Autism Classification Using Topological Features and Deep Learning - A Cautionary Tale

Archit Rathore[1], Sourabh Palande[1], Jeffrey S. Anderson[1], Brandon A. Zielinski[1], P. Thomas Fletcher[2], and Bei Wang[1]

[1]University of Utah [2]University of Virginia

*This work was supported in part by NSF IIS 1513616 and NIH R01EB022876

Overview

- Explore utility of topological features in ASD classification using rs-fMRI
- Topological features represented as:
  - Persistence Diagrams (PD), Persistence Images (PI) and Persistence Landscapes (PL)
- Classification using SVMs, random forests and neural networks
- Augmenting topological features with functional correlations provides best accuracy
- Evaluate statistical significance of improvement in accuracy using permutation test
- Improvement in classification accuracy due to topological features is not always statistically significant
- **A cautionary tale to the practitioners regarding the limited discriminative power of topological features derived from fMRI data for the classification of autism.**

Dataset and Preprocessing

- Craddock 200 (CC200), Craddock 400 (CC400) from ABIDE Preprocessed dataset
- 505 ASD and 530 typically developing control (TDC) subjects
- 200 × 200 or 400 × 400 connectivity matrices per subject (Pearson correlation)
- Map connectivity matrix M to point cloud X with pairwise distance d_X

$$d_X(x, y) = \sqrt{1 - M(x, y)}$$

Persistent Homology

- Measures the evolution of topological features across varying scale (α)
- Topological features: dim-0 (connected components), dim-1 (tunnels), dim-2 (voids)
- As α changes, birth (b) and death (d) of features is tracked
- Represented as Barcodes or PDs. PIs, PLs Derived from PDs
- Space of PDs can be endowed with distance metrics, kernels

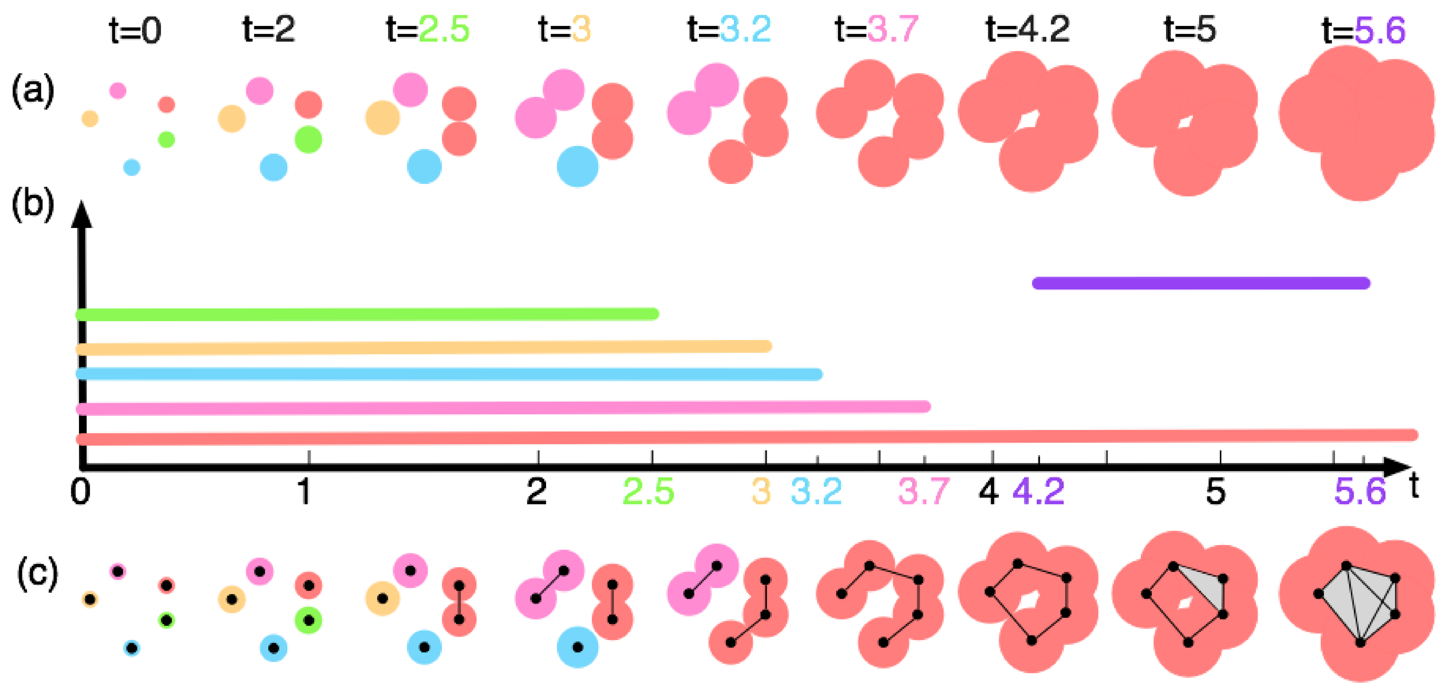
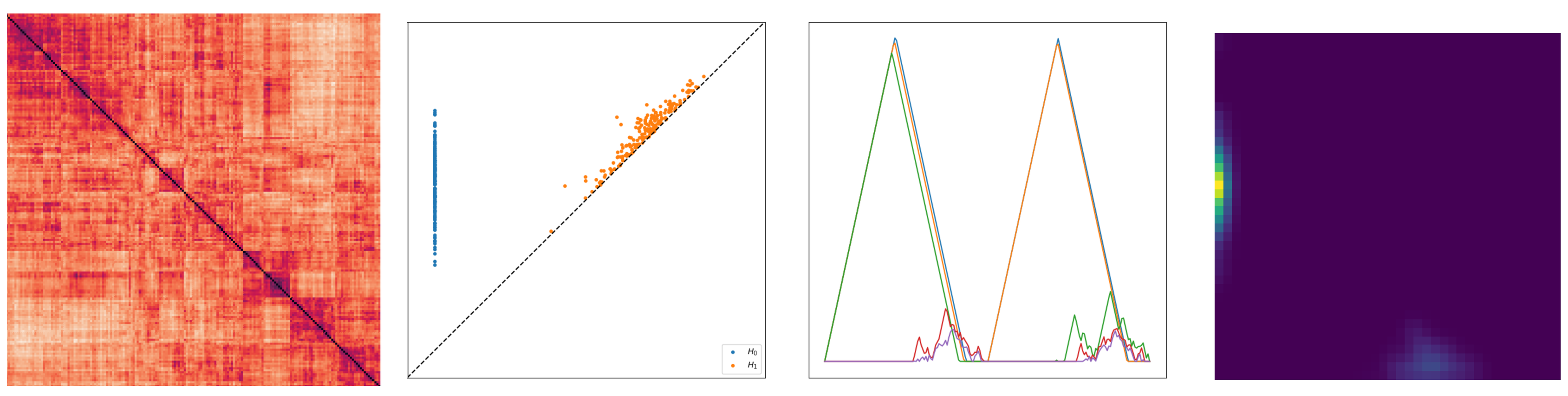


Figure 1: Computing persistent homology of a point cloud in $\mathbb{R}^2$



Correlation Matrix | Persistence Diagram | Persistence Landscape | Persistence Image

Conclusions

- Modest improvement in classification on combining topological and correlation features
- Topological features capture some information that is not captured by correlations
- Results affected by heterogeneity of ASD, fMRI acquisition strategy

Method

- Flattened correlation matrix mapped as vector inputs for a branch of the network
- Persistence diagrams are passed through *projection layer* as described by Hofer et al [1]
- Projection layer maps persistence diagrams to vector space
- Experiments with 3, 5 and 7 layered networks

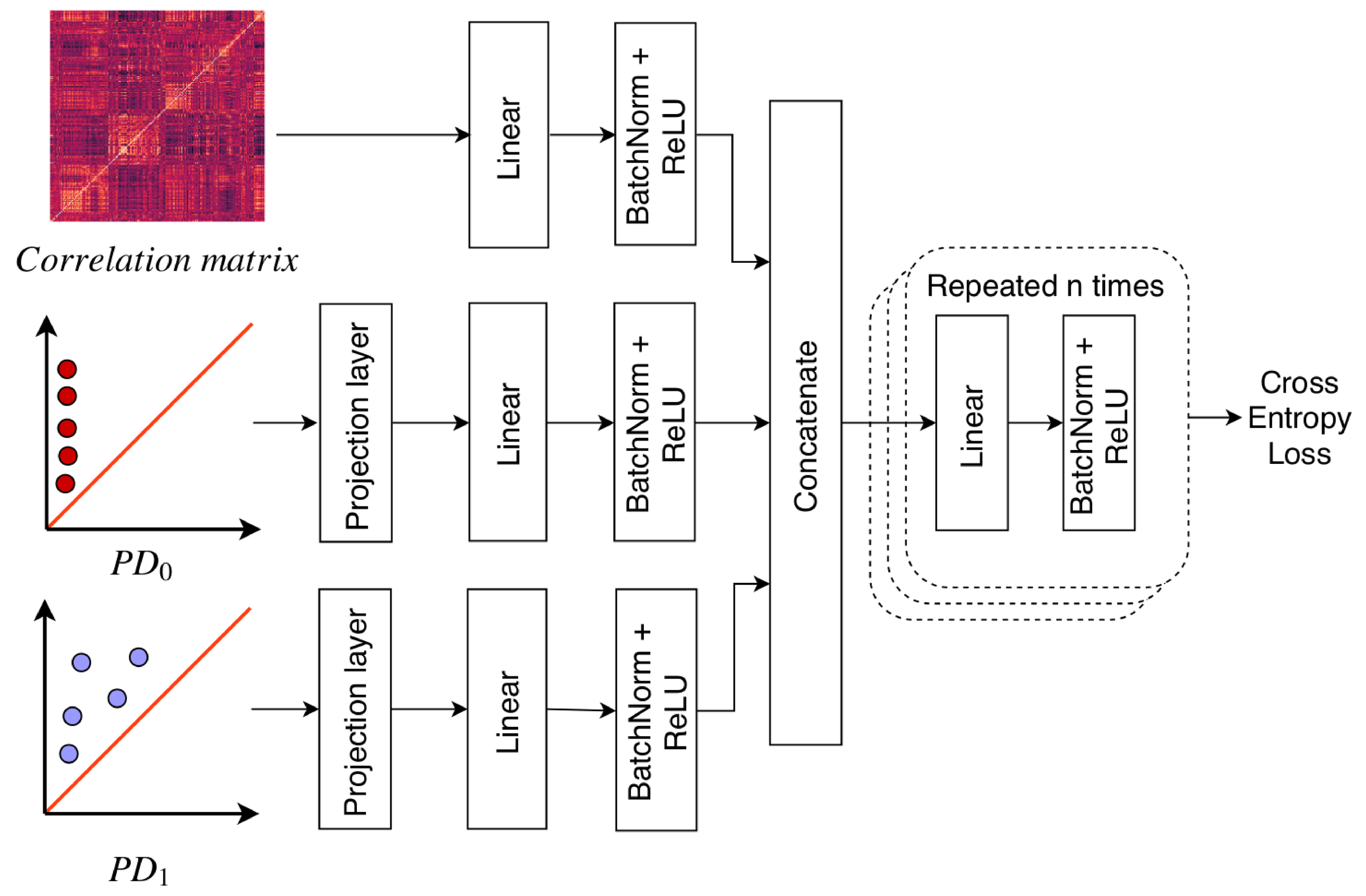


Figure 2: Neural network architecture that combines correlation and topological features

- Projection layer: $S_{\boldsymbol{\mu},\boldsymbol{\sigma},\nu} = \sum_{(x,y)\in\rho(A)} s_{\boldsymbol{\mu},\boldsymbol{\sigma},\nu}(x, y)$, where

$$s_{\boldsymbol{\mu},\boldsymbol{\sigma},\nu}(x, y) = \begin{cases} e^{-\sigma_x^2(x-\mu_x)^2 - \sigma_y^2(y-\mu_y)^2} & y \in [\nu, \infty) \\ e^{-\sigma_x^2(x-\mu_x)^2 - \sigma_y^2(\ln\frac{y}{\nu}+\nu-\mu_y)^2} & y \in (0, \nu) \\ 0, & y = 0 \end{cases}$$

- $(x, y) = (death + birth, death - birth)$ - $\pi/4$ rotation of PD
- $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ - learned parameters assigning importance to different regions of the diagram
- Compare the above model with SVM and Random Forest on correlation features
- Kernel-SVM defined on PD - persistence scale space kernel [2]. PL and PI are vector representations and directly usable with SVM and random forests

Results

Model	CC200	CC400	Model	CC200	CC400	Model	CC200	CC400
SVM_{Corr}	65.41	66.33	-	-	-	-	-	-
RF_{Corr}	64.81	63.92	-	-	-	-	-	-
$NN3_{Corr}$	68.35	63.92	-	-	-	-	-	-
$NN5_{Corr}$	68.46	65.58	-	-	-	-	-	-
$NN7_{Corr}$	67.10	62.06	-	-	-	-	-	-
SVM_{PD}	53.03	53.69	SVM_{PI}	54.54	53.76	SVM_{PL}	53.03	53.69
RF_{PD}	-	-	RF_{PI}	52.25	53.04	RF_{PL}	52.51	53.12
$NN3_{PD}$	56.06	55.90	$NN3_{PI}$	58.56	56.10	$NN3_{PL}$	55.36	54.24
$NN5_{PD}$	56.15	56.04	$NN5_{PI}$	59.09	57.39	$NN5_{PL}$	55.18	55.72
$NN7_{PD}$	55.48	54.33	$NN7_{PI}$	56.75	55.58	$NN7_{PL}$	54.85	53.67
$SVM_{PD+Corr}$	65.86	63.36	$SVM_{PI+Corr}$	64.25	62.68	$SVM_{PL+Corr}$	65.65	64.12
NN3 $_{PD+Corr}$	**69.19**	**67.84**	**NN3** $_{PI+Corr}$	**67.2**	**67.02**	**NN3** $_{PL+Corr}$	**68.5**	**66.76**
$NN5_{PD+Corr}$	68.20	66.03	$NN5_{PI+Corr}$	66.87	66.23	$NN5_{PL+Corr}$	67.45	66.48
$NN7_{PD+Corr}$	64.47	61.25	$NN7_{PI+Corr}$	65.10	64.16	$NN7_{PL+Corr}$	67.02	65.26

Figure 3: Mean classification accuracy using various classifiers and feature combinations

	RF_{Corr}	SVM_{Corr}	$SVM_{PD+Corr}$	$NN3_{Corr}$
SVM_{Corr}	0.1502			
$SVM_{PD+Corr}$	0.1943	0.4213		
$NN3_{Corr}$	**0.0461**	**0.0480**	0.0631	
$NN3_{PD+Corr}$	**0.0406**	**0.0446**	**0.0414**	0.1894
	RF_{Corr}	SVM_{Corr}	$SVM_{PI+Corr}$	$NN3_{Corr}$
$SVM_{PI+Corr}$	0.1943	0.4213		
$NN3_{Corr}$	-	-	**0.0420**	
$NN3_{PI+Corr}$	**0.0493**	0.0763	0.0734	0.7432
	RF_{Corr}	SVM_{Corr}	$SVM_{PL+Corr}$	$NN3_{Corr}$
$SVM_{PL+Corr}$	0.1623	0.3513		
$NN3_{Corr}$	-	-	0.0581	
$NN3_{PL+Corr}$	**0.0467**	0.0683	0.0717	0.3524

Figure 4: Statistical significance of improvements in classification accuracy, comparing each row method against each column method, captured by p-values

References

[1] Christoph Hofer et al. "Deep learning with topological signatures". In: *Advances in Neural Information Processing Systems* (2017), pp. 1634–1644.

[2] Jan Reininghaus et al. "A Stable Multi-Scale Kernel for Topological Machine Learning". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 4741–4748.