



# The Importance of Anti-correlations in Graph Theory Based Classification of Autism Spectrum Disorder

Amirali Kazeminejad<sup>1,2\*</sup> and Roberto C. Sotero<sup>2,3</sup>

<sup>1</sup> Department of Biomedical Engineering, University of Calgary, Calgary, AB, Canada, <sup>2</sup> Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada, <sup>3</sup> Department of Radiology, University of Calgary, Calgary, AB, Canada

## OPEN ACCESS

### Edited by:

Reza Lashgari,  
Institute for Research in Fundamental  
Sciences, Iran

### Reviewed by:

Vahid Salari,  
University of the Basque Country,  
Spain  
Reza Khosrowabadi,  
Shahid Beheshti University, Iran

### \*Correspondence:

Amirali Kazeminejad  
amirali.kazeminejad@ucalgary.ca

### Specialty section:

This article was submitted to  
Brain Imaging Methods,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 09 May 2019

**Accepted:** 02 June 2020

**Published:** 07 August 2020

### Citation:

Kazeminejad A and Sotero RC  
(2020) The Importance  
of Anti-correlations in Graph Theory  
Based Classification of Autism  
Spectrum Disorder.  
*Front. Neurosci.* 14:676.  
doi: 10.3389/fnins.2020.00676

With the release of the multi-site Autism Brain Imaging Data Exchange, many researchers have applied machine learning methods to distinguish between healthy subjects and autistic individuals by using features extracted from resting state functional MRI data. An important part of applying machine learning to this problem is extracting these features. Specifically, whether to include negative correlations between brain region activities as relevant features and how best to define these features. For the second question, the graph theoretical properties of the brain network may provide a reasonable answer. In this study, we investigated the first issue by comparing three different approaches. These included using the positive correlation matrix (comprising only the positive values of the original correlation matrix), the absolute value of the correlation matrix, or the anticorrelation matrix (comprising only the negative correlation values) as the starting point for extracting relevant features using graph theory. We then trained a multi-layer perceptron in a leave-one-site out manner in which the data from a single site was left out as testing data and the model was trained on the data from the other sites. Our results show that on average, using graph features extracted from the anti-correlation matrix led to the highest accuracy and AUC scores. This suggests that anti-correlations should not simply be discarded as they may include useful information that would aid the classification task. We also show that adding the PCA transformation of the original correlation matrix to the feature space leads to an increase in accuracy.

**Keywords:** autism spectrum disorder, fMRI, machine learning, graph theory, anti-correlations, neural networks, GSR

## INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that is growing in prevalence in recent years (Zablotsky et al., 2015). While it is usually diagnosed by carefully monitoring a child's behavioral development (Barbarese et al., 2006), recent studies have shown that brain imaging can also be used to aid in that diagnosis by identifying underlying differences between the ASD and Healthy Control (HC) Brain (Dichter, 2012; Bos et al., 2014).

Functional Magnetic Resonance Imaging (fMRI) is a widely used tool for such studies due to its high spatial resolution. It monitors the changes in the Blood Oxygen Level Dependent (BOLD) signal which indirectly measures the neuronal activity (Sotero and Trujillo-Barreto, 2007). By computing functional connectivity measures between BOLD signals from different brain areas,

researchers can examine the human brain at a network level. A variant of the technique called Resting State fMRI (rs-fMRI) has been widely used to examine brain networks while subjects are at rest (Fox and Greicius, 2010; van den Heuvel and Hulshoff Pol, 2010). Graph theory is one of the more novel methods being used for the network-level analysis of the brain. It provides a mathematical framework for quantifying network characteristics and quantitatively analyzing the differences between different brain networks (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010). Machine learning is another relatively new technique that is being applied to rs-fMRI data to extract insights such as important biomarkers as well as to develop novel algorithms with the hope of automatically diagnosing brain disorders from medical imaging data (Pereira et al., 2009; Lee et al., 2013). The network characteristics that is provided by graph theory has been used as the input to machine learning algorithms to identify diseases such as Alzheimer's (Khazaei et al., 2016), Parkinson's (Kazeminejad et al., 2017), and Autism (Kazeminejad and Sotero, 2019b).

One of the variables that can affect the results of the mentioned analysis is how the graph construction step was performed. After the initial preprocessing of the rs-fMRI data, a correlation matrix is created by calculating the correlation between the activation time-series of different brain regions. The correlation matrix is then transformed into a sparse binary matrix representing the existence connections between different regions. This transformation is usually performed by a thresholding step in which only a percentage of the strongest correlations are kept (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010). However, this step ignores the anti-correlations which may be biologically relevant (Fox et al., 2005). This problem will be even more severe if the preprocessing pipeline includes the regression of the global mean signal (GSR) from the time-series, known to result in the removal of motion, cardiac and respiratory signals, which may result in the presence of more anticorrelations in the correlation matrix (Murphy and Fox, 2017). The use of GSR is a controversial matter in the field. Although there is evidence suggesting that the anticorrelations introduced through GSR have no biological basis (Murphy et al., 2009), a recent study has shown using GSR in the preprocessing pipeline for rs-fMRI leads to better prediction accuracies of behavioral measures (Li et al., 2019).

There has been evidence of network-level changes in the ASD brain compared to a HC brain (Redcay et al., 2013; Rudie et al., 2013; Bos et al., 2014). Therefore, using graph theory to extract features for classification of ASD is likely to provide good results. This claim was previously examined by training a SVM classifier on graph theoretical features to classify between ASD and HC (Kazeminejad and Sotero, 2019b). However, the methodology of that paper was limited to using only the positive correlations in order to construct the brain graph, potentially ignoring some informative connections.

The release of the Autism Brain Imaging Data Exchange I (ABIDE I) dataset (Di Martino et al., 2014) allowed researchers to examine ASD in large sample sizes. Nielsen et al. (2013) conducted one of the earliest classification studies on ABIDE and were able to achieve an accuracy of 60% over all samples.

More recently, Heinsfeld et al. (2018) were able to achieve an accuracy of 70%, evaluated by 10 fold cross-validation, on the entire dataset by utilizing neural networks and transfer learning. They also reported the leave-one-site out performance of their model averaging at 65% accuracy. Plitt et al. (2015) achieved a higher accuracy, 69.71% when using only 178 subjects from the ABIDE I dataset. Another study by Khosla et al. (2018) used 3D convolutional neural networks to achieve cross validated accuracies as high as 71.7% on a subset of the ABIDE I dataset.

