Predictive Models for Autism Spectrum Disorder Based on Multiple Cortical Features

Yun Jiao and Zuhong Lu
Key Laboratory of Child Development and Learning Science
School of Biological Science and Medical Engineering
Southeast University, Nanjing, China, 210096
Telephone: +86-25-83795664

***Abstract*—Autism spectrum disorder (ASD) is a neurodevelopmental disorder with a wide phenotypic range, often affecting personality and communication. Previous surface-based morphometry (SBM) work only used cortical thickness to generate classification model between ASD and normal control. The goals of this study were twofold: 1) to construct predictive models for ASD, based on combination any of these five cortical measurements (cortical thicknesses (T), mean curvature (MC), Gaussian curvature (GC), folding index (FI), and curvature index(CI)) extracted from SBM, 2) and to compare these models. Our study included 22 subjects with ASD (mean age 9.2±2.1 years) and 16 volunteer controls (mean age 10.0±1.9 years). Using SBM, we obtained T, MC, GC, FI and CI for 66 brain structures for each subject. Then, we combined any of these five cortical measurements and obtained 31 combinations as classification inputs. To generate predictive models, we employed three machine-learning techniques: support vector machines (SVMs), functional trees (FTs), and logistic model trees (LMTs). We found that "thickness + mean curvature"-based classification model was modest superior to that based on thickness-based features when LMT employed , and curvature only provide limited information to thickness in ASD predictive model. Our result suggested that ASD may be cortical thickness abnormal disorder rather than cortical curvature abnormal disorder.**

***Index Terms*—Autism spectrum disorder; Surface-based morphometry; Cortical thickness; Cortical curvature; Predictive model**

I. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with a prevalence of approximately in 1 in 150 children [1]. Children with ASD have abnormal social behavior, impaired communication and language skills, and repetitive/stereotyped behavior [2].

Previous voxel-based morphometry (VBM) studies, which measures voxel-wise gray- and whitematter volume changes across the entire brain, have suggested that children with ASD have subtle structural changes in many brain structures, including the frontal lobe, parietal lobe, hippocampus, amygdala, cerebellum and brain stem [3], [4], [5].

However, the intrinsic topology of the cerebral cortex is that of a 2-D sheet with a highly folded and curved geometry of gray matter, which lies between cerebrospinal fluid white matter [6]. As a result, surface-based morphometry (SBM) may reflect intrinsic cortical topography better than widely used VBM [7].

There have been several SBM studies of ASD. For example, Hadjikhani et al. [8] reported thickness differences in the mirror-neuron system and other areas involved in social cognition in individuals with ASD. Nordahl et al. [9] demonstrated cortical folding abnormalities in individuals with ASD, primarily in the left operculum, bilateral parietal operculum, and bilateral intraparietal sulcus. Our previous work [7] found that children with ASD had decreased cortical thickness in the left and right pars triangularis, left medial orbitofrontal gyrus, left parahippocampal gyrus, and left frontal pole, and increased cortical thickness in the left caudal anterior cingulate and left precuneus.

SBM can derive features such as regional gray-matter thickness and regional surface area [10], as well as mean curvature (MC), Gaussian curvature (GC), folding index (FI), and curvature index (CI) (details were described in Methods part). These features may provide complementary information to thickness. So far as we know, there was no study focus on analysis the multiple contributions of these cortical features to ASD.

Currently, ASD is diagnosed based on behavioral criteria. Given the SBM findings described above, an MR-based diagnostic model holds the promise of enhancing, perhaps complementing, behavioral assessment [7]. Toward this end, Singh et al. [11] developed a diagnostic model generated by the LPboost-based algorithm to distinguish autistic children from control subjects, based on voxel-wise cortical thickness, based on approximately 40,000 points for each subject; they reported 89% classification accuracy based on cross-validation. The principal limitation of their work was basing the feature dimension reduction step on all samples outside cross-validation. Our previous work [7] generated a ASD predictive model based on logistic model tree using reginal cortical thickness, and obtain a 87% accuracy classify based on cross-validation. So far as we know, there also was no study focus on generate predictive model base on multiple cortical features.

In this study, we test the hypothesis that predictive models can distinguish children with ASD from controls based on multiple cortical features (cortical thickness (T), MC, GC, FI, and CI), and the models by combining any of these cortical features can be more accuracy than models generated by single feature of cortical thickness. To test this hypothesis, we first computed 5 cortical features of 66 structures defined on a

978-1-61284-181-6/11/$26.00 ©2011 IEEE

brain atlas, for each subject. We then applied three data-mining approaches to generate four diagnostic models based on combination of these 5 features. Finally, we compared all performance metrics of cortical feature-based predictive models.

