Identifying Autism from Resting-State fMRI Using Long Short-Term Memory Networks

Nicha C. Dvornek, Pamela Ventola, Kevin A. Pelphrey, and James S. Duncan

MLMI 2017

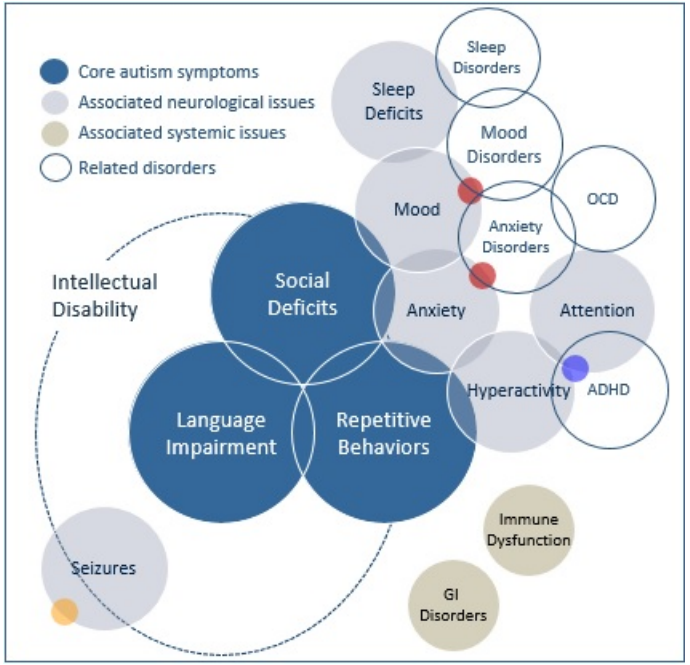
Quebec City, Canada

September 10, 2017

Yale SCHOOL OF MEDICINE

Motivation: Derive objective biomarkers for autism spectrum disorder (ASD)

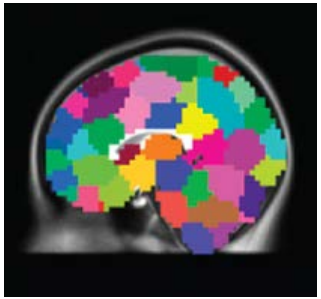
- ASD: Neurological developmental disorder
  - Prevalence in U.S.: 1 in 68 children
- Resting-state fMRI (rsfMRI) used to study pathophysiology
- High impact of imaging biomarkers on:
  - Understanding causes of ASD
  - Diagnosis
  - Design of therapies
  - Monitoring/predicting treatment outcomes



www.autismspeaks.org

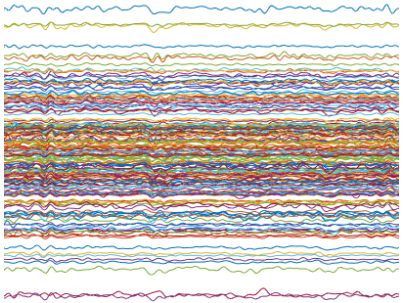
Standard approach: Classify ASD/typical control using rsfMRI connectivity