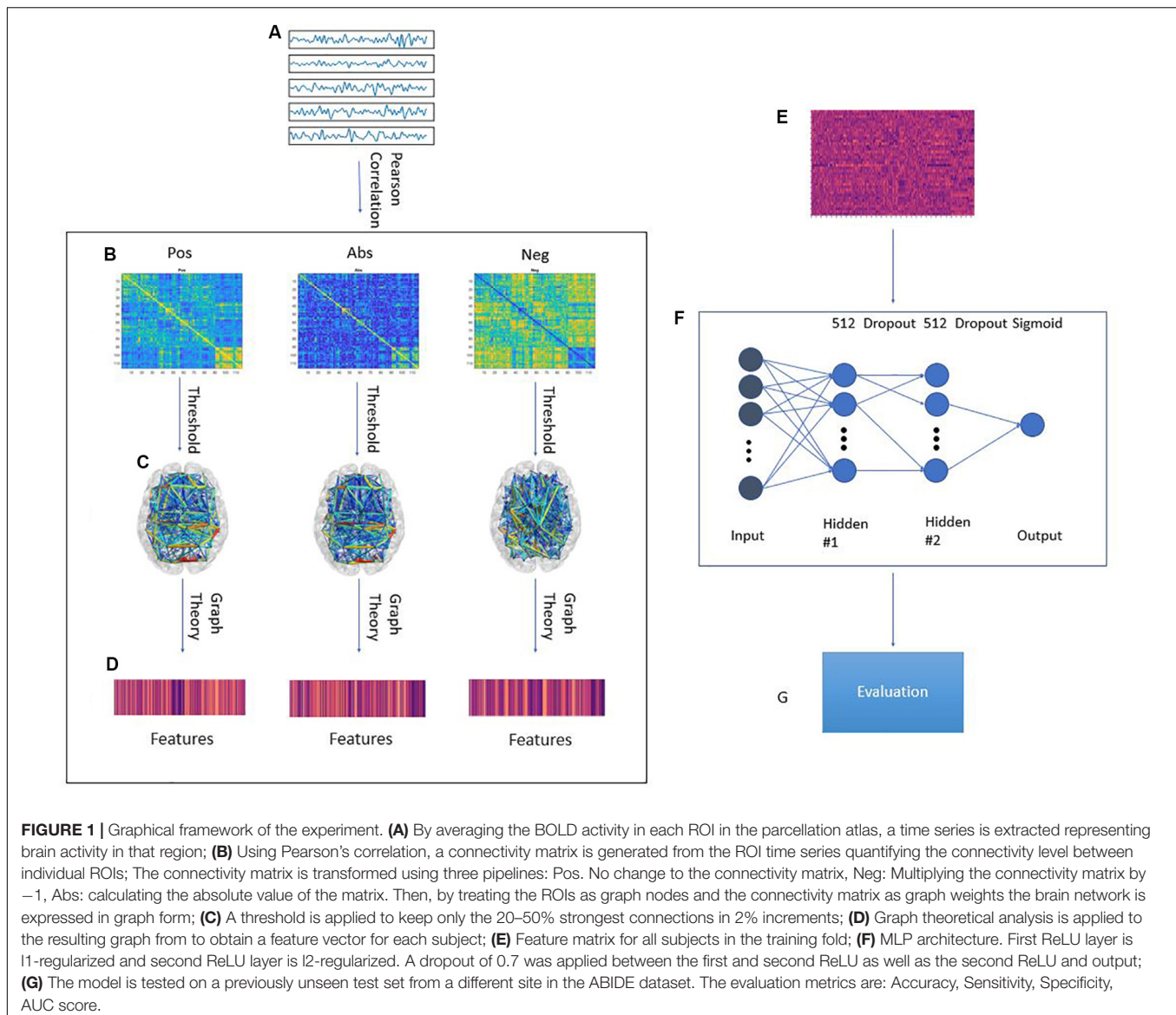
A recent study by Hallquist and Hillary (2019), examining 106 papers (with only 2 papers focusing on ASD) using graph theoretical measures, shows that 79.2% of graph theoretical studies either did not specify how they handled negative correlations or discarded them. Another 9.4% used absolute values of the correlation matrix. However, this is mostly an arbitrary choice.

In this paper, we study the effect of different approaches to handling anti-correlations for the classification accuracy of a machine learning model on the ABIDE dataset. We trained a regularized deep learning neural network using features extracted from transforming the correlation matrix using three pipelines: The positive correlation pipeline which does not change the matrix, the anti-correlation pipeline which prioritizes negative correlations, and the absolute value pipeline which disregards the sign of the correlations. These pipelines are explained further in **Figure 1**. Our model was evaluated using a leave-one-site out approach. Our results suggest that on average, the anti-correlation pipeline results in better accuracy and area under curve (AUC) score with a small loss in sensitivity. Interestingly, adding the PCA transformation of the original correlation matrix increased the accuracy, sensitivity, specificity and AUC score for all pipelines.

## MATERIALS AND METHODS

### Data and Preprocessing

This study used a publicly available dataset from the ABIDE Preprocessed Initiative (Cameron et al., 2013a). To ensure that our results were not affected by any custom preprocessing pipeline, we used the preprocessed data provided by ABIDE in the C-PAC (Cameron et al., 2013b) pipeline. The preprocessing included the following steps. The Analysis of Functional Neuro Images (AFNI) (Cox, 1996) software was used for removing the skull from the images. The brain was segmented into three tissues using FMRIB Software Library (FSL) (Smith et al., 2004). The images were then normalized to the MNI 152 stereotactic space (Mazziotta et al., 2001; Grabner et al., 2006) using Advanced Normalization Tools (ANTs) (Avants et al., 2011). Functional preprocessing included motion and slice-timing correction as well as the normalization of voxel intensity. Nuisance signal regression included 24 parameters for head motion, CompCor (Behzadi et al., 2007) with 5 principal components for tissue signal in Cerebrospinal fluid and white matter, linear and quadratic trends for Low-frequency drifts and a global bandpass filter (0.01–0.1 Hz). GSR was also applied to remove the global mean from the signals. These images were then co-registered to



their anatomical counterpart by FSL. They were then normalized to the MNI 152 space using ANTs. The average voxel activity in each Region of Interest (ROI) of the Craddock 200 atlas (Craddock et al., 2012) was then extracted as the time-series for that region. Furthermore, in order to replicate our results on the Craddock 200 atlas, we reran the pipelines on the anatomical automatic labeling (AAL) (Tzourio-Mazoyer et al., 2002) atlas. Overall, this study had 1,035 subjects, the demographics of whom are outlined in **Table 1**.

