Prediction of severity and treatment outcome for ASD from fMRI

Juntang Zhuang[1], **Nicha C. Dvornek**[2,3], **Xiaoxiao Li**[1], **Pamela Ventola**[2], **James S. Duncan**[1,3,4]

[1]Biomedical Engineering, Yale University, New Haven, CT USA

[2]Child Study Center, Yale University, New Haven, CT USA

[3]Radiology & Biomedical Imaging, Yale School of Medicine, New Haven, CT USA

[4]Electrical Engineering, Yale University, New Haven, CT USA

Abstract

Autism spectrum disorder (ASD) is a complex neurodevelop-mental syndrome. Early diagnosis and precise treatment are essential for ASD patients. Although researchers have built many analytical models, there has been limited progress in accurate predictive models for early diagnosis. In this project, we aim to build an accurate model to predict treatment outcome and ASD severity from early stage functional magnetic resonance imaging (fMRI) scans. The difficulty in building large databases of patients who have received specific treatments and the high dimensionality of medical image analysis problems are challenges in this work. We propose a generic and accurate two-level approach for high-dimensional regression problems in medical image analysis. First, we perform region-level feature selection using a predefined brain parcellation. Based on the assumption that voxels within one region in the brain have similar values, for each region we use the bootstrapped mean of voxels within it as a feature. In this way, the dimension of data is reduced from number of voxels to number of regions. Then we detect predictive regions by various feature selection methods. Second, we extract voxels within selected regions, and perform voxel-level feature selection. To use this model in both linear and non-linear cases with limited training examples, we apply two-level elastic net regression and random forest (RF) models respectively. To validate accuracy and robustness of this approach, we perform experiments on both task-fMRI and resting state fMRI datasets. Furthermore, we visualize the influence of each region, and show that the results match well with other findings.

Keywords

fMRI; ASD; predictive model

1 Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental syndrome characterized by impaired social interaction, difficulty in communication and repetitive behavior. ASD is most commonly diagnosed with a behavioral test [1], however, the behavioral test is insufficient to understand the mechanism of ASD. Functional magnetic resonance imaging (fMRI) has been widely used in research on brain diseases and has the potential to reveal brain malfunctions in ASD.

Behavior based treatment is a widely used therapy for ASD, and Pivotal Response Treatment (PRT) is empirically-supported [2]. PRT addresses core deficits in social motivation to improve social communication skills. Such therapies require large time commitments and lifestyle changes. However, an individual's response to PRT and other behavioral treatments vary, yet treatment is mainly assigned by trial and error. Therefore, prediction of treatment outcome during early stages is essential.

fMRI measures blood oxygenation level dependent (BOLD) signal and reflects brain activity. Recent studies have applied fMRI in classification of ASD and identifying biomarkers for ASD [3]. Although some regions are found to have higher linear correlations with certain types of ASD severity scores, the correlation coefficient is typically low (below 0.5). Moreover, most prior studies apply *analytical* models, and lack *pre-dictive* accuracy.

The goal of our work is to build accurate *predictive* models for fMRI images. To deal with the high dimensionality of the medical image regression problem, we propose a two-level modeling approach: 1) region-level feature selection, and 2) voxel-level feature selection. In this paper, we demonstrate predictive models for PRT treatment outcomes and ASD severity, and validate robustness of this approach in both task fMRI and resting state fMRI datasets. Furthermore, we analyze feature importance and identify potential biomarkers for ASD.

2 Methods

2.1 Two-level modeling approach

Dimensionality of medical images (i.e., the number of voxels) is far higher than the number of subjects in most medical studies. The high dimensionality causes inaccuracy in variable selection and affects modeling performance. However, medical images are typically locally smooth, and voxels are not independent of each other. This enables us to perform the following two-level feature selection as shown in Fig. 1. The proposed procedure first selects important features at the region level, then performs feature selection at the voxel level. Our generic approach can be used with both linear and non-linear models.

Region-level modeling and variable selection—Based on brain atlas research, we assume that voxels within the same region of a brain parcellation have similar values. Therefore, we use the bootstrapped (sample with replacement) mean for each region as a feature, reducing dimension of data from number of voxels to number of regions. Then we can perform feature selection on this new dataset, where each predictor variable represents a region.