Brain Parcellation

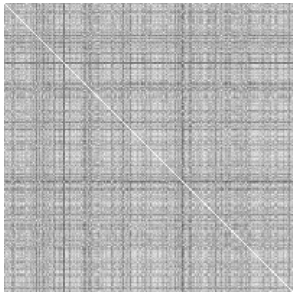


From Craddock et al., Nature Methods 2013

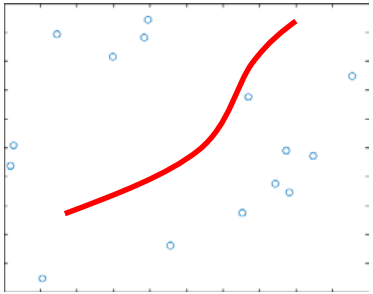
Resting-State fMRI Mean Time Series



Connectivity Matrix



Machine Learning Algorithm

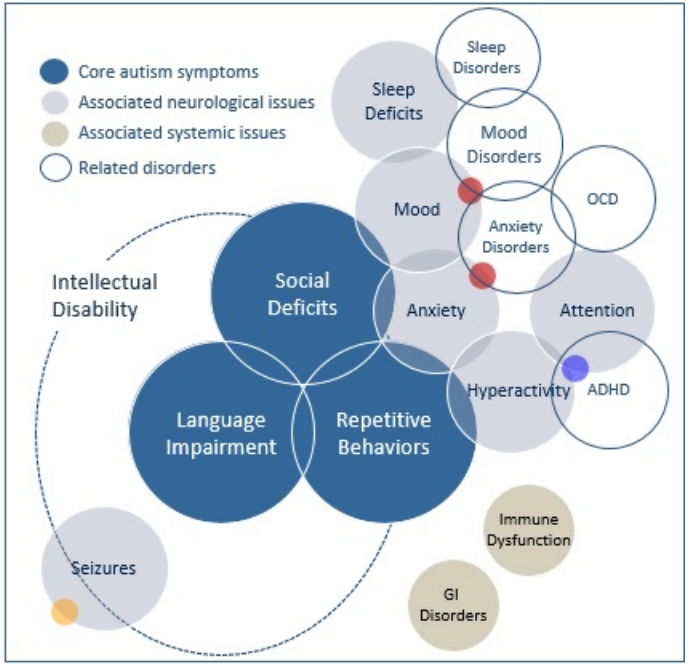


Popular approaches:
Support Vector Machines
Random Forest
Ridge Regression

*Biomarkers: pairwise connections important for classification

Challenge: Heterogeneity of ASD

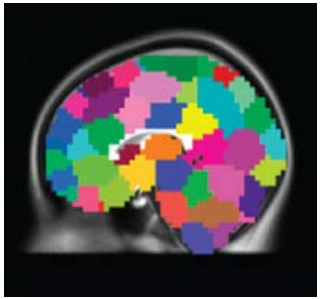
- Wide range of symptoms and severity
- Studies using small (N<100), homogeneous datasets:
  - \+ High accuracy (>70%) within sample
  - – Poor generalization
- Studies using large, heterogeneous datasets:
  - \+ Better representation of population
  - – Low accuracy (~60%), likely due to heterogeneity



www.autismspeaks.org

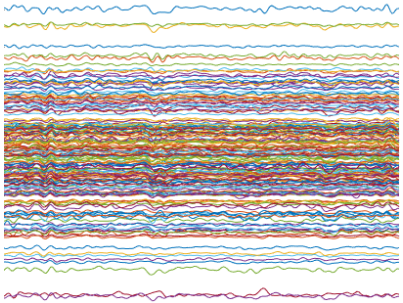
Standard approach: Classify ASD/typical normal using rsfMRI connectivity

Brain Parcellation

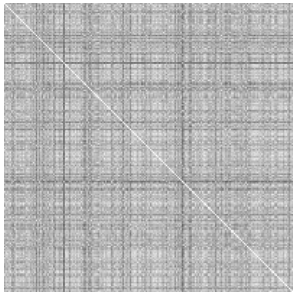


From Craddock et al., Nature Methods 2013

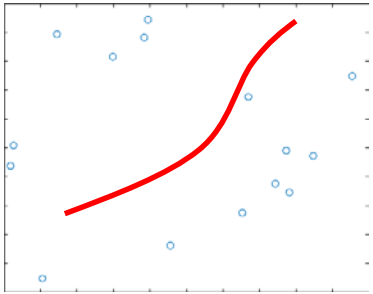
Resting-State fMRI Mean Time Series



Connectivity Matrix



Machine Learning Algorithm

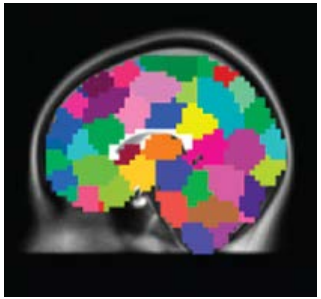


Popular approaches:
Support Vector Machines
Random Forest
Ridge Regression

*Biomarkers: pairwise connections important for classification

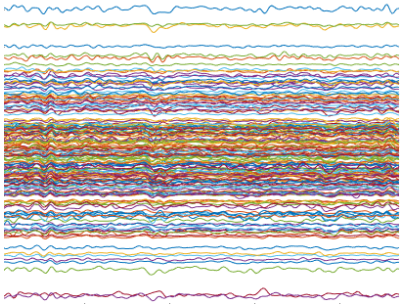
Proposed approach: Classify ASD/typical control directly from rsfMRI time-series

