarking work[41] selected 871 subjects, after visual quality inspection of the data, while our own selection includes 884 subjects. Subject selection is a source of (often positive) bias in the results; therefore, we have excluded from Table 1 references such as the recursive feature selection on 532 subjects,[81] the time series clustering approach tested on 814 subjects,[82] and others that report results on ABIDE subsamples of 209,[83] 365,[31] 182,[84] 211,[85] and 119[86] subjects.

3. Materials and Methods

In this section, we will first introduce the dataset used, then we comment on the brain parcellations and functional connectivity measures considered. The next sections describe the classifier building methods employed, and the feature extraction and feature selection methods examined. Finally, we comment on the performance measures selected to report results.

3.1. *The dataset*

The dataset analyzed in the study is extracted from the Autism Brain Imaging Data Exchange (ABIDE)[14,15] providing a publicly available dataset of rsfMRI acquisitions of subjects diagnosed with ASD and TD, data of 1112 subjects. This dataset collects data from 10 sites as detailed in Table 2. We have excluded cases with diagnosis as Asperger or PDD-NOS according to the fourth Diagnostic and Statistical Manual of Mental Disorders (DSM IV TR). We have selected and processed the 884 subjects (ASD $n = 408$, TD $n = 476$) that underwent the DSM IV TR test as shown in Table 2.

In order to have a fair comparison with the published literature, we have resorted to the preprocessed acquisitions which are available as part of the Pre-processed Connectome Project (http://preprocessed-connectomes-project.org/abide/).[87] The raw rs-fMRI data have been processed using the Configurable Pipeline for the Analysis of Connectomes (C-PAC) (http://fcp-indi.github.io/) in order to obtain the corrected and spatially normalized rs-fMRI volumes. C-PAC applies skull striping, slice timing correction, motion correction, global mean

intensity normalization, nuisance signal regression, band-pass filtering (0.01–0.1 Hz) and registration of fMRI images to standard anatomical MNI space.

3.2. *Brain parcellations*

The parcellations presented in Table 3 were applied in order to obtain the region representative time series for each of the regions of the selected parcellations. As discussed in Ref. 27, there are several approaches to the definition of the brain parcellation which may lead to significant differences in the computational experiments.[41,46] On one hand, we consider in this paper the anatomically guided parcellations such as the Tailarach and Tournoux (TT), Eickhoff-Zilles,[88] Harvard Oxford (HO),[68] and the Automated Anatomical Labeling (AAL)[69,70] defined from the brain segmentation of selected control populations. On the other hand, we consider parcellations are produced from the segmentation of the rs-fMRI time series of the brain volume using clustering techniques, such as the Dosenbach[89] and Craddok[90] parcellations.

3.3. *Connectivity matrices*

Recalling Fig. 1, the first step of our computational pipeline is the estimation of the connectivity matrices. We have considered five similarity metrics to build the connectivity matrices from the time series representatives of the brain parcellations which are available from the nilearn python package (https://nilearn.github.io/modules/generated/nilearn.connectome.ConnectivityMeasure.html).
The root of these computations is the robust estimation of the covariance matrix of the time series. We use the Ledoit–Wolf shrinkage estimator[91] following methodological recommendations in Refs. 46 and 41. We consider the following connectivity measures:

Table 3. Number of regions of the parcellations used in this study.

Atlas	#ROIs
Automated Anatomical Labeling (AAL)	116
Eickhoff-Zilles (EZ)	116
Harvard–Oxford (HO)	110
Talarach and Tournoux (TT)	110
Dosenbach 160	160
Craddock 200 (CC200)	200

- The covariance matrix computed using the Ledoit–Wolf shrinkage estimator.[91]
- The Pearson Correlation Coefficient (PCC)[92] among each pair of ROI time series, which is computed as the normalization of the covariance matrix.[93]
- The precision computed as the inverse of the covariance matrix.
- The partial correlation obtained regressing out all other connections for each pair of regions.[94]
- The tangent space representation of the matrices obtained by whitening them.[95]

Hence, for each subject in the ABIDE dataset and brain parcellation, we have five different connectivity matrices as input for the feature extraction and classifier cross-validation.

3.4. *Classifier model building methods*

We have applied classifier building methods that are available from the open- and free-source Python library scikit-learn v0.22 (https://scikit-learn.org/).[96] Specifically, we have used the following:

- Random Forest (RF)[97] is a popular ensemble method that combines by majority voting the response from a committee of decision trees[98] trained upon bootstrapped versions of the training data. Moreover, the variables used to compute each node split are randomly selected.
- K-Nearest Neighbor (KNN)[99] is the basic non-parametric classifier building approach where the test sample class is assigned by majority voting among the class labels of the K closest training samples according to the Euclidean distance.
- Gaussian Naive Bayes (GNB)[99] assumes the statistical independence of the features, so that the classifier can be built as an aggregation of one dimensional not interacting classifiers modeled by a loose mixture of Gaussians.
- Support Vector Classifier (SVC)[100,101] which look for the maximum margin hyperplane discriminating the sample into two classes solving a linear programming problem on the relevance of the samples to this class boundary. We use the linear kernel version because its response is more stable, needs less parameter tuning, and is more efficient computationally. In

some instances, we carry out a variable selection procedure based on their statistical significance in an ANOVA analysis.[102] We consider both sparse (ℓ_1) and nonsparse (ℓ_2) regularization terms. We test the two implementations available from scikit-learn, based respectively on libsvm (https://www.csie.ntu.edu.tw/cjlin/libsvm/) and liblinear (https://www.csie.ntu.edu.tw/cjlin/liblinear/) libraries.

- Logistic regression (LR)[103–105] is the classical approach that models the probability of the binary classes by a logistic linear function, enabling linear regression solvers to cope with classification problems. We apply both sparse (ℓ_1) and nonsparse (ℓ_2) regularizations.
- Least absolute shrinkage and selection operator (LASSO)[106] is a sparse (ℓ_1) regularized regression method that performs simultaneously variable selection and regularization.
- Ridge Classifier (RC)[107,108] which treats the classification problem as a straightforward regression in the $[-1, 1]$ interval with a penalty on the size of the coefficients.
- Bayesian Ridge Classifier (BRC)[109] performs the ridge regression in a Bayesian framework modeling the priors of the coefficients as a spherical Gaussian distribution whose parameters follow prior Gamma distributions. Model fit and hyperparameter estimation is carried out concurrently allowing for better adaptability to the data at hand.
- Multi-Layer Perceptron (MLP)[110] implements the classical artificial neural network architecture with sigmoid activation functions in the hidden and output layers trained by backpropagation of the error at the output layer. We apply both the adam and the L-BFGS solvers. We explore MLP architectures with 5 and 10 hidden layers in order to assess the impact of different hierarchical representational depths.

Additionally, we report a limited series experiments on the application of CNNs[110,111] over the connectivity matrices obtained with the diverse parcellations and connectivity measures. CNNs carry out induction of feature extraction filters at diverse abstraction levels; hence, no feature extraction has been included in the experiment. We have used the MATLAB implementation of CNNs, publishing the code and the data in zenodo (https://zenodo.org/record/4121200). The number of the experiments is limited by available computing resources.