Beyond dimension reduction, representing each region with the bootstrapped mean of its voxel values decreases correlation between predictor variables. Another potential benefit is to increase sample size. We can generate many artificial training examples from one real training example by repeatedly bootstrapping each region. Since this generates correlated training samples, repeated bootstrapping can only be used in models that are robust to sample correlation.

Voxel-level modeling and variable selection—Region-level feature selection preserves predictive regions. However, representing all voxels within a region as one number is too coarse, and may affect model accuracy. Therefore, we extract all voxels within the selected regions, perform spatial down-sampling by a factor of 4, and apply feature selection on voxels.

Pipeline repetition—Due to the randomness in bootstrapping for region level modeling, we repeat the whole process. For each of the four models (linear and non-linear models at region-level and voxel-level respectively), we average outcomes to generate stable predictions.

2.2 Linear and non-linear models

We can apply any model in the approach proposed in section 2.1. To instantiate a generic approach for both linear and non-linear cases, we train elastic net regression and random forest (RF) independently, both trained at two levels.

Variable selection with elastic net regression—Elastic net is a linear model with both l_1 and l_2 penalty to perform variable selection and shrinkage regularization [4]. Given predictor variables X and targets y, the model is formalized as

$$\hat{\beta} = \operatorname{argmin}_{\beta}\left\{ \left\| y - X\beta \right\|^2 + \lambda\left[\alpha\|\beta\|_1 + (1 - \alpha)\|\beta\|^2 \right] \right\} \tag{1}$$

where $0 \leq \alpha \leq 1$, α controls the proportion of regularization on l_1 and l_2 term of estimated coefficients, and λ controls the amplitude of regularization. l_2 penalty is shrinkage regularization and improves robustness of the model. l_1 penalty controls sparsity of the model. By choosing proper parameters, irrelevant variables will have coefficients equal to 0, enabling variable selection.

Variable selection with random forest—Random forest is a powerful model for both regression and classification problems and can deal with interaction between variables and high dimensionality [5]. Although random forest can handle medium-high dimensional problems, it's insufficient to handle ultra-high dimensional medical image problems. Therefore, two-level variable selection is still essential.

Conventional variable selection technique for random forest builds a predictive model with forward stepwise feature selection [6]. For high dimensional problems, it is computationally intensive. Therefore, we use a similar thresholding method to perform fast variable selection as in [7] (Fig. 2). We generate noise ("shadow") variables from a Gaussian distribution independent of target variables. Shadow variables are added to the original data matrix, and a random forest is trained on the new data matrix. The random forest model calculates the importance of each variable. A predictive variable should have higher importance than noise variables. A threshold is calculated as:

$$Thres = \operatorname{median}\left(VI_i^{shadow}\right) \quad s.t. \quad VI_i^{shadow} > 0, \quad i = 1, 2, \ldots n \tag{2}$$

where n is the total number of shadow variables, and VI_i^{shadow} is the importance measure for the ith shadow variable. We use permutation accuracy importance measure [6] in this experiment. The threshold is calculated using positive shadow variable importance because permutation accuracy importance can be negative. We use the median to make a conservative threshold, because even noise variables can have high importance in high-dimensional problems, due to randomness of the model. After variable selection, we build a gradient boosted regression tree model based on selected variables.

2.3   Visualization of each variable's influence

To achieve both predictability and interpretability, we use the following methods to visualize influence of each region in the brain. For linear models, we plot the linear coefficients map.

For non-linear models, we visualize the influence based on the partial dependence plot. The partial dependence plot shows the dependence between target and predictor variables, marginalizing over all other features [8],

$$D_l(z_l) = \mathbb{E}_x(\widehat{F}(x)\,|\,z_l) = \int \widehat{F}(x)p(z_{\backslash l}\,|\,z_l)dz_{\backslash l} \qquad (3)$$

where $D_l(z_l)$ is the partial dependence function for variable z_l, $\widehat{F}(x)$ is the trained model, $z_{\backslash l}$ is the set of variables except z_l, $p(z_{\backslash l}\,|\,z_l)$ is the distribution of $z_{\backslash l}$ given z_l. Each $D_l(z_l)$ is calculated as a sequence varying with z_l in practice.

