

Comparing Autoregressive and Network Features for Classification of Depression and Anxiety

Sebastian F. Ruf^{1,2}, Md Navid Akbar¹, Susan Whitfield-Gabrieli², Deniz Erdoğmuş¹

Abstract—Autocorrelation in functional MRI (fMRI) time series has been studied for decades, mostly considered as noise in the time series which is removed via prewhitening with an autoregressive model. Recent results suggest that the coefficients of an autoregressive model fit to fMRI data may provide an indicator of underlying brain activity, suggesting that prewhitening could be removing important diagnostic information. This paper explores the explanatory value of these autoregressive features extracted from fMRI by considering the use of these features in a classification task. As a point of comparison, functional network based features are extracted from the same data and used in the same classification task. We find that in most cases, network based features provide better classification accuracy. However, using principal component analysis to combine network based features and autoregressive features for classification based on a support vector machine provides improved classification accuracy compared to single features or network features, suggesting that when properly combined there may be additional information to be gained from autoregressive features.

I. INTRODUCTION

The analysis of functional MRI (fMRI) is a challenging time series analysis problem, with many potential sources of noise including scanner drift, physiological artifacts, and subject motion [1]. fMRI time series are known to exhibit autocorrelation, which are at least partially due to noise [2], [3]. Addressing autocorrelation in fMRI time series remains an area of research as autocorrelation between subsequent time points affects the reliability [4] and invalidates the assumptions [5] of the General Linear Model, a widely used task based fMRI analysis framework. Autocorrelation has also been shown to affect measures of functional connectivity in resting state fMRI [6], [7].

Recently Arbabshirani et. al found that autocorrelation at a voxel level, as captured by a lag 1 autoregressive (AR) model fitted to a voxel level time series, shows differences based on whether subjects are performing a task or not and based on whether the subject had a diagnosis of schizophrenia [8]. While this paper takes the AR approach in Arbabshirani et. al as a starting point for investigation, it is important to note that the AR approach shares commonalities with many others in the fMRI literature.

At a voxel level, brain activity has been studied using local signal characteristics such as the Amplitude of Low

Frequency Fluctuations (ALFF) [9] and scale invariance in the power spectrum [10]. When considering how regions of the brain interact, there are a number of effective connectivity [11] approaches which model interactions between regions based on temporal information in the fMRI time course, such as Dynamic Causal Modeling [12], Granger Causality [13], and Vector Autoregressive Models [14].

However, when understanding alterations of brain function in conditions such as depression, machine learning approaches typically use functional connectivity [15]–[18]. He et al. combined multiset canonical correlation analysis and joint independent component analysis with a support vector machine (SVM) classifier to obtain a 99% accuracy in discriminating between two classes of depression [16]. Similar binary classification studies introduce a sparse low rank model and Fisher score feature selection to obtain a 95% accuracy [17], and a time-varying dynamic functional connectivity to achieve a 99% accuracy [18], both involving SVM classifiers. More recent advances have been towards adopting neural networks, such as Jun et al. using graph convolutional networks to obtain a 79% accuracy [19].

Currently, measures of functional connectivity and effective connectivity are used independently in the literature. To properly understand the advantages and disadvantages of these approaches requires comparing behavioral outcomes with multiple measures calculated on the same data set. Here we take a necessary first step to understand the utility of autoregressive features by comparing them to standard measures of functional connectivity in fMRI in performing a depression and anxiety classification task.

This paper is organized as follows. In Section II we describe the relevant methodology, including preparation of the data in Section II-A, the network based and autoregressive features in Section II-B, and the selected classification pipelines in Section II-C. The results of the evaluation of the classification pipelines are shown in Section III and the paper concludes in Section IV.

II. METHODS

A. Data Preprocessing

This paper considers a subset of the subjects from the Boston Adolescent Neuroimaging of Depression and Anxiety study [20], [21]. We refer the reader to [20] for details on the image acquisition. All experimental procedures involving human subjects were approved by the MIT Ethical Review Board. These data were preprocessed according to the Human Connectome Project minimal preprocessing guidelines [22]. There are 200 subjects, of which 63 are Controls, 64 are

*This work was not supported by any organization

¹S. Ruf, M. N. Akbar, and D. Erdoğmuş are with Cognitive Systems Laboratory, Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115, USA. sebastianruf@ece.neu.edu

² S. Ruf and S. Whitfield-Gabrieli are with the Department of Psychology, Northeastern University, Boston, MA 02115, USA.