Brain Parcellation

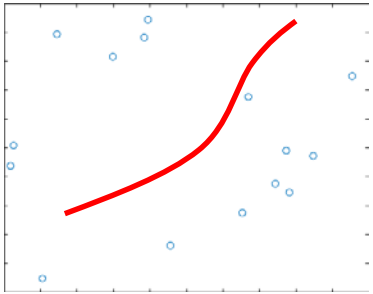


From Craddock et al., Nature Methods 2013

Resting-State fMRI Mean Time Series



Machine Learning Algorithm



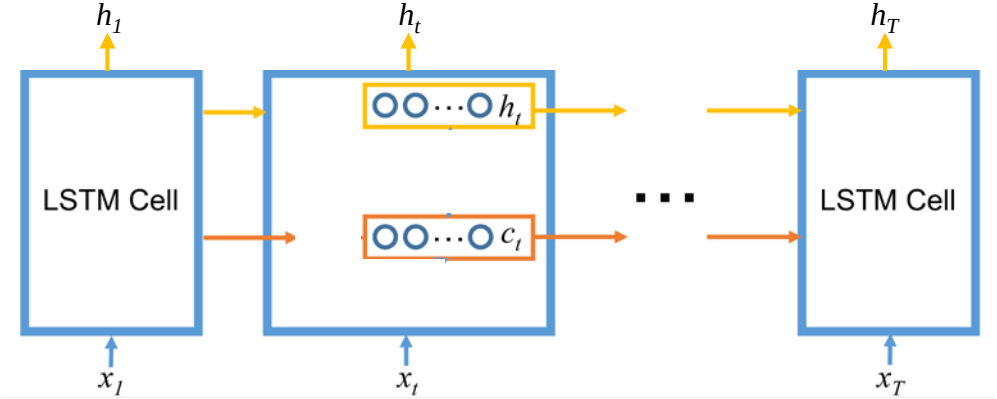
Our approach: Recurrent Neural Network with Long Short-Term Memory (LSTM)

