Derivation of Functional Networks with Independent Component Analysis

Vanessa Sochat
Goals

identify biomarkers of disease
use data-driven methods
work with publicly available data
predict disease state
Goals

Are there significant differences in the default mode network (DMN) between ADHD and Control?
## Data

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typically-Developing Controls</td>
<td>99</td>
</tr>
<tr>
<td>ADHD Combined</td>
<td>77</td>
</tr>
<tr>
<td>ADHD Inattentive</td>
<td>44</td>
</tr>
<tr>
<td>ADHD Hyperactive</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>222</strong></td>
</tr>
</tbody>
</table>
Data

<table>
<thead>
<tr>
<th>Typically-Developing Controls</th>
<th>99</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Gender (M / F)</td>
</tr>
<tr>
<td>TD Control</td>
<td>69</td>
</tr>
<tr>
<td>ADHD Combined</td>
<td>62</td>
</tr>
<tr>
<td>ADHD Inattentive</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>162</td>
</tr>
</tbody>
</table>
Data

<table>
<thead>
<tr>
<th></th>
<th>Anatomical</th>
<th>Functional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>1.3 x 1.0 x 1.3 mm</td>
<td>3.0 x 3.0 x 4.0 mm</td>
</tr>
<tr>
<td>Scan Time</td>
<td>8:07 minutes</td>
<td>6:00 minutes</td>
</tr>
<tr>
<td>TR / TE (ms)</td>
<td>2530 / 3.25</td>
<td>2000 / 15</td>
</tr>
<tr>
<td>Slices</td>
<td>162</td>
<td>33 slices</td>
</tr>
<tr>
<td></td>
<td>Interleaved 4mm</td>
<td></td>
</tr>
</tbody>
</table>

MRI parameters:

- Anatomical:
  - Resolution: 1.3 x 1.0 x 1.3 mm
  - Scan Time: 8:07 minutes
  - TR / TE (ms): 2530 / 3.25
  - Slices: 162
  - Interleaved 4mm

- Functional:
  - Resolution: 3.0 x 3.0 x 4.0 mm
  - Scan Time: 6:00 minutes
  - TR / TE (ms): 2000 / 15
  - Slices: 33

MRI images and data files:

- **Anatomical MRI**:
  - `l4.1-fslinfo_mprage_noface.nii`
  - `data_type`: INT16
  - `dim1`: 128
  - `dim2`: 256
  - `dim3`: 256
  - `dim4`: 1
  - `datatype`: 4
  - `pixdim1`: 1.3300018311
  - `pixdim2`: 1.0000000000
  - `pixdim3`: 1.0000000000
  - `pixdim4`: 0.0000000000
  - `cal_max`: 0.0000
  - `cal_min`: 0.0000
  - `file_type`: NIFTI-1+

- **Functional MRI**:
  - `l4.1-fslinfo_rest.nii.gz`
  - `data_type`: FLOAT32
  - `dim1`: 64
  - `dim2`: 80
  - `dim3`: 33
  - `dim4`: 176
  - `datatype`: 16
  - `pixdim1`: 3.0000000000
  - `pixdim2`: 2.9999942780
  - `pixdim3`: 4.0000028610
  - `pixdim4`: 2.0000000000
  - `cal_max`: 0.0000
  - `cal_min`: 0.0000
  - `file_type`: NIFTI-1+
Independent Component Analysis

FMRI data = spatial maps
Independent Component Analysis

\[ \text{FMRI data} \rightarrow \text{components} \rightarrow \text{spatial maps} \]

\[ \text{FMRI data} \rightarrow \text{spatial maps} \]
This python submission script allows for either single checking quality analysis after ICA, group ICA analysis, or dual regression based on a group ICA. The script sets up a user specified experiment folder, and multiple ICA can be done from one common set of single subject ICA run experiment folder. # Preprocessing

# USER DEFINED VARIABLES
TR=2.0 # TR
HPF=150 # highpass frequency filter cutoff
LPF=.3 # lowpass frequency filter cutoff
SKERN=6.0 # Smoothing kernel (PARSKE)
BBTHRESH=10 # Brain background threshold
MATLABRUNTIME="/home/vsochat/software/matlab/Matlab64bitRuntime" # Lower filter threshold (0.03)
# FILTR=0.08 # Upper filter threshold
# BBTHRESH=10 # Brain background threshold
# MATLABRUNTIME = path to Matlab Runtime

# VARIABLES DEFINED AT RUNTIME
# OUTPUT=$1 # Output folder (not yet created)
# SETUP=$2 # Full path to raw rest data
# PINF=$3 # Full path to raw anatomical data runtime

# make sure output directory was made by submission script.
if [ ! -d "$OUTPUT" ]; then
    mkdir -p $OUTPUT
    SETUP
fi

# Check once more for functional data by submission script.
if [ ! -f "$FUNCADATA" ]; then echo "Cannot find functional data file "$FUNCADATA". Exiting." exit 32
fi

# Check once more for anatomical data...
if [ ! -f "$SANADATA" ]; then echo "Cannot find anatomical data file "$SANADATA". Exiting." exit 32

# Prepare functional mask...

# Go to output directory then cd $OUTPUT
# make sure output directory was made by submission script.

# If you can use the _ica.txt output file (from ICA) under
# ICA_durs.txt is a csv file with list of full paths to
# along with the output go to

# Run qualify analysis...

# Prepare functional mask...
Preprocessing

FSL 4.1

DATA INPUT
- MOTION CORRECTION
- SKULL REMOVAL
- SMOOTHING
- BANDPASS FILTERING

PREPROCESS

FUNCTIONAL
ANATOMICAL
- SKULL REMOVAL

REGISTRATION
NORMALIZATION

QUALITY ANALYSIS
- GROUP ICA ANALYSIS

INDIVIDUAL

GROUP NETWORKS
Quality Analysis

MOTION CORRECTION

BANDPASS FILTER

NOISE REDUCTION
Quality Analysis

MOTION CORRECTION

BANDPASS FILTER

NOISE REDUCTION

ica+ Motion Report

Experiment: /scratch/users/vsochat/Project/NYUALL
QA Report Name: NYUALL
Total Subjects: 216
Flagged Subjects: 38
Passing Subjects: 178
Rotational Motion Benchmark (deg): 2.0
Translational Motion Benchmark (mm): 2.0

Flagged Subject Summary

rotation y: /scratch/users/vsochat/Project/NYUALL/ica0010003.ica
rotation x: /scratch/users/vsochat/Project/NYUALL/ica0010005.ica
rotation x: /scratch/users/vsochat/Project/NYUALL/ica0010014.ica
translation z: /scratch/users/vsochat/Project/NYUALL/ica0010015.ica
rotation x: /scratch/users/vsochat/Project/NYUALL/ica0010018.ica
rotation x: /scratch/users/vsochat/Project/NYUALL/ica0010025.ica
translation z: /scratch/users/vsochat/Project/NYUALL/ica0010030.ica
rotation x: /scratch/users/vsochat/Project/NYUALL/ica0010032.ica
rotation x: /scratch/users/vsochat/Project/NYUALL/ica0010041.ica
Quality Analysis

MOTION CORRECTION

BANDPASS FILTER

NOISE REDUCTION

Flagged Subjects

0010003.ica

NEFLIRT estimated rotations (radians)

NEFLIRT estimated translations (mm)

NEFLIRT estimated mean displacement (mm)

rotation x: /scratch/users/vsochat/Project/NYUALL/ica/0010041.ica
Derive Single Subject Networks

\[ \mathcal{L} \leq R \]

**MELODIC Component 5**

2.49% of explained variance; 2.12% of total variance
Derive Single Subject Networks

1.71 % of explained variance; 1.46 % of total variance
Derive Single Subject Networks

MELODIC Component 46

1.21% of explained variance; 1.03% of total variance
Derive Single Subject Networks
Derive Group Networks

PCA estimates

Eigenspectrum Analysis

- ordered Eigenvalues
- % of expl. variance

Number of included components

Range: 0.00 to 1.00
Derive Group Networks

PCA estimates

Eigenspectrum Analysis

Subject/Session modes

Component No.

Number of included components
Derive Group Networks

PCA estimation

MELODIC Component 26

0.64% of explained variance; 0.44% of total variance
Template Matching Procedure
Template Matching Procedure

```
#!/usr/bin/env python2

***

pyMatch Component --> Template Matching in Python

This python script uses MRtools to read in a nifti image template (Data class), and match a list of user specified images to the template (Match class). Usage could be to match components (individual functional networks) from an individual lca analysis to the most similar component from a group lca (group networks) analysis, OR to match a group significant result from one modality to group networks. The images used for matching will first be filtered (getting rid of high frequency components) and then matched.

Inputs should be as follows:

**INPUT:**

-h, --help Print this usage
-s --subs= Single column text file w/ list of subject (or group) folders containing components
-t --template= The template image to match, such as a group network
-l --images= Single column text file w/ a list of component images in folders
-o --output= Name of output folder. If not specified, will use cwd

If you input a list of subjects longer than one, keep in mind that each should have the corresponding component images in the designated folder. Whether 3D or 4D, the first timepoint will be used by default to extract data. If an image's first timepoint is empty, the script will try the second. If the second ts also empty, it will exit with error, because there is something wrong with your template or image!

**USAGE:** python pymatch.py --subs=sublist.txt --template=/path/to/image.nii.gz --images=imgelist.txt --output=/path/for/oufile

Intended usage is for one template for 1+ subjects/groups with a list of component images.

**OUTPUT:** (template_name)_bestcomps.txt and (template_name)_beststats.txt w/ top 3 components for each subject/group

***

_author_ = "Vanessa Sochat (vsochat@stanford.edu)"
__version__ = "$Revision: 1.0 $"```
Template Matching Procedure

```
#!/usr/bin/env python2

pyMatch Component -> T
MRtools.py - Python tools for image manipulation and filtration
This python script uses
a list of user specified MRtools.Data: Translate between images of different dimensions and formats
components (individual MRtools.Filter: Determine goodness of an input image and a frequency timeseries
most similar component MRtools.Match: Return match score for two MRtools Data objects
significant result from input will first be filtered
Inputs should be as follows:
INPUT:
-h, --help Print this
-s --subs= Single SAMPLE USAGE: Make sure script is somewhere on your path
-t --template= The template
-l --images= Single To use Data class:
-o --output= Name of
If you input a list of corresponding component:
empty, the script will
error, because there is
no template to use with:
To use Filter class with Group Network Image:
>> import MRtools
>> Image = MRtools.Data('myimage.nii.gz')
>> Filter = MRtools.Filter()
>> Filter.isGoodImage('timeseries.txt','frequency.txt')

USAGE: python pymatch.py

Intended usage is for
To use Match class with Image:
>> import MRtools
>> Template = MRtools.Data('myimage.nii.gz')
>> Match = MRtools.Match(Template)
>> Match.setIndexCrit("">0")
>> Match.genIndexMNI()

_author_ = "Vanessa S"
_version_ = "$Revision: 1.0 $"
```
Statistical Analysis

1 Sample T-Tests
Statistical Analysis

2 Sample T-Tests
Statistical Analysis

2 Sample T-Tests
Goals

Are there significant differences in the default mode network (DMN) between ADHD and Control?
Support Vector Machines (SVM)

\[
\begin{align*}
\min_{w,b,\xi} & \quad \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \\
\text{subject to} & \quad y_i (w^T \phi(x_i) + b) \geq 1 - \xi_i, \\
& \quad \xi_i \geq 0.
\end{align*}
\]
Feature Extraction

![SPM8 (vanessa): Graphics](image)

Statistics: p-values adjusted for search volume

<table>
<thead>
<tr>
<th>Section</th>
<th>Cluster-Size</th>
<th>cluster-size</th>
<th>mm x mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0001</td>
<td>0.000, 0.000, 994, 0.000</td>
<td>0.000, 0.000, 12, 0.000</td>
<td></td>
</tr>
<tr>
<td>0.0002</td>
<td>0.000, 0.000, 491, 0.000</td>
<td>0.000, 0.000, 12, 0.000</td>
<td></td>
</tr>
<tr>
<td>0.0003</td>
<td>0.000, 0.000, 25, 0.000</td>
<td>0.000, 0.000, 12, 0.000</td>
<td></td>
</tr>
</tbody>
</table>

Table shows 3 local maxima more than 0.0004 apart.

- Height threshold: T = 4.88, p = 0.000
- SPM threshold: k = 0 (FWE)
- Extent threshold: k = 0 (FWE)
- Expected voxel intensity: 23.016 ± 1.274
- Intensity: 1

Crosshair Position:
- mm: 20.45, 25.55
- vx: 23.016 ± 1.274
- Intensity: 1

File: /sampleTT245_b0s_b0s
- Dimensions: 45 x 54 x 45
- Data type: uint16
- Intensity: 3.05185e-05 x spm - algebra

Voxels:
- 4 x 4 x 4
- Origin: 23.53, 32.519
- Dir Cos: 1.000, 0.000, 0.000
- Intensity: 0.000, 0.000, 0.000
- Intensity: 0.000, 0.000, 0.000
- Intensity: 0.000, 0.000, 0.000

Reorient images: Reset

Full Volume: Hide Crosshairs
World Space: bilinear interp
Auto Window: Add Blobs
Feature Extraction
Data Transformation

Z score (no transform)
Data Transformation

quadratic
Data Transformation

square root
Data Transformation
Simple Correlations

Correlations for Log Z Scores, ADHD vs Control

[Diagram showing correlation matrix with ADHD and CONTROL axes ranging from 20 to 180]
Simple Correlations

Correlations for Quadratic Z Scores, ADHD vs Control
Simple Correlations
Simple Correlations

Correlations for Absolute Value Z Scores, ADHD vs Control

ADHD

CONTROL

20

40

60

80

100

120

140

160

180

20

40

60

80

100

120

140

160

180

ADHD

CONTROL
Simple Correlations

Correlations for Sigmoid Z Scores, ADHD vs Control
Future Work

Use filtered data masked with DMN
Find attention network
Explore components
Long term: ICA → Diagnosis
Thank You!