Encoding Brain Networks Through Geodesic Clustering of Functional Connectivity for Multiple Sclerosis Classification

Muhammad Abubakar Yamin, Paola Valsasina, Michael Dayan, Sebastiano Vascon, Jacopo Tessadori, Massimo Filippi, Vittorio Murino, Maria A. Rocca and Diego Sona.


Department of Environmental Sciences, Informatics and Statistics, Ca’ Foscari University, Venice, Italy

Neuroinformatics Laboratory, Fondazione Bruno Kessler, Trento, Italy

Department of Computer Science, Universit`a di Verona, Verona, Italy

Vita-Salute San Raffaele University, Milan, Italy
Outline

• Goal and Motivation

• Method
  • Datasets
  • The Pipeline
  • Background: Manifold Representation Of SPD Matrices
  • Geodesic Clustering
  • Feature Extraction And Classification
  • Experiments

• Results on Datasets

• Conclusions
Goal and Motivation

Goal
• Encoding of brain functional connectivity (FC) data to discriminate between healthy controls (HC) and multiple-sclerosis (MS) patients

Motivation
• The FC analysis is based on graphs comparison, which is usually done by Euclidean distance (ED)
• Use of ED is sub-optimal because it does not capture the real geometry of manifold of symmetric positive definite (SPD) matrices
• A Better choice is to exploit the geometrical nature of SPD matrices on Riemannian manifold.
Method (Datasets)

**Dataset:** Private dataset collected at the Neuroimaging Research Unit (Hospital San Raffaele, Milan, Italy)
- Resting state-functional magnetic resonance imaging (rs-fMRI)
- 33 HC and 72 multiple-sclerosis (MS) patients (age matched)
- 37 relapsing-remitting (RRMS) and 35 Progressive (PMS)
- FC matrix size is 90x90, based on AAL atlas computed using covariance.
The Pipeline

Functional Connectivity → Manifold Representation → Geodesic Metric on Riemannian Manifold

Support Vector Machines

Distance Feature Vector
**Method** (Background: Manifold Representation Of SPD Matrices)

Covariance based connectivity matrices \( \{ \Sigma_i \} \)
- Set of symmetric & positive semi-definite matrices.
- Represent the functional connectomes showing both positively and negatively synchronous connections.
- Can be easily made SPD with a small regularization \( \hat{\Sigma} = \Sigma + \lambda I \),

\[ x^T \Sigma x \geq 0 \quad \forall x \neq 0 \in \mathbb{R} \]

Symmetric Positive Definite (SPD) Riemannian Manifold

\[ \mathcal{M} \]

\( L_1 \rightarrow L_2 \rightarrow L_3 \rightarrow L_n \)
Method (Background: Geodesic Analysis on Manifold)

- Euclidean distance is sub-optimal because it does not capture the real geometry of manifold of SPD matrices.

- Use of Geodesic distance is proposed which better define the distance along with manifold.

- Log-E distance eq.(1) \[1\] and geodesic mean in the closed form eq.(2) \[2\].

\[Log – Euclidean Distance: \quad d_L(\Sigma_i, \Sigma_j) = \| \log \Sigma_i - \log \Sigma_j \| \quad (1)\]

\[Geodesic Mean: \quad \Sigma_L = \exp \left\{ \arg \inf \sum_{i=1}^{n} \| \log \Sigma_i - \log \Sigma \|^2 \right\} = \exp \left\{ \frac{1}{n} \sum_{i=1}^{n} \log \Sigma_i \right\} \quad (2)\]
**Method** (Geodesic K-means Clustering of SPD Matrices)

- **Aim:**
  - Cluster FC matrices into homogeneous groups of subjects.

- **Underlying assumption:**
  - Alterations in brain connections grasped by the clusters.