3.5. *Feature extraction/selection*

We have considered several dimensional reduction procedures which are either feature extraction or feature selection techniques. Feature extraction usually involves some transformation of the feature space where the meaning of the original variables is lost unless there is some backprojection transformation. Feature selection preserves some of the original variables discarding others. As feature extraction techniques, we have applied the following ones available in the scikit-learn Python package:

- Probabilistic Principal Component Analysis (PCA)[112] is a probabilistic approach to the estimation of the eigen decomposition of the feature vectors covariance matrix instead of the conventional singular value decomposition (SVD) approach. A maximum likelihood approach is followed for this estimation under the assumption of a Gaussian multivariate model.
- Isometric Mapping (Isomap)[113] looks for a low dimensional embedding of the feature space which preserves the geodesic distances among the data samples. It involves the search for the nearest neighbors, the shortest-path search between samples, and the computation of the partial eigendecomposition.
- Local Linear Embedding (LNE) is manifold learning approach that can be assimilated to a sequence of PCA transformations, which try to benefit from and enhance the local linear structure of the data.[114]
- Multi-Dimensional Scaling (MDS)[115] looks for a dimensional reduction of the feature space such that the relative ordering of the distances between samples in the original space is preserved in the reduced dimension space.
- Factor Analysis (FA)[116,117] tries to explain the observed variables as a linear model of unseen latent variables. The conventional approach assumes a Gaussian prior for the distribution of the latent variables. Changing the prior distribution gives way to diverse algorithms.

- Connection selection by PCC: We compute the PCC[92] between each connection value in the functional connectivity matrix across subjects and the class label variable valued $\{-1, 1\}$. We compute the empirical distribution of the absolute values of these connection-label correlations. Finally, we select the connections falling in the set upper percentile, discarding low correlated connections.

All feature extraction/selection procedures are estimated on the training dataset; then the computed transformations or variable selections are applied to the test data in each cross-validation fold; hence, we have to re-estimate the feature extraction/selection operators for each cross-validation folds.

3.6. *Classification performance report*

We carry out 100 repetitions of the 10-fold cross-validation, where feature extraction/selection parameters are always computed only on the training dataset and applied in the test dataset to avoid double dipping issues.[79,80] Most of the papers in the literature report the accuracy (Acc) of the classification results averaged after the repetition of 10-fold cross-validation experiments. Accuracy is computed as $\text{Acc} = (\text{TP} + \text{TN})/N$ where N is the number of test samples, and TP and TN are the number of correct positive and negative predictions on the test set, respectively. Some papers[41] report the area under the receiver operating curve (ROC) (AUC)[118] as a more general and robust measure of the classifier performance. The accuracy is determined by the actual decision threshold applied to classify the test samples, while the ROC plots the balance of false positives (FP) versus TP across the entire range of decision threshold values. Similar to Ref. 41, we report best results of the median, 5% and 95% percentile values of the cross-validation repetitions results instead of the average value as a better description of their distribution. We plot the densities of the median AUC results across the repetitions of cross-validation experiments in order to visualize their distribution for different pipeline choices. The plots use the density() function in R that generates smooth curves that are not always bounded in the interval $[0, 1]$. To provide a quantitative ranking of the choices, we carry out one-sided nonparametric Wilcoxon's rank sum tests among all pairwise combinations of choices for each pipeline module. For each test, we consider the results of all cross-validation repetitions with all possible choice combinations for the remaining pipeline modules. We present the p-values of these tests in tables organized as follows: for each table entry, the null hypothesis is that the median AUC of the row choice is greater than that of the column choice. We specify (row $\geq$ column) at each table caption as a reminder to the reader.

4. Results and Discussion

We have aggregated the results and discussion into sections regarding the effect of the brain parcellation, the functional connectivity matrix estimation, the classifier building and the feature extraction process. Finally, the best results are compared with the state of the art of Table 1. Before proceeding with the detailed discussion of the effect of each pipeline module, we note an effect that is common to all of them: All the presented distribution density approximations of the median AUC have a big peak around the value 0.5, which is equivalent to random choice. This is a clear indication of the difficulty of the problem. Most pipeline combinations are poor performing and results are quite unstable in general, with big variations between cross-validation repetitions. We think that this is the most salient empirical demonstration of the data heterogeneity and the need for careful design of large scale data collection efforts. Data heterogeneity is due to site differences on data capture devices and procedures, as well as implementation of diagnostic criteria. Another source of heterogeneity is the openness of the diagnostic criteria leading to the inclusion of subjects with widely diverse cognitive signatures. Clustering analysis[119] of data from a mentalizing task has revealed the existence of at least six well-differentiated subgroups in a large sample of ASD and controls. Another demonstrated source of heterogeneity is sex, which has been proven to have a significant effect on the neurobiology of autism.[4]

4.1. *Effect of the brain parcellation*

Figure 2 shows the density plots corresponding to the aggregation of the median AUC results per brain parcellation used. As expected, these distributions are not Gaussian shaped, and some of them are markedly multimodal. In the case of AAL, the distribution is

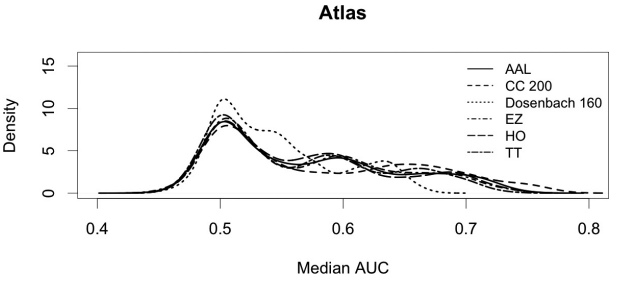

Fig. 2. Density plots of the median AUC results achieved from the different brain parcellations tested in the experiments.

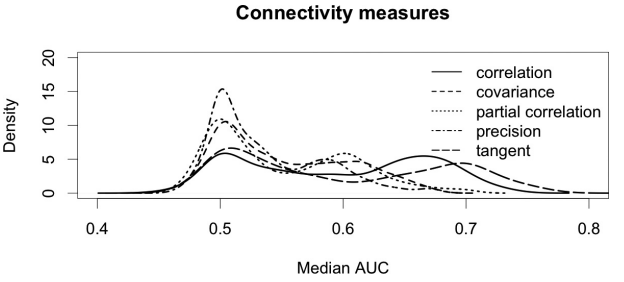

Fig. 3. Density plots of the median AUC results achieved from the different measures used to build the connectivity matrices.

pretty close to a uniform distribution. Most parcellation distributions have a big peak at the 0.5 value of the median AUC with a low tail of values above 0.7. We use one-sided Wilcoxon's rank sum test to assess quantitatively the improvement of results achieved with each parcellation. Table 4 shows the p-values of paired comparisons among the parcellations. It is quite apparent that the AAL parcellation improves over all others, followed by the CC200 parcellation. However, the maximum median AUC is greater for CC200 parcellation (0.767). The worse results are obtained from the Dosenbach parcellation, which has the greatest concentration of results around AUC = 0.5. These findings are quite interesting since the AAL parcellation has a direct anatomical interpretation, allowing results of feature selection to be reported as anatomical biomarkers naturally.

4.2. *Effect of the connectivity matrix estimation*

Figure 3 shows the distribution plots of the cross-validation repetitions median AUC aggregated by the kind of approach applied to compute the connectivity matrix per individual. It can be appreciated that these distributions are bimodal, with a high peak in 0.5. The tangent measure has the greatest second peak, around 0.7, consequently having the greatest maximum value of the median AUC. We use the one-sided Wilcoxon's rank sum test for a quantitative comparison shown in Table 5. The PCC-based connectivity and the tangent space connectivity allow to achieve much better results than the others, as reflected in the p-values reported in Table 5. Tangent space connectivity has a slightly significant improvement ($p = 0.059$) over the correlation-based connectivity, which is reflected in the best median AUC achieved (0.76).