Classifier	Dim. Reduction	Autoregression	Positive Strength	Negative Strength	Clustering	Auto+Pos	All Network	All Features
SVC	None	0.540	0.617	0.467	0.655	0.501	0.524	0.503
	PCA	0.506	0.558	0.427	0.483	0.430	0.563	0.620
	Best K MI	0.543	0.657	0.484	0.447	0.524	0.467	0.558
XGBoost	None	0.464	0.466	0.580	0.444	0.484	0.504	0.540
	PCA	0.637	0.353	0.466	0.483	0.561	0.315	0.558
	Best K MI	0.484	0.481	0.524	0.483	0.521	0.543	0.466

TABLE I: Balanced Accuracy on the test set for each combination of initial features, dimensionality reduction, and final classifier. Entries are bold if the classifier performed best in the comparisons described in the text and italicized if the classifier performed best over all attempted feature sets.

Classifier	Dim. Reduction	Autoregression	Positive Strength	Negative Strength	Clustering	Auto+Pos	All Network	All Features
SVC	None	0.671	1.596	-0.425	2.026	0.030	0.299	0.045
	PCA	0.070	1.031	-1.275	-0.340	-0.888	0.745	1.415
	Best K MI	0.535	1.915	-0.238	-0.766	0.299	-0.425	1.031
XGBoost	None	-0.674	-0.507	1.037	-1.792	-0.238	0.057	0.671
	PCA	1.741	-2.370	-0.507	-0.340	0.781	-3.887	1.031
	Best K MI	-0.238	-0.994	0.299	-0.340	0.345	0.535	-0.507

TABLE II: Number of standard errors away from a Balanced Accuracy of 0.5, which corresponds to only choosing one class on the test set for each combination of initial features, dimensionality reduction, and final classifier. Standard error was estimated via jackknife resampling. Entries are bold and italicized to match Table I.

diagnosed as Anxious, and 73 are diagnosed with depression or comorbid anxiety and depression. For the purposes of this comparison, we treat all patients as a single group, resulting in 137 patients and 63 controls.

Following a standard network based approach [23], pre-processed time series data was extracted in the 360-node Glasser Atlas [24]. Data was also standardized so that the mean is 0 and the standard deviation is 1 within each session to allow multiple sessions to be concatenated.

In order to calculate network metrics, a fully connected functional connectivity matrix was computed. In this network, the nodes correspond to areas of the brain as delineated in the Glasser atlas. The edges are measures of functional connectivity, i.e. the edge which connects node i and node j , w_{ij} , is the Fischer transformed correlation coefficient between the time series at the node i and node j .

B. Features

In order to compare the use of autoregression based features with network based features, we consider four sets of features which are extracted at node level, i.e. their dimension is determined by the number of nodes in the parcellation $n = 360$. The four features are the lag-1 autoregression coefficient $a \in \mathbb{R}^n$, the positive strength $p \in \mathbb{R}^n$, the negative strength $n \in \mathbb{R}^n$, and the clustering coefficient $c \in \mathbb{R}^n$.

Autoregressive Feature: To follow [8] and for the purpose of initial comparison, we consider here an AR model at lag 1. Inspection of the partial autocorrelation function suggests that for many of the nodal level time series this model order is optimal and for all the time series the lag one coefficient is the majority contribution to the model. Figure 1 shows the partial autocorrelation function for two regions of an example subject. Figure 1a shows the pattern common to most regions across subjects, with almost the full contribution coming at lag 1. Figure 1b shows that for some regions, higher order models would be required to adequately capture the autocorrelation structure. It bears exploring in future work,

how to deal with time series that require a higher order model as this would generate a heterogeneous number of features per region. If the activity of node i at time t is given by $x_{i,t}$ then the lag one AR model is given by $x_{i,t} = \phi_i x_{i,t-1} + \alpha_i$. As the time series has been standardized the offset α_i is expected to be near zero and is ignored. The lag one AR model was fit with Maximum Likelihood via the statsmodels [25] AutoReg function and the resulting AR feature vector a satisfies $a_i = \phi_i$.