*Biomarkers: extracted networks from LSTM weights

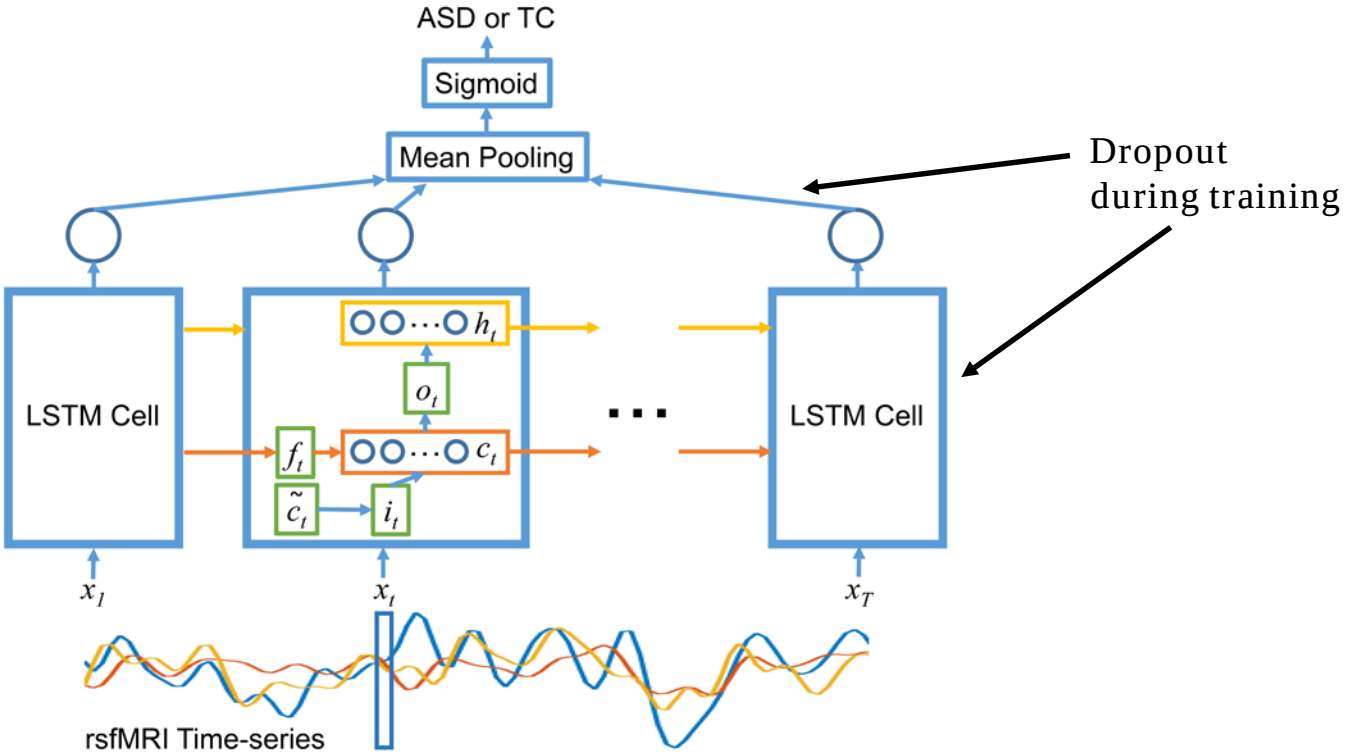
Why LSTMs?

- Designed for modeling temporal data
  - LSTM cell takes data and previous cell information ("memory") as inputs
- 2 types of recurrent information
  - Hidden state (output)
  - Cell state – helps learn long-term dependencies

$$
\begin{aligned}
i_t &= \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \\
f_t &= \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \\
\tilde{c}_t &= \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \\
c_t &= i_t * \tilde{c}_t + f_t * c_{t-1} \\
o_t &= \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \\
h_t &= o_t * \tanh\left(c_t\right)
\end{aligned}
$$

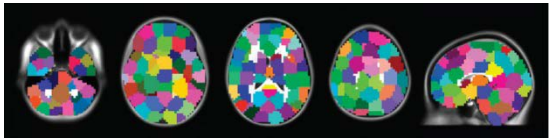


Network Architecture

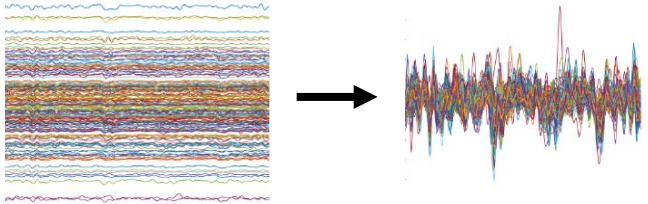


Dataset and Preprocessing

- Autism Brain Imaging Data Exchange (ABIDE) Dataset I
  - 539 ASD + 573 typical controls
  - Phenotypic + neuroimaging data (resting-state fMRI, structural MRI)
- Preprocessed by the Preprocessed Connectomes Project:
  - Connectome Computation System pipeline (1100 subjects)
  - Craddock 200 atlas (functional homogeneity)
- ROI mean time-series preprocessing
  - Normalized to percent signal change
  - Resample to 2 second interval