## Network Construction, Graph Extraction and Feature Extraction

To construct the brain network, the timeseries for each atlas ROI where correlated, using Pearson correlation, with the other regions. The strengths of these correlations were used as the strengths of the connection between different ROIs. In graph

terms, this represents a fully connected graph with each of the nodes being in the center of the corresponding ROIs and each edge weight is the correlation between the two nodes on the opposite ends of the edge. Three different pipelines were obtained based on selecting only the positive values, only the negative values, or taking the absolute value of the correlation matrix. These graphs were then subjected to a thresholding step in which only the 20–50% strongest connections were set to one and the other edges were discarded. This threshold was incremented in 2% steps. This resulted in sparse binary graphs. This step was done because binary graphs have been shown to have more easily defined null models and are more easily characterizable (Rubinov and Sporns, 2010). Several measures of integration (characteristic path length and efficiency), segregation (clustering coefficient and transitivity), centrality (betweenness centrality, eigenvector centrality, participation coefficient and within module z-score) and resilience (assortativity) (Bullmore and Sporns, 2009) were

TABLE 1 | Subject demographics.

Group	Age Mean (SD)				Count			
	ASD		HC		ASD		HC	
	F	M	F	M	F	M	F	M
<b>Sex</b>								
<b>Site</b>								
CALTECH	24.35 (6.91)	28.27 (11.08)	22.77 (4.31)	29.51 (11.83)	4	15	4	14
CMU	27.00 (6.93)	26.18 (5.88)	26.00 (5.29)	27.10 (6.12)	3	11	3	10
KKI	9.61 (1.85)	10.12 (1.38)	9.49 (0.81)	10.22 (1.23)	4	16	8	20
LEUVEN	14.00 (1.41)	18.18 (5.09)	13.56 (0.75)	19.01 (5.05)	3	26	5	29
MAX_MUN	35.33 (8.50)	24.76 (15.27)	32.00 (X)	24.37 (8.81)	3	21	1	27
NYU	18.62 (9.14)	14.14 (6.59)	15.37 (6.45)	15.76 (6.10)	10	65	26	74
OHSU	–	11.43 (2.18)	–	10.09 (1.12)	–	–	–	14
OLIN	19.00 (5.00)	16.06 (3.04)	17.50 (3.54)	16.54 (3.78)	3	16	2	13
PITT	12.60 (0.95)	20.03 (7.38)	16.07 (3.60)	19.36 (6.97)	4	25	4	23
SBL	–	35.00 (10.43)	–	33.73 (6.61)	–	15	–	15
SDSU	12.44 (X)	14.89 (1.70)	13.26 (2.69)	14.58 (1.46)	1	13	6	16
STANFORD	10.02 (1.63)	9.99 (1.68)	9.01 (1.01)	10.19 (1.66)	4	15	4	16
TRINITY	–	16.81 (3.17)	–	17.08 (3.77)	–	22	–	25
UCLA	12.46 (3.42)	13.08 (2.34)	12.82 (0.89)	13.03 (2.04)	6	48	6	38
UM	13.53 (3.12)	13.11 (2.30)	14.78 (2.82)	14.82 (3.85)	9	57	18	56
USM	–	23.46 (8.33)	–	21.29 (8.35)	–	46	–	25
YALE	12.88 (3.00)	12.70 (3.14)	13.52 (2.64)	12.34 (2.79)	8	20	8	20

M: Male, F: Female, ASD: Autism spectrum disorder, HC: Healthy control, SD: Standard deviation. Some sites did not have any female individuals, these have been identified with a “–.” The term “X” for SD means that there were not enough samples to calculate SD.

then extracted from the binary brain network graph. These steps were done using the Python libraries `brainconn` (<sup>1</sup>FIU-Neuro/brainconn) and `networkx`. This resulted in 3 feature sets of 1,404 features for each group. A breakdown of these features can be found in section 6 of the **Supplementary Material**. The main body of this paper focuses mainly on only using the graph features in the analysis. In the **Supplementary Material**, we report the results of adding the PCA transformation (Bartholomew, 2010) of the unaltered correlation matrix to the graph features. These principle components were calculated for each individual and we have 600 components per individual.

## Leave-One-Site-Out Cross-Validation

The structure of the ABIDE I dataset allows for an interesting cross-validation approach that captures the multi-site nature of it. Our data was divided into 17 cross-validation sets. In each one, one site was used as the test set and the other 16 sites were used as the training set.

## Model Training and Evaluation

In this study, we used a multilayer perceptron with two hidden layers. The model was implemented with Keras and using the Tensorflow 1.13 backend (Khosla et al., 2018). This model was chosen due to its ability to construct data driven features before using them for the final classification task. Both layers consisted of 512 rectified linear neurons. Due to the limited number of

samples, the model was heavily regularized. The first hidden layer is subjected to L1 regularization in order to force the network to have some feature selection capabilities. A dropout layer is then applied before the second hidden layer. L2 regularization was used in the second hidden layer. Finally, a dropout layer was included before the output layer. The output layer is a single neuron which is activated by a sigmoid function.

The model training process is as follows. First, each feature was standardized by removing the mean and scaling to unit variance. Then the training data was ran through the neural network. The model used binary cross-entropy as its loss function and an Adam optimizer with  $\beta_1 = 0.99$ ,  $\beta_2 = 0.01$ , and an initial learning rate of  $l = 0.0002$ . The learning rate was decayed by a factor of 0.5 base on the validation loss calculated on a random 10% of the training data withheld during the process. The model parameters were tuned on only one of the left out sites (PITT) to ensure low information leak. The trained model was then evaluated on the test set. We repeated the above steps 5 times for each Leave-One-Site-Out (LOSO) fold and report the average accuracy, sensitivity, specificity and AUC score over the 5 repetitions. It is worth noting that the healthy subjects were given a label of 1 and the ASD subjects were given label 0. Thus, sensitivity should be interpreted as the percentage of HCs correctly identified. Likewise, specificity is the percentage of ASDs correctly identified.

We used two approaches to aggregate the results of each left out site. First, we report the overall average metric for each of these sites as “average.” We also report a weighted average by using the sample size of each left out site as the weight of that site.

