Original Research

A Convolutional Neural Network Model to Differentiate Attention Deficit Hyperactivity Disorder and Autism Spectrum Disorder Based on the Resting State fMRI Data

Azadeh Mozhdehfarahbakhsh[1,2], Saman Chitsazian[3], Prasun Chakrabarti[4], KS Jagannatha Rao[5], Babak Kateb[6*], Mohammad Nami[1,2, 5,6,8*]

[1]DANA Brain Health Institute, Iranian Neuroscience Society-Fars Chapter, Shiraz, Iran
[2]Senses Cultural Foundation, Sacramento, CA, USA
[3]Department of Computer Science, Faculty of Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran
[4]Provost, Techno India NJR Institute of Technology, Udaipur 313003, Rajasthan, India
[5]Centre for Neuroscience, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT AIP), Panama City, Panama
[6]Director, National Center for NanoBioElectoronics; Director, Brain Technology and Innovation Park; President,
Brain Mapping Foundation; Chairman and CEO, Society for brain mapping and Therapeutics and, Chairman of Neuroscience20-G20 summit, Los Angeles, CA, USA
[7]Neuroscience Laboratory, NSL (Brain, Cognition and Behavior), Department of Neuroscience, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran
[8]Inclusive Brain Health, Swiss Alternative Medicine, Geneva, Switzerland

*Corresponding authors:
Mohammad Nami MD, PhD, Department of Neuroscience, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz-7134814336- Iran, Email: torabinami@sums.ac.ir and Babak Kateb MD, Society for Brain Mapping and Therapeutics (SBMT), Los Angeles, CA, USA, Email: babak.kateb@worldbrainmapping.org

Abstract

Background and Objectives: Attention Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD) are the two most common neurodevelopmental disorders often with overlapping symptoms. Misdiagnosis of these disorders is the leading cause of a variety of problems including inappropriate interventions and improper treatment outcome. Over the last few years, resting state functional magnetic Resonance imaging (rs-fMRI) has received clinical attention among other beneficial brain scan techniques to extract functional connectivity in the brain. However, extracting useful information by human observation is prone to errors.

Material and Methods: The above unmet need prompted us to design the present investigation to construct a convolutional neural network model with 12 layers architecture in rsFMRI data aiming to differentiate the two conditions. The rs-fMRI data was collected from the ADHD-200 and ABIDE to feed into a convolutional neural network. Over the preprocessing phase, we have removed undesirable data and coordinated the remaining to MSDL atlas to recruit 39 regions of the brain.

Results: Ultimately, out results obtained a 0.92 accuracy, an AUC of 0.97 and loss of 0.17 in classification and discrimination of ADHD and ASD.

Conclusion: Though cross-validity with larger datasets is deemed required, the results obtained from the present investigation suggest that convolutional neural network may serve as a beneficial tool to differentiate ADHD and ASD from relatively small fMRI datasets. This further highlights the potential application of deep neural networks for serving the above purpose.

Keywords*: ADHD, ASD, Convolutional Neural Network, rs-fMRI

1. Introduction

Autism spectrum disorder (ASD) and attention-deficit/hyperactivity disorder (ADHD) are complex neurodevelopmental disorders [1] that their symptoms considerably have overlap. [2-4]
People who have hallmarks of both autism and ADHD often even face more serious challenges. They suffer from greater impairments in social and cognitive matters. Finding a way for differentiating two conditions accurately can have a profound effect on patients' life. This will lead to improved therapies and interventions.

In recent years, neuroimaging technology has become a point of discussion within the field of neuroscience. Such technology includes functional Magnetic Resonance Imaging (fMRI) [5,6], electroencephalography (EEG)

[7], Magnetoencephalography (MEG), and functional Near Infrared Spectroscopy (fNIRS) [8]. Especially, because of its powerful special resolusion, fMRI has been vastly applied in the research of neuroscience.

