Discovering Alzheimer's disease subtypes with imaging and genetic signatures via multi-view weakly-supervised deep clustering

Zhijian Yang, Junhao Wen, Ahmed Abdulkadir, Yuhan Cui, Guray Erus, Elizabeth Mamourian et al.

Corresponding Author: Christos Davatzikos

Artificial Intelligence in Biomedical Imaging Laboratory (AIBIL), Center for and Data Science for Integrated Diagnostics (AI2D), University of Pennsylvania, Philadelphia, USA

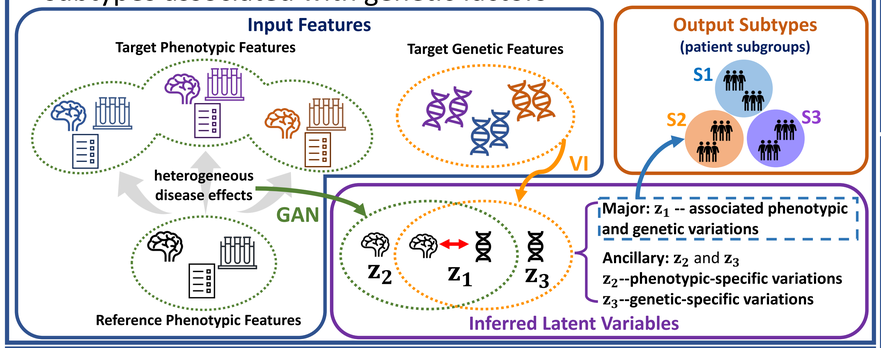
Graduate Group in Applied Mathematics and Computational Science, University of Pennsylvania, Philadelphia, PA, USA.

Background: Heterogeneity of AD and machine learning

- Alzheimer's disease and distinct co-pathology causes varying patterns of brain atrophy
- Increasing evidence suggests associations between genetic variants and heterogeneous imaging patterns in brain diseases
- **Machine learning methods** could dissect neuroanatomical heterogeneity and identify genetically-explained disease subtypes with distinct brain phenotypes.

Gene-SGAN Method

- **Gene-SGAN** aims to identify genetically driven disease subtypes from phenotypic and genetic features
- Learn one-to-many mappings through GAN → Capture the heterogeneous brain change patterns related to disease
- Guidance from genetic features through VI → disentangle linked and unlinked phenotypic and genetic features → phenotypic
- subtypes associated with genetic factors

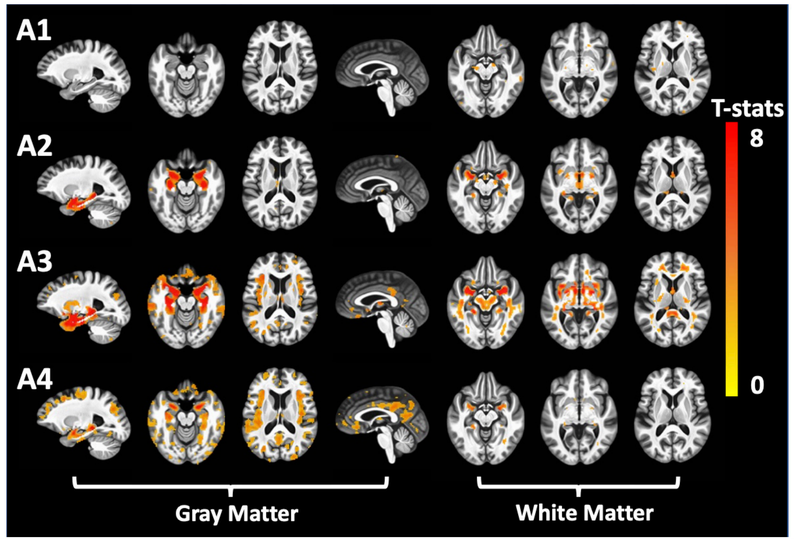


Application of Gene-SGAN to ADNI Dataset

- **Participants:** All CN (Reference) & MCI/AD (Target) participants with WGS data from ADNI were used for model training.
- **Phenotypic features**: imaging ROIs; **Genetic features**: AD-associated SNPs collected from GWAS-Catalog

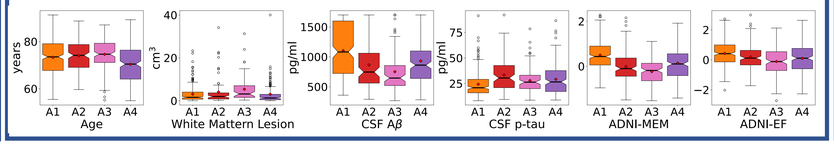
Four subtypes show distinct atrophy patterns

- **A1:** Minimal atrophy (no significant difference from control)
- **A2**: MTL-predominant atrophy
- **A3**: Severe widespread atrophy, including MTL
- **A4**: dominant cortical atrophy , sparing MTL



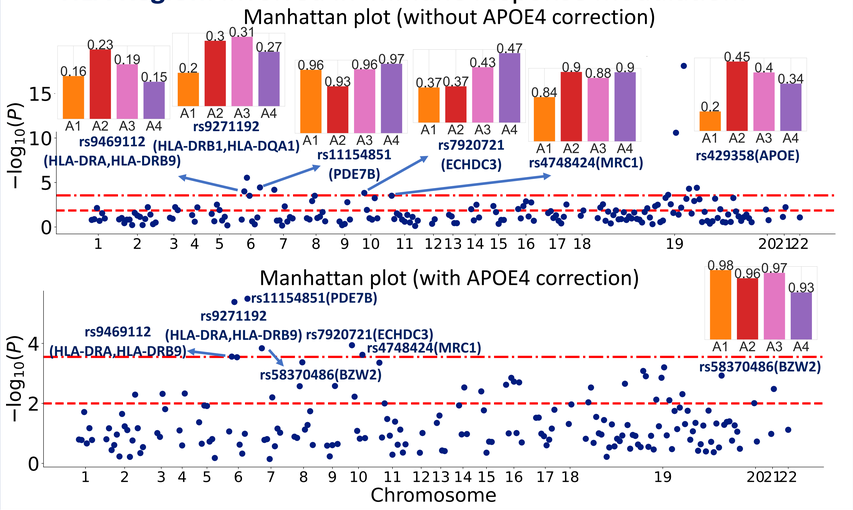
Demographic and Clinical Variables

- **A1**: best cognitive performance and relatively normal CSF-Aβ and CSF-pTau
- **A2:** the most abnormal CSF-pTau
- **A3**: the worst cognitive performance, highest WMH, most abnormal CSF-Aβ
- **A4**: significantly younger than other groups.



Genetic Associations

- The four subtypes reveal significant differences in seven known AD-related genetic variants.
  - **rs429358**: most significant genetic risk factor of AD
  - **rs7920721**: associated with AD among Non-APOE carriers
  - **HLA Region**: involved in immune response modulation.



Differences in CSF/Plasma Biomarkers

- These biomarkers are related to several biological mechanisms contributing to the heterogeneity of AD:
  - Hemostatic functions of the blood-brain barrier Microglial activation or proliferation; Aβ degradation and clearance.