<sup>1</sup><https://github.com>

<sup>2</sup><https://networkx.github.io/>



## Statistical Evaluation

We performed two statistical analyses to compare model results for each left out site between different pipelines. First, we compared the average of the 5 trials performance of the different models for each pipeline for each site against each other using a Welch's *t*-test (Welch, 1947). The average was calculated by computing the mean of each metric for each trial over the different thresholds. This was corrected for multiple comparisons using the Holms-Sidak method (Cardillo, 2006). We also compared model performance in a threshold-wise manner, comparing each threshold results for each pipeline against the same threshold results in another pipeline using the same test. Furthermore, we also report the *p*-value of the mean of each performance metrics between the three pipelines by conducting a paired *t*-test based on the results of individual sites. The significance level is set to  $p < 0.05$  and all tests are two-tailed.

## RESULTS

In order to manage the different thresholds, we report the average performance measures and their standard deviation across the different thresholds. This applies to all following paragraphs unless explicitly specified. The effects of this choice are discussed further in section "Discussion."

Our results in **Table 2** show that, on average, the negative correlation pipeline can achieve higher accuracies and AUC score than the other two pipelines. The average sensitivity of the model remains comparable over the pipelines while being slightly lower for the absolute value pipeline. The model was able to achieve a higher specificity, increasing its ability to correctly identify ASDs. This is interesting because it suggests different feature extraction pipelines will allow some flexibility between interchanging sensitivity and specificity.

The same results held when adding the PCA transformation of the original correlation matrix as input features. However, the overall performance of the model was improved. Suggesting that information from the original correlation matrix supplements the information available in the graph features (**Supplementary Table 1**). To test whether these observations were dependent on the choice of the atlas, we applied the same methodology on a different atlas. When using the AAL parcellation, the negative pipeline performed the best as well. Thus, strengthening the hypothesis that this result holds across different parcellation atlases (**Supplementary Table 2**). Another variable in the dataset was gender. The negative correlation pipeline was still the best performing pipeline on average when only modeling the male subjects. However, the difference between the absolute correlation and negative correlation pipelines were negligible (AUC of 0.59 vs 0.592). Interestingly, overall model performance was lowered when the training and test data only included from male subjects (**Supplementary Table 3**).

As evident in the performance tables, there is a high variability in model performance between sites. One possible explanation for this may be related to the age of the subjects in that site. To investigate this, we plotted the average AUC scores of the 5 trials for each site against their age means. A quadratic regression

was applied on this data and is shown in **Figure 2**. In the case of added PCA features, the sites at opposing ends of the age mean scale performed worse than the sites closer to the mean. This phenomenon was less pronounced in the case of using only graph features with it being almost non-existent in the case of the positive correlation pipeline. Furthermore, we report the R-squared values and *F*-test *p*-values of these plots. While most regression curves show statistical significance, the quadratic regression is shown to be a poor fit for the results from the positive pipeline when using only graph features.

Our results show that using different pipelines may play a significant role in the classification outcome for certain sites. **Figure 3** shows how these pipelines compare across different site when the same threshold is applied. We report statistically significant differences between the specificity of the Yale site model between the absolute value and positive pipelines as well as between the anticorrelation and absolute value pipeline. Such significant differences can be spotted elsewhere for the PITT site in the case of absolute value vs anticorrelation pipeline as well as to a more limited scale across different sites and pipelines. This suggests that between site differences such as subject demographics and equipment differences could be used to inform the choice of which pipeline to use.

**Supplementary Tables 4.1** and **4.2** show the *p*-values of between pipeline differences compared to the anti-correlation pipeline for the graph-features-only method. When comparing the best models, based on what threshold led to the highest AUC score, trained with only graph features for each site, the negative pipeline achieved a higher AUC than the other pipelines except in the following sites. OLIN, OHSU, NYU and SBL for the unchanged pipeline and OHSU and NYU for the absolute value pipeline. Of these sites, none were significantly different ( $p < 0.05$ ) between the unchanged and anticorrelation pipelines. For accuracy, the same comparison holds between the anticorrelation and unchanged pipelines. OSHU ( $p = 0.042$ ) and NYU ( $p = 0.043$ ) results were the only significantly different sites. The absolute value pipeline was able to achieve better accuracies than the negative pipeline for the following site: OLIN, OHSU, NYU. Likewise, the AAL atlas results did not show many statistical significances in the site accuracies between different pipelines.

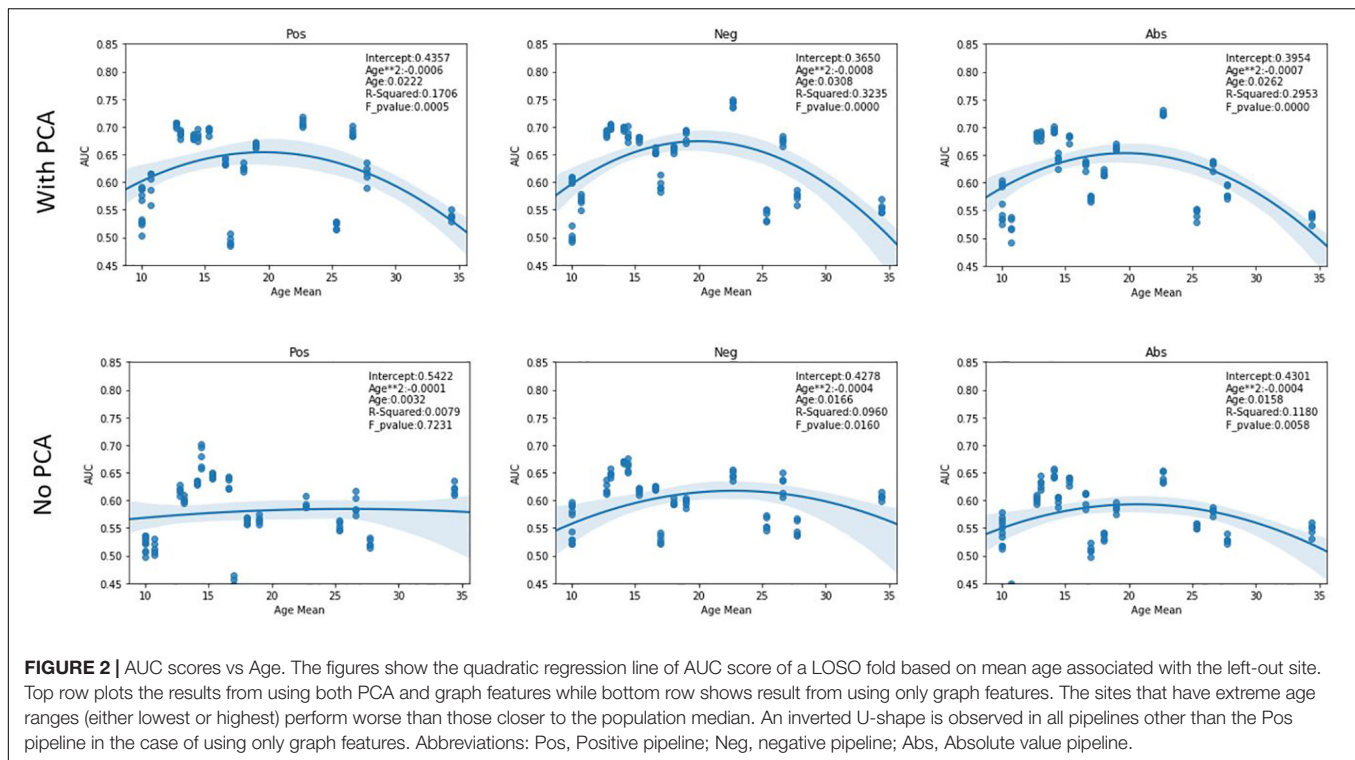
As it may not be fair to compare different thresholds for each site, we also performed threshold matched comparisons for the graph-only craddock-200 atlas results. A paired *t*-test was performed on the performance metrics of each pipeline pair for each site. **Figure 3** shows these comparisons in a heatmap style plot. Interestingly, no site showed consistent difference in performance (AUC) over the tested threshold range, suggesting that this value should be treated as a parameter than can affect the accuracy of machine learning predictions using graph theory and tuned for the task at hand.