Deep Learning is one of the parts of machine learning technique which is different from traditional existing techniques and concerned with algorithms inspired by the structure and function of the brain neural network. Deep learning is capable of learning complex structures and achieving high levels of accuracy. It can acquire optimal representation through nonlinear transformation of large amounts of raw data. In the medical field, it has drawn attention in the field of radiology, and is employed for classification, diagnosis, risk factor analysis, prognosis, and prediction of treatment response. In the field of neuroimaging, it is used for classification and diagnosis of neurologic conditions such as stroke, neurodegenerative disorders, and psychiatric disorders. [9–16]

2. Issue Statement and Background Knowledge

A large and incrementing body of literature has investigated distinctiveness of neurodevelopmental disorders through machine learning and deep learning algorithms on neuroimaging data. A study of machine learning algorithm on fMRI for diagnosis of ASD by Moon et al (2019) has undertaken. The Results emerged from the data were the integrated sensitivity of 0.69, specificity to be 0.66, and AUC/pAUC of 0.71. [17]

Li et al (2018) for the purpose of comparison of HC and ASD developed a Multi-channel CNN on fMRI data, they reported 76.24 accuracy for the model. [6] In another study which set out to introduce Locally Linear Embedding for extracting neural activity information using BOLD time-series by Sidhu G et al. (2019) found diagnostic performance (>80%) for 11 datasets including patients with schizophrenia, ADHD, ASD, and NC. [19]

In the same vein, Xiao et al (2019) have developed in their study a methodology for the introduction of validate diagnostic classification using DL on the full brain frequency from resting-state fMRI data of 198 school-aged ASD children. The average diagnostic accuracy was found to be 96.26%, sensitivity of 98.03%, and specificity to be 93.62%.[20] Similarly, Aghdam et al (2019) found that by using CNN for resting state fMRI data of 5–10 years old ASD children, maximum diagnostic accuracy achieved 0.7273, sensitivity was 0.712, and specificity was 0.7348. [21] Xu et al (2019) draw our attention to a study in which fNIRS was used to examine whether ASD and typically developing children could be differentiated using a multi-layer neural network that combines CNN and gate recurrent unit. As a result, the model was able to discriminate between the ASD and NC groups with 90% accuracy, 81.6% sensitivity, and 98.5% specificity. [22]

While some research have highlighted the implications of neuroimaging and deep learning to differentially diagnose ASD and ADHD, none has suggested a deep learning model for the distinction of ASD and ADHD based on the rs-fMRI data across regions of interest as defined by the MSDL atlas.
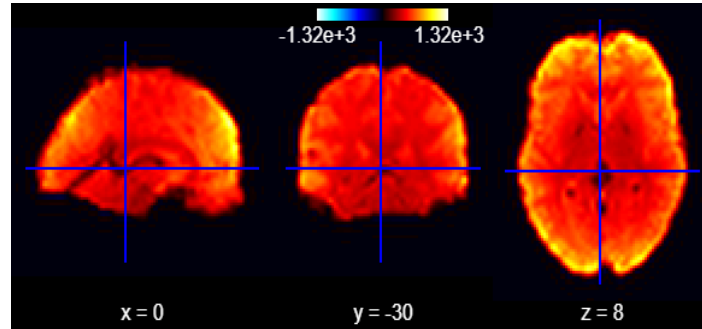
3. Research Methodology

3.1. Datasets

The preprocessed version of rs-fMRI data for 40 subjects subset of ADHD200 dataset were made available and wrapped up by Nilearn. Similarly, 40 preprocessed subject of ABIDE provided by the preprocessed connectome projects (PCP) were employed.
for More details are available at the ADHD-200 site (http://fcon_1000.projects.nitrc.org/indi/adhd200/) and ABIDE site(https://www.frontiersin.org/10.3389/conf.fninf.2013.09.00041/event_abstract/).
Sample of individuals was as follows: age between 8-22 years, right handed male and female.

Eighty 3D functional MRI scans of each subject's brain in different sizes were given in 4D spatio-temporal NIFTI (Neuroimaging Informatics Technology Initiative) format. Figure 1 shows functional scans of subject's brain.