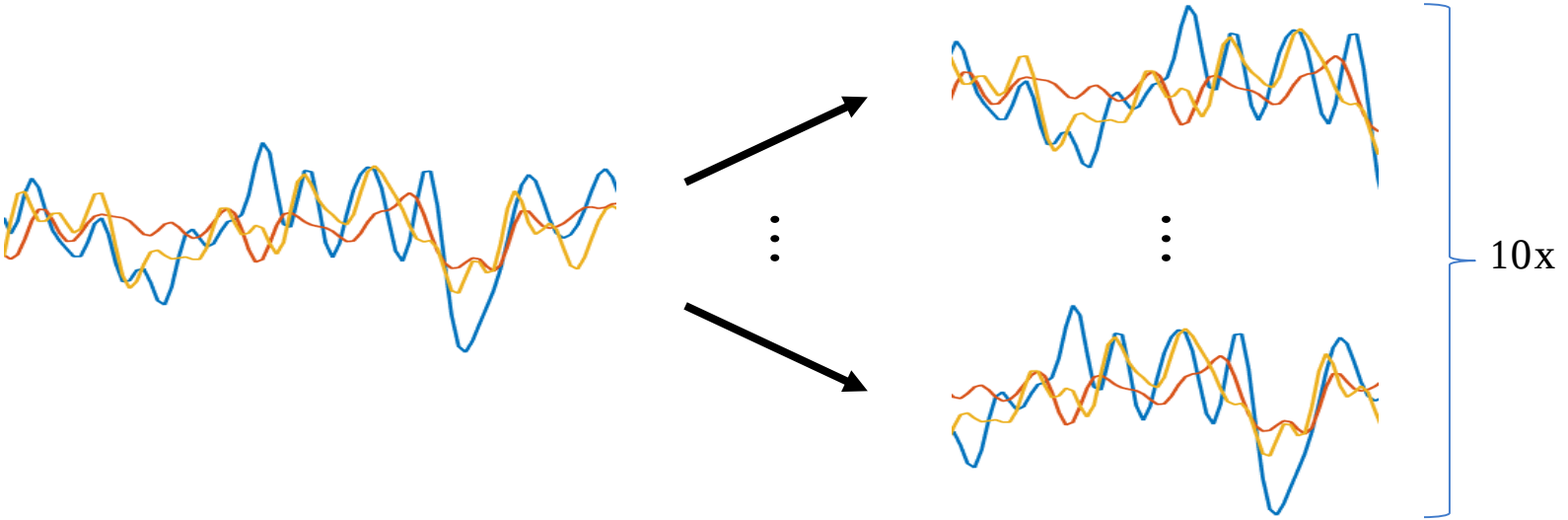


Craddock et al., Nature Methods 2013



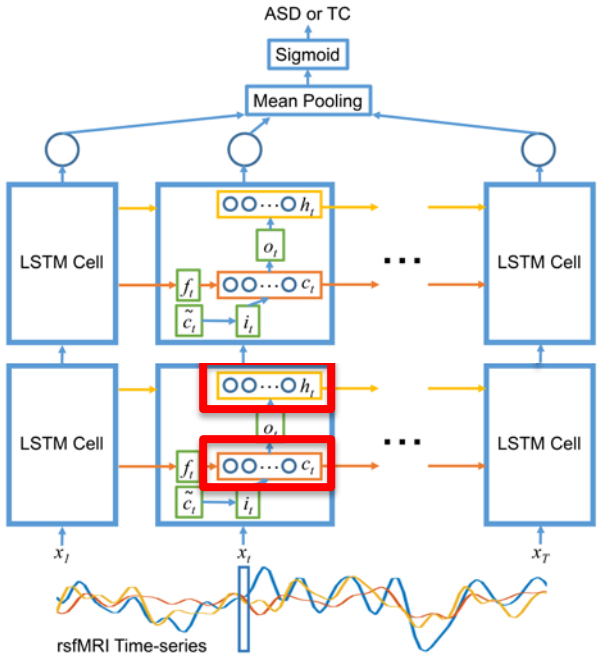
Data Augmentation

- Crop input time courses to fixed sequence length ($T = 90$) from random starting points
- 10 random crops / subject → $N = 11000$ samples