Another paired *t*-test was performed comparing the out of site fold performance of the three pipelines with just graph features (**Table 3**). The negative correlation pipeline had statistically different accuracy and AUC metrics when compared to the absolute value pipeline. No other significant differences were found between the pipelines.

**TABLE 2** | Results with only graph features.

	ACC			AUC			SEN			SPE		
	Abs	Neg	Pos	Abs	Neg	Pos	Abs	Neg	Pos	Abs	Neg	Pos
CALTECH	0.527 (0.032)	0.547 (0.034)	0.519 (0.029)	0.529 (0.031)	0.55 (0.034)	0.523 (0.029)	0.592 (0.04)	0.671 (0.055)	0.675 (0.07)	0.466 (0.058)	0.429 (0.06)	0.372 (0.065)
CMU	0.575 (0.044)	0.623 (0.052)	0.584 (0.044)	0.581 (0.043)	0.629 (0.053)	0.592 (0.043)	0.723 (0.052)	0.778 (0.083)	0.823 (0.076)	0.438 (0.064)	0.479 (0.054)	0.362 (0.101)
KKI	0.55 (0.037)	0.549 (0.035)	0.491 (0.035)	0.564 (0.036)	0.586 (0.034)	0.512 (0.041)	0.481 (0.052)	0.364 (0.056)	0.389 (0.061)	0.646 (0.052)	0.807 (0.055)	0.635 (0.107)
LEUVEN	0.551 (0.026)	0.608 (0.038)	0.581 (0.04)	0.535 (0.025)	0.596 (0.037)	0.563 (0.042)	0.734 (0.05)	0.744 (0.051)	0.785 (0.038)	0.337 (0.035)	0.448 (0.043)	0.342 (0.066)
MAX_MUN	0.556 (0.024)	0.562 (0.033)	0.557 (0.033)	0.554 (0.024)	0.558 (0.034)	0.554 (0.033)	0.582 (0.063)	0.612 (0.034)	0.602 (0.047)	0.526 (0.059)	0.505 (0.061)	0.505 (0.053)
NYU	0.635 (0.021)	0.615 (0.019)	0.644 (0.019)	0.636 (0.02)	0.618 (0.018)	0.645 (0.019)	0.626 (0.033)	0.602 (0.032)	0.641 (0.028)	0.646 (0.026)	0.633 (0.02)	0.648 (0.025)
OHSU	0.44 (0.038)	0.403 (0.047)	0.516 (0.045)	0.442 (0.039)	0.4 (0.048)	0.515 (0.041)	0.412 (0.047)	0.442 (0.051)	0.532 (0.141)	0.473 (0.074)	0.357 (0.06)	0.497 (0.111)
OLIN	0.619 (0.034)	0.625 (0.042)	0.643 (0.043)	0.603 (0.035)	0.623 (0.043)	0.633 (0.043)	0.463 (0.05)	0.608 (0.074)	0.542 (0.057)	0.743 (0.039)	0.638 (0.051)	0.723 (0.054)
PITT	0.586 (0.029)	0.592 (0.034)	0.562 (0.018)	0.586 (0.029)	0.596 (0.033)	0.565 (0.018)	0.578 (0.048)	0.723 (0.042)	0.653 (0.04)	0.594 (0.046)	0.47 (0.046)	0.476 (0.043)
SBL	0.548 (0.03)	0.607 (0.036)	0.62 (0.04)	0.548 (0.03)	0.607 (0.036)	0.62 (0.04)	0.838 (0.045)	0.793 (0.031)	0.761 (0.056)	0.257 (0.076)	0.42 (0.07)	0.48 (0.058)
SDSU	0.603 (0.027)	0.648 (0.052)	0.668 (0.024)	0.6 (0.028)	0.662 (0.058)	0.68 (0.023)	0.613 (0.039)	0.597 (0.045)	0.63 (0.038)	0.587 (0.051)	0.728 (0.091)	0.729 (0.036)
STANFORD	0.52 (0.03)	0.522 (0.035)	0.527 (0.015)	0.525 (0.03)	0.528 (0.035)	0.533 (0.014)	0.346 (0.069)	0.286 (0.054)	0.299 (0.035)	0.703 (0.063)	0.77 (0.039)	0.766 (0.028)
TRINITY	0.512 (0.032)	0.535 (0.029)	0.444 (0.043)	0.51 (0.032)	0.53 (0.03)	0.44 (0.044)	0.541 (0.053)	0.609 (0.048)	0.494 (0.05)	0.479 (0.044)	0.451 (0.065)	0.387 (0.06)
UCLA	0.628 (0.025)	0.646 (0.024)	0.592 (0.024)	0.631 (0.026)	0.648 (0.023)	0.601 (0.022)	0.656 (0.048)	0.676 (0.027)	0.691 (0.03)	0.605 (0.033)	0.621 (0.036)	0.511 (0.044)
UM	0.644 (0.019)	0.664 (0.017)	0.629 (0.028)	0.647 (0.02)	0.67 (0.017)	0.632 (0.028)	0.584 (0.026)	0.58 (0.029)	0.57 (0.038)	0.711 (0.035)	0.759 (0.029)	0.695 (0.022)
USM	0.618 (0.019)	0.607 (0.042)	0.555 (0.042)	0.642 (0.019)	0.646 (0.04)	0.593 (0.035)	0.725 (0.031)	0.779 (0.043)	0.721 (0.031)	0.56 (0.027)	0.513 (0.051)	0.465 (0.062)
YALE	0.602 (0.023)	0.622 (0.031)	0.618 (0.026)	0.602 (0.023)	0.622 (0.031)	0.618 (0.026)	0.768 (0.035)	0.59 (0.044)	0.655 (0.062)	0.436 (0.048)	0.654 (0.044)	0.581 (0.041)
Mean	0.571	0.587	0.574	0.573	0.592	0.578	0.604	0.615	0.615	0.542	0.569	0.540
Weighted Mean	0.592	0.604	0.587	0.594	0.609	0.592	0.614	0.619	0.622	0.574	0.599	0.561