Figure 1. The fMRI brain scan

3.2.   Data preparation for feeding into a deep neural network model

The deep neural network and preprocessing steps were written in libraries including Keras, Tensorflow, Nilearn and Scikit-Learn.
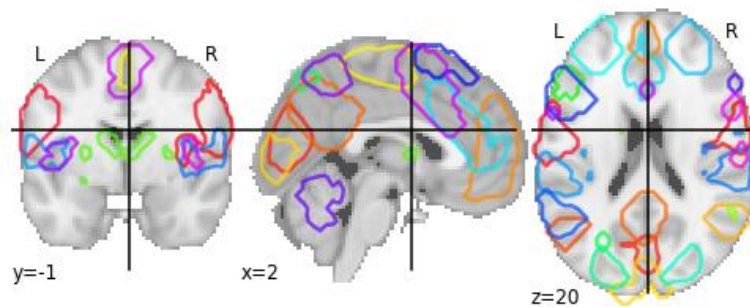
fMRI images are 4D matrices representing the three-dimensional space of voxels and time. Nilearn was applied for the purpose of loading and providing the 4D datasets for machine learning.

In the direction of removing the irrelevant data, masks were applied. On the side of gaining regions of interest in this study MSDL atlas (multi-subject dictionary learning) was adopted which defines a set of 39 ROIs throughout the brain. The atlas was transformed into the fMRI data by a fitted masker. In the interest of normalizing the data a masker, that extracted a 2D array, was constructed by Nilearn.

A primary concern is that timestamps for the datasets might not be uniform across the whole dataset. However, machine learning algorithms need a uniform shape throughout subjects in terms of homogenous scanning length. It was decided that the best solution to adopt for this purpose is utilizing padding via appending zeros to the end of each subject's scan.

After converting brain images to a data matrix, we can apply machine-learning to classify and differentiate Autism and ADHD disorders.

Figure 2. applying regions of interest using MSDL atlas
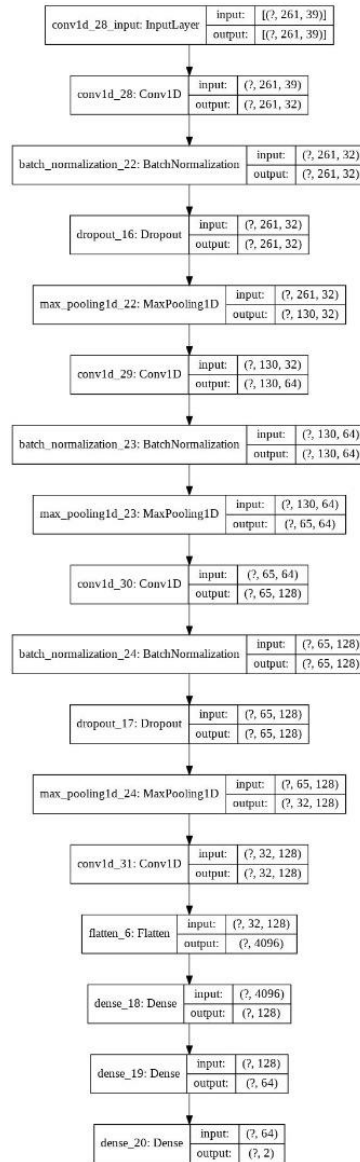

3.2.1.   Architecture of CNN Model

Over the past decade some research in Convolutional Neural Networks (CNN) has emphasized on fMRI data due to extracting automatically accurate and valid features [23].