II. Methods

A. Participants

Participants in this study, aged 6-15 years, consisted of two groups: 22 children with ASD (mean age, 9.2 $\pm$ 2.1 years), and 16 volunteer control subjects (VC) (mean age, 10.0 $\pm$ 1.9). Children with ASD and control subjects were group-matched on age, sex, full-scale IQ, handedness, weight, height, and socioeconomic status. All participants with ASD were recruited by the Child Mental Health Research Center of Nanjing Brain Hospital. The diagnosis of ASD was based on the criteria of the fourth edition of Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [2], the Autism Diagnostic Inventory-Revised (ADI-R) [12], and the Childhood Autism Rating Scale (CARS) [13]. IQ scores were obtained using the Wechsler Intelligence Scale for Children-II (Chinese version). Volunteer control subjects were recruited from the local community in Nanjing. None of the VC subjects had a history of axis I or II psychiatric disorder. Exclusion criteria for all subjects included history of seizure, head trauma, genetic or neurological disorder, major medical problem, and fullscale IQ less than 70. Each participant's parents gave informed consent, and all research procedures were approved by Institutional Review Board of Nanjing Brain Hospital of Nanjing Medical University.

B. Magnetic resonance imaging protocol

We acquired MR images at Nanjing Brain Hospital, on a 1.5-Tesla Signa GE instrument (NVi, General Electric Medical System, Milwaukee, WI), using a standard quadrature head coil. For the purposes of this study, we acquired high-resolution images for surface-based analysis based on a T1-weighted three-dimensional spoiled gradient-echo sequence with the following parameters: TR = 9.9 ms, TE = 2.0 ms; flip angle = 15; FOV = 24 cm; slice thickness = 2.0 mm; in-plane resolution = 0.94$\times$0.94 mm; matrix = 256$\times$256; number of slices = 132 contiguous; and number of excitations = 1.0. MR images were reviewed by an experienced radiologist for quality. We excluded subjects with poor MR image quality.

C. Image precessing

SBM arthogram provides us with an array of cortical topographical measurements [6]. The definitions of these descriptors are as follows:

1) Average thickness (T) is the average of the thicknesses for each vertex in the region [14];

2) Mean curvature (MC) is the average mean curvature (the average of the principal curvatures at each vertex) for each region;

3) Gaussian curvature is the average Gaussian curvature (GC) (the product of the principal curvatures at each vertex) for each region;

4) Folding index (FI) is the Van Essen folding index for the region (i.e., the integral of the product of the maximum principal curvature and the difference between maximum and minimum curvature divided by 4k) [15];

5) Curvature index (CI) is the Van Essen curvature index for the region (i.e., the integral across all regions of positive intrinsic curvature divided by 4k) [15].

It should be noted that the above cortical topographical measurements computed by the SBM arthogram are based on the surfaces created by the segmentations and parcellations. We used a cortical atlas that is based on statistics computed from manually labeled cortical regions; this atlas divides cerebral cortex (cerebellum and brain stem were removed and not analyzed in FreeSurfer) into 66 structures (33 structures for each hemisphere) [16]. The processes of above cortical features are detail in [6], [16]

D. Data mining

A classification model includes two components: the structural form of the model, S, and model parameters, θ.

The categories of combination any of the five features (T, MC, GC, FI, and CI) were $C_5^1+C_5^2+C_5^3+C_5^4+C_5^5 = 31$. We respectively used 31 combinations of these cortical features as classification inputs. As described in the previous work [7], we divided the cerebral cortical surface into 66 regions, including both left and right hemispheres, according to surface-base labeling. The input data to the diagnostic model generation methods therefore consisted of 38 instances of 66 predictor variables with the group-membership variable assuming the label "ASD" or "VC".

To avoid model-generation bias, we applied three machine-learning methodsłsupport vector machines (SVMs), functional trees (FTs), and logistic model trees (LMTs)łto generate diagnostic models. We selected SVMs [17] because they have been shown empirically to achieve good generalization performance on a wide variety of classification problems [18].We selected both FT and LMT approaches because they can be seen as advanced versions of decision trees, and performed good in our previous thickness models [7]. The model generated by FT or LMT is declarative and simple to interpret, and more accurate than that generated by standard version decision trees, such as those generated by C4.5. FT is based on standard tree-based algorithms, such as CART and C4.5, yet FT has greater classification accuracy than these approaches [19]. LMT is based on the concept of a model tree, which is similar to an ordinary regression tree, except that it constructs a piece-wiselinear, rather than a piecewise-constant, approximation to the target function; LMT builds a single tree containing only relevant attributes for classification, is more accurate than C4.5 and standalone logistic regression, and does not require tuning of parameters [20].