ACC: ACCURACY; SEN: SENSITIVITY; SPE: SPECIFICITY; AUC: Area under curve score; Pos: Unchanged pipeline; Neg: Anti-correlation pipeline; Abs: Absolute value pipeline.



Although not the main goal of this study, our model was able to, based on the weighted accuracy, perform on-par with the deep learning model described in Heinsfeld et al. (2018) when using the graph and PCA features. The sites CALTECH, KKI, MAX\_MUN, OHSU, SBL, STANFORD and TRINITY showcased lower accuracy in our model.

Section 4 of the **Supplementary Material** reports the LOSO results of 3 other models: Logistic regression, random forest classification and gaussian support vector classifiers. These tables show that our neural network model was able to outperform these 3 classical machine learning algorithms in most sites.

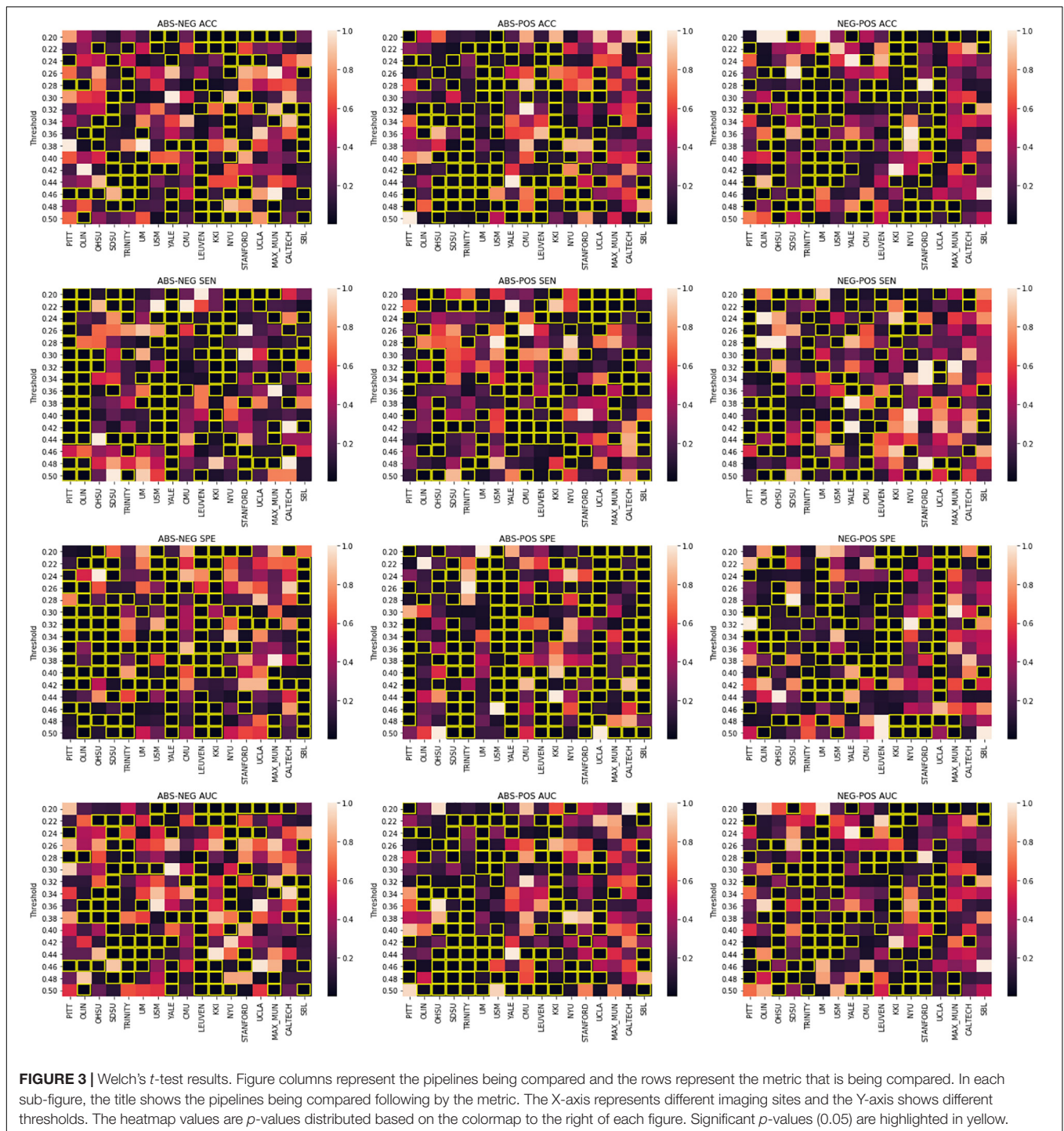
## DISCUSSION

Our results consistently showed that, on average, using graph theoretical features from the negative pipeline increases the MLP model accuracy in distinguishing between HC and ASD. This suggests that anti-correlations in the brain network contain important information that helps our model in distinguishing between ASD and HC. One possible reason for the higher performance of the anti-correlation method could be that contrary to some previous studies (Murphy et al., 2009; Murphy and Fox, 2017), GSR introduced anti-correlations may in fact stem from real neurological basis and are not spurious (Khosla et al., 2018). Unfortunately, as we do not know of any gold standards to test this hypothesis in real data, it should only be interpreted as a speculation made based on the results of this study.

Previous studies have concluded that the use of absolute value graph metrics in the presence of GSR may be compromised (Bartholomew, 2010). This was attributed to the fact that in the presence of GSR, the topologies of the anti-correlation and correlation matrix will be mixed when using an absolute value pipeline as the anti-correlations have comparable magnitude to the correlations. However, our model trained with the absolute value graph features was able to perform on-par with the positive correlation pipeline.