4.3. *Effect of the classifier building method*

We have selected several classifiers to carry out the cross-validation experiments, some exploratory analysis (not reported here) of their performance results was carried out in order to select model building representatives for the comparison here. Figure 4 presents the plots of the densities of the selected classifiers, where two groups of classifiers can be easily identified visually, one group of less performing classifiers whose mass of results is centered around

Table 4. One-sided (row $\geq$ column) Wilcoxon's rank sum test p-values between median AUC results achieved from the different parcellations used to extract representative time series for the connectivity matrices.

	AAL	CC200	D160	EZ	HO	TT	Max AUC
AAL		2.06e-15	2.404e-68	1.632e-30	2.127e-31	1.340e-36	0.753
CC200	1		6.098e-30	4.169e-06	1.164e-05	2.089e-10	0.767
D160	1	1		1	1	1	0.669
EZ	1	1	4.032e-14		4.721e-01	2.372e-02	0.748
HO	1	1	8.746e-15	5.278e-01		2.602e-02	0.739
TT	1	1	1.032e-08	9.762e-01	9.739e-01		0.734

Table 5. One-sided (row ≥ column) Wilcoxon's rank sum test p-values between median AUC results achieved from the different measures used to build the connectivity matrices: cv = covariance, pc = partial correlation, p = precision, t = tangent, c = correlation, max = maximum median AUC achieved.

	cv	pc	p	t	c	Max AUC
cv	—	0.011	2.2e-16	1	1	0.67704
pc	0.988	—	8.69e-10	1	1	0.70
p	1	1	—	1	1	0.67
t	2.2e-16	2.2e-16	2.2e-16	—	0.059	0.76
c	2.2e-16	2.2e-16	2.2e-16	0.9408	—	0.74

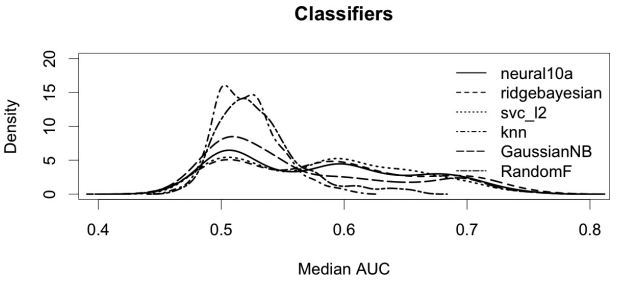

Fig. 4. Density plots of the median AUC results achieved from the different classifiers tested in the experiments. Neural10a =10 hidden layers MLP trained with adam procedure, SVC L2 = nonsparse SVC, b_ridge = Bayesian ridge regression, GNB = Gaussian naive Bayes, RF = random forest.

median AUC = 0.5, and the other that achieve better responses. The results of the one-sided Wilcoxon's rank sum test in Table 6 provide confirmation of the qualitative identification of two groups of classifiers. Top performing are sparse SVC ℓ_2, ridge classifier and the MLP with 10 hidden layers. Among them, the sparse SVC has an almost significant improvement over the other two. These results are in agreement with state-of-the-art results.

4.4. *Effect of the feature extraction/selection*

Feature extraction consists of a data space transformation where the new variables lose the meaning of the original space, i.e. the anatomical localization of the effects. Feature selection (such as the PCC-based connection selection) preserves the meaning of the original variables, because the selected variables are not transformed. Feature selection is the preferred approach for the medical researchers because they can explain and compare the found biomarkers in the framework of the medical literature.

We have explored the effect of the feature extraction and feature selection procedures described above. The Isometric Map has been discarded as its results were far worse than any other feature extraction method. For the other methods, we have made an exploration of the performance achieved when varying the number of features retained, finding slight significant improvements leading to specific selections for each approach that are compared in Table 7 using the one-sided Wilcoxon's rank sum test as the density functions plotted in Fig. 5 are far from Gaussian in most cases (exception made of MDS

Table 6. One-sided (row ≥ column) Wilcoxon's rank sum test p-values between median AUC results achieved by the diverse kind of classifiers experimented with. Neural10a =10 hidden layers MLP trained with adam procedure, SVC ℓ_2 = nonsparse SVC, b_ridge = Bayesian ridge regression, GNB = Gaussian naive Bayes, RF = random forest.

	Neural10a	SVC ℓ_2	b_ridge	kNN	RF	GNB	Max AUC
Neural10a		0.963	0.817	2.547e-47	1.481e-06	2.143e-30	0.75
SVC ℓ_2	0.036		0.192	5.789e-62	0.841e-11	1.264e-43	0.761
b_ridge	0.182	0.807		1.438e-61	1.407e-09	1.594e-42	0.761
kNN	1	1	1		1	1	0.61
RF	1	1	1	7.701e-20		2.036e-08	0.756
GNB	1	1	1	6.782e-07	1		0.66

Table 7. One-sided (row≥column) Wilcoxon's rank sum test p-values between median AUC results achieved from the different best versions of the feature extraction algorithms. PCA2 = PCA retaining only half of the features, MDS2000, fa2000 = MDS, FA retaining 2000 features, LNE3 = LNE retaining one-third of the transformed features.

	PCA2	MDS2000	LNE3	p90	fa2000	Max AUC
PCA2		1.248e-116	8.525e-55	0.0029	0.078	0.75
MDS2000	1		1	1	1	0.55
LNE3	1	4.878e-66		1	1	0.62
fa200	0.921	1.577e-122	6.900e-59	0.0833		0.76

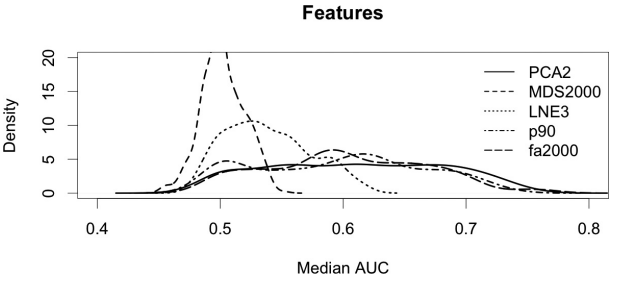

Fig. 5. Density plots of the median AUC results achieved from the different feature extraction approaches. PCA2 = PCA retaining only half of the features, MDS2000, fa2000 = MDS, FA retaining 2000 features, LNE3 = LNE retaining one-third of the features, p90 = PCC selection 90% percentile.

which appears to be almost Gaussian). PCA2 has an almost uniform distribution in the interval $[0.5, 0.7]$ of AUC values. Regarding Table 7, PCA2 improves significantly over the other procedures, although the significance of the improvement over FA is short.

4.5. *Best results*

One of the conclusions that can be extracted from the previous sections is that finding good performing pipelines requires exploration of many computational choices where most of them will not achieve good results. Here we have selected the best performing pipelines found by exhaustive search over our experimental results, some of them improving over most of the results reported in the state of the art of Table 1. Table 8 gives the best median AUC scores found, together with the 5% and 95% percentiles of the cross-validation results in our experiments for comparison with the most comprehensive exploration of results to date,[41] where the best reported results are 0.66, 0.711, and 0.756 for the 5%, 50%, and 95% percentiles of median AUC distribution across all repetitions of the cross-validation experiments. This comparison shows the impact of feature extraction approaches to enhance classification results. Because many of the results in the literature are reported in terms of accuracy, we include here the corresponding accuracy tables. Table 9 gives the instances with the best accuracy results of our experiments, comparing favorably with the results gathered in Table 1. Some recent results[48,52] have been achieved using brain parcellations that are not accessible; hence, direct comparison against them is not possible for us.

Table 8. Best median AUC scores found in cross-validation repetitions, with corresponding settings (parcellation, feature extraction, classifier, and connectivity measure) that achieved it.