We evaluated the diagnostic models generated by these three machine-learning algorithms based on 10-fold cross-validation

TABLE I
CLASSIFICATION ACCURACIES FOR COMBINATION ANY OF CORTICAL FEATURES

Features	SVM (%)	FT (%)	LMT (%)
Thickness (T)	82	84	87*
Mean Curvature (MC)	50	58	50
Guassian Curvature (GC)	45	45	42
Folding Index (FI)	50	66	50
Curvature Index (CI)	63	55	53
T+MC	74	76	89*
T+GC	71	76	84
T+FI	74	71	74
T+CI	84	79	79
MC+GC	47	45	50
MC+FI	53	66	66
MC+CI	61	61	50
GC+FI	50	61	58
GC+CI	58	53	53
FI+CI	66	53	50
T+MC+GC	63	76	74
T+MC+FI	82	84	82
T+MC+CI	84	76	82
T+GC+FI	68	79	76
T+GC+CI	68	79	79
T+FI+CI	74	79	82
MC+GC+FI	55	63	50
MC+GC+CI	55	50	53
MC+FI+CI	58	63	58
GC+FI+CI	58	50	55
T+MC+GC+FI	79	79	79
T+MC+GC+CI	71	74	74
T+MC+FI+CI	79	79	79
T+GC+FI+CI	74	74	74
MC+GC+FI+CI	61	61	53
ALL	74	76	76

(CV), because 10-fold CV is commonly used in machine-learning areas.

III. RESULTS

Demographic characteristics of subjects in the ASD and VC groups were listed in our previous work [7]. We did not find significant differences between groups in gender (χ^2- test p-value=0.67), age (t-test p-value=0.25), WISC full-scale IQ (t-test p-value=0.44), weight (t-test p-value=0.12), or height (t-test p-value=0.16).

There are 31 categories of combination of five cortical features (See first columns of Table I). The classification results of SVM, FT, and LMT list in second, third and forth columns of table I, respectively. All accuracies were based on 10-fold CV.

We found that the combination of T and MC obtained the modest highest accuracies of 89% when using LMT, however, other combination didn't get higher accuracies than only thickness (87%) used as classification inputs.

TABLE II
PREDICTIVE VARIABLES SELECTED BY LMT USING COMBINATION OF THICKNESS (T) AND MEAN CURVATURE (MC)

Cortical structures	Left	Right
pars triangularis	T*	T*
para hippocampal	T*	-
superior frontal	-	T
caudal anterior cingulate	T*	MC
medial orbitofrontal	T*	-
precuneus	T*	-
frontal pole	T*	
pars orbitalis	-	MC

TABLE III
AVERAGE MEAN CURVATURE OF STRUCTURES SELECTED BY LMT FOR T + MC COMBINATION CLASSIFICATION, AND TWO-SAMPLE T-TEST p-VALUES (NOT ADJUSTED FOR MULTIPLE COMPARISONS) TO DETERMINE WHETHER SIGNIFICANT DIFFERENCES EXIST BETWEEN GROUPS. R: RIGHT

Structures	ASD	VC	p-value
R caudal anterior cingulate	0.16±0.01	0.17±0.02	0.32
R pars orbitalis	0.21±0.03	0.19±0.02	0.03

The model of combination of T and MC by LMT had 10 predictive variables: 8 thickness features and 2 curvature features Details list in Table II. In Table II, the variables marked with "*" were also the predictive variables when only use thickness feature in LMT.

Table III shows average mean curvature of structures selected by LMT for T + MC combination classification, and two-sample t-test p-values (not adjusted for multiple comparisons) to determine whether significant differences exist between groups.

We computed the two-sample t statistic to determine whether the mean curvature were significant between groups. We only found one structure – pars orbitalis – were significant with p-value of 0.03. In our previous work, we also computed the two-sample t statistic of thickness, and there were 17 structures significant between groups [7].

IV. DISCUSSION

To our knowledge, there has been no previous attempt to compare multiple cortical features (T, MC, GC, FI, and CI) classification of ASD. We found that T + MC classification was more accurate than any other combination classification.