The low number of significantly different within-site metrics between the three pipelines may be attributed to the way they were calculated. Each LOSO fold was run 5 times in order to offset any effects of random weight initialization and random training-validation set splits. Increasing the number of trials should add to the strength of this test by capturing the metric distributions more accurately. Another possible explanation is that there is enough information available in all pipelines for most of these sites.

In the case of added PCA, plotting the accuracies of the negative pipeline against number of participants did not provide further insight into why some sites performed worst. Plotting the same against the age of the participants revealed that there may be a link between age and the ability of our model to make accurate predictions. Age has been previously shown to affect network properties of subject within the ABIDE dataset (Henry et al., 2018). KKI and STANFORD are the two sites with the lowest average ASD and HC age. Of the 4 sites with the highest average ASD age, 3 of them, CALTECH, MAX\_MUN and SBL showed lower accuracies. Interestingly, CMU, did not follow this behavior. However, the standard deviation of the age of CMU's



**FIGURE 3 |** Welch's *t*-test results. Figure columns represent the pipelines being compared and the rows represent the metric that is being compared. In each sub-figure, the title shows the pipelines being compared following by the metric. The X-axis represents different imaging sites and the Y-axis shows different thresholds. The heatmap values are *p*-values distributed based on the colormap to the right of each figure. Significant *p*-values (0.05) are highlighted in yellow.

participants was lower than those of the other 3. The larger age SD may have hindered the algorithm's ability to perform well for the sites with high average age.

Interestingly, the model trained only on the graph features was able to perform relatively well on SBL despite that site having the highest age mean and performed worst on OHSU, the site with the least number of subjects. While this could have happened purely by chance, it is an interesting avenue for

further insights about the nature of our model. Furthermore, these models were more robust to the effects of age as shown in Figure 2 suggesting age-related information is better captured in the graph theoretical features than the PCA transformation of the correlation matrix. However, this conclusion assumes that a quadratic regression fits the performance results of the graph only pipelines. The positive pipeline does not show a significant *f*-test result, thus perhaps a higher order regression would fit its



**TABLE 3** | Paired *t*-test results (*p*-values) between the average metrics of the specified pipelines, corrected by the holm-sidak method.

	ACC	SEN	SPE	AUC
POS-NEG	0.451	0.966	0.367	0.382
POS-ABS	0.846	0.867	0.931	0.645
NEG-ABS	0.086	0.867	0.490	0.028

POS: positive correlation pipeline; NEG: anti-correlation pipeline; ABS: absolute correlation pipeline; ACC, SEN, SPE, AUC like above tables. Green cells show statistical significance.

results better. During our experiments, a third-degree regression did in fact show significant *f*-test values across all members of the graph only pipelines.

Another intriguing observation about our results is that some thresholds and pipelines were better suited for classifying specific sites. This suggests that an ensemble model, such as training different models on each of these combinations and assigning the most selected label for each subject between these models, may achieve better performance on this dataset.

In this study we utilized a leave-one-site out cross validation approach in order to capture the multi-site nature of the ABIDE I dataset. We believe this approach leads to a better estimate of how this model will perform in the real world as it allows for better generalization across previously unseen imaging sites and protocols.

Using graph theoretical measures to analyze and classify neurological and neurodevelopmental diseases is not a novel idea. However, the practice of thresholding the connectivity matrix in order to produce a binary graph introduces an artificial bias toward some of the correlations. While this bias can be avoided by using the absolute value of the correlation matrix, this may disregard the important information that the positive correlation and anti-correlation networks hold. We have shown that, after running a GSR pipeline, focusing on the anti-correlation network may lead to better classification performance in the case of ASD vs HC. We also showed that by adding non-graph features to the anti-correlation graph features, classification metrics are improved suggesting that a different process may be needed to accurately capture the graph properties of both the positive and negative correlation networks.

One limitation in the present study may be the use of multiple threshold values in constructing the functional graph. Here we reported the average the results of all thresholds for each site. Another approach may be to report results from the threshold resulting in the highest AUC for each site. This would lead to a significant increase in model performance (achieving 70% weighted accuracy when using graph and PCA features). However, the negative correlation pipeline still outperformed the other pipeline.

Another limitation of this study was the inclusion of all age ranges in our dataset. As different sites have different demographics, this may affect the power of our model. Furthermore, identifying ASD is most important in earlier stages of life. The ABIDE I dataset included only a small number of children under the age of 10 (146 subjects) spread across multiple sites and none below the age of 5. Thus, the

results presented in this study may not generalize to studies on younger children.

The ABIDE I data exhibits numerous inherent variability due to its multi-site nature. Here we presented a deep learning model that was able to navigate the intricacies of this data and generalize over multiple sites by using graph theoretical features. Our model was able to, on average, perform on-par with previously reported deep learning models using the same number of subjects.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

We used the data collected as part of the ABIDE database and complied with everything that they have asked to be included in any manuscript using that data. The original ethics statement form the (Di Martino et al., 2014) manuscript is as follows: All contributions were based on studies approved by local IRBs, and data were fully anonymized (removing all 18 HIPAA protected health information identifiers, and face information from structural images). All data distributed were visually inspected prior to release.

## AUTHOR CONTRIBUTIONS

AK designed the experiments including writing the experiment code as well as writing the manuscript. RS provided advice and guidance on the experimental design and contributed to many manuscript revisions and final proofreading. Both authors contributed to the article and approved the submitted version.

## FUNDING

This work was partially funded by the Department of Biomedical Engineering Research and Academic Excellence awards at the University of Calgary.

## ACKNOWLEDGMENTS

Part of the code for the preprocessing pipeline was reused from the open source code provided by the Heinsfeld et al. manuscript. This work could not be completed without the immense contribution of open sourcing the ABIDE datasets. This manuscript has been released as a Pre-Print at bioRxiv (Kazeminejad and Sotero, 2019a).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2020.00676/full#supplementary-material>