Settings				AUC percentiles		
Parcel.	Feat. extr.	Classifier	Conn. meas.	5%	Median	95%
cc200	fa	Logistic ℓ_1	Tangent	0.733	0.765	0.803
cc200	pca	Logistic ℓ_2	Tangent	0.739	0.765	0.801
cc200	fa	SVC ℓ_1	Tangent	0.739	0.764	0.803
cc200	fa1000	Logistic ℓ_1	Tangent	0.733	0.765	0.803
cc200	pca	SVC ℓ_1	Tangent	0.735	0.767	0.805
cc200	fa2000	Logistic ℓ_1	Tangent	0.733	0.765	0.803
cc200	fa2000	SVC ℓ_1	Tangent	0.739	0.764	0.803
Best reported to date[41]				0.66	0.711	0.756

Table 9. Best Acc scores found in cross-validation, with corresponding settings (parcellation, feature selection, classifier, and connectivity measure) that achieved it.

Settings				Acc percentiles		
Parcel.	Feat. ext.	Classifier	Conn. meas.	5%	Median	95%
cc200	fa	Logistic ℓ_1	Tangent	66.9	69.9	72.3
cc200	fa	SVC ℓ_1	Tangent	65.7	69.9	72.8
cc200	fa1000	Logistic ℓ_1	Tangent	66.9	69.9	72.3
cc200	fa1000	SVC ℓ_1	Tangent	65.7	69.9	72.8
cc200	fa2000	Logistic ℓ_1	Tangent	66.9	69.9	72.3
cc200	fa2000	SVC ℓ_1	Tangent	65.7	69.9	72.8
cc200	pca	Logistic ℓ_2	Tangent	67.3	70.5	72.1
cc200	pca	SVC ℓ_1	Tangent	65.9	70.1	72.1
Best reported to date[52]				—	—	77

4.6. *Deep learning results*

As suggested by a reviewer, given the deep learning techniques extreme success in computer vision applications it seems mandatory to test the most successful approach, namely CNNs,[111] on the combinations of brain parcellations and connectivity measures. Table 10 gives the best results achieved by each CNN topology after 10 repetitions of 10-fold cross-validation with each setting. One of the difficulties of the application of deep learning approaches is the finding the optimal topology of the network, which can be very tricky. For our experiments, we have followed the strategy of increasing the depth of the network and changing the size of the filters following a pyramid structure, broader filters at the bottom layers and smaller ones at the top layers. We have also tested ensembles of CNNs, though not very big for lack of computational resources. We found that adding layers provided some improvements, reaching kind of overfitting situation when we applied a five layers topology. Using an ensemble of 11 CNNs provided a small improvement, lack of computational resources and time prevented experimentation with larger ensembles. The use of a pyramidal strategy in the definition of the filters did not provide significant improvements. Comparison with results in Table 9 shows that the examined CNN topologies do not provide any improvement over conventional feature selection and classification methods. This observation does not preclude the existence some specific CNN topology that improves over conventional approaches on this dataset; however, the ingenuity and computational resources to find it is beyond our current capabilities. Another observation from the

Table 10. Results of explored CNN topologies. We report median accuracy, brain parcellation (parcel.) and connectivity measure (conn. meas.) with best results. n@m denotes a convolution layer with n filters of size m. full denotes full connectivity layer. Output is always a softmax of two units. E denotes the number of CNNs in an ensemble.

Settings			Acc percentiles		
CNN topology	Parcel.	Conn. meas.	5%	Median	95%
20@5, full	TT	Correlation	48.30	54.55	62.50
20@5,20@5,full	EZ	Correlation	51,14	60.23	67,61
20@5,20@5,20@5,full	HO	Correlation	54.55	63.64	72.73
20@5,20@5,20@5,20@5,full	AAL	Correlation	55.93	64.04	72.16
20@5,20@5,20@5,20@5,20@3,full	AAL	Correlation	53.93	61.36	68.18
20@5,20@5,20@5,20@5,full, $E = 11$	AAL	Correlation	57.95	66.29	74.58
20@9,20@7,20@5,full	HO	Correlation	55.68	64.77	72.73
20@11,20@59,20@7,20@5,full	AAL	Correlation	55.68	64.77	71.19
20@11,20@59,20@7,20@5,full, $E = 11$	AAL	Correlation	58.43	66.29	74.01

results in Table 10 is that the 5–95% percentile interval is much larger than in Table 9, likely due to the stochastic learning characteristics.

Besides our experience reported above, the literature has several examples of attempts to apply deep learning to ASD prediction on the ABIDE dataset. In order to discuss comparative results, we face the issue of the diversity of the underlying pipeline selections and deep learning design peculiarities. We have not found in the literature an exhaustive exploration of deep learning approaches over the brain parcellations and connectivity measures comparable to ours. However, we have shown that they have quite significant effect on the predictive performance. For instance, experiments involving a large ensamble of 300 CNNs[52] was carried out on a very specific irreproducible brain parcellation and connectivity matrix constructions, with ad hoc simplified CNN topologies found after a long trial and error process with little success (Acc = 67). A greater computational (irreproducible) tour de force of training an ensemble of 3D CNN applied on the normalized rs-fMRI data[48] provided a small improvement (Acc = 72). The use of recurrent networks such as the LSTM[42] did not achieve better results than the conventional approaches (Acc = 68%). Graph convolutional networks (GCN)[35,38] did not provide significant improvement over conventional results in Table 9 achieving the best result Acc = 70 adding ancillary information to the connectivity data. Even using ensembles of GCN[39] did not add significant benefits. Using autoencoders for feature extraction combined with conventional MLP classifier provided one of the best reported results.[55]

5. Conclusions

The predictive approach to the analysis of brain connectivity from rs-fMRI data is gaining importance in recent studies. In this approach, brain connectivity biomarkers are confirmed by the predictive performance in the classification between target populations. Up to this date, there is no comprehensive study of the impact of the choices that can be made while building the machine learning pipelines; hence, we have carried out a comprehensive assessment on the ABIDE I dataset, finding that some feature extraction procedures provide a boost on the performance of the classifiers across several connectivity matrix building approaches, namely the classical principal component analysis (PCA) and factor analysis (FA). A key issue is a reproducibility of the results that depends on the availability of the data and the precise computational resources to other researchers in the community. For this reason, we emphasize the public availability of the data and programming resources used for this study via github (https://github.com/mmscnet/Impact-feature-extraction-in-Autism) and zenodo (https://zenodo.org/record/4121200) repositories.

Future work will be addressed to the extension of the computational experiments to the full extent of the ABIDE II dataset. Other connectomics datasets collecting subjects and controls from connectivity analysis regarding other diseases will also be considered. Additionally, innovative machine learning approaches[120–123] will be explored. The instability of the validation of the predictive approaches in many instances of neuroscience datasets is an issue of methodological concern. In the case of the ABIDE dataset, the sources of this instability are the heterogeneity of the subjects, diagnostic criteria implementation, and the data capture differences among sites. Future work will explore the relevance of novel validation approaches such as the works underlying the statistical agnostic mapping[124] to provide more robust performance predictions leading to better grounded biomarker identification.

Acknowledgments

This work has been partially supported by the FEDER funds through MINECO project TIN2017-85827-P. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 777720.

References

1. C. Lord, T. S. Brugha, T. Charman, J. Cusack, G. Dumas, T. Frazier, E. J. H. Jones, R. M. Jones, A. Pickles, M. W. State, J. L. Taylor and J. Veenstra-VanderWeele, Autism spectrum disorder, *Nat. Rev. Dis. Prime.* **6**(1) (2020) 5.
2. S. Bhat, U. R. Acharya, H. Adeli, G. M. Bairy and A. Adeli, Autism: Cause factors, early diagnosis and therapies, *Rev. Neurosci.* **25**(6) (2014) 841–850.
3. D. L. Christensen, K. Van Naarden Braun, J. Baio, *et al.*, Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years

— Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. MMWR Surveill Summ 65 (No. SS-13);2018:1–23.