According to **Figure 3** that reveals the scheme of Convolutional Neural Network, the architecture of CNN in this study totally had 12 layers consisting of four convolutional layers, the flatten layer, two dropout layers, two fully connected layers and an output layer. Three BatchNormalization were applied. Sigmoid function
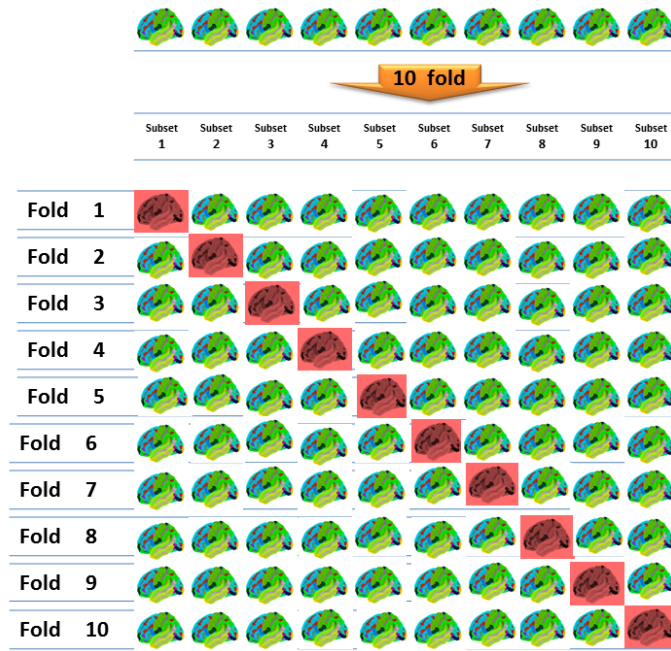
was utilized as the activation function in the output layer. The convolutional layers used kernels of [3×3]. Furthermore, MaxPooling and padding were employed with the convolution so that the output feature maps keep the size of the input. The nonlinear activation function of the convolutional layers and fully connected layers were ReLu function. For training and testing, the normalized data was fed into the constructed deep neural network. The total number of trainable parameters of the present CNN was 617218.

Figure 3. *Architecture of CNN method*

3.3.   K-Fold Cross validation:

Cross validation was to evaluate the significance of deep learning results across the whole data set and examine the performance of the predictive model. K-fold cross validation uses a complete dataset for training and validation. K-fold partition of the complete dataset is constructed so that K-1 folds are for training and the rest of data separated for validation. In this study, we applied the "K values of 10" in both methods to make bias small. Figure 4 illustrates the basic structure of the "k-fold cross validation method".
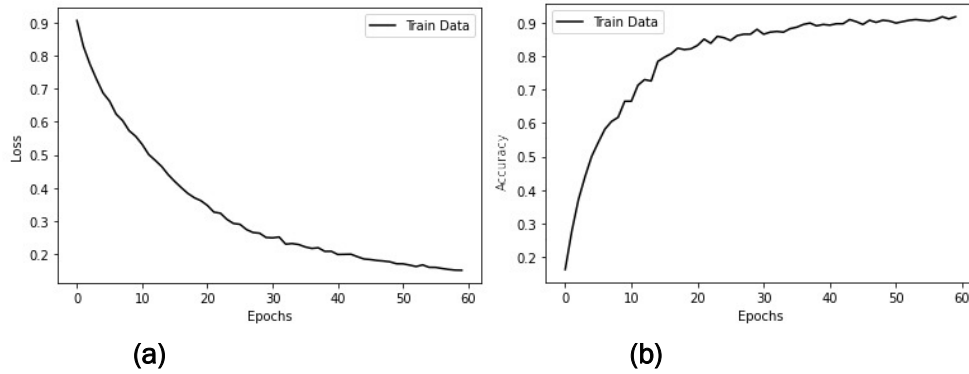
Figure 4. A 10-fold validation

4. Results

The proposed CNN model was implemented to ADHD-200 and ABIDE datasets. Python programming language and the Keras library was adopted with Tensorflow to create the CNN network. Furthermore, as mentioned above explanation, conducive to assess and improve validation of the model, the 10-fold cross validation was used for CNN model. The classification's results are presented as follows:

4.1. Experimental Results (CNN)

Figures 5(a) and 5(b) indicate the progress of training in CNN model, and the corresponding fluctuations in accuracy and loss metrics. Various evaluation criteria were investigated including: accuracy, AUC and Loss. 10-Fold Mean for loss, accuracy, AUC in CNN model after 60 epochs achieved 0.1795, 0.9261, 0.971 respectively.
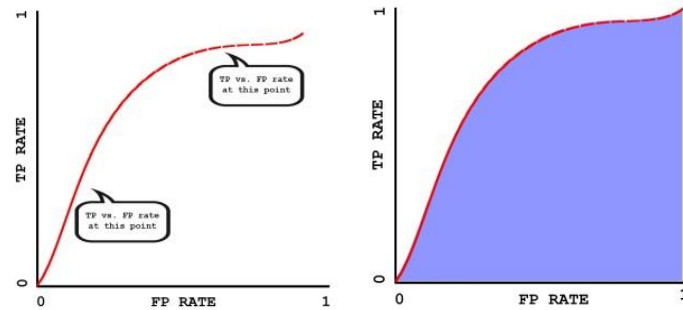
Figure 5. Accuracy and Loss of CNN model

(a)   (b)

AUC - ROC curve is a measurement for classification model performance at various threshold settings. It indicates how much the model correctly classifies the classes and distinguishes between them. High AUC is showing that the model is genuinely found classes as they are. By analogy, the Higher the AUC, the better the model distinguishes between patients with ADHD or ASD classes. This curve plots two parameters i.e. the True Positive and False Positive rates.

Figures 6 and 7 illustrate some of the main characteristics of the ROC curve. It simply plots True positives vs. false positives at different classes. By decreasing the classification threshold, our model classifies more samples as positive, and increasing that means higher False Positives and True Positives.
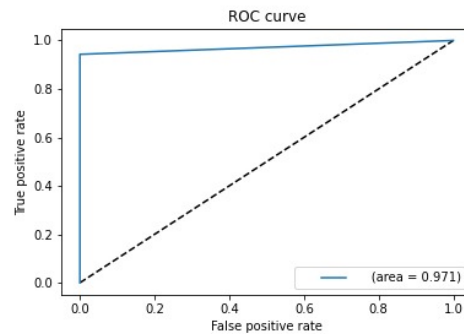
AUC or "Area under the ROC Curve." measures the whole two-dimensional area underneath the ROC from (0,0) to (1,1).

Figure 6. AUC provides a combined measure of performance across all possible classification thresholds.

AUC is a classification-threshold-invariant parameter. It measures the quality performance of the model's presentation in predictions irrespective of what classification threshold is chosen for the model. In the achieved results as has shown, it seems our predictive model has been successfully found an acceptable classification threshold that can leverage the model performance.

Figure 7. diagrams of AUC for CNN

5. Discussion

Two common neurodevelopmental diseases, which affect people of all ages, are ADHD and ASD that have received considerable critical attention. The two diseases have similarities in symptomatology so that it is difficult to differentiate them. Traditionally, doctors try to diagnose the disorders by performing examinations. Nevertheless, this method relies too heavily on clinical examinations and questionnaires (self-perceived assessment) that is prone to human error. Preliminary work on differentiating ASD and ADHD was undertaken using distinct brain patterns [24]. Regardless of the urgent need for diagnosis of effective biomarkers, detecting them has been insufficient and challenging. Recently, researchers have shown an increased interest in deep learning to extract useful information and biomarkers. Numerous studies have attempted to develop deep learning model to distinguish ADHD from healthy people [6,25–28] and techniques to detect ASD from non-ASD. [13,20,21]

Apart from that, there are few studies which investigated machine learning in neuroimaging data to differentiate ASD and ADHD using ABIDE and ADHD200 data sets. [29]

6. Conclusion

To our best knowledge, this is the first study to adopt CNN in rs-fMRI data aiming to develop a model to distinguish these disorders and coordinating MSDL atlas to determine the corresponding regions of interest. The information which has automatically been extracted from neuroimaging, is thusfar found to be potentially capaple in differentiating the two disease conditions.