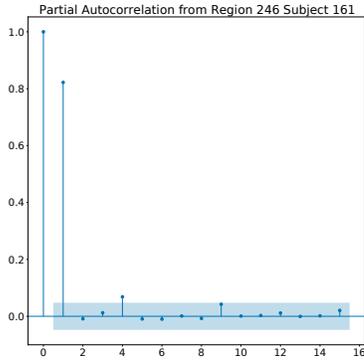
Strength The functional connectivity network when calculated is both signed and weighted. The edges of node i are defined as $w_{ij}, \forall j \in \{1, \dots, n\}$. To preserve the sign of the underlying network, the positive strength, p , and the negative strength, n , are calculated for each node. For node i the positive and negative strength are defined as $p_i = \sum_j \max(0, w_{ij})$ and $n_i = \sum_j \max(0, -w_{ij})$.

Clustering Coefficient Clustering coefficient is a measure which captures the local network structure by counting the number of triangles a node is part of [26], [27]. The clustering coefficient feature is calculated based on a variant of the clustering coefficient that was specifically developed for correlation networks as it takes into account both the sign and the weight of the underlying edges [28], as implemented in the Brain Connectivity Toolbox [29]. The weighted, signed correlation coefficient satisfies:

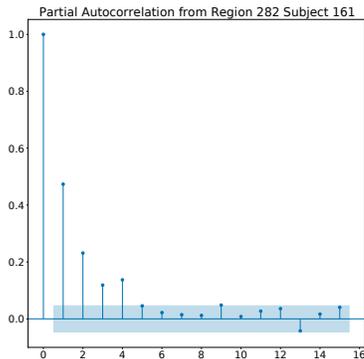
$$c_i = \frac{\sum_{j,q} (w_{j,i} w_{i,q} w_{j,q})}{\sum_{j \neq q} |w_{j,i} w_{i,q}|}$$

C. Classification Pipeline

First, the $N = 200$ samples were split 80 : 20 into a train and test set using class labels to ensure roughly equal distribution of patients and controls in each group. The resulting features from the 160 subjects in the train set were then standardized based on the train set, passed into one of three dimensionality reduction steps, and then one



(a)



(b)

Fig. 1: Partial Autocorrelation plot from two regions of Subject 161. 1a shows the more common pattern across regions, where a lag of 1 captures most of the autocorrelation. 1b shows that some regions would require multiple lags to account for the autocorrelation in the region time series.

of the two classifiers was trained. 10-fold cross validation (CV) was performed on the train set for hyperparameter optimization of the various stages of the pipeline. Due to the imbalanced nature of the dataset, the performance of pipelines was evaluated based on balanced accuracy:

$$\frac{1}{2} \left(\frac{tp}{tp + fn} + \frac{tn}{tn + fp} \right),$$

where tp is number of true positives, tn is number of true negatives, fp is number of false positive, fn is number of false negatives. A balanced accuracy of 0.5 is equivalent to always selecting one class, which would occur if the classifier only selected the majority class. The standard error of the balanced accuracy on the test set was calculated via jackknife resampling [30]. Unless otherwise noted, the algorithms are as implemented in scikit-learn [31].

The three dimensionality reduction steps were to either perform no dimensionality reduction, to perform Principal Component Analysis (PCA), or to use univariate Mutual Information (MI) to select informative features. Both the

number of components for PCA and the number of features to retain based on MI were selected from $\{2, 4, 6, 8, 10\}$.

The two classifiers are a Support Vector Machine (SVM) [32] and XGBoost (XGB) [33]. The support vector machine had as potential parameters, a linear or a radial basis function kernel as well as an $l2$ regularization parameter from $\{10, 1, \frac{1}{10}, \frac{1}{100}, \frac{1}{1000}\}$. XGBoost was tuned based on the max depth of trees, with possible parameters $\{1, 3, 5, 7\}$.

III. RESULTS

A comparison of the balanced accuracy on the test set for the various features, classifiers, and dimensionality reduction techniques is shown in Table I for the balanced accuracy and Table II for the number of standard errors away from a balanced accuracy of 0.5. Two comparisons were made to assess the impact of these autoregressive features. First, the four potential feature vectors were compared individually to see how the classifier would perform using only that feature. Then a comparison was made between performance with combinations of features, including the autoregressive features combined with the positive strength features as well as the combination of the three network features and the combination of all four features.

The classification performance of single features is shown in the Autoregression, Positive Strength, Negative Strength, and Clustering Coefficient columns of Table I. In all except the PCA+XGB classification pipelines, network based features provided the best performance. The individual network feature that provided the best performance varied by pipeline, with the positive strength providing the best balanced accuracy on 2 of 6 pipelines, negative strength providing best balanced accuracy on 2 of 6, and clustering coefficient on 1.