4. M.-C. Lai, M. V. Lombardo, J. Suckling, A. N. V. Ruigrok, B. Chakrabarti, C. Ecker, S. C. L. Deoni, M. C. Craig, D. G. M. Murphy, E. T. Bullmore, M. A. Consortium and S. Baron-Cohen, Biological sex affects the neurobiology of autism, *Brain* **136** (2013) 2799–2815.
5. M.-C. Lai, M. V. Lombardo, B. Auyeung, B. Chakrabarti and S. Baron-Cohen, Sex/gender differences and autism: Setting the scene for future research, *J. Am. Acad. Child Adolesc. Psychiatry* **54** (2015) 11–24.
6. M.-C. Lai, M. V. Lombardo and S. Baron-Cohen, Autism, *The Lancet* **383**(9920) (2013) 896–910.
7. L. Mottron and D. Bzdok, Autism spectrum heterogeneity: Fact or artifact? *Mol. Psychiatry* **7** (2020) 3178–3185.
8. C. Lord, E. Petkova, V. Hus, W. Gan, F. Lu, D. M. Martin, O. Ousley, L. Guy, R. Bernier, J. Gerdts, M. Algermissen, A. Whitaker, J. S. Sutcliffe, Z. Warren, A. Klin, C. Saulnier, E. Hanson, R. Hundley, J. Piggot, E. Fombonne, M. Steiman, J. Miles, S. M. Kanne, R. P. Goin-Kochel, S. U. Peters, E. H. Cook, S. Guter, J. Tjernagel, L. A. Green-Snyder, S. Bishop, A. Esler, K. Gotham, R. Luyster, F. Miller, J. Olson, J. Richler and S. Risi, A multisite study of the clinical diagnosis of different autism spectrum disorders, *Arch. Gen. Psychiatry* **69** (2012) 306–313.
9. S. Bhat, U. R. Acharya, H. Adeli, G. M. Bairy and A. Adeli, Automated diagnosis of autism: In search of a mathematical marker, *Rev. Neurosci.* **25**(6) (2014) 851–861.
10. T. Chen, Y. Chen, M. Yuan, M. Gerstein, T. Li, H. Liang, T. Froehlich and L. Lu, The development of a practical artificial intelligence tool for diagnosing and evaluating autism spectrum disorder: Multicenter study, *JMIR Med. Inform.* **8** (2020) e15767.
11. H. S. Nogay and H. Adeli, Machine learning (ML) for the diagnosis of autism spectrum disorder (ASD) using brain imaging, *Reviews in the Neurosciences* **31**(8) (2020) 825–841
12. F. Segovia, R. Holt, M. Spencer, J. M. Górriz, J. Ramírez, C. G. Puntonet, C. Phillips, L. Chura, S. Baron-Cohen and J. Suckling, Identifying endophenotypes of autism: A multivariate approach, *Front. Comput. Neurosci.* **8** (2014) 60.
13. L. Q. Uddin, D. R. Dajani, W. Voorhies, H. Bednarz and R. K. Kana, Progress and roadblocks in the search for brain-based biomarkers of autism and attention-deficit/hyperactivity disorder, *Translat. Psychiatry* **7**(8) (2017) e1218–e1218.
14. A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, M. Assaf, S. Y. Bookheimer, M. Dapretto, B. Deen, S. Delmonte, I. Dinstein, B. Ertl-Wagner, D. A. Fair, L. Gallagher, D. P. Kennedy, C. L. Keown, C. Keysers, J. E. Lainhart, C. Lord, B. Luna, V. Menon, N. J. Minshew, C. S. Monk, S. Mueller, R.-A. Müller, M. B. Nebel, J. T. Nigg, K. O'Hearn, K. A. Pelphrey, S. J. Peltier, J. D. Rudie, S. Sunaert, M. Thioux, J. M. Tyszka, L. Q. Uddin, J. S. Verhoeven, N. Wenderoth, J. L. Wiggins, S. H. Mostofsky and M. P. Milham, The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism, *Mol. Psychiatry* **19** (2013) 659–667.
15. A. Di Martino, D. O'Connor, B. Chen, K. Alaerts, J. S. Anderson, M. Assaf, J. H. Balsters, L. Baxter, A. Beggiato, S. Bernaerts, L. M. E. Blanken, S. Y. Bookheimer, B. B. Braden, L. Byrge, F. X. Castellanos, M. Dapretto, R. Delorme, D. A. Fair, I. Fishman, J. Fitzgerald, L. Gallagher, R. J. J. Keehn, D. P. Kennedy, J. E. Lainhart, B. Luna, S. H. Mostofsky, R.-A. Müller, M. B. Nebel, J. T. Nigg, K. O'Hearn, M. Solomon, R. Toro, C. J. Vaidya, N. Wenderoth, T. White, R. C. Craddock, C. Lord, B. Leventhal and M. P. Milham, Enhancing studies of the connectome in autism using the autism brain imaging data exchange II, *Sci. Data* **4** (2017) 170010.
16. B. G. Becker, T. Klein and C. Wachinger, Gaussian process uncertainty in age estimation as a measure of brain abnormality, *NeuroImage* **175** (2018) 246–258.
17. T. Wolfers, D. L. Floris, R. Dinga, D. van Rooij, C. Isakoglou, S. M. Kia, M. Zabihi, A. Llera, R. Chowdanayaka, V. J. Kumar, H. Peng, C. Laidi, D. Batalle, R. Dimitrova, T. Charman, E. Loth, M.-C. Lai, E. Jones, S. Baumeister, C. Moessnang, T. Banaschewski, C. Ecker, G. Dumas, J. O'Muircheartaigh, D. Murphy, J. K. Buitelaar, A. F. Marquand and C. F. Beckmann, From pattern classification to stratification: Towards conceptualizing the heterogeneity of autism spectrum disorder, *Neurosci. Biobehav. Rev.* **104** (2019) 240–254.
18. M. Ahmadlou, H. Adeli and A. Adeli, Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder, *Phys. A Stat. Mech. Appl.* **391**(20) (2012) 4720–4726.
19. M. Ahmadlou and H. Adeli, Complexity of weighted graph: A new technique to investigate structural complexity of brain activities with applications to aging and autism, *Neurosci. Lett.* **650** (2017) 103–108.
20. M. Ahmadlou, H. Adeli and A. Adeli, Fuzzy synchronization likelihood-wavelet methodology for diagnosis of autism spectrum disorder, *J. Neurosci. Meth.* **211**(2) (2012) 203–209.
21. E. Puerto, J. Aguilar, C. López and D. Chávez, Using multilayer fuzzy cognitive maps to diagnose