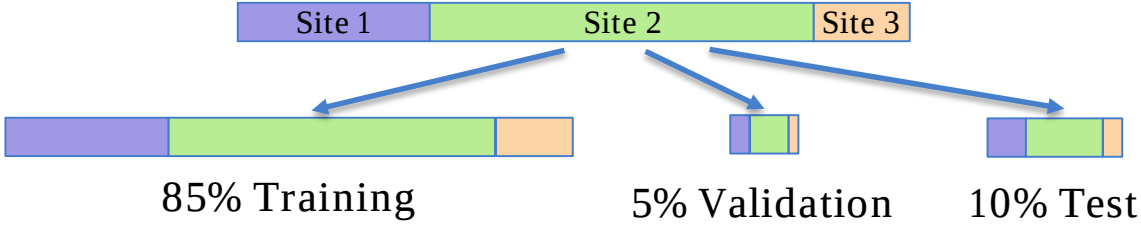
Experimental Methods

- Implementation in Keras
  - Loss function: Binary cross-entropy
  - Optimizer: Adadelta
  - Dropout rate: 0.5
- Network variations tested
  - Removing data augmentation or dropout
  - # LSTM nodes: 8, 16, 32, 64
  - Using only LSTM output at final timestep
  - 2 LSTM layers



Algorithm Evaluation

- 10-fold cross-validation, stratified across sites



- Compare to previous studies trained on ABIDE rsfMRI
- Accuracy measures
  - Sequence accuracy
  - Subject accuracy: Average score of all sequences per subject
  - Difference from chance:
    (Model accuracy) – (Accuracy by random assignment)

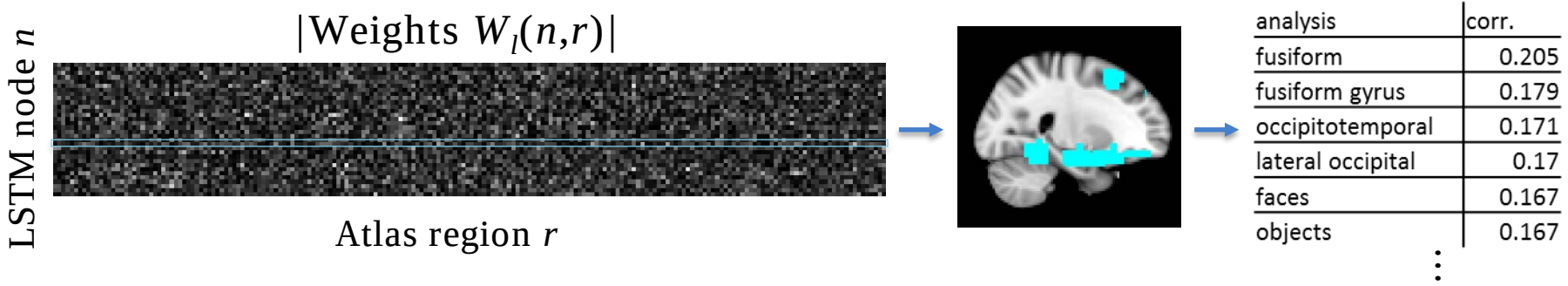
Classification Accuracy

Classification Method	Validation Method	Number of Subjects	Mean (SD) Sequence Accuracy (%)	Difference from Chance (%)	Mean (SD) Subject Accuracy (%)	Difference from Chance (%)
Plitt et al. [13]	CV10	178	-	-	69.7	19.7
Chen et al. [3]	Train/Val	252	-	-	66	16
Abraham et al. [1]	CV10	871	-	-	66.9 (2.7)	13.2
Nielsen et al. [12]	LOO	964	-	-	60.0	6.4
Ghiassian et al. [9]	Train/Val	1111	-	-	59.2	7.6
LSTM8	CV10	1100	65.6 (4.1)	13.7	66.7 (5.3)	14.8
LSTM16	CV10	1100	65.3 (4.8)	13.3	66.8 (5.4)	14.9
LSTM32	CV10	1100	**66.8 (4.5)**	**14.9**	**68.5 (5.5)**†	**16.6**
LSTM64	CV10	1100	65.8 (3.8)	13.9	67.5 (4.4)†	15.5
LSTM32_NoAug	CV10	1100	-	-	61.4 (4.5)*	9.5
LSTM32_NoDrop	CV10	1100	59.7 (2.3)	7.7	61.8 (4.0)*	9.9
LSTM32_Last	CV10	1100	62.2 (3.3)	10.3	64.5 (4.5)†	12.5
LSTM32x2	CV10	1100	66.3 (4.2)	14.4	67.5 (5.0)†	15.5