The influence of each variable is stored in a sequence. For visualization, we summarize the influence of a variable (Influence$_l$) by calculating the variance of $D_l(z_l)$ to measure the amplitude of its influence, and the sign of its correlation with z_l to show if it has a positive or negative influence on targets:

$$\text{Influence}_l = \text{Sign}(\text{corr}(D_l(z_l), z_l))\text{Var}(D_l(z_l)). \qquad (4)$$

3   Experiments and results

3.1   Task-fMRI experiment

Ninteen children with ASD participated in 16 weeks of PRT treatment, with pretreatment and post-treatment social responsiveness scale (SRS) scores [9], and pretreatment autism diagnostic observation schedule (ADOS) [10] scores measured. Each child underwent a pre-treatment baseline task fMRI scan (BOLD, TR = 2000ms, TE = 25ms, flip angle = 60°, slice thickness = 4.00mm, voxel size 3.44×3.44×4mm^3) and a structural MRI scan (T1-weighted MPRAGE sequence, TR = 1900ms, TE=2.96ms, flip angle = 9°, slice thickness = 1.00mm, voxel size = $1 \times 1 \times 1 mm^3$) on a Siemens MAGNETOM Trio TIM 3T scanner.

During the fMRI scan, coherent (BIO) and scrambled (SCRAM) point-light biological motion movies were presented to participants in alternating blocks with 24s duration [11]. The fMRI data were processed using FSL v5.0.8 in the following pipeline: a) motion correction with MCFLIRT, b) interleaved slice timing correction, c) BET brain extraction, d)

grand mean intensity normalization for the whole four-dimensional data set, e) spatial smoothing with $5mm$ FWHM, f) denoising with ICA-AROMA, g) nuisance regression for white matter and CSF, h) high-pass temporal filtering.

The timing of the corresponding blocks (BIO and SCRAM) was convolved with the default gamma function (phase=0s, sd=3s, mean lag=6s) with temporal derivatives. Participant-level t-statistics for contrast BIO>SCRAM were calculated for each voxel with first level analysis. This 3D t-statistic image is the input to the proposed approach. The input image is parcellated into 268 regions using the atlas from group-wise analysis [12].

We tested the approach on three target scores using leave-one-out cross validation: pre-treatment SRS score, pre-treatment standardized ADOS score [13], and treatment outcome defined as the difference between pre-treatment and post-treatment SRS score. For elastic net regression, we used nested cross-validation to select parameters ($\lambda \in \{0.001, 0.01, 0.1\}$, a ranging from 0.1 to 0.9 with a stepsize of 0.1). Other parameters were set according to computation capability. For random forest models, we set tree number as 2000. For region level modeling, each region was represented as the mean of 2000 bootstrapped samples from its voxels. For gradient boosted tree model after feature selection with random forest, the number of trees was set as 500. The whole process was repeated 100 times and averaged. All models were implemented in MATLAB, with default parameters except as noted above. Neurological functions of selected regions were decoded with Neurosynth [14]. For each experiment, results (linear correlation between predictions and measurements r, uncorrected p-value, root mean square error $RMSE$) of the best model are shown in Fig. 3 and 4.

3.2 Resting state fMRI experiment

We performed similar experiments on the ABIDE dataset [15] using the UM and USM sites with five-fold cross validation. We selected male subjects diagnosed with ASD, resulting in 51 patients from UM and 13 patients from USM. We built models to predict the ADOS Gotham total score from voxel-mirrored homotopic connectivity images [16]. We set parameters the same as in section 3.1. Results are shown in Fig. 5.

3.3 Result analysis

Training and validation datasets were independent for all experiments. The proposed two-level approach accurately selected predictive features, while elastic net and RF directly applied to the whole-brain image failed to generate predictive results in all experiments (correlation between predictions and measurements < 0.1). The proposed approach generated very high predictive accuracy on various datasets and different scores, achieving better accuracy than state-of-the-art.

For SRS scores, we found no predictive models in the literature. Kaiser et al. reported regions of correlation r=0.502 [11] in *analytical* modeling, while our *predictive* model achieved r=0.45 (Fig. 3(a)).

For standardized ADOS score, the best result in literature achieves r=0.51 between predictions and measurements with 156 subjects based on cortical thickness [17]. Our model achieved r=0.50 with 19 patients (Fig. 3(c)) based on fMRI.

For raw ADOS score, Björnsdotter et al. found no significant correlation with brain responses in fMRI scan [18]. Predictive models based on structural MRI achieved correlation of r=0.362 between predictions and measurements [19]. In our experiment with resting-state fMRI, we achieved correlation r=0.40 (Fig. 5).

