Identifying Multimodal Imaging-Driven Subtypes in Mild Cognitive Impairment using Deep Multiview Learning

Yixue Feng[1], Mansu Kim[1], Kefei Liu[1], Andrew J. Saykin[2,3], Jason H. Moore[1], Qi Long[1], Li Shen[1]

[1]University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA, [2]Indiana University School of Medicine, Indianapolis, IN, USA, [3]Indiana Alzheimer Disease Center, Indianapolis, IN, USA

alzheimer's association AAIC 21 POSTER

Background

- **Multimodal neuroimaging** data can provide complementary information that a single modality cannot about neurodegenerative diseases such as Alzheimer's disease (AD).
- **Deep Generalized Canonical Correlation Analysis (DGCCA)** is able to learn a shared feature representation from different views of data by applying non-linear transformation using neural network.
- We utilize DGCCA to extract maximally correlated components from 3 modalities of neuroimaging data to identify potential **imaging-driven MCI subtypes**.

Materials & Methods

We study 308 Mild Cognitive Impairment (MCI) participants (195 early MCI and 113 late MCI) from the **Alzheimer's Disease Neuroimaging Initiative (ADNI)**, each with voxel level features from **FDG** PET, amyloid PET (**AV45**) and structural MRI processed using voxel-based morphometry (**VBM**).

- Six experimental settings are designed to compare **single modality** features with multiview methods (GCCA and DGCCA). **Agglomerative clustering** was used to generate 2 subtypes with features from each experiment.
- To investigate differences between the subtypes, Wilcoxon rank-sum tests are conducted on **5 cognitive assessments and 6 brain volume measures** at the baseline, from the ADNI QT-PAD dataset http://www.pi4cs.org/qt-pad-challenge.

	# Features	CH Score (↑)	Silhouette (↑)	AMI Score
Exp 1 - Concat	348	145.531	0.287	0.032
Exp 2 - VBM	116	182.839	0.308	0.008
Exp 3 - AV45	116	322.853	0.431	0.02
Exp 4 - FDG	116	144.537	0.251	0.028
Exp 5 - GCCA	94	2.908	0.038	-0.002
Exp 6 - DGCCA	20	133.704	0.303	0.039

Table 1. **Clustering evaluation for 6 experiments.** Higher CH score and Silhouette score indicates better defined clusters. AMI score computes the adjusted mutual information between cluster assignment and the original diagnosis groups (EMCI & LMCI). AMI close to 0 means two set of assignment are independent, while value close to 1 means they are identical.

Experiment Design

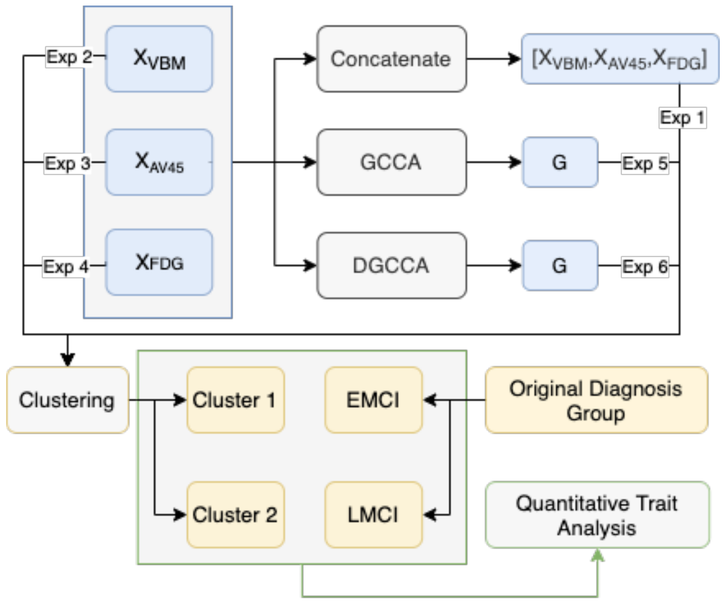

Figure 1. *Flowchart for 6 experiments.*

Results

Multiview Learning

Across the two multiview methods, the shared representation (denoted **G**) learnt from DGCCA explains **68.57% variance with 20 features**, while that from GCCA explains 68.66% variance with 94 features, with both learning from 116 ROI features from 3 imaging modalities.

Clustering

After applying agglomerative clustering to generate MCI subtypes for 6 experiments, we evaluate them using the Calinski-Harabasz (CH) score, Silhouette score and adjusted mutual information (AMI), see **Table 1**.

All experiments produce clusters with low AMI scores, meaning that the imaging-driven subtypes are different from the original EMCI/LMCI groups. AV45 generates the best defined clusters (high CH and Silhouette scores), where DGCCA generates clusters with quality comparable to single modality with a smaller set of features.

Results (Cont.)

Clusters from FDG and DGCCA features show differential measure in all cognitive assessment and brain volume measures where DGCCA learns from multimodal data.

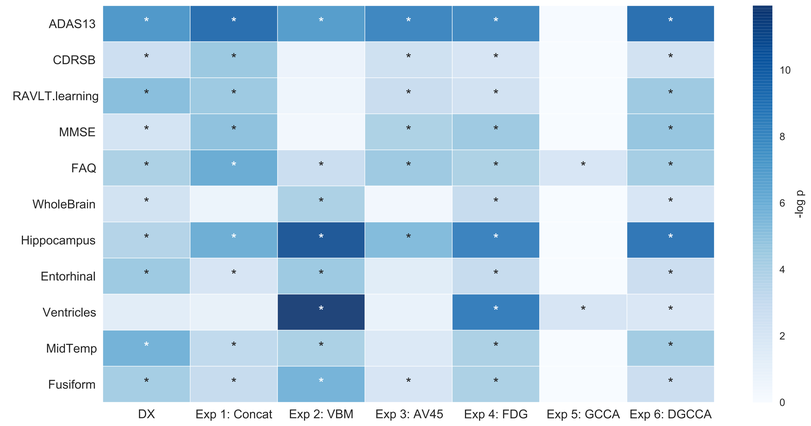

Figure 2. *QT-PAD Data Analysis*. Heatmap of -log(p) of the rank sum test. Significant entries are marked with "*",

Conclusions

DGCCA can learn maximally correlated features from multimodal neuroimaging data with reduced dimensionality, and explain more variance than its linear counterpart GCCA. These imaging-driven MCI subtypes are distinct from the currently diagnosis with differential measures in cognitive assessments and brain volumes, by incorporating complementary information from 3 imaging modalities. DGCCA shows to be an effective feature learning method, and this multiview learning framework can identify potentially novel MCI subtypes to facilitate early detection of AD.

Acknowledgements

This work was support in part by National Institute of Health R01 LM013463, U01 AG068057 and RF1 AG063481; and National Science Foundation IIS 1837964. Data used in preparation of this article were obtained from the Alzheimer's disease neuroimaging initiative (ADNI) database adni.loni.usc.edu.

Author contact *: Yixue Feng (wendyfyx@seas.upenn.edu), Li Shen (li.shen@pennmedicine.upenn.edu)*