We used three machine-learning methods to generate classifiers, in order to avoid bias with respect to the functional form of the classifier. The results listed in Table I suggest that:

1) The accuracies were poor when only use curvature feature (MC, GC, FI, and CI), otherwise, the performance was good when only use thickness feature. Also, the accuracies were better when thickness participating classification inputs. Since, ASD patients had more thickness abnormalities rather than curvature abnormalities.

2) The accuracy was modest better when using T + MC combination in LMT. This mean curvature may provide extra

but limited information to thickness in ASD. As a result, ASD may have less curvature abnormalities.

3) The accuracies were even worse when combination three or more features. The "overfitting" may cause this results.

4) LMT obtained better result in both thickness and T + MC classification. So, LMT may be suitable for this problem, but can not aviod "overfitting" problems.

In Table II, there were 10 predictive variables selected by LMT in T + MC classification: thickness-based variables were far more than curvature-based variables. In two-sample t statistic, significant thickness abnormal structures were also far more than significant curvature abnormal structures. These were another evidences that ASD may be thickness abnormal disorder rather than curvature abnormal disorder.

All the thickness abnormal structures were described in our previous work [7]. Here, we were only focus on two curvature abnormal structures. Right caudal anterior cingulate and right pars orbitalis were selected by LMT in T+MC classification as mean curvature features. Also, curvature of right pars orbitalis was significant different between groups. Pars orbitalis (orbital part of inferior frontal gyrus) is the part of the inferior frontal gyrus named opercularis because it covers part of the insula. The similar finding was reported by Nordahl et al. [9] which demonstrated cortical folding abnormalities in individuals with ASD, primarily in the left operculum, bilateral parietal operculum, and bilateral intraparietal sulcus. Kemper and Bauman [21] found the caudal anterior cingulate appeared unusually coarse in a study that included six cases of autism. In our studies, the right caudal anterior cingulate was selected by LMT in T+MC classification as mean curvature features, but was not significant in two-sample t-test between groups.

However, few curvature abnormalities were found in our study demonstrated that ASD may be thickness abnormal disorders, and curvature abnormalities play less important roles in discriminating ASD and normal controls.

One of the limitations of this study is the small sample sizes of both the ASD and VC groups. Small sample size leads to decreased precision in estimates of various properties of the population, adversely influences diagnostic model generation and testing, and makes performance estimation of diagnostic models difficult. Thus, we might find that increasing sample size would lead us to change our opinions about the relative merits of the four machine-learning approaches that we applied. Cross-validation only partly alleviates the error-estimation problem; it is known to overestimate classification accuracy. Although other machine-learning researchers have used cross-validation with similar sample sizes (e.g., Kloppel et al. [22] and Singh et al. [11]); the ultimate way to refine classification models, and to verify these results, is to increase sample size. We plan to accrue additional subjects, and to repeat these analyses (and thereby refine our classification models) as our sample sizes increase.

Image quality may be another limitation of any MR-based analysis. We excluded subjects with poor MR image quality, and all MR images were reviewed by an experienced radiologist for quality. We also correct for motion. We would expect that the noise introduced by poor image quality would cause our algorithms to find few or no features that distinguish ASD from control subjects, and that the resulting models would manifest poor classification performance, which was not the case.

V. Conclusion

To our knowledge, this study represents the first attempt to classify autism using multiple cortical measurements extracted from SBM, and to compare classification results between those combination features. The principal contribution of our work is our determination that "thickness+curvature"-based classification results in more accurate classification than thickness-based classification. The machine-learning approach that performed best, LMT, generated a simple and declarative diagnostic model. We found curvature only provide limited information to thickness in ASD predictive model. Our results suggest that ASD patients may have more abnormal in cortical thickness than in cortical curvature.

Acknowledgment

Yun Jiao was supported by the China Scholarship Council (Project number 2008101370), the National Natural Science foundation of China (Project number 30570655), and the Scientific Research Foundation of Graduate School of Southeast University (Project number YBJJ1011). Dr. Lu was supported by the National Natural Science foundation of China, Project No. 30570655.

The authors would like to thank Drs. Xiaoyan Ke and Kangkang Chu for subjects recruitment and image acquirement.

References

[1] D. G. Amaral, C. M. Schumann, and C. W. Nordahl, "Neuroanatomy of autism," *Trends in Neurosciences*, vol. 31, no. 3, pp. 137–145, MAR 2008.

