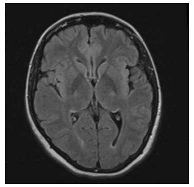
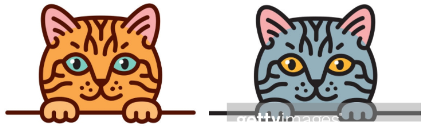
Heterogeneity Learning for Identifying Brain Disorder from Resting-state fMRI

Jiating Zhu (Sophie)

Problem



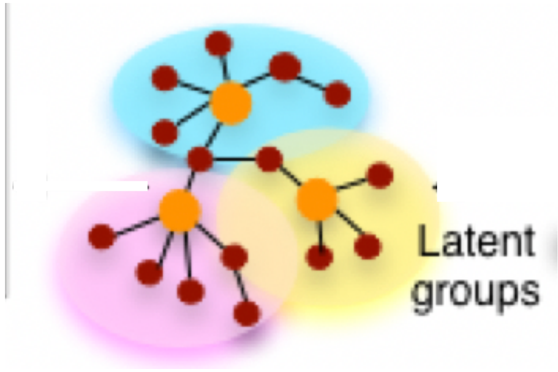
- **Diagnostic** methods based on **objectively** measurable **neuroimaging** data are desirable.
- Most of the methods rely on the assumption that both diagnostic category and typical control samples have **within-group homogeneity.**



Healthy samples not homogeneous

Identify the disorder by similarity measurement

Suitable?



within-group heterogentiy

(this finding is consistent with existing works)

Challenges

- **Within-group heterogeneity**

  Individual difference.

  Subjects in control group are not similar.
- **Overlap**

  Individual difference greater than group difference

  similarity measurement ✖
- **Brain data:** relatively small amount of data samples. Each sample has large amount of data.

  popular deep learning methods require large amount of data samples.

Main idea

Clustering

within-group heterogentiy → within-cluster homogenity

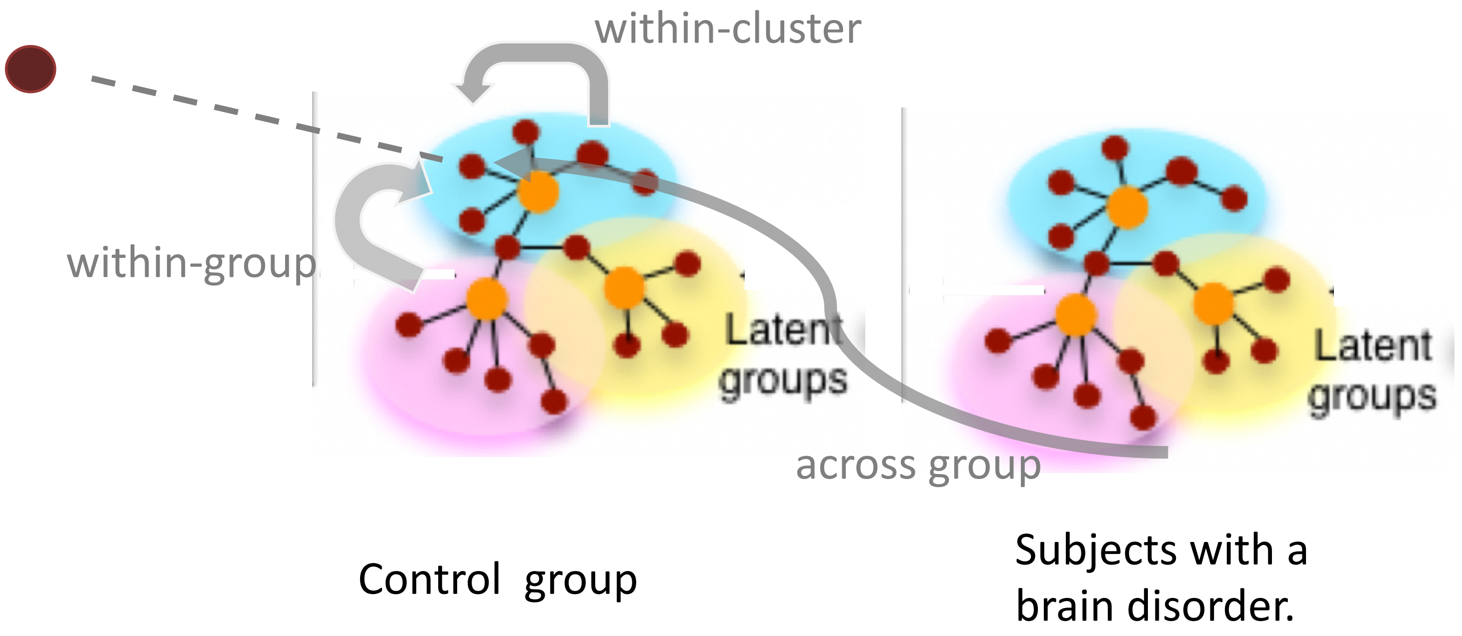
Individiual differences ↓ → Overlap ↓ → Disordered VS Control ↑

