Towards Autism detection on brain structural MRI scans using deep unsupervised learning models

Mélanie Garcia[1,2], **Jean-Marc Orgogozo**[1], **Clare Kelly**[2], **Margaux Luck**[3]
[1] HyperCube Institute, Paris
[2] Trinity College Dublin
[3] University of Montréal, Mila

Correspondence: melanie.garcia@institut-hypercube.org

Abstract

Autism Spectrum Disorder (ASD) is a relatively common neurodevelopmental condition that for which we currently lack any objective biomarkers. The study of patient brain MRI data has the potential to reveal regions of dysfunction that may serve as biomarkers to supplement current clinician-based diagnoses. In this paper, we propose a method that enhances the diagnosis of ASD by compressing structural MRI from the open science Autism Brain Imaging Database Exchange (ABIDE, 892 ASD, 972 non-ASD) to obtain a representation of the brain that is relevant for the prediction of ASD using unsupervised deep learning models. Our experimental evaluation demonstrates promising performance on the task of automated ASD diagnosis on ABIDE.

1 Introduction

Autism Spectrum Disorder (ASD) is a relatively common neurodevelopmental condition that for which we currently lack any objective biomarkers [Goldani et al., 2014]. Indeed, diagnosing ASD requires the expert integration of observations from the family, school educators, and an appropriate medical team. Yet early identification is crucial to facilitate early intervention and better long-term outcomes [Reichow & Wolery, 2009]. Quantifiable brain imaging metrics may help us to diagnose ASD earlier and more accurately. In particular, we study structural brain MRI data in an effort to highlight key morphology differences between individuals which may correlate with abnormal cognitive, sensory, or motor function [Chen et al., 2011; Jiao et al., 2010; Blackmon et al., 2016].

Several studies have investigated the use of handcrafted extracted correlated features with autism from MRI scans in a machine learning framework to build a diagnosis model for ASD learned from small set of labeled data as in [Di Martino & al., 2014]. However, the subjectivity of the feature extraction procedures as well as the size of the cohorts have lead to many conflicting results showing that the non heterogeneity of the data used and the lack of consistency of the features extraction procedures may affect model performances and limit comparison across studies [Chen et al., 2011]. To reduce the subjectivity of the feature extraction procedure Heinsfeld et al. [2018] used stacked denoising autoencoders [Vincent et al., 2010] for learning low-dimensional representations of functional MRI scans and used these representations for training a classifier for ASD diagnosis showing significant improvement when compared to a classifier only trained on extracted features.

In this study, we propose to investigate new ways of assessing ASD using structural MRI (rather than functional MRI) and an using deep unsupervised learning models on data from the world-wide multi-site Autism Brain Imaging Data Exchange (ABIDE) I [Di Martino & al., 2014] & II [Di Martno & al., 2017].

Medical Imaging meets NeurIPS. NeurIPS 2019 Workshop.

To our knowledge this is the first attempt to use such deep learning models on structural MRI scans on ABIDE. Our experimental evaluation demonstrates promising performance on the task of automated ASD diagnosis.

2 Experiments/Methods

2.1 Data

The data used in this study are structural MRIs and phenotypic data from the Autism Brain Data Exchange I [Di Martino & al., 2014] and II [Di Martno & al., 2017]. This initiative has aggregated functional and structural brain imaging data collected from laboratories around the world to investigate the neural basis of Autism. We used the Configurable Pipeline for the Analysis of Connectomes (C-PAC) [Craddock et al., 2013] as a tool for the preprocessing step to make MRI comparable between each other: we transformed structural MRIs with deobliquing, reorienting, skull-stripping, intensity normalization and registration with the MNI 152 template [Fonov et al., 2011, 2009].

We used 1607 (86% of the dataset) MRI scans to train our unsupervised deep learning models and 275 (14%) for the holdout test set. In both sets there were the same proportion of data from ABIDE I and ABIDE II, and the same proportion of autism cases. The training and test sets were contributed by different data collection sites. Images were grid sampled with a window size of (32, 32, 32), each volume sample corresponded to a location on the brain and was input of our models. There were 36 sampled locations per image.