autism spectrum disorder, *Appl. Soft Comput.* **75** (2019) 58–71.
22. A. d'Anjou, M. Graña and F. Moutet, Brief survey about the search for biomarkers and computer aided diagnosis of autism spectrum disorder, Tech. Report, http://doi.org/10.5281/zenodo.1408202, Zenodo (2018).
23. A. M. Pagnozzi, E. Conti, S. Calderoni, J. Fripp and S. E. Rose, A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective, *Int. J. Develop. Neurosci.* **71** (2018) 68–82.
24. S. H. Ameis and M. Catani, Altered white matter connectivity as a neural substrate for social impairment in autism spectrum disorder, *Cortex* **62** (2015) 158–181. Special issue: The clinical anatomy of the limbic lobe and connected structures.
25. X. Di, A. Azeez, X. Li, E. Haque and B. B. Biswal, Disrupted focal white matter integrity in autism spectrum disorder: A voxel-based meta-analysis of diffusion tensor imaging studies, *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **82** (2018) 242–248.
26. J. H. Balsters, D. Mantini and N. Wenderoth, Connectivity-based parcellation reveals distinct cortico-striatal connectivity fingerprints in autism spectrum disorder, *NeuroImage* **170** (2018) 412–423, Segmenting the Brain.
27. S. Arslan, S. I. Ktena, A. Makropoulos, E. C. Robinson, D. Rueckert and S. Parisot, Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex, *NeuroImage* **170** (2018) 5–30, Segmenting the Brain.
28. N. D. Woodward and C. J. Cascio, Resting-state functional connectivity in psychiatric disorders, *JAMA Psychiatry* **72** (2015) 743–744.
29. M. Graña, L. Ozaeta and D. Chyzhyk, Resting state effective connectivity allows auditory hallucination discrimination, *Int. J. Neural Syst.* **27**(5) (2017) 1750019, PMID: 28274168.
30. P. Aggarwal and A. Gupta, Multivariate graph learning for detecting aberrant connectivity of dynamic brain networks in autism, *Med. Image Anal.* **56** (2019) 11–25.
31. Z. Fu, Y. Tu, X. Di, Y. Du, J. Sui, B. B. Biswal, Z. Zhang, N. de Lacy and V. Calhoun, Transient increased thalamic-sensory connectivity and decreased whole-brain dynamism in autism, *NeuroImage* **190** (2019) 191–204, Mapping diseased brains.
32. A. Andriamananjara, R. Muntari and A. Crimi, Overlaps in brain dynamic functional connectivity between schizophrenia and autism spectrum disorder, *Sci. Afr.* **2** (2019) e00019.
33. D. Mastrovito, C. Hanson and S. J. Hanson, Differences in atypical resting-state effective connectivity distinguish autism from schizophrenia, *NeuroImage Clin.* **18** (2018) 367–376.
34. M. A. Reiter, L. E. Mash, A. C. Linke, C. H. Fong, I. Fishman and R.-A. Müller, Distinct patterns of atypical functional connectivity in lower-functioning autism, *Biol. Psychiatry Cogn. Neurosci. Neuroimag.* **4**(3) (2019) 251–259.
35. S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker and D. Rueckert, Disease prediction using graph convolutional networks: Application to autism spectrum disorder and Alzheimer's disease, *Med. Image Anal.* **48** (2018) 117–130.
36. S. I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. Lee, B. Glocker and D. Rueckert, Distance metric learning using graph convolutional networks: Application to functional brain networks (2017), arXiv:1703.02161v2.
37. A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz and F. Meneguzzi, Identification of autism spectrum disorder using deep learning and the ABIDE dataset, *NeuroImage Clin.* **17** (2018) 16–23.
38. S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. G. Moreno, B. Glocker and D. Rueckert, Spectral graph convolutions for population-based disease prediction (2017), arXiv:1703.03020.
39. R. Anirudh and J. J. Thiagarajan, Bootstrapping graph convolutional neural networks for autism spectrum disorder classification, in *ICASSP 2019 — 2019 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)* (Brighton, United Kingdom, 2019), pp. 3197–3201.
40. M. N. Parikh, H. Li and L. He, Enhancing diagnosis of autism with optimized machine learning models and personal characteristic data, *Front. Comput. Neurosci.* **13** (2019) 9.
41. K. Dadi, M. Rahim, A. Abraham, D. Chyzhyk, M. Milham, B. Thirion and G. Varoquaux, Benchmarking functional connectome-based predictive models for resting-state fMRI, *NeuroImage* **192** (2019) 115–134.
42. N. C. Dvornek, P. Ventola, K. A. Pelphrey and J. S. Duncan, Identifying autism from resting-state FMRI using long short-term memory networks, in *Machine Learning in Medical Imaging*, eds. Q. Wang, Y. Shi, H.-I. Suk and K. Suzuki (Springer International Publishing, Cham, 2017), pp. 362–370.
43. R. Tejwani, A. Liska, H. You, J. Reinen and P. Das, Autism classification using brain functional connectivity dynamics and machine learning (2017), arXiv:1712.08041.
44. J. Leonard, J. Flournoy, C. P. L. de los Angeles and K. Whitaker, How much motion is too much motion? determining motion thresholds by sample size for reproducibility in developmental resting-state MRI, *Res. Ideas Outcom.* **3** (2017) e12569.
45. S. Rane, E. Jolly, A. Park, H. Jang and C. Craddock, Developing predictive imaging biomarkers using whole-brain classifiers: Application to the

ABIDE I dataset, *Res. Ideas Outcom.* **3** (2017) e12733.