Such preliminary results support the possibility that machine learning classifiers are appropriate to be used as timely, cost-effective, and accurate screening measures with the possibility to be installed on mobile platforms and provide a quick and accurate risk prediction model.

References

1. Kushki A, Anagnostou E, Hammill C, et al. Examining overlap and homogeneity in ASD, ADHD, and OCD: a data-driven, diagnosis-agnostic approach. *Transl Psychiatry*. 2019;9(1):318. doi:10.1038/s41398-019-0631-2
2. Handen BL, Aman MG, Arnold LE, et al. Atomoxetine, Parent Training, and Their Combination in Children With Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. *J Am Acad Child Adolesc Psychiatry*. 2015;54(11):905-915. doi:10.1016/j.jaac.2015.08.013
3. Randomized, controlled, crossover trial of methylphenidate in pervasive developmental disorders with hyperactivity. *Arch Gen Psychiatry*. 2005;62(11):1266-1274. doi:10.1001/archpsyc.62.11.1266
4. Rommelse N, Visser J, Hartman C. Differentiating between ADHD and ASD in childhood: some directions for practitioners. *Eur Child Adolesc Psychiatry*. Published online 2018. doi:10.1007/s00787-018-1165-5
5. Liu F, Guo W, Liu L, et al. Abnormal amplitude low-frequency oscillations in medication-naive, first-episode patients with major depressive disorder: A resting-state fMRI study. *J Affect Disord*. Published online 2013. doi:10.1016/j.jad.2012.10.001

6.   Li G, Liu M, Sun Q, Shen D, Wang L. Early diagnosis of autism disease by multi-channel CNNs. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. ; 2018. doi:10.1007/978-3-030-00919-9_35

7.   Zhang H, Su J, Wang Q, Liu Y, Good L, Pascual JM. Predicting seizure by modeling synaptic plasticity based on EEG signals - a case study of inherited epilepsy. *Commun Nonlinear Sci Numer Simul*. Published online 2018. doi:10.1016/j.cnsns.2017.08.020

8.   Vergotte G, Torre K, Chirumamilla VC, et al. Dynamics of the human brain network revealed by time-frequency effective connectivity in fNIRS. *Biomed Opt Express*. Published online 2017. doi:10.1364/boe.8.005326

9.   Song J-W, Yoon N-R, Jang S-M, Lee G-Y, Kim B-N. Neuroimaging-Based Deep Learning in Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. *Soa--ch'ongsonyon chongsin uihak = J child Adolesc psychiatry*. 2020;31(3):97-104. doi:10.5765/jkacap.200021

10.   Kamal H, Lopez V, Sheth SA. Machine learning in acute ischemic stroke neuroimaging. *Front Neurol*. Published online 2018. doi:10.3389/fneur.2018.00945

11.   Vieira S, Pinaya WHL, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci Biobehav Rev*. Published online 2017. doi:10.1016/j.neubiorev.2017.01.002

12.   Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: A review of machine learning applications. *NeuroImage Clin*. Published online 2018. doi:10.1016/j.nicl.2018.08.019

13.   Aarssen, L. W., & Crimi, L. (2016). Legacy, leisure and the 'work hard—Play hard' hypothesis. The Open Psychology Journal 9. Retrieved from aarssenl@queensu.ca, Abdollahi, A. (2013). Political psychology of the death terror. In S. J. Sinclair, D. Antonius, S. J. Sinclair (Ed), & D. Antonius (Ed) (Eds.) T political psychology of terrorism fears. (pp. 213–226). https://doi. org/10. 1093/acprof:oso/9780199925926. 003. 001., Abdollahi, A., Pyszczynski, T., Maxfield, M., & Luszczynska, A. (2011). Posttraumatic stress reactions as a disruption in anxiety-buffer functioning: Dissociation and responses to mortality salience as predictors of severity of posttraumatic symptoms. Psy 329–341. Retrieved from tpyszczy@uccs.edu, et al. Managing terror: Strategies Generation Y college students use for coping with homeland security threats. *Diss Abstr Int Sect B Sci Eng*. Published online 2017.

14.   Davatzikos C. Machine learning in neuroimaging: Progress and challenges. *Neuroimage*. 2019;197:652-656. doi:10.1016/j.neuroimage.2018.10.003

15.   Zaharchuk G, Gong E, Wintermark M, Rubin D, Langlotz CP. Deep learning in neuroradiology. *Am J Neuroradiol*. Published online 2018. doi:10.3174/ajnr.A5543

16.   Plis SM, Hjelm DR, Slakhutdinov R, et al. Deep learning for neuroimaging: A validation study. *Front Neurosci*. Published online 2014. doi:10.3389/fnins.2014.00229

17.   Moon SJ, Hwang J, Kana R, Torous J, Kim JW. 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 Heal*. 2019;6(12):e14108. doi:10.2196/14108

18.   Wang XH, Jiao Y, Li L. Identifying individuals with attention deficit hyperactivity disorder based on temporal variability of dynamic functional connectivity. *Sci Rep*. Published online 2018. doi:10.1038/s41598-018-30308-w

19.   Sidhu G. Locally Linear Embedding and fMRI Feature Selection in Psychiatric Classification. *IEEE J Transl Eng Heal Med*. 2019;7:2200211. doi:10.1109/JTEHM.2019.2936348

20.   Xiao Z, Wu J, Wang C, Jia N, Yang X. Computer-aided diagnosis of school-aged children with ASD using full frequency bands and enhanced SAE: A multi-institution study. *Exp Ther Med*. Published online 2019. doi:10.3892/etm.2019.7448

21.   Aghdam MA, Sharifi A, Pedram MM. Diagnosis of Autism Spectrum Disorders in Young Children Based on Resting-State Functional Magnetic Resonance Imaging Data Using Convolutional Neural Networks. *J Digit Imaging*. Published online 2019. doi:10.1007/s10278-019-00196-1

22.   Xu L, Geng X, He X, Li J, Yu J. Prediction in Autism by Deep Learning Short-Time Spontaneous Hemodynamic Fluctuations. *Front Neurosci*. Published online 2019. doi:10.3389/fnins.2019.01120

23.   Meszlényi RJ, Buza K, Vidnyánszky Z. Resting state fMRI functional connectivity-based classification using a convolutional neural network architecture. *Front Neuroinform*. Published online 2017. doi:10.3389/fninf.2017.00061

24.   Albajara Sáenz A, Septier M, Van Schuerbeek P, et al. ADHD and ASD: distinct brain patterns of inhibition-related activation? *Transl Psychiatry*. 2020;10(1):24. doi:10.1038/s41398-020-0707-z

25.   Yoo JH, Kim JI, Kim B-N, Jeong B. Exploring characteristic features of attention-deficit/hyperactivity disorder: findings from multi-modal MRI and candidate genetic data. *Brain Imaging Behav*. 2020;14(6):2132-2147. doi:10.1007/s11682-019-00164-x

26.   Deshpande G, Wang P, Rangaprakash D, Wilamowski B. Fully Connected Cascade Artificial Neural Network Architecture for Attention Deficit Hyperactivity Disorder Classification From Functional Magnetic Resonance Imaging Data. *IEEE Trans Cybern*. 2015;45(12):2668-2679. doi:10.1109/TCYB.2014.2379621

27.   Kuang D, Guo X, An X, Zhao Y, He L. Discrimination of ADHD Based on fMRI Data with Deep Belief Network. In: *ICIC*. ; 2014.

28.   A.J. Hao C.H. Yin BLH. Discrimination of ADHD children based on Deep Bayesian Network. Published online 2015:6 .--6 .

29.   Jung M, Tu Y, Park J, et al. Surface-based shared and distinct resting functional connectivity in attention-deficit hyperactivity disorder and autism spectrum disorder. *Br J Psychiatry*. Published online 2019. doi:10.1192/bjp.2018.248