A Hierarchical Feature Extraction Pipeline using Resting-state fMRI for Autism Classification

Qian Wang, Jongwoo Choi, Xiaofu He*

*Assistant Professor of Clinical Neurobiology, Department of Psychiatry
Faculty Member, Data Science Institute, Columbia University
Email: xh2170@cumc.Columbia.edu
Contents

• Background
• Our method
• Summary
What is Autism Spectrum Disorder?

- Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder with core deficits in social communication and interaction as well as stereotypic behaviors.
- Prevalence: 1 in 59 children had a diagnosis of ASD by age 8 in 2014 (estimates are for 4 years prior to the report date)

Diagnosis of ASD

- Diagnostic and Statistical Manual for Mental Disorders (DSM), published by the American Psychiatric Association (APA). The first version (DSM-I) was published in 1952 and the most recent version (DSM-5) was published in 2013.
- “According to the DSM-5, a guide created by the APA used to diagnose mental disorders, people with ASD have:
  - Difficulty with communication and interaction with other people
  - Restricted interests and repetitive behaviors
  - Symptoms that hurt the person’s ability to function properly in school, work, and other areas of life”

Challenges of diagnosis of mental disorders

• Subjective: highly depends on assessments and psychiatrists experience

• Sensitivity (true positive rate) and specificity (true negative rate)
Can we objectively diagnose patients with mental disorders?

- All medical conditions except mental disorders can be objectively diagnosed by instruments
  - e.g., heart disease can be diagnosed by electrocardiograms (ECG/EKG). Similarly, blood sugar levels can be used for diagnosing patients with diabetes.

- Can a patient with mental disorders be objectively diagnosed by brain scans?
  - Not yet!
Brain scans are usually used to identify brain biomarkers which would be used for differentiating healthy controls from patients, e.g., abnormal activity in motor cortex can be linked to ADHD patients.

ADHD: attention-deficit/hyperactivity disorder  Adapt from Stefan Posse, Real-Time Functional MRI, 2012
Classical methods for mental disorder classification

• Statistical methods
  – e.g., identify abnormal brain regions in ASD, such as thalamus and amygdala (Wee et al. 2014)
  – Problems: heterogeneity of ASD, Comorbidity in ASD (Masi et al. 2017)

• Machine Learning methods
  – Data-driven, such as logistic regression, SVM, Random Forest etc. (Abraham et al. 2014)
  – Limitations: Relatively small sample size (Wolfers et al. 2015). Also, impacted by the complex non-linear changes in volumetric (Courchesne et al. 2011a), morphometric (Ecker et al. 2014), and connectivity (Betzl et al. 2014) measures across the lifespan.
Workflow for classification of brain imaging data

Adapted from https://ars.els-cdn.com/content/image/1-s2.0-S1053811910014163-gr1.jpg
Our method: preprocessing

• Structural MRI (sMRI) data: a set of structural features have been extracted for each subject with normalized brain volume computed using subcortical segmentation, and cortical thickness and area for right and left hemisphere of Freesurfer output.

• Resting State functional MRI (rs-fMRI) data: create a functional connectivity map for each subject from the rs-fMRI signals based on different brain atlases, such as bootstrap analysis of stable clusters (BASC) parcellations (Bellec 2010), Ncuts parcellations (Craddock 2012) and Power atlas (Power 2011).

Structural MRI data preprocessing

http://surfer.nmr.mgh.harvard.edu/fswiki/Tutorials
Rs-fMRI data preprocessing

- Single voxel time series
- Time series Extraction
- Correlation Matrix

ROIs distributed across the cortex and subcortical structures with Power atlas (264 ROIs)

http://nilearn.github.io/auto_examples/  https://www.fil.ion.ucl.ac.uk/spm/course/slides16oct/
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5624035/figure/F1/
Our method: feature selection

• Hierarchical feature selection pipeline consists of two steps: for the first level, the cortical thickness & area and sub-cortex volume features of sMRI and the functional connectivity features of rs-fMRI are extracted; for the second level, pattern-based features of functional connectivity are extracted from the first level features.

• Those hierarchical features are combined together and are put into random forest for feature selection, finally the top 205 features are extracted for classification.
Hierarchical feature Extraction

sMRI Raw Data

First Level

Merged Cortical + Subcortical

Subcortical volume + Cortical thickness and area

Use ROI volume as computed from aparc (more accurate)

Correlation Matrix Features

Pattern-based Features

http://nilearn.github.io/auto_examples/
Pattern-based Features

Spearman's correlation results of pattern-based features from Autism patients and Healthy controls:

For class 0: correlation=0.92, pvalue=4.5e-27
For class 1: correlation=0.82, pvalue=1.6e-16
For class 2: correlation=0.80, pvalue=1.1e-15
For class 3: correlation=0.69, pvalue=3.6e-10

GLCM illustration comes from https://www.researchgate.net/publication/277313083_638563
Hierarchical feature Selection

sMRI Data → First level features → Second level features → Feature Combination

rs-fMRI Data → First level features → Second level features

Top 205 features

N_1 feature → N_2 feature → N_3 feature → N_4 feature

Class N → Class O → Class M → Class N

MAJORITY VOTING

FINAL CLASS

https://blog.quantinsti.com/random-forest-algorithm-in-python/
Our method: classification

- **Method 1:** Feature selection + classification

---

**Random Forest Feature Selection**

**Logistic Regression Classifier**

---

Method 2: Convolutional Neural Network

- We proposed a CNN model, where row-by-row and column-by-column filters form the first two convolution layers. We then added a fully connected layer followed by an output softmax layer that outputs probability of each class. For instance, for a functional connectivity matrix with size K*K (row * col), we used N filters with kernel size of 1*K and K*1.

In the experiment, the parameters are set: K=(64, 122, 197 respectively). Please see details in our poster (Poster Session 2, Poster#5: Classification of Autism Spectrum Disorder Based on Brain Imaging Using Convolutional Neural Networks).
Datasets

• Data we used for this project is available at https://paris-saclay-cds.github.io/autism_challenge/ and initially published for competition: Imaging-psychiatry challenge: predicting autism (IMPAC). The data contained 1150 subjects (601 health controls and 549 ASD patients). Age range: 5~64.
Results

- Method 1: Feature selection + classification (cv=8)

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First-level</td>
<td>First + Second level</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>0.78</td>
<td>0.75</td>
</tr>
<tr>
<td>SVM</td>
<td>0.58</td>
<td>0.77</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>0.75</td>
<td>0.73</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>0.72</td>
<td>0.71</td>
</tr>
</tbody>
</table>

- Method 2: We trained a classifier using the proposed CNN model for each BASC atlas separately and combined all three classifiers through majority vote (Ensemble). (Training 920 (80%) vs Testing 230 (20%))

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASC 64</td>
<td>0.7</td>
<td>0.69</td>
<td>0.72</td>
<td>0.73</td>
</tr>
<tr>
<td>BASC 122</td>
<td>0.7</td>
<td>0.69</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>BASC 197</td>
<td>0.64</td>
<td>0.7</td>
<td>0.68</td>
<td>0.74</td>
</tr>
<tr>
<td>Combined</td>
<td>0.72</td>
<td>0.72</td>
<td>0.75</td>
<td>0.78</td>
</tr>
</tbody>
</table>

ROC curve for BASC 64, BASC 122, BASC 197 and Ensemble Classifier
Summary

• Challenges:
  – Small sample size (MRI is too expensive!)
  – Standardization of data acquisition, preprocessing, and atlas selection.
  – Test-retest reliability

• Future work
  – Unsupervised machine learning
  – Interpretable machine learning
Thank you!