46. A. Abraham, M. P. Milham, A. D. Martino, R. C. Craddock, D. Samaras, B. Thirion and G. Varoquaux, Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example, *NeuroImage* **147** (2017) 736–745.
47. C. Singh, B. Wang and Y. Qi, A constrained, weighted-L1 minimization approach for joint discovery of heterogeneous neural connectivity graphs (2017).
48. M. Khosla, K. Jamison, A. Kuceyeski and M. R. Sabuncu, Ensemble learning with 3d convolutional neural networks for functional connectome-based prediction, *NeuroImage* **199** (2019) 651–662.
49. C. J. Brown, S. P. Miller, B. G. Booth, J. G. Zwicker, R. E. Grunau, A. R. Synnes, V. Chau and G. Hamarneh, Predictive connectome subnetwork extraction with anatomical and connectivity priors, *Comput. Med. Imag. Graph.* **71** (2019) 67–78.
50. Q. Li, B. Becker, X. Jiang, Z. Zhao, Q. Zhang, S. Yao and K. M. Kendrick, Decreased interhemispheric functional connectivity rather than corpus callosum volume as a potential biomarker for autism spectrum disorder, *Cortex* **119** (2019) 258–266.
51. S. I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. Lee, B. Glocker and D. Rueckert, Metric learning with spectral graph convolutions on brain connectivity networks, *NeuroImage* **169** (2018) 431–442.
52. M. Leming, J. M. Górriz and J. Suckling, Ensemble deep learning on large, mixed-site FMRI datasets in autism and other tasks, *Int. J. Neural Syst.* **30**(7) (2020) 2050012.
53. A. Brahim and N. Farrugia, Graph Fourier transform of FMRI temporal signals based on an averaged structural connectome for the classification of neuroimaging, *Artif. Intell. Med.* **106** (2020) 101870.
54. V. Subbaraju, M. B. Suresh, S. Sundaram and S. Narasimhan, Identifying differences in brain activities and an accurate detection of autism spectrum disorder using resting state functional-magnetic resonance imaging: A spatial filtering approach, *Med. Image Anal.* **35** (2017) 375–389.
55. M. Rakić, M. Cabezas, K. Kushibar, A. Oliver and X. Lladó, Improving the detection of autism spectrum disorder by combining structural and functional MRI information, *NeuroImage Clin.* **25** (2020) 102181.
56. M. Wang, D. Zhang, J. Huang, P. Yap, D. Shen and M. Liu, Identifying autism spectrum disorder with multi-site FMRI via low-rank domain adaptation, *IEEE Trans. Med. Imag.* **39**(3) (2020) 644–655.
57. J. Gorriz, J. Ramirez, A. Ortiz, F. J. Martinez-Murcia, F. Segovia, J. Suckling, M. Leming, Y.-D. Zhang, J. Alvarez-Sanchez, G. Bologna, P. Bonomini, F. Casado, D. Charte, F. Charte, R. Contreras, A. Cuesta-Infante, R. Duro, A. Fernandez-Caballero, E. Fernandez-Jover, P. Gomez-Vilda, M. Graña, F. Herrera, R. Iglesias, A. Lekova, J. de Lope, E. Lopez-Rubio, R. Martinez-Tomas, M. A. Molina-Cabello, A. S. Montemayor, P. Novais, D. Palacios-Alonso, J. J. Pantrigo, B. R. Payne, F. de la Paz Lopez, M. A. Pinninghoff, M. Rincon, J. Santos, K. Thurnhofer-Hemsi, A. Tsanas, R. Varela and J. M. Ferrandez, Artificial intelligence within the interplay between natural and artificial computation: Advances in data science, trends and applications, *Neurocomputing* **410** (2020) 237–270.
58. S. J. Moon, J. Hwang, R. Kana, J. Torous and J. W. Kim, Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: Systematic review and meta-analysis of brain magnetic resonance imaging studies, *JMIR Ment. Health* **6** (2019) e14108.
59. J. M. Górriz, J. Ramírez, F. Segovia, F. J. Martínez, M.-C. Lai, M. V. Lombardo, S. Baron-Cohen, MRC AIMS Consortium and J. Suckling, A machine learning approach to reveal the NeuroPhenotypes of autisms, *Int. J. Neural Syst.* **29** (2019) 1850058.
60. V. Subbaraju, S. Sundaram, S. Narasimhan and M. B. Suresh, Accurate detection of autism spectrum disorder from structural MRI using extended metacognitive radial basis function network, *Expert Syst. Appl.* **42**(22) (2015) 8775–8790.
61. H. Shahamat and M. Saniee Abadeh, Brain MRI analysis using a deep learning based evolutionary approach, *Neural Netw.* **126** (2020) 218–234.
62. F. J. Martinez-Murcia, M.-C. Lai, J. M. Gorriz, J. Ramirez, A. M. H. Young, S. C. L. Deoni, C. Ecker, M. V. Lombardo, MRC AIMS Consortium, S. Baron-Cohen, D. G. M. Murphy, E. T. Bullmore and J. Suckling, On the brain structure heterogeneity of autism: Parsing out acquisition site effects with significance-weighted principal component analysis, *Hum. Brain Mapp.* **38**(3) (2017) 1208–1223.
63. N. Yahata, J. Morimoto, R. Hashimoto, G. Lisi, K. Shibata, Y. Kawakubo, H. Kuwabara, M. Kuroda, T. Yamada, F. Megumi, H. Imamizu, J. Náñez Sr, H. Takahashi, Y. Okamoto, K. Kasai, N. Kato, Y. Sasaki, T. Watanabe and M. Kawato, A small number of abnormal brain connections predicts adult autism spectrum disorder, *Nat. Commun.* **7**(1) (2016) 11254.
64. L. Q. Uddin, K. Supekar, C. J. Lynch, A. Khouzam, J. Phillips, C. Feinstein, S. Ryali and V. Menon, Salience network–based classification and prediction of symptom severity in children with autism, *JAMA Psychiatry* **70** (2013) 869–879.
65. R. X. Smith, K. Jann, M. Dapretto and D. J. J. Wang, Imbalance of functional connectivity

and temporal entropy in resting-state networks in autism spectrum disorder: A machine learning approach, *Front. Neurosci.* **12** (2018) 869.

66. X. Guo, K. C. Dominick, A. A. Minai, H. Li, C. A. Erickson and L. J. Lu, Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method, *Front. Neurosci.* **11** (2017) 460.
67. X. Li, Y. Gu, N. Dvornek, L. H. Staib, P. Ventola and J. S. Duncan, Multi-site FMRI analysis using privacy-preserving federated learning and domain adaptation: Abide results, *Med. Image Anal.* **65** (2020) 101765.
68. R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert and R. J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, *NeuroImage* **31**(3) (2006) 968–980.
69. E. T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng and M. Joliot, Automated anatomical labelling atlas 3, *NeuroImage* **206** (2020) 116189.
70. N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer and M. Joliot, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, *NeuroImage* **15**(1) (2002) 273–289.
71. J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar and S. E. Petersen, Functional network organization of the human brain, *Neuron* **72**(4) (2011) 665–678.
72. M. A. Molina-Cabello, R. M. Luque-Baena, E. Lopez-Rubio and K. Thurnhofer-Hemsi, Vehicle type detection by ensembles of convolutional neural networks operating on super resolved images, *Integr. Comput. Aid. Eng.* **25** (2018) 321–333.
73. P. Wang and X. Bai, Regional parallel structure based CNN for thermal infrared face identification, *Integr. Comput. Aid. Eng.* **25** (2018) 247–260.
74. J. F. Torres, A. Galicia, A. Troncoso and F. Martinez-Alvarez, A scalable approach based on deep learning for big data time series forecasting, *Integr. Comput. Aid. Eng.* **25** (2018) 335–348.
75. F. J. Vera-Olmos, E. Pardo, H. Melero and N. Malpica, Deepeye: Deep convolutional network for pupil detection in real environments, *Integr. Comput. Aid. Eng.* **26** (2019) 85–95.
76. T. Yang, C. Cappelle, Y. Ruichek and M. El Bagdouri, Multi-object tracking with discriminant correlation filter based deep learning tracker, *Integr. Comput. Aid. Eng.* **26** (2019) 273–284.
77. K. Maeda, S. Takahashi, T. Ogawa and M. Haseyama, Convolutional sparse coding-based deep random vector functional link network for distress classification of road structures, *Comput. Aid. Civil Infrastruct. Eng.* **34**(8) (2019) 654–676.
78. S. Bang, S. Park, H. Kim and H. Kim, Encoder–decoder network for pixel-level road crack detection in black-box images, *Comput. Aid. Civil Infrastruct. Eng.* **34**(8) (2019) 713–727.
79. K. S. Button, Double-dipping revisited, *Nat. Neurosci.* **22**(5) (2019) 688–690.
80. N. Kriegeskorte, W. K. Simmons, P. S. F. Bellgowan and C. I. Baker, Circular analysis in systems neuroscience: The dangers of double dipping, *Nat. Neurosci.* **12**(5) (2009) 535–540.
81. C. Wang, Z. Xiao and J. Wu, Functional connectivity-based classification of autism and control using SVM-RFECV on RS-FMRI data, *Phys. Medica* **65** (2019) 99–105.
82. M. C. Vidal, J. R. Sato, J. B. Balardin, D. Y. Takahashi and A. Fujita, ANOCVA in r: A software to compare clusters between groups and its application to the study of autism spectrum disorder, *Front. Neurosci.* **11** (2017) 16.
83. J. C. Beer, H. J. Aizenstein, S. J. Anderson and R. T. Krafty, Incorporating prior information with fused sparse group lasso: Application to prediction of clinical measures from neuroimages, *Biometrics* **75**(4) (2019) 1299–1309.
84. Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang and J. Liu, Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier, *Neurocomputing* **324** (2019) 63–68, Deep Learning for Biological/Clinical Data.
85. S.-J. Hong, R. Vos de Wael, R. A. I. Bethlehem, S. Lariviere, C. Paquola, S. L. Valk, M. P. Milham, A. Di Martino, D. S. Margulies, J. Smallwood and B. C. Bernhardt, Atypical functional connectome hierarchy in autism, *Nat. Commun.* **10**(1) (2019) 1022.
86. K. Martínez, M. Martínez-García, L. Marcos-Vidal, J. Janssen, F. X. Castellanos, C. Pretus, Ó. Villarroya, L. Pina-Camacho, C. M. Díaz-Caneja, M. Parellada, C. Arango, M. Desco, J. Sepulcre and S. Carmona, Sensory-to-cognitive systems integration is associated with clinical severity in autism spectrum disorder, *J. Am. Acad. Child Adoles. Psychiatry* **59**(3) (2019) 422–433.
87. C. Cameron, B. Yassine, C. Carlton, C. Francois, E. Alan, J. András, K. Budhachandra, L. John, L. Qingyang, M. Michael, Y. Chaogan and B. Pierre, The neuro bureau preprocessing initiative: Open sharing of preprocessed neuroimaging data and derivatives, *Front. Neuroinform.* **7** (2013).
88. S. B. Eickhoff, T. Paus, S. Caspers, M.-H. Grosbras, A. C. Evans, K. Zilles and K. Amunts, Assignment of functional activations to probabilistic cytoarchitectonic areas revisited, *NeuroImage* **36**(3) (2007) 511–521.

