Global PCA of Local Moments

With Applications to MRI Segmentation

Jacob M. Maronge[1, 2], John Muschelli[3], Ciprian M. Crainiceanu[3]

[1]University of Wisconsin - Madison, Department of Statistics

[2]University of Wisconsin - Madison, Department of Biostatistics and Medical Informatics

[3]Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Contact Information:
Jacob M. Maronge
Department of Statistics
University of Wisconsin - Madison
1300 University Ave, Madison, WI, 53706

Email: jmmaronge@gmail.com
Website: https://jmmaronge.github.io

Abstract

We are interested in describing the information contained in local neighborhoods, and higher moments of local neighborhoods, of complex multimodal imaging techniques at the population level. This is problematic because of the size of medical imaging data. We propose a simple, computationally-efficient approach for representing the variation in multimodal images using the spatial information contained in all local neighborhoods across multiple subjects. This method achieves 3 goals: 1) decomposes the observed variability images at the population level; 2) describes and quantifies the main directions of variation; 3) uses these directions of variation to improve segmentation and studies of association with health outcomes. To achieve this, we efficiently decompose the observed variation in local neighborhood moments. In order to assess the quality of this method we show results using the 2015 Ischemic Stroke Lesion Segmentation (ISLES) Challenge.

Introduction

Every image can be vectorized. However, its meaning, interpretation, and complexity is encapsulated in the collection of all neighborhoods of all locations. More precisely, every image with V voxels can be represented as a $V \times V$ matrix, where every row represents a location in the image and every column reprents a particular position in the neighborhood of that location; e.g., the first column could be the neighbor just above the location, the second column could be the neighbor to the left. Such matrices are very large and store information inefficiently, but they provide a useful theoretical framework for representation of imaging information. Here we propose to exploit this theoretical framework to introduce simple methods to quantify the variation in multimodal images based on the shared information across local spatial neighborhoods and subjects.

Here we consider the case when multimodal imaging is available and images are spatially co-registered within-subject, but not necessarily across subjects. Such data are routinely available in MRI studies, where FLAIR, T1-weighted and T2-weighted images are collected for multiple subjects. Our goals are:

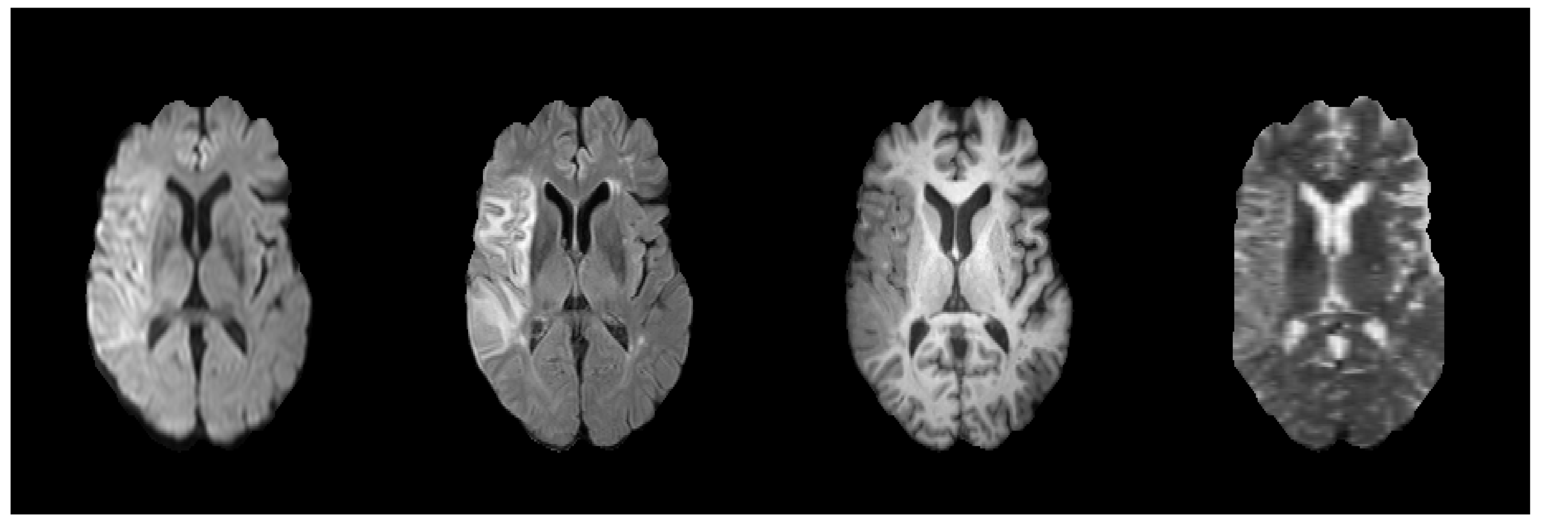

Figure 1: From left to right: DWI, FLAIR, T1w, and T2w images for Subject 05, 2015 ISLES Challenge

1. Decompose observed variability images at the population level
2. Describe and quantify the main directions of variation
3. Use these directions of variation to improve segmentation and studies of association with health outcomes.

Materials and Methods

To achieve these objectives we propose to decompose the local variability of various moments of the image intensities: the image intensity, image intensity square, and so on. To build intuition, consider the case when we only have one image per subject, which contains V_i voxels for subject i, though ideas will extend to the case when there are more images. The main idea is to represent the image as a $V_i \times K$ matrix, where every row corresponds to a voxel and contains the image intensity of the K neighbors of that particular voxel given a specified ordering. For presentation and computational purposes we work with neighborhoods of size $K = 27$ (adjacent neighbors only in 3D), but methods can easily be applied to any neighborhood size. The second idea is to consider matrices obtained by taking the higher order moments of the original image and unfold them using the same procedure as the one for image intensities. Thus, if we consider all moments up to the fourth moment then the matrix representation of the image will be $V_i \times (4K)$ dimensional.

Consider the following mock example:

