Heterogeneous Data Mining for Neurological Disorder Identification

Presenter: Bokai Cao
Outline

• Introduction

• Tensor Imaging Analysis

• Brain Network Analysis

• Multi-view Feature Analysis

• Future Work


Outline

• Introduction
• Tensor Imaging Analysis
• Brain Network Analysis
• Multi-view Feature Analysis
• Future Work
Introduction

• Brain Disorders
  • HIV, AD, ADHD

• Neuroimaging Techniques
  • fMRI, DTI, PET, EEG

• Data Representations
  • Tensor, graph, vector

What is the data about?
Where does the data come from?
What does the data look like?
Introduction

• Brain Disorders
  • HIV infection on brain
  • Bipolar disorder
  • Alzheimer's disease (AD)
  • Attention deficit hyperactivity disorder (ADHD)
• Schizophrenia
Introduction

• Neuroimaging Techniques

• Functional magnetic resonance imaging (fMRI)
• Diffusion tensor imaging (DTI)
• Positron emission tomography (PET)
• Electroencephalogram (EEG)
Introduction

• Data Representations
  
  • Tensor: raw images
  
  • Graph: brain networks
  
  • Vector: multi-view features
Outline

• Introduction

• **Tensor Imaging Analysis**

• Brain Network Analysis

• Multi-view Feature Analysis

• Future Work
Tensor Imaging Analysis

• Tensor Data in Neuroimaging

A voxel is the smallest three-dimensional point volume referenced in a neuroimaging of the brain. Typically, 256*256*256.

vector $\mathbf{x}$

$1$st-order tensor

matrix $\mathbf{X}$

$2$nd-order tensor

tensor $\mathcal{X}$

$3$rd-order tensor
Tensor Imaging Analysis

• Brain Network Discovery
  1. Node discovery
  2. Edge discovery
  3. Network verification

Tensor Imaging Analysis

• Brain Network Discovery

• Tensor factorization

Tensor Imaging Analysis

• Brain Network Discovery

• Contributions:
  (1) simultaneously discover nodes and edges
  (2) facilitate better understanding about interaction mechanism between brain regions

• Drawback: fail to leverage relationships between neuroimages and their associated labels

Tensor Imaging Analysis

- Supervised Tensor Learning
- Classification (review)

features: \(<\text{age, weight}>\)

label: diseased (positive) or healthy (negative)

Tensor Imaging Analysis

- Supervised Tensor Learning
- Tensor classification
  1. High dimensionality

\[ m=256 \times 256 \times 256 = 16,777,216 \]

Tensor Imaging Analysis

- Supervised Tensor Learning
  - Tensor classification
    1. High dimensionality
    2. Structural complexity

Tensor Imaging Analysis

• Supervised Tensor Learning

• Tensor classification

  1. High dimensionality
  2. Structural complexity
  3. Nonlinear separability

Tensor Imaging Analysis

- Supervised Tensor Learning
- Support tensor machine

**primary form**

$$\min_{W,b,\xi} \frac{1}{2} \|W\|_F^2 + C \sum_{i=1}^{n} \xi_i$$

s.t. $y_i(\langle W, X_i \rangle + b) \geq 1 - \xi_i$

$\xi_i \geq 0, \forall i = 1, \cdots, n.$

**dual form**

$$\max_{\alpha_1, \cdots, \alpha_n} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \langle \phi(X_i), \phi(X_j) \rangle$$

s.t. $\sum_{i=1}^{n} \alpha_i y_i = 0$

$0 \leq \alpha_i \leq C, \forall i = 1, \cdots, n.$

Tensor Imaging Analysis

• Tensor Data in Neuroimaging

• Brain Network Analysis
  • Node discovery, edge discovery, network verification, tensor factorization

• Supervised Tensor Learning
  • Tensor classification, support tensor machine
Outline