89. 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-Schlaggar, K. A. Barnes, J. W. Dubis, E. Feczko, R. S. Coalson, J. R. Pruett, D. M. Barch, S. E. Petersen and B. L. Schlaggar, Prediction of individual brain maturity using FMRI, *Science* **329**(5997) (2010) 1358–1361.
90. R. C. Craddock, G. James, P. E. Holtzheimer III, X. P. Hu and H. S. Mayberg, A whole brain FMRI atlas generated via spatially constrained spectral clustering, *Hum. Brain Mapp.* **33**(8) (2012) 1914–1928.
91. O. Ledoit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivar. Anal.* **88**(2) (2004) 365–411.
92. K. Pearson, Notes on regression and inheritance in the case of two parents, *Proc. Royal Soc. London* **58** (1895) 240–242.
93. A. Zalesky, A. Fornito and E. Bullmore, On the use of correlation as a measure of network connectivity, *NeuroImage* **60**(4) (2012) 2096–2106.
94. S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey and M. W. Woolrich, Network modelling methods for FMRI, *NeuroImage* **54**(2) (2011) 875–891.
95. G. Varoquaux, F. Baronnet, A. Kleinschmidt, P. Fillard and B. Thirion, Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, T. Jiang, N. Navab, J. P. W. Pluim and M. A. Viergever (eds.), MICCAI 2010. Lecture Notes in Computer Science, Vol. 6361. Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-15705-9_25.
96. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, Scikit-learn: Machine learning in python *JMLR* **12** (2011) 2825–2830.
97. L. Breiman, Random forests, *Mach. Learn.* **45(1)** (2001) 5–32.
98. L. Breiman, J. Friedman, R. Olshen and C. Stone, *Classification and Regression Trees* (Wadsworth, Belmont, CA, 1984).
99. R. T. T. Hastie and J. Friedman, *Elements of Statistical Learning* (Springer, 2009).
100. N. Cristianini and E. Ricci, Support vector machines, in *Encyclopedia of Algorithms*, M. Y. Kao (eds), Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30162-4_415.
101. C. Cortes and V. Vapnik, Support-vector network, *Mach. Learn.* **20** (1995) 273–297.
102. D. R. Cox, *Principles of Statistical Inference* (Cambridge University Press, Cambridge New York, 2006).
103. D. Freedman, *Statistical Models: Theory and Practice* (Cambridge University Press, 2009).
104. F.-L. Hsiang-Fu Yu and C. Huang, Dual coordinate descent methods for logistic regression and maximum entropy models., *Mach. Learn.* **85**(1–2) (2011) 41–75.
105. S. Walker and D. Duncan, Estimation of the probability of an event as a function of several independent variables, *Biometrika* **54** (1967) 167–178.
106. R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (methodological)* **58**(1) (1996) 267–88.
107. A. Hoerl, Application of ridge analysis to regression problems, *Chemical Engineering Progress* **58**(3) (1962) 54–59.
108. A. Tikhonov and V. Arsenin, *Solution of Ill-posed Problems* (Winston & Sons, Washington, 1977).
109. M. Tipping, Sparse bayesian learning and the relevance vector machine, *Journal of Machine Learning Research* **1** (2001) 211–244.
110. G. Hinton, Connectionist learning procedures, *Artificial intelligence* **40** (1989) 185–234.
111. Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, Gradient-based learning applied to document recognition, *Proceedings of the IEEE* **86** (November 1998) 2278–2324.
112. M. Tipping and C. Bishop, Probabilistic principal component analysis, *Journal of the Royal Statistical Society Series B* **61**(Part 3) (2002) 611–622.
113. J. B. Tenenbaum, V. D. Silva and J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, *Science* **290**(5500) (2000) 2319–2323.
114. S. T. Roweis and L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science* **290**(5500) (2000) 2323–2326.
115. I. Borg and G. P., *Modern Multidimensional Scaling — Theory and Applications* (Springer Series in Statistics, 1997).
116. R. Cattell, *Factor analysis.* (Harper., New York, 1952).
117. R. Cattell, *Use of Factor Analysis in Behavioral and Life Sciences* (Plenum, New York, 1978).
118. T. Fawcett, Roc graphs with instance-varying costs, *Pattern Recognition Letters* **27**(8) (2006) 882–891, ROC Analysis in Pattern Recognition.
119. M. V. Lombardo, M.-C. Lai, B. Auyeung, R. J. Holt, C. Allison, P. Smith, B. Chakrabarti, A. N. V. Ruigrok, J. Suckling, E. T. Bullmore, A. J. Bailey, S. Baron-Cohen, P. F. Bolton, E. T. Bullmore, S. Carrington, M. Catani, M. C. Craig, E. M. Daly, S. C. L. Deoni, C. Ecker, F. Happé, J. Henty, P. Jezzard, P. Johnston, D. K. Jones, M.

V. Lombardo, A. Madden, D. Mullins, C. M. Murphy, D. G. M. Murphy, G. Pasco, A. N. V. Ruigrok, S. A. Sadek, D. Spain, R. Stewart, S. J. Wheelwright, S. C. Williams, C. Ellie Wilson, M. C. Craig, D. G. M. Murphy and M. A. Consortium, Unsupervised data-driven stratification of mentalizing heterogeneity in autism, *Scientific Reports* **6**(1) (2016) p. 35333.

120. M. Ahmadlou and H. Adeli, Enhanced probabilistic neural network with local decision circles: A robust classifier, *Integrated Computer-Aided Engineering* **17** (2010) 197–210.
121. M. H. Rafiei and H. Adeli, A new neural dynamic classification algorithm, *IEEE Transactions on Neural Networks and Learning Systems* **28**(12) (2017) 3074–3083.
122. D. R. Pereira, M. A. Piteri, A. Souza, J. P. Papa and H. Adeli, Fema: A finite element machine for fast learning, *Neural Computing and Applications* **32**(10) (2020) 6393–6404.
123. K. M. R. Alam, N. Siddique and H. Adeli, A dynamic ensemble learning algorithm for neural networks, *Neural Computing and Applications* **32**(12) (2020) 8675–8690.
124. J. Gorriz, C. Jimenez-Mesa, R. Romero-Garcia, F. Segovia, J. Ramirez, D. Castillo-Barnes, F. Martinez-Murcia, A. Ortiz, D. Salas-Gonzalez, I. Illan, C. Puntonet, D. Lopez-Garcia, M. Gomez-Rio and J. Suckling, Statistical agnostic mapping: A framework in neuroimaging based on concentration inequalities, *Information Fusion* **66** (2021) 198–212.