[2] A. Gmitrowicz and A. Kucharska, "[developmental disorders in the fourth edition of the american classification: diagnostic and statistical manual of mental disorders (dsm iv – optional book)]," *Psychiatr Pol*, vol. 28, no. 5, pp. 509–21, Sep-Oct 1994.

[3] D. H. Geschwind, "Advances in autism," *Annu Rev Med*, vol. 60, pp. 367–380, Feb. 2009 2009.

[4] R. A. Muller, "The study of autism as a distributed disorder," *Mental Retardation and Developmental Disabilities Research Reviews*, vol. 13, no. 1, pp. 85–95, 2007.

[5] B. F. Sparks, S. D. Friedman, D. W. Shaw, E. H. Aylward, D. Echelard, A. A. Artru, K. R. Maravilla, J. N. Giedd, J. Munson, G. Dawson, and S. R. Dager, "Brain structural abnormalities in young children with autism spectrum disorder," *Neurology*, vol. 59, no. 2, pp. 184–92, Jul 23 2002.

[6] B. Fischl, M. I. Sereno, and A. M. Dale, "Cortical surface-based analysis. ii: Inflation, flattening, and a surface-based coordinate system," *Neuroimage*, vol. 9, no. 2, pp. 195–207, Feb 1999.

[7] Y. Jiao, R. Chen, X. Y. Ke, K. K. Chu, Z. H. Lu, and E. H. Herskovits, "Predictive models of autism spectrum disorder based on brain regional cortical thickness," *Neuroimage*, vol. 50, no. 2, pp. 589–599, APR 1 2010.

[8] N. Hadjikhani, R. M. Joseph, J. Snyder, and H. Tager-Flusberg, "Abnormal activation of the social brain during face perception in autism," *Human Brain Mapping*, vol. 28, no. 5, pp. 441–449, MAY 2007.

[9] C. W. Nordahl, D. Dierker, I. Mostafavi, C. M. Schumann, S. M. Rivera, D. G. Amaral, and D. C. Van Essen, "Cortical folding abnormalities in autism revealed by surface-based morphometry," *Journal of Neuroscience*, vol. 27, no. 43, pp. 11 725–11 735, OCT 24 2007.
[10] N. L. Voets, M. G. Hough, G. Douaud, P. M. Matthews, A. James, L. Winmill, P. Webster, and S. Smith, "Evidence for abnormalities of cortical development in adolescent-onset schizophrenia," *Neuroimage*, vol. 43, no. 4, pp. 665–75, Dec 2008.
[11] V. Singh, L. Mukherjee, and M. K. Chung, "Cortical surface thickness as a classifier: boosting for autism classification," *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, vol. 11, no. Pt 1, pp. 999–1007, 2008.
[12] C. Lord, M. Rutter, and A. Le Couteur, "Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders," *J Autism Dev Disord*, vol. 24, no. 5, pp. 659–85, Oct 1994.
[13] E. Schopler, R. J. Reichler, R. F. DeVellis, and K. Daly, "Toward objective classification of childhood autism: Childhood autism rating scale (cars)," *J Autism Dev Disord*, vol. 10, no. 1, pp. 91–103, Mar 1980.
[14] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 20, pp. 11 050–11 055, SEP 26 2000.
[15] D. C. V. Essen and H. A. Drury, "Structural and functional analyses of human cerebral cortex using a surface-based atlas." *J Neurosci*, vol. 17, no. 18, pp. 7079–7102, Sep 1997.
[16] R. S. Desikan, F. Segonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, and R. J. Killiany, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *NeuroImage*, vol. 31, no. 3, pp. 968–980, 2006/7/1/ 2006.
[17] J. C. Platt, "Advances in kernel methods: support vector learning," pp. 185–208, 1999.
[18] P. Sajda, "Machine learning for detection and diagnosis of disease," *Annual Review of Biomedical Engineering*, vol. 8, pp. 537–565, 2006.
[19] J. Gama, "Functional trees," *Machine Learning*, vol. 55, no. 3, pp. 219–250, JUN 2004.
[20] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine Learning: Ecml 2003*, vol. 2837, pp. 241–252, 2003.
[21] T. L. Kemper and M. L. Bauman, "The contribution of neuropathologic studies to the understanding of autism," *Neurol Clin*, vol. 11, no. 1, pp. 175–87, Feb 1993.
[22] S. Kloppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, J. Jack, C. R., J. Ashburner, and R. S. Frackowiak, "Automatic classification of mr scans in alzheimer's disease," *Brain*, vol. 131, no. Pt 3, pp. 681–9, Mar 2008.