To predict treatment outcome from baseline fMRI scan, Dvornek et al. achieved correlation r=0.83 between predictions and measurements [20]. We achieved r=0.71 (Fig. 3(b)). However, the study by Dvornek et al. takes pre-selected regions as input and loses interpretability because it does not perform region selection. In contrast, our proposed approach takes a whole-brain image as input and can select predictive regions for interpretation and biomarker selection. Furthermore, the proposed approach is generic and any non-linear model (including Dvornek's method) can be applied.

Neurosynth decoder results (Fig. 4(d), Fig. 5 right figure) show that selected regions match the literature [11]. The selected regions are slightly different across experiments due to different tasks, datasets and target measures. Many regions are shared across experiments, such as prefrontal cortex and visual cortex.

4 Conclusion

We propose a generic approach to build *predictive* models based on fMRI images. To deal with high-dimensionality, we perform two-level variable selection: region-level modeling, and voxel-level modeling. This generic approach includes elastic net and random forest models to fit both linear and non-linear cases. The proposed approach is tested on both task-fMRI and resting-state fMRI, and validated on different scores. The proposed predictive approach achieves higher correlation than state-of-the-art predictive modeling in many experiments. Overall, the proposed approach is generic, accurate, and achieves both predictability and interpretability.

Acknowledgement

References

1. Baird G et al., "Diagnosis of autism," BMJ, 2003.

2. Koegel LK et al., "Pivotal response intervention i: Overview of approach," TASH, 1999.

3. Anderson JS et al., "Functional connectivity magnetic resonance imaging classification of autism," Brain, 2011.

4. Zou H et al., "Regularization and variable selection via the elastic net," *J. Royal Stat. Soc.*, 2005.

5. Liaw A et al., "Classification and regression by randomforest," R news, 2002.

6. Genuer R et al., "Variable selection using random forests," Pattern Recognition Letters, 2010.

7. Zhuang J et al., "Prediction of pivotal response treatment outcome with task fmri using random forest and variable selection," in ISBI, 2018.

8. Friedman JH, "Greedy function approximation: a gradient boosting machine," Annals of statistics, 2001.

9. Bruni TP, "Test review: Social responsiveness scale–second edition (srs-2)," 2014.

10. Lord C et al., "The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism," J Autism Dev Disord, 2000.

11. Kaiser MD et al., "Neural signatures of autism," Proc Natl Acad Sci U S A, 2010.

12. Shen X et al., "Groupwise whole-brain parcellation from resting-state fmri data for network node identification," Neuroimage, 2013.

13. Gotham K et al., "Standardizing ados scores for a measure of severity in autism spectrum disorders," J Autism Dev Disord, 2009.

14. Yarkoni T et al., "Large-scale automated synthesis of human functional neuroimaging data," Nature methods, 2011.

15. Martino A Di et al., "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism," Mol. Psychiatry, 2014.

16. Zuo X-N et al., "Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy," Journal of Neuroscience, 2010.

17. Moradi Elaheh et al., "Predicting symptom severity in autism spectrum disorder based on cortical thickness measures in agglomerative data," NeuroImage, 2017.

18. Björnsdotter M et al., "Evaluation of quantified social perception circuit activity as a neurobiological marker of autism spectrum disorder," JAMA psychiatry, 2016.

19. Sato JR et al., "Inter-regional cortical thickness correlations are associated with autistic symptoms: a machine-learning approach," JJ PSYCHIAT RES, 2013.

20. Dvornek NC et al., "Prediction of autism treatment response from baseline fmri using random forests and tree bagging," Multimodal Learn Clin Decis Support, 2016.
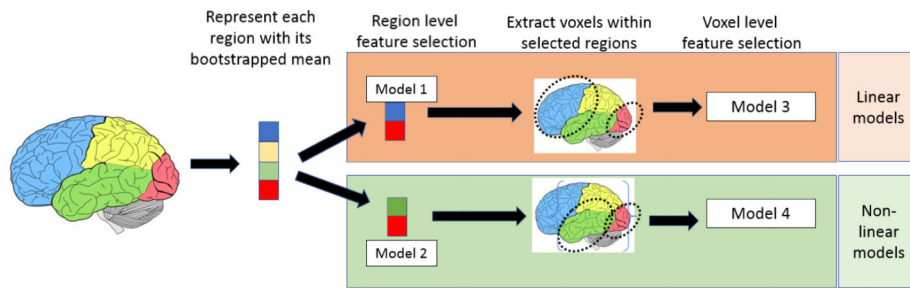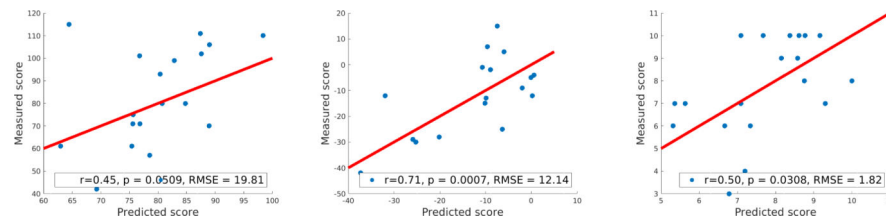
Fig.1:
Flowchart of the proposed approach. Each column represents a stage of the approach, region-level and voxel-level models. Top row shows linear models (e.g. elastic net regression), bottom row shows non-linear models (e.g. random forest).
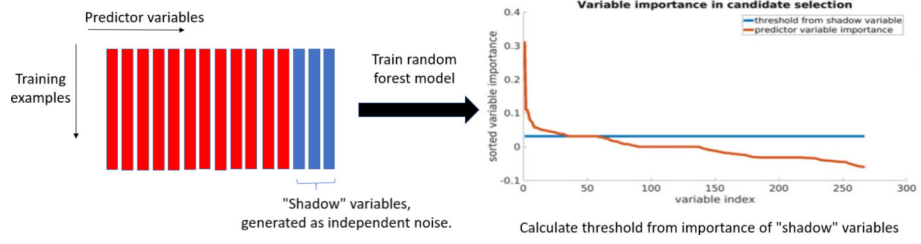
Fig.2:
Flowchart of variable selection with random forest and "shadow" method.

Fig.3:

Results for various scores predicted from task fMRI, red lines are reference lines of perfect prediction $y = x$.

(a) Region-level elastic net model for pre-treatment SRS

(b) Voxel-level random forest model for change of SRS after treatment

(c) Voxel-level random forest model for standardized ADOS
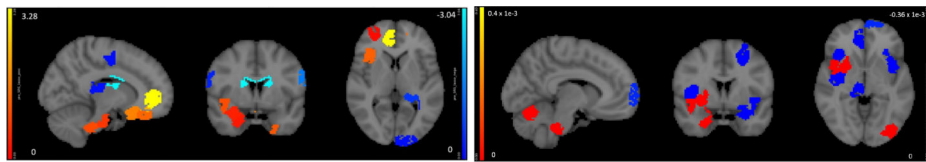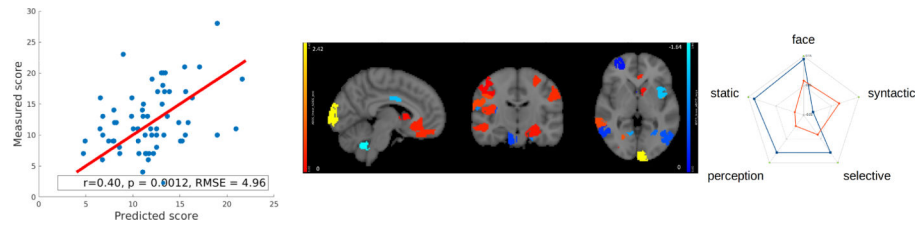
(a) Region-level elastic net model for pre-treatment SRS

(b) Voxel-level random forest model for change of SRS after treatment
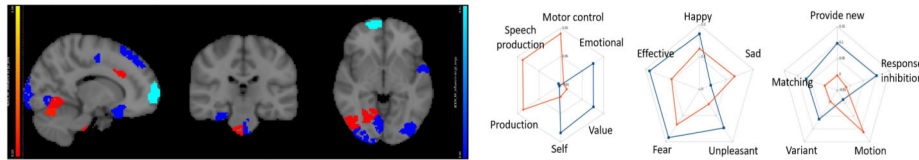
(c) Voxel-level random forest model for standardized ADOS

(d) From left to right: results of Neurosynth decoder for model in (a) (b) (c).

Fig.4:

Regions are colored in red for positive influence and blue for negative influence. (a-c): Influence of regions for various scores based on task fMRI. (d): Functions decoded by Neurosynth.

Fig.5:

Left: Results of region-level elastic net regression model for resting-state fMRI experiment. Middle: Linear coefficients of model. Right: Functions decoded by Neurosynth, red for positive regions, blue for negative regions.