• Introduction

• Tensor Imaging Analysis

• **Brain Network Analysis**

• Multi-view Feature Analysis

• Future Work
Brain Network Analysis

• Brain Network Data

  • Nodes: brain regions
    e.g., insula, hippocampus, thalamus

  • fMRI links: correlations between the functional activities of brain regions
  • DTI links: number of neural fibers connecting different brain regions
Brain Network Analysis

- Kernel Learning on Graphs

Brain Network Analysis

- Kernel Learning on Graphs
  1. Extracting clustering coefficients
  2. Selecting the most discriminative features
  3. Constructing kernel matrices
  4. Training classifiers
Brain Network Analysis

- Kernel Learning on Graphs

Brain Network Analysis

• Kernel Learning on Graphs

• Contributions:
  (1) integrate complementary DTI and fMRI
  (2) achieve 96% classification accuracy
  (3) select the most discriminative brain regions

• Drawbacks:
  (1) connectivity structures are blinded
  (2) interpretability is limited

Brain Network Analysis

- Subgraph Pattern Mining

Brain Network Analysis

• Subgraph Pattern Mining
  - Binary links: \{0,1\}
    e.g. gSpan [Yan and Han 2002]
  - fMRI links: [-1,1]
  - DTI links: \(N(0,1,2,\ldots,1e6,\ldots)\)

Brain Network Analysis

• Subgraph Pattern Mining

• Mining uncertain graphs

Brain Network Analysis

• Brain Network Data

• Kernel Learning on Graphs
  • Extracting clustering coefficients

• Subgraph Pattern Mining
  • Mining uncertain graphs
Outline

- Introduction
- Tensor Imaging Analysis
- Brain Network Analysis
- Multi-view Feature Analysis
- Future Work
Multi-view Feature Analysis

- Multi-view Data in Medical Studies

Limited subjects available yet introducing a large number of measurements

Multi-view learning + feature selection
Multi-view Feature Analysis

• Modeling View Correlations
  • Tensor and AAL features from MRI images
  • Demographic information: age and gender
  • Genetic information

Multi-view Feature Analysis

- Modeling View Correlations

Multi-view Feature Analysis

• Modeling View Correlations

Multi-view Feature Analysis

• Modeling View Correlations

• Contributions:
  (1) integrate different types of features
  (2) identify biomarkers (brain regions) from multiple data sources

• Drawbacks: fail to explicitly consider correlations between features

Multi-view Feature Analysis

• Modeling Feature Correlations

• Vector-based method

Multi-view Feature Analysis

- Modeling Feature Correlations
- Tensor-based method

Multi-view Feature Analysis

- Modeling Feature Correlations
  - Dual method

Multi-view Feature Analysis

• Modeling Feature Correlations

• Wrapper-based feature selection

Multi-view Feature Analysis

- Modeling Feature Correlations

Multi-view Feature Analysis

- Multi-view Data in Medical Studies
- Modeling View Correlations
  - Multi-kernel method
- Modeling Feature Correlations
  - Vector-based method, tensor-based method, dual method
Outline

- Introduction
- Tensor Imaging Analysis
- Brain Network Analysis
- Multi-view Feature Analysis
- **Future Work**
Future Work

• Learning on Different Data Representations

neuroimaging experiments → tensor data

brain network data

other sources, e.g., clinical and serologic experiments

multi-view data

heterogeneous data fusion

single modality
Future Work

• Integrating Multiple Imaging Modalities

  • fMRI: functional connections
    multiple thresholds [Jie. et al. 2014], multi-spectrum [Wee et al. 2012]

  • DTI: structural connections
    multiple physiological parameters: fiber count, fractional anisotropy (FA), mean diffusivity (MD), and principal diffusivities [Wee et al. 2011]
Future Work

- Mining Bioinformatics Information Networks
Future Work

• BRAIN Initiative  http://www.whitehouse.gov/BRAIN

President Obama is making over $300 million investments in the “BRAIN” Initiative to revolutionize our understanding of the human mind and uncover new ways to treat, prevent, and cure brain disorders like Alzheimer’s, schizophrenia, autism, epilepsy, and traumatic brain injury.
Thank you.!