- K-means clustering was implemented using geodesic distance and geodesic mean

- **Drawbacks in traditional K-mean?**
  - Need to pick ‘K’,
  - Sensitive to initialization
  - Sensitive to outliers
**Method** (Geodesic Dominant Set Clustering of SPD Matrices)

**AIM**
Cluster FC matrices into homogeneous groups of subjects.

**Dominant set clustering**:
- Graph theoretic concept [3]
- Computes well separated and compact subset of nodes (dominant sets (DS))
- Extraction of DS is sequential (one by one)
- Can be solved using game dynamic e.g. replicator dynamics [4] \( \text{maximize} \ x^T A x \)
- Clusters are more similar inside and less similar to outside.
- Data in form of similarity matrix
  \[ S(i,j) = 1 - \frac{d_L(i,j)}{\max(d_L)} \]
- No prior information on number of clusters (since we extract them sequentially).
- Leaves clutter elements unassigned
Method (Feature Extraction and Classification)

- Due to high dimensionality of FC matrices (90x90), feature encoding is needed.
- Using cluster centroid as dictionary
- Building vector representation by computing geodesic distance from each cluster centroid.
- Using this vector as feature set to classifier (Support Vector Machine, SVM).
Experiments

- For k-means Number of clusters were chosen to be variable between K=2−15.
- To avoid double dipping
  - 5-fold cross validation
  - Training folds for clustering & extracting training features
  - Test fold for computing test feature vector.
- Repeating 5-fold cross validation 100 time and taking mean of accuracies.
- Permutation test on labels (To check the significance of results).
- For Comparison, same analysis is performed using ED
Results

- Average of 100 iterations of 5-folds cross validations of experiments:
- DS always extract 6 or 7 clusters, so for comparison we implement K-mean for K= 2–15
- HC vs MS: Accuracy 73.94 %
Results

- Average of 100 iterations of 5-folds cross validations of experiments:
- DS always extract 6 or 7 clusters, so for comparison we implement K-mean for K= 2–15
- HC vs MS: Accuracy 73.94 %
- HC vs RRMS: Accuracy 72.51%

TABLE II
AVERAGE CONFUSION MATRIX OF CLASSIFICATION RESULTS FOR THE PROPOSED MEAN GEODESIC DS CLUSTERING APPROACH AND BEST OF GEODESIC k-MEANS CLUSTERING FOR HC VS. RRMS.

<table>
<thead>
<tr>
<th>Geodesic Dominant-Set</th>
<th>Geodesic k-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Class</td>
<td>Predicted Class</td>
</tr>
<tr>
<td></td>
<td>HC</td>
</tr>
<tr>
<td></td>
<td>HC</td>
</tr>
<tr>
<td>Actual Class</td>
<td>21.68</td>
</tr>
<tr>
<td>Class</td>
<td>7.92</td>
</tr>
</tbody>
</table>

HC vs RRMS Cross Validation Results

![Graph showing cross-validation results for HC vs RRMS with accuracy values and box plots for different cluster numbers.]}
Results

- Average of 100 iterations of 5-folds cross validations of experiments:
- DS always extract 6 or 7 clusters, so for comparison we implement K-mean for K= 2–15
- HC vs MS: Accuracy 73.94 %
- HC vs RRMS: Accuracy 72.51%
- HC vs PMS: Accuracy 84.06 %
- Geodesic Clustering gives superior results and also Geodesic Dominant–Set clustering is always better in performance as compared to K-Mean clustering.
Results

- Average of 100 iterations of 5-folds cross validations of experiments:
- DS always extract 6 or 7 clusters, so for comparison we implement K–mean for K= 2-15
- HC vs MS: Accuracy 73.94 %
- HC vs RRMS: Accuracy 72.51%
- HC vs PMS: Accuracy 84.06 %
- Geodesic Clustering gives superior results and also Geodesic Dominant–Set clustering is always better in performance as compared to K–Mean clustering.

Significance of permutation test
- HC vs MS, P_value <0.0005
- HC vs RRMS, P_value <0.05
- HC vs PMS, P_value <0.0005
Conclusion

- **Neuroscientific**
  - Alteration in brain is helpful in discriminating between HC and patients affected with different phenotype of MS.

- **Computational**
  - Proper data representation allow an effective exploitation on the manifold space.
  - Geodesic method-based clustering gives superior results.
  - Specific encoding of FC matrices leads to good performance in discriminating task.


Thank you!