†Significant difference between sequence and subject accuracies *Significant difference compared to LSTM32

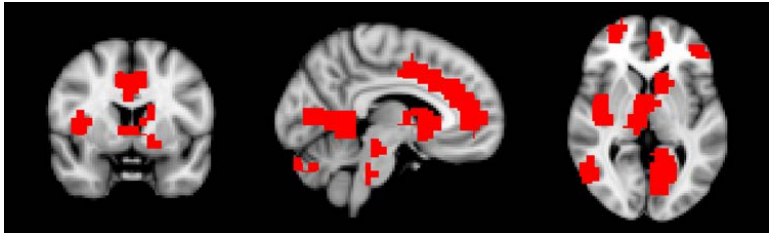
Model Interpretation

- Large weights $W_l(n,r)$ denote atlas region r has strong influence on LSTM node n for layer l
- Create binary mask of regions with large weights (> Mean + 2xSD)
- Input mask of important regions into Neurosynth
  - Meta-analysis correlating 10,000 fMRI studies with 3000 descriptors

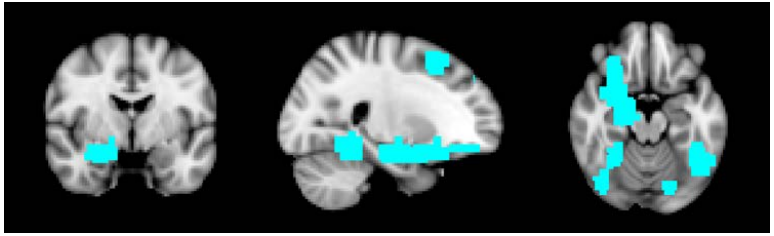


analysis	corr.
fusiform	0.205
fusiform gyrus	0.179
occipitotemporal	0.171
lateral occipital	0.17
faces	0.167
objects	0.167
⋮	

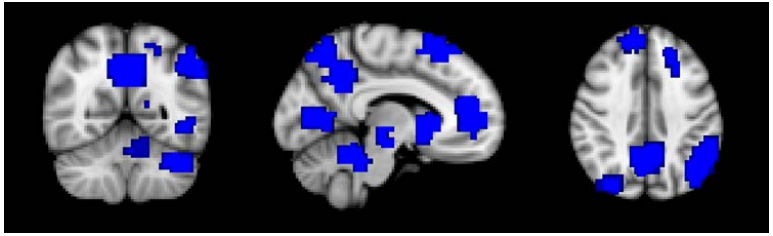
Heavily weighted regions in one LSTM node encode potential brain networks



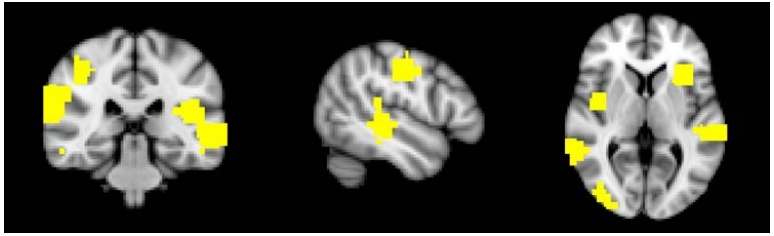
Pain, reward, anticipation, incentive



Faces, objects, word form, emotional, visual



Default mode, reward, listening, mental states



Listening, sounds, theory of mind, social

Conclusions and Future Work

- Contributions
  - LSTM-based model for classifying ASD/typical control directly from rsfMRI time-series
  - Higher classification accuracy than previous methods using rsfMRI connectivity as inputs, tested on majority of ABIDE I dataset
  - Model weights interpreted as meaningful brain networks important for classification
- Future work
  - Use fMRI image sequence as input?
  - Incorporate demographic information
  - Combine structural MRI information

Thank you!

- Dr. Ventola and Dr. Pelphrey (ASD network interpretation)
- NIH Grants T32 MH18268 and R01 NS035193

- Contact: nicha.dvornek@yale.edu