In the single feature comparisons, the positive strength outperformed the autoregressive features in 4 of 6 classification pipelines considered, all except the PCA+XGB and MI+XGB. To explore whether the combination of positive strength and autoregressive features would allow for greater performance, the two sets of features were input together into the classification pipelines. As shown in the Auto+Pos column of Table I, in all cases the balanced accuracy of the combined features under-performed compared to the individual features. While this behavior is to be expected in the absence of any dimensionality reduction, as an increase in the dimensionality of the input is likely to introduce more noise and hence decrease the accuracy, it is surprising that this holds even in the case of the PCA based pipelines.

The final comparison was between the classification performance on the three network based features compared to all four potential features. In this case, the inclusion of autoregressive features improved performance as compared to the three network features in 4 of 6 pipelines, all except the SVC and MI+XGB. The PCA+SVC pipeline using all features was the only case where not only did inclusion of the autoregressive features improve performance but the performance on combined features outperformed single features and, as shown in Table II, to a level greater than 1 standard error above choosing the majority class.

It is worth noting, based on Table II, that many of the classifiers achieve performance that is within ± 1 standard error of the chance level. This points to both the difficulty of the underlying classification problem and the need for further exploration of these classification pipelines before considering the neuroscientific implications of these results.

IV. CONCLUSION

In this paper, we considered whether autoregressive features would increase performance in a depression/anxiety classification task. For most classification pipelines considered, network based features provide better classification accuracy than autoregressive features, though in one pipeline the combination of network and autoregressive features provided the best performance. This shows that the classifier and dimensionality reduction step changes the impact of incorporating autoregressive features. Fully understanding the utility of using autoregressive features to study the brain will require further exploration.