## REFERENCES

- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54, 2033–2044. doi: 10.1016/j.neuroimage.2010.09.025
- Barbaredi, W. J., Katusic, S. K., and Voigt, R. G. (2006). Autism: a review of the state of the science for pediatric primary health care clinicians. *Arch. Pediatr. Adolesc. Med.* 160, 1167–1175.
- Bartholomew, D. J. (2010). *Principal Components Analysis in International Encyclopedia of Education*. Berlin: Springer.
- Behzadi, Y., Restom, K., Liu, J., and Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37, 90–101. doi: 10.1016/j.neuroimage.2007.04.042
- Bos, D. J., van Raalten, T. R., Oranje, B., Smits, A. R., Kobussen, N. A., van Belle, J., et al. (2014). Developmental differences in higher-order resting-state networks in Autism Spectrum disorder. *Neuroimage Clin.* 4, 820–827. doi: 10.1016/j.nicl.2014.05.007
- Bullmore, E., and Sporns, O. (2009). Erratum: complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 312–312. doi: 10.1038/nrn2618
- Cameron, C., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakob, A., et al. (2013a). The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. *Front. Neuroinform.* 7:41. doi: 10.3389/conf.fninf.2013.09.00041
- Cameron, C., Sikka, S., Cheung, B., Khanuja, R., Ghosh, S. S., Yan, C., et al. (2013b). Towards automated analysis of connectomes: the configurable pipeline for the analysis of connectomes (C-PAC). *Front. Neuroinform.* 7:42. doi: 10.3389/conf.fninf.2013.09.00042
- Cardillo, G. (2006). *Holm-Sidak T-Test: A Routine For Multiple T-Test Comparisons*. Available online at: <http://www.mathworks.com/matlabcentral/fileexchange/12786>
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Craddock, R. C., James, G. A., Holtzheimer, P. E., Hu, X. P., and Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* 33, 1914–1928. doi: 10.1002/hbm.21333
- Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667.
- Dichter, G. S. (2012). Functional magnetic resonance imaging of autism spectrum disorders. *Dialog. Clin. Neurosci.* 14, 319–351.
- Fox, M. D., and Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Front. Syst. Neurosci.* 4:19. doi: 10.3389/conf.fninf.2013.09.00019
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102
- Grabner, G., Janke, A. L., Budge, M. M., Smith, D., Pruessner, J., and Collins, D. L. (2006). *Symmetric Atlasing and Model Based Segmentation: An Application to the Hippocampus in Older Adults*. Berlin: Springer.
- Hallquist, M. N., and Hillary, F. G. (2019). Graph theory approaches to functional network organization in brain disorders: a critique for a brave new small-world. *Netw. Neurosci.* 3, 1–26. doi: 10.1162/netn\_a\_00054
- Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., and Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *Neuroimage Clin.* 17, 16–23. doi: 10.1016/j.nicl.2017.08.017
- Henry, T. R., Dichter, G. S., and Gates, K. (2018). Age and gender effects on intrinsic connectivity in autism using functional integration and segregation. *Biol. Psychiatry Cogn. Neurosci. Neuroimage.* 3, 414–422. doi: 10.1016/j.bpsc.2017.10.006
- Kazeminejad, A., Golbabaie, S., and Soltanian-Zadeh, H. (2017). “Graph theoretical metrics and machine learning for diagnosis of Parkinson’s disease using rs-fMRI,” in *Proceedings of the 2017 Artificial Intelligence and Signal Processing Conference (AISP)*, Shiraz.
- Kazeminejad, A., and Sotero, R. (2019a). The importance of anti-correlations in graph theory based classification of autism spectrum disorder. *bioRxiv* [Preprint]. Available online at: <https://www.biorxiv.org/content/10.1101/557512v1> (accessed May 10, 2019).
- Kazeminejad, A., and Sotero, R. C. (2019b). Topological properties of resting-state fMRI functional networks improve machine learning-based autism classification. *Front. Neurosci.* 12:1018. doi: 10.3389/conf.fninf.2013.09.01018
- Khazaei, A., Ebrahimzadeh, A., and Babajani-Feremi, A. (2016). Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer’s disease. *Brain Imaging Behav.* 10, 799–817. doi: 10.1007/s11682-015-9448-7
- Khosla, M., Jamison, K., Kuceyeski, A., and Sabuncu, M. R. (2018). “3D convolutional neural networks for classification of functional connectomes,” *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. 137–145.
- Lee, M. H., Smyser, C. D., and Shimony, J. S. (2013). Resting-state fMRI: a review of methods and clinical applications. *AJNR Am. J. Neuroradiol.* 34, 1866–1872. doi: 10.3174/ajnr.a3263
- Li, J., Kong, R., Liègeois, R., Orban, C., Tan, Y., Sun, N., et al. (2019). Global signal regression strengthens association between resting-state functional connectivity and behavior. *Neuroimage* 196, 126–141. doi: 10.1016/j.neuroimage.2019.04.016
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., et al. (2001). A four-dimensional probabilistic atlas of the human brain. *J. Am. Med. Assoc. Inform. Assoc.* 8, 401–430.
- Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44, 893–905. doi: 10.1016/j.neuroimage.2008.09.036
- Murphy, K., and Fox, M. D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154, 169–173. doi: 10.1016/j.neuroimage.2016.11.052
- Nielsen, J. A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2013). Multisite functional connectivity MRI classification of autism: ABIDE results. *Front. Hum. Neurosci.* 7:599. doi: 10.3389/conf.fninf.2013.09.0599
- Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199–S209.
- Plitt, M., Barnes, K. A., and Martin, A. (2015). Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *YNI* 7, 359–366. doi: 10.1016/j.nicl.2014.12.013
- Redcay, E., Moran, J. M., Mavros, P. L., Tager-Flusberg, H., Gabrieli, J. D. E., and Whitfield-Gabrieli, S. (2013). Intrinsic functional network organization in high-functioning adolescents with autism spectrum disorder. *Front. Hum. Neurosci.* 7:573. doi: 10.3389/conf.fninf.2013.09.0573
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Rudie, J. D., Brown, J. A., Beck-Pancer, D., Hernandez, L. M., Dennis, E. L., Thompson, P. M., et al. (2013). Altered functional and structural brain network organization in autism. *Neuroimage Clin.* 2, 79–94. doi: 10.1016/j.nicl.2012.11.006
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Berg, H. J., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219.
- Sotero, R. C., and Trujillo-Barreto, N. J. (2007). Modelling the role of excitatory and inhibitory neuronal activity in the generation of the BOLD signal. *Neuroimage* 35, 149–165. doi: 10.1016/j.neuroimage.2006.10.027
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM Using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

- van den Heuvel, M. P., and Hulshoff Pol, H. E. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* 20, 519–534. doi: 10.1016/j.euroneuro.2010.03.008
- Welch, B. L. (1947). The generalization of 'student's' problem when several different population variances are involved. *Biometrika* 34, 28–35. doi: 10.1093/biomet/34.1-2.28
- Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., and Blumberg, S. J. (2015). Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 national health interview survey. *Natl. Health Stat. Rep.* 87, 1–20.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2020 Kazeminejad and Sotero. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*