Objective

- Given an individual fMRI
- We can represent the subject by a score vector

Divergence follow different distributions when observed from different latent group

- Indentify the category by identifying which cluster it belongs to.



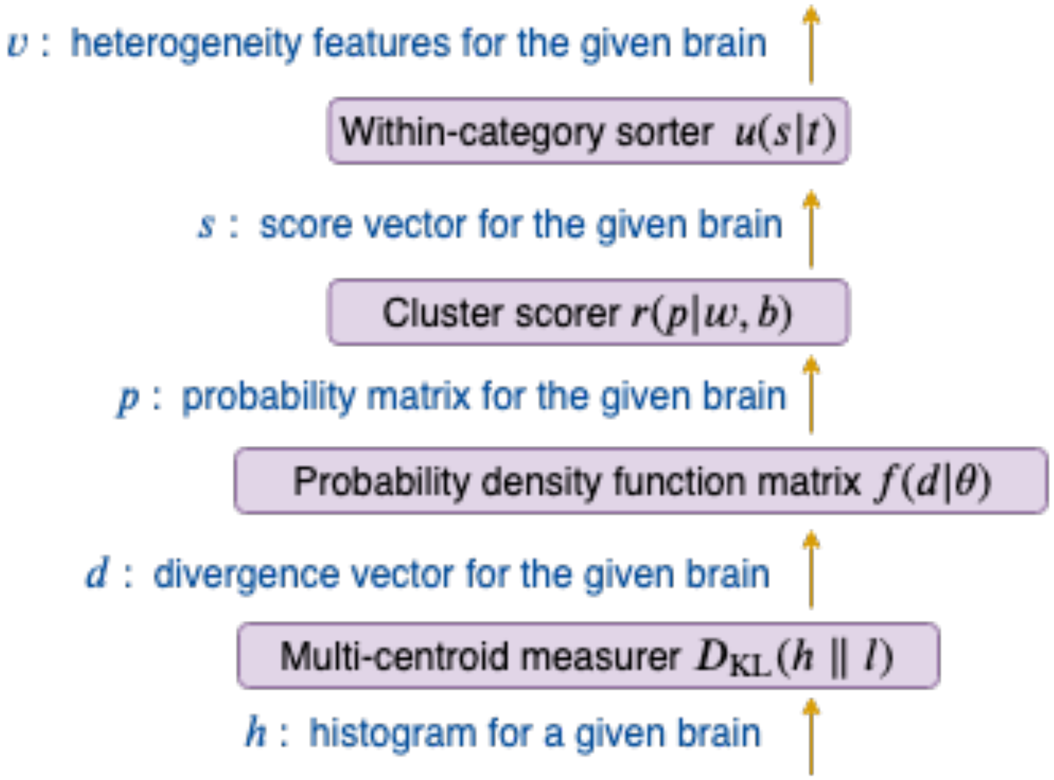
Model

Novelty of Solution:
We define the heterogeneity learning problem as a problem to construct a parameters **representational function** Fr(h|β) that is **suitable for learning**, so that each brain can represented by a vector with the **heterogeneity features**.