REFERENCES

- [1] T. T. Liu, "Noise contributions to the fMRI signal: An overview," *NeuroImage*, vol. 143, pp. 141–151.
- [2] K. J. Friston, P. Jezzard, and R. Turner, "Analysis of functional MRI time-series," *Human Brain Mapping*, vol. 1, no. 2, pp. 153–171.
- [3] E. Bullmore, M. Brammer, S. C. R. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham, "Statistical methods of estimation and inference for functional MR image analysis," *Magnetic Resonance in Medicine*, vol. 35, no. 2, pp. 261–277.
- [4] W. Olszowy, J. Aston, C. Rua, and G. B. Williams, "Accurate autocorrelation modeling substantially improves fMRI reliability," *Nature Communications*, vol. 10, no. 1, p. 1220.
- [5] M. M. Monti, "Statistical Analysis of fMRI Time-Series: A Critical Review of the GLM Approach," *Frontiers in Human Neuroscience*, vol. 5.
- [6] M. R. Arbabshirani, E. Damaraju, R. Phlypo, S. Plis, E. Allen, S. Ma, D. Mathalon, A. Preda, J. G. Vaidya, T. Adali, and V. D. Calhoun, "Impact of autocorrelation on functional connectivity," *NeuroImage*, vol. 102, pp. 294–308.
- [7] H. Honari, A. S. Choe, J. J. Pekar, and M. A. Lindquist, "Investigating the impact of autocorrelation on time-varying connectivity," *NeuroImage*, vol. 197, pp. 37–48.
- [8] M. R. Arbabshirani, A. Preda, J. G. Vaidya, S. G. Potkin, G. Pearson, J. Voyvodic, D. Mathalon, T. van Erp, A. Michael, K. A. Kiehl *et al.*, "Autoconnectivity: A new perspective on human brain function," *Journal of Neuroscience Methods*, vol. 323, pp. 68–76, 2019.
- [9] H. Yang, X.-Y. Long, Y. Yang, H. Yan, C.-Z. Zhu, X.-P. Zhou, Y.-F. Zang, and Q.-Y. Gong, "Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional MRI," *NeuroImage*, vol. 36, no. 1, pp. 144–152.
- [10] D. Tolkunov, D. Rubin, and L. Mujica-Parodi, "Power spectrum scale invariance quantifies limbic dysregulation in trait anxious adults using fMRI: Adapting methods optimized for characterizing autonomic dysregulation to neural dynamic timeseries," *NeuroImage*, vol. 50, no. 1, p. 72. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2819748/>
- [11] K. J. Friston, "Functional and Effective Connectivity: A Review," *Brain Connectivity*, vol. 1, no. 1, pp. 13–36.
- [12] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302.
- [13] A. K. Seth, A. B. Barrett, and L. Barnett, "Granger Causality Analysis in Neuroscience and Neuroimaging," *The Journal of Neuroscience*, vol. 35, no. 8, pp. 3293–3297.
- [14] R. Liégeois, J. Li, R. Kong, C. Orban, D. Van De Ville, T. Ge, M. R. Sabuncu, and B. T. T. Yeo, "Resting brain dynamics at different timescales capture distinct aspects of human behavior," vol. 10, no. 1, p. 2317. [Online]. Available: <http://www.nature.com/articles/s41467-019-10317-7>
- [15] L.-L. Zeng, H. Shen, L. Liu, L. Wang, B. Li, P. Fang, Z. Zhou, Y. Li, and D. Hu, "Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis," *Brain*, vol. 135, no. 5, pp. 1498–1507, 2012.
- [16] H. He, J. Sui, Y. Du, Q. Yu, D. Lin, W. C. Drevets, J. B. Savitz, J. Yang, T. A. Victor, and V. D. Calhoun, "Co-altered functional networks and brain structure in unmedicated patients with bipolar and major depressive disorders," *Brain Structure and Function*, vol. 222, no. 9, pp. 4051–4064, 2017.
- [17] X. Wang, Y. Ren, and W. Zhang, "Depression disorder classification of fmri data using sparse low-rank functional brain network and graph-based features," *Computational and mathematical methods in medicine*, vol. 2017, 2017.
- [18] B. Yan, X. Xu, M. Liu, K. Zheng, J. Liu, J. Li, L. Wei, B. Zhang, H. Lu, and B. Li, "Quantitative Identification of Major Depression Based on Resting-State Dynamic Functional Connectivity: A Machine Learning Approach," 2020.
- [19] E. Jun, K. S. Na, W. Kang, J. Lee, H. I. Suk, and B. J. Ham, "Identifying resting-state effective connectivity abnormalities in drug-naïve major depressive disorder diagnosis via graph convolutional networks," *Human Brain Mapping*, vol. 41, no. 17, pp. 4997–5014, 2020.
- [20] V. Siless, N. A. Hubbard, R. Jones, J. Wang, N. Lo, C. C. C. Bauer, M. Goncalves, I. Frosch, D. Norton, G. Vergara, K. Conroy, F. V. De Souza, I. M. Rosso, A. H. Wickham, E. A. Cosby, M. Pinaire, D. Hirshfeld-Becker, D. A. Pizzagalli, A. Henin, S. G. Hofmann, R. P. Auerbach, S. Ghosh, J. Gabrieli, S. Whitfield-Gabrieli, and A. Yendiki, "Image acquisition and quality assurance in the Boston Adolescent Neuroimaging of Depression and Anxiety study," *NeuroImage: Clinical*, vol. 26, p. 102242, Jan. 2020.
- [21] N. A. Hubbard, V. Siless, I. R. Frosch, M. Goncalves, N. Lo, J. Wang, C. C. C. Bauer, K. Conroy, E. Cosby, A. Hay, R. Jones, M. Pinaire, F. Vaz De Souza, G. Vergara, S. Ghosh, A. Henin, D. R. Hirshfeld-Becker, S. G. Hofmann, I. M. Rosso, R. P. Auerbach, D. A. Pizzagalli, A. Yendiki, J. D. E. Gabrieli, and S. Whitfield-Gabrieli, "Brain function and clinical characterization in the Boston adolescent neuroimaging of depression and anxiety study," *NeuroImage: Clinical*, vol. 27, p. 102240.
- [22] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni *et al.*, "The minimal preprocessing pipelines for the human connectome project," *NeuroImage*, vol. 80, pp. 105–124, 2013.
- [23] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of Brain Network Analysis*. Academic Press.
- [24] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, S. M. Smith, and D. C. Van Essen, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171–178.
- [25] S. Seabold and J. Perktold, "statsmodels: Econometric and statistical modeling with python," in *9th Python in Science Conference*, 2010.
- [26] M. Newman, *Networks: An Introduction*, 1st ed. Oxford University Press.
- [27] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442.
- [28] G. Costantini and M. Perugini, "Generalization of Clustering Coefficients to Signed Correlation Networks," *PLOS ONE*, vol. 9, no. 2, p. e88669.
- [29] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069.
- [30] B. Efron and T. Hastie, *Computer age statistical inference*. Cambridge University Press, 2016, vol. 5.
- [31] 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," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [32] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [33] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794.