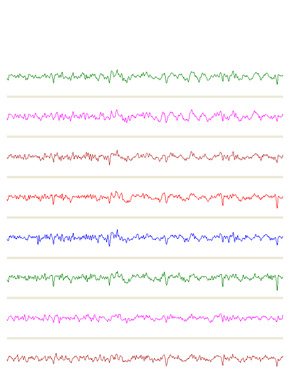
Interpretable and Lightweight Machine Learning Approach for Autism Classification Using Biomarkers Derived from Multi-trial Resting EEG

Michelle Su[1,2], Deeksha M. Shama[2,3], Archana Venkataraman[2,3]

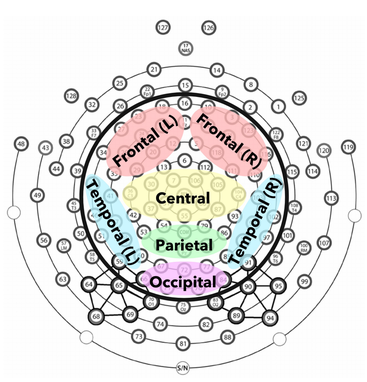
Lincoln-Sudbury Regional High School, 390 Lincoln Rd, Sudbury, MA[1]; Department of Electrical and Computer Engineering, Boston University, 8 St. Mary's St, Boston, MA[2]; Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD[3]

INTRODUCTION

- Autism Spectrum Disorder (ASD): heterogenous, hard to identify biomarkers for diagnosis
- Electroencephalography (EEG): valuable tool for biomarker identification—non-invasive, high temporal resolution, affordable



Subset of EEG signals
(Figure from Biopac)



EEG head map
(Template from Neuhaus *et. al.*[1])

GOALS

- Predict ASD from EEG using machine learning tools
- Explore distributional features across EEG trials
- Identify important features for ASD diagnosis

METHODS

1) DATA[1]

Group	Size	Male/Female	Age (years)
Autism (ASD)	94	55/39	12.5±2.9
Control (CON)	96	49/47	12.9±2.8

3) FEATURE EXTRACTION

- Get Power Spectral Density (PSD) at each electrode and frequency band using Welch's estimate {hamming, 512 point}

Two Feature Aggregation Types

Mean-EEG	Across-trial
PSD of Mean EEG (conventional)	PSD of EEG for each trial
Mean, std. deviation, skew, & kurtosis of PSD in each frequency band	Mean, std. deviation, skew, & kurtosis of PSD in each frequency band
# Electrodes × # Within-band features × # Frequency bands = 128 x 5 x 4 = 2560 features	Mean and std. deviation of band features <u>across trials</u>
	2 x 2560 = 5210 features

2) PREPROCESSING

- Low pass (<100 Hz) and Notch (60Hz) filters
- Wavelet threshold and bad segment rejection
- ICA and bad channel rejection[1]

4) TRAINING AND TESTING

- Four classifiers trained on the two standardized feature types separately
  1. Logistic regression (LR)
  2. Random forest without bootstrap (RF)
  3. Kernel-support vector machine (SVM)
  4. An artificial neural network (ANN)
- Stratified 5-fold cross-validation
  - Repeated 5x
- Evaluation: accuracy, F1 score, precision, recall, specificity, AUROC, AUPRC metrics
- Feature importance: determined by magnitude of mean feature weights in cross-validation across-trial LR models

RESULTS

REPEATED 5-FOLD CROSS-VALIDATION MODEL PERFORMANCE

Model	Features	Accuracy	F1 score	Precision	Recall	Specificity	AUROC	AUPRC
LR	Across-trial	**0.678±0.083**	**0.678±0.083**	**0.674±0.101**	0.693±0.108	0.664±0.123	**0.748±0.073**	**0.751±0.086**
	Mean-EEG	0.608±0.060	0.597±0.088	0.601±0.082	0.604±0.117	0.61±0.085	0.648±0.070	0.678±0.084
RF	Across-trial	0.660±0.085	0.646±0.102	0.665±0.112	0.649±0.147	**0.681±0.121**	0.742±0.082	0.744±0.089
	Mean-EEG	0.614±0.068	0.589±0.076	0.626±0.095	0.571±0.109	0.662±0.115	0.672±0.092	0.693±0.088
SVM	Across-trial	0.659±0.079	0.656±0.089	0.659±0.101	0.679±0.146	0.654±0.106	0.734±0.087	0.743±0.093
	Mean-EEG	0.577±0.064	0.532±0.087	0.608±0.136	0.509±0.161	0.669±0.146	0.608±0.146	0.619±0.102
ANN	Across-trial	0.654±0.088	0.682±0.099	0.650±0.108	**0.751±0.156**	0.590±0.160	0.721±0.095	0.709±0.084
	Mean-EEG	0.622±0.079	0.616±0.112	0.600±0.112	0.661±0.160	0.564±0.158	0.651±0.108	0.683±0.098

Table 1: Metrics from all experiments. Significant differences ($p<0.05$) from two-sample t-tests between the two feature types are highlighted in green. (No significant differences between model choices found.)

REPEATED 5-FOLD CROSS-VALIDATION ACCURACY

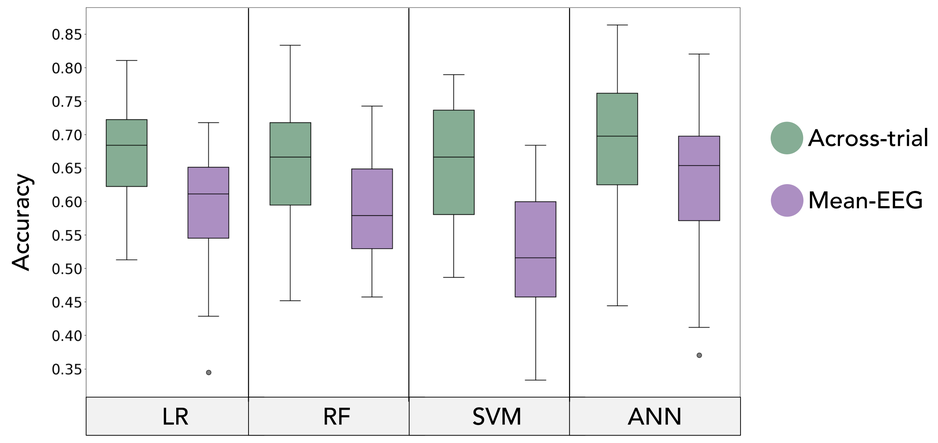
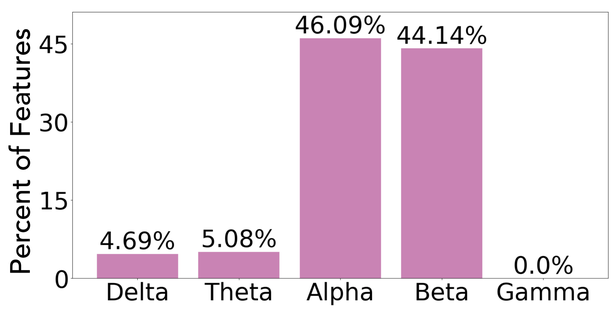


Fig 1: Accuracy of all models using the two feature types. Across-trial feature models show higher accuracy.

DISTRIBUTION OF TOP 5% MOST IMPORTANT FEATURES



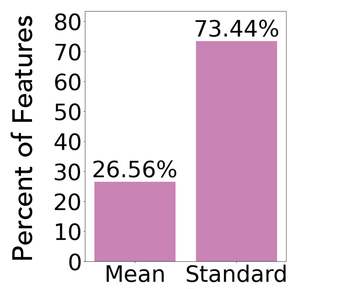


Fig 2: Frequency band distribution (left) and across-trial feature distribution (right) of the top 5% most important features. P-values from chi-square goodness-of-fit tests with the null hypothesis of a uniform distribution.

SIGNIFICANTLY DIFFERENT BRAIN REGIONS BETWEEN GROUPS

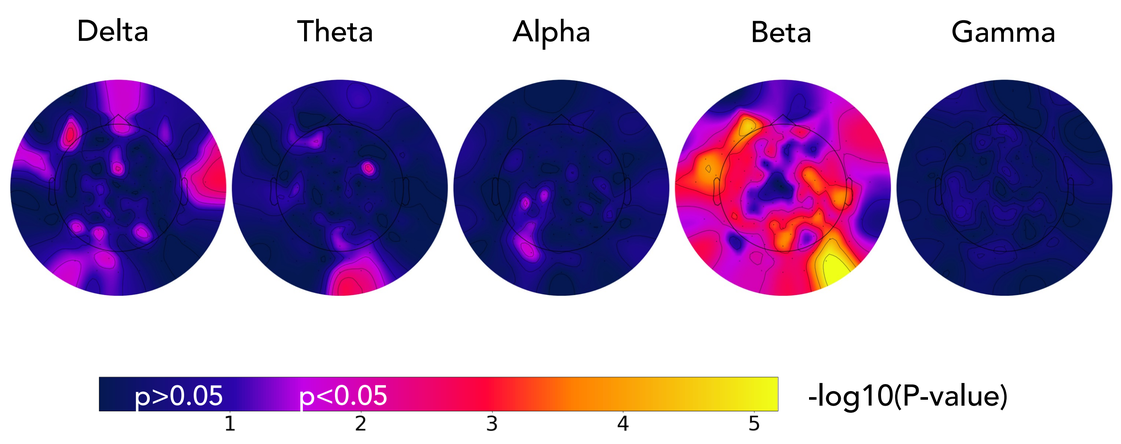


Fig 3: Topological head plots of significant regions based on two-sample t-tests for the difference between ASD and CON in the mean of summed across-trial features adjusted by LR model weights.

DISCUSSION AND CONCLUSION

- Features recording variability across trials are more important
  - Trial-to-trial variability differs between ASD and CON groups[2]
- The choice of machine learning model is not significant
- Beta frequency: highest diagnostic value in all brain regions
- Occipital and frontal temporal regions: higher diagnostic value across multiple frequency bands
- Future work:
  - Incorporate features from task-EEG data
  - Use deep learning for feature extraction

ACKNOWLEDGEMENTS

We thank the Autism Center For Excellence at the University of Virginia and Seattle Children's Research Institute for acquiring the data. Thank you to Prof. Archana Venkataraman and Deeksha M. Shama for providing unwavering support and guidance. Thank you to BU RISE for this incredible opportunity.

REFERENCES

(1) Neuhaus, E.; Lowry, S. J.; Santhosh, M.; Kresse, A.; Edwards, L. A.; Keller, J.; Libsack, E. J.; Kang, V. Y.; Naples, A.; Jack, A.; Jeste, S.; McPartland, J. C.; Aylward, E.; Bernier, R.; Bookheimer, S.; Dapretto, M.; Van Horn, J. D.; Pelphrey, K.; Webb, S. J. Resting State EEG in Youth with ASD: Age, Sex, and Relation to Phenotype. Journal of Neurodevelopmental Disorders **2021**, 13 (1). https://doi.org/10.1186/s11689-021-09390-1.

(2) Milne, E. Increased Intra-Participant Variability in Children with Autistic Spectrum Disorders: Evidence from Single-Trial Analysis of Evoked EEG. Frontiers in Psychology **2011**, 2. https://doi.org/10.3389/fpsyg.2011.00051.