2.2 Method and architecture

Our model is designed such that learning is split into two parts: a feature extractor and a classifier. The feature extractor is a Variational Autoencoder (VAE) [Kingma & Welling, 2013; Rezende et al., 2014] or with Adversarial Learned Inference (ALI) [Dumoulin et al., 2016]. We used a common architecture for the two networks, inspired from the one used in the maxout network study [Goodfellow et al.], which was also tested in the ALI study [Dumoulin et al., 2016]. Specifically, the discriminator consists of convolution and max pooling blocks, followed by maxout layers.

The classifier is the interpretation with respect to ASD prediction. For each brain location, we trained different classifiers on latent data with a binary target autistic / non-autistic and compared them: elascticnet penalized logistic regression with a L1 ratio of 0.5 or a one-layer perceptron with 10 neurons. This resulted in a probability and the prediction of a class for each location which was used to train a second classifier. The second classifiers trained on the first predictions were a simple majority rule classifier, an elasticnet regression with a L1 ratio of 0.5, or a one-layer perceptron with 5 neurons.

3 Results

We trained ALI and VAE on five epochs on the train set, with a validation step on 10% of the set every 10 iterations. According to the training loss values, the two algorithms reached an asymptote. We used the ROC AUC score and accuracy as metrics to evaluate our models.

The first classifier on each cerebral location enabled us to find locations responsible for the prediction of ASD. We noticed that the highlighted brain regions vary according to the pipeline of feature extraction used, which includes the unsupervised deep learning algorithm and the first linear or non-linear classifier. With VAE as MRI compressor, we noticed that the locations involved in ASD predictions corresponded to the frontal lobe known to be linked with planning and attention, and to the occipital lobe known to be linked with vision recognition. With ALI as MRI compressor, the occipital lobe was again a salient region for predicting ASD, but also the parietal lobe known to be linked with touch. To obtain more accurate results, it would probably be necessary to review the grid sampling step and use multi-scale algorithms.

For the second classifier, particularly for the logistic regression and for the one-layer perceptron, we reduced the training set to the 10 brain sampled locations inferring the highest standard-deviated predictions, according to each combination of extractor and classifier. The threshold for the majority rule was more than half of the locations predicting ASD. The logistic regression was trained with

4-folds cross-validation to optimize the strength of regularization, and a final refit was made on the whole train set to get the coefficients. We inferred predictions on the whole test set and detailed the scores on subsets corresponding to each data provider in the test set.

The resulting scores of performance are shown in figure 1 and we can observe significant variations between pipelines and the providing site concerned. The best accuracy and ROC AUC scores for every site provide a state-of-the-art for detection of Autism with with brain structural MRI on multi-site study using ABIDE [Di Martino & al., 2014; Di Martno & al., 2017]. Our results showed that data collected at different centers yielded widely varying results. Further analysis is needed to understand and control this variance.

Unsupervised extractor	VAE						ALI					
Classifier 1	Elasticnet			OLP			Elasticnet			OLP		
Classifier 2	MR	ER	OLP	MR	ER	OLP	MR	ER	OLP	MR	ER	OLP
Train	56 %	0,64 (59%)	0,65 (59%)	53 %	0,73 (66%)	0,74 (66%)	55 %	0,66 (60%)	0,66 (61%)	53 %	0,57 (55%)	0,60 (55%)
Test	54 %	0,51 (51%)	0,51 (51%)	50 %	0,5 (47%)	0,51 (50%)	44 %	0,50 (51%)	0,51 (51%)	51 %	0,49 (49%)	0,48 (49%)
Per site :												
ABIDE II, Inst. Pasteur and R. Debré hospital 1	52 %	0,40 (46%)	0,40 (46%)	**62 %**	0,41 (52%)	0,45 (54%)	38 %	**0,63 (52%)**	0,63 (52%)	**62 %**	0,47 (38%)	0,41 (38%)
ABIDE I, Social Brain Lab	63,3 %	0,76 (67%)	**0,76 (70%)**	50 %	0,56 (47%)	0,55 (50%)	47 %	0,56 (47%)	0,56 (43%)	50 %	0,50 (50%)	0,45 (50%)
ABIDEII, Univ. of California Davis 1	54 %	0,47 (46%)	0,46 (46%)	46 %	0,42 (39%)	0,46 (46%)	43 %	0,48 (46%)	0,48 (46%)	46 %	**0,61 (57%)**	0,45 (57%)
ABIDE I, Univ. of California Los Angeles, sample 2	38 %	0,57 (54%)	0,57 (50%)	50 %	0,53 (50%)	0,46 (50%)	46 %	0,46 (50%)	0,46 (46%)	50 %	**0,72 (62%)**	**0,79 (50%)**
ABIDE I, Univ. of California Los Angeles, sample 1	57 %	0,45 (50%)	0,45 (50%)	43 %	0,55 (50%)	0,57 (51%)	60 %	0,57 (60%)	0,57 (60%)	43 %	0,56 (54%)	**0,62 (57%)**
ABIDE II, Erasmus MC 1	53 %	0,57 (51%)	0,57 (51%)	51 %	0,45 (43%)	0,48 (47%)	27 %	0,40 (43%)	0,41 (45%)	**57 %**	0,42 (43%)	0,46 (41%)

Each case : ROC AUC score (Accuracy %)
MR = Majority Rule, accuracy only
ER = Elasticnet Regression
OLP = One-Layer Perceptron

Figure 1: Resulting scores from the different pipelines of brain MRI scan compression and binary classification on the presence of ASD; ROC AUC score (accuracy score%) are given for the whole train and test set, and detailed for each site in the test set.

4 Limitations/Conclusion

This work is a proof of concept that deep learning could be useful for ASD diagnosis, and could raise new paths of research on finding Autism neurological markers. This pipeline on brain structural MRI to detect Autism give promising results, providing weak locations of the part of the brain involved. Nevertheless, these findings are very preliminary and should be compared with more classic models as baselines. The approach could also be improved by taking into account the data collection sites and other modalities like functional MRI during training.

5 Acknowledgement

We would like to thank Martin Weiss our colleague at Mila for reviewing and editing this manuscript.

References

Blackmon, K., Ben-Avi, E., Wang, X., Pardoe, H. R., Di Martino, A., Halgren, E., Devinsky, O., Thesen, T., and Kuzniecky, R. Periventricular white matter abnormalities and restricted repetitive behavior in autism spectrum disorder. *NeuroImage: Clinical*, 10:36–45, 2016.

Chen, R., Jiao, Y., and Herskovits, E. H. Structural mri in autism spectrum disorder. *Pediatr Res*, 69 (5 Pt 2):63R–8R, 2011.

Craddock, C., Sikka, S., Cheung, B., S. Ghosh, S., Khanuja, R., Yan, C., Li, Q., Lurie, D., Vogelstein, J., Burns, R., Colcombe, S., Mennes, M., Kelly, C., Di Martino, A., Castellanos, F. X., and Milham, M. Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac). *Neuroinformatics*, 2013.

Di Martino, A. and al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 2014.

Di Martno, A. and al. Enhancing studies of the connectome in autism using the autism brain imaging data exchange ii. *Scientific Data*, 2017.

Dumoulin, V., Belghazi, I., Poole, B., Mastropietro, O., Lamb, A., Arjovsky, M., and Courville, A. Adversarially learned inference. *arXiv preprint arXiv:1606.00704*, 2016.

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., Group, B. D. C., et al. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*, 54(1):313–327, 2011.

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C., and Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, (47):S102, 2009.

Goldani, A. A., Downs, S. R., Widjaja, F., Lawton, B., and Hendren, R. L. Biomarkers in autism. *Frontiers in psychiatry*, 5:100, 2014.

Goodfellow, I. J., Warde-Farley, D., Courville, A., and Bengio, Y. Maxout networks.

Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., and Meneguzzi, F. Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: Clinical*, 17: 16–23, 2018.

Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z., and Herskovits, E. H. Predictive models of autism spectrum disorder based on brain regional cortical thickness. *Neuroimage*, 50(2):589–599, 2010.

Kingma, D. P. and Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Reichow, B. and Wolery, M. Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the ucla young autism project model. *Journal of Autism and Developmental Disorders*, 39(1), 2009.

Rezende, D. J., Mohamed, S., and Wierstra, D. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P.-A. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(Dec):3371–3408, 2010.