Method:
We propose a model for constructing the representational function in four steps.

Why:

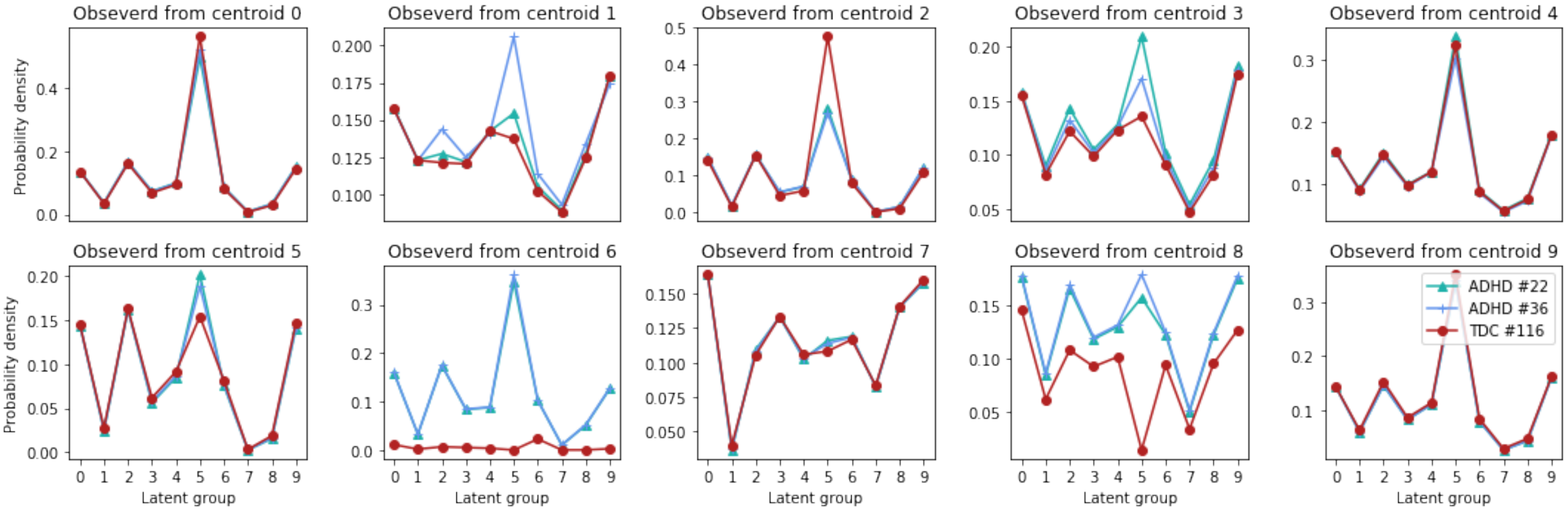
- heterogeneity learning model can capture the **heterogeneity distribution** with categories of both clinical and control samples.
- By relieving the heterogeneity, the **overlapping** crux can also be relieved, thus leading to a promising progress in brain disorder classification tasks.



Experiment
- Overlap relieved

1. Patterns are different

2. Subjects with ADHD has more similar patten than the one in control group



Patterns observed from differenet centroids

Progress & Plan

Without clustering - 25% overlap
With clustering - 5% overlap
Less 20% overlap

- Achieved above 90% distinguishable subjects on take one out experiment.
  - The individual to be tested participated in estimating probability matrix
- Other Evaluation - Prediction task
  - Designed algorithm (CNN approach) for cluster scorer.
  - Current difficulty: overfitting in test case

Plan

- Try different clustering approach
- Try Extended dataset/ different dataset
  - **Current** studies (such as deep learning methods on region-based functional networks) -> roughly around 70% accuracy. We would like to **rise the diagnostic accuracy** to 80%.
- Evaluation method other than prediction

Thank You !