

Towards Quantifying Psychiatric Diagnosis using Machine Learning Algorithms and Big fMRI Data

Fahad Saeed

Department of Computer Science,
Western Michigan University, Kalamazoo MI 49008
Email: fahad.saeed@wmich.edu

Abstract

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common brain disorders among children and is very difficult to diagnose using current methods. Similarly other mental disorders are subject to the same systematic errors with sufficient evidence of diagnostic errors as well as over-prescribing of drugs due to misdiagnosis. For most mental health disorders there is no quantitative method that will inform the presence or absence of a given mental disorder. We argue that definitive and quantitative diagnostic tests are necessary for ADHD and other mental disorders. To this end, big data Functional Magnetic Resonance Imaging (fMRI) and machine learning algorithms can be instrumental in changing the way psychiatric disorders are diagnosed and treated. We briefly discuss our recent research efforts and future directions for a quantitative gold standard tests for psychiatric diagnosis.

Keywords: Machine Learning and Data Mining; ADHD; fMRI; Big Data; High Performance Computing;

Current and Future Research Directions

It is widely known that defining and diagnosing mental disorders is a difficult process due to overlapping nature of symptoms, and lack of a biological test that can serve as a definite and quantified gold standard [1]. To this end, Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common brain disorders among children and relies on the identification of abnormal mental characteristics. ADHD is notoriously difficult to diagnose, especially in children, with sufficient evidence of diagnostic errors as well as over-prescribing of drugs due to misdiagnosis [2]. The current psychiatric diagnosis is based purely on behavioural observation (DSM-5/ICD-10) but lacks biological and/or genetic validity and is prone to errors [1]. For most mental health disorders definitive and quantitative diagnostic tests which can detect the presence or the absence of a specific psychiatric disorder(s) are non-existent [1, 3].

Functional Magnetic Resonance Imaging (fMRI) is a non-invasive technique for studying the brain functional activities and is based on Blood Oxygen Level Dependent (BOLD) contrast [4]. During the fMRI scan, a series of images are taken using a scanner while the subject to a specific task such as resting, or doing various pre-determined tasks. The result of brain scanning is a series of low resolution images over time which shows the activity of the brain and can allow us to develop a highly informative brain connectome. There is evidence that ADHD, Bipolar Disorder (BD) and Schizophrenia have characteristics that differ in the regional and global connectivity of the brain when studied under resting state Functional Magnetic Resonance Imaging or fMRI [5, 1, 3].

As data-scientists, along with our colleagues from psychiatry and neuroscience, the overarching question that we have been posing is as follows: *Can we diagnose a person with ADHD (or other*

mental disorders) using fMRI scans using techniques from machine learning and novel algorithmic designs?. If possible, such a test will provide a definitive quantitative 'gold standard' test for diagnosing ADHD and other disorders.

In this endeavor machine-learning, especially deep-learning algorithms, have the potential to show exceptional promise [6, 7, 8, 9]. To this end, we have been successful in developing a machine learning algorithm that allow us to classify fMRI ADHD scans from normal healthy brain scans *without* using any demographic information. Our recently proposed technique is based on computing similarity between two multivariate time series along with k-Nearest-Neighbor classifier. We designed a model selection scheme called *J-Eros* which is able to pick the optimum value of k for k-Nearest-Neighbor from the training data. Our results show a 20% increase in accuracy, with superior sensitivity and specificity, as compared to the state of the art algorithms in classifying ADHD using open-data ADHD-200 that is available to the community [10]. We expect to introduce deep-learning based machine learning algorithms which will help us quantify psychiatric evaluation for ADHD, Bipolar Disorder, and Schizophrenia.

Compute time for analyzing big fMRI data is a bottleneck in introducing such techniques in a clinical setting [11, 12]. High performance computing (HPC) techniques will help ease the computational load in big fMRI data analysis for clinical and diagnostic purposes [13]. Among parallel computing techniques, Graphic Processing Unit (GPU) architecture provides a high degree of parallelism while being less expensive. Therefore, the focus of our research has been towards designing GPU based parallel computing algorithms for analyzing big fMRI data and considering pairwise relations between all *voxels* (the smallest cubic unit in fMRI data) without spatial or temporal constraint. In this way, voxels from different regions of brain which have similar spatiotemporal behaviour could be placed in the same cluster. The resulting clusters are regions which are functionally homogeneous and would give a more system-wide view of brain connectivity. This will allow neuroscientists to study the regional and global connectivity of the brain in much fine-grained detail than is currently possible. To this end, we have recently presented a GPU based strategy that allows us to calculate the functional connectivity using Pearson correlations for big fMRI data [4].

Such quantified psychiatric evaluations based on data science, big data, and algorithms can increase the accuracy of diagnosis, prognosis, and treatments of difficult to assess mental disorders. It will be an immense service for mankind if we are successful in developing a quantitative framework for big fMRI data to identify brain disorders from healthy brains. These algorithms will play a vital role in peering into mental disorders and will be akin to the role that microscope played for diagnostic medicine in the early 19th century.

Acknowledgements

This material is based on the work that was partially supported by National Science Foundation grants NSF CAREER ACI-1651724 and NSF CRII CCF-1464268.

References

- [1] N. C. C. for Mental Health (UK *et al.*, "Attention deficit hyperactivity disorder: diagnosis and management of adhd in children, young people and adults," British Psychological Society, 2018.
- [2] A. J. Dunlop and L. K. Newman, "Adhd and psychostimulantsoverdiagnosis and overprescription," *The Medical Journal of Australia*, vol. 204, no. 4, p. 139, 2016.

- [3] N. C. C. for Mental Health (UK *et al.*, “Bipolar disorder: The nice guideline on the assessment and management of bipolar disorder in adults, children and young people in primary and secondary care,” British Psychological Society, 2018.
- [4] T. Eslami and F. Saeed, “Fast-gpu-pcc: A gpu-based technique to compute pairwise pearsons correlation coefficients for time series datafmri study,” *High-Throughput*, vol. 7, no. 2, 2018.
- [5] M. Argyelan, T. Ikuta, P. DeRosse, R. J. Braga, K. E. Burdick, M. John, P. B. Kingsley, A. K. Malhotra, and P. R. Szeszko, “Resting-state fmri connectivity impairment in schizophrenia and bipolar disorder,” *Schizophrenia bulletin*, vol. 40, no. 1, pp. 100–110, 2013.
- [6] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, “Applications of deep learning and reinforcement learning to biological data,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–17, 2018.
- [7] B. Shi, Y. Chen, P. Zhang, C. D. Smith, J. Liu, A. D. N. Initiative, *et al.*, “Nonlinear feature transformation and deep fusion for alzheimer’s disease staging analysis,” *Pattern recognition*, vol. 63, pp. 487–498, 2017.
- [8] H.-I. Suk, S.-W. Lee, D. Shen, A. D. N. Initiative, *et al.*, “Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis,” *NeuroImage*, vol. 101, pp. 569–582, 2014.
- [9] S. Sarraf, G. Tofighi, *et al.*, “Deepad: Alzheimer s disease classification via deep convolutional neural networks using mri and fmri,” *bioRxiv*, p. 070441, 2016.
- [10] T. Eslami and F. Saeed, “Similarity based classification of adhd using singular value decomposition,” in *Proceedings of ACM International Conference on Computing Frontiers*, 2018.
- [11] S. Chen, L. Huang, H. Qiu, M. B. Nebel, S. H. Mostofsky, J. J. Pekar, M. A. Lindquist, A. Eloyan, and B. S. Caffo, “Parallel group independent component analysis for massive fmri data sets,” *PloS one*, vol. 12, no. 3, p. e0173496, 2017.
- [12] M. F. Stollenga, W. Byeon, M. Liwicki, and J. Schmidhuber, “Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation,” in *Advances in neural information processing systems*, pp. 2998–3006, 2015.
- [13] A. Phinyomark, E. Ibáñez-Marcelo, and G. Petri, “Resting-state fmri functional connectivity: Big data preprocessing pipelines and topological data analysis,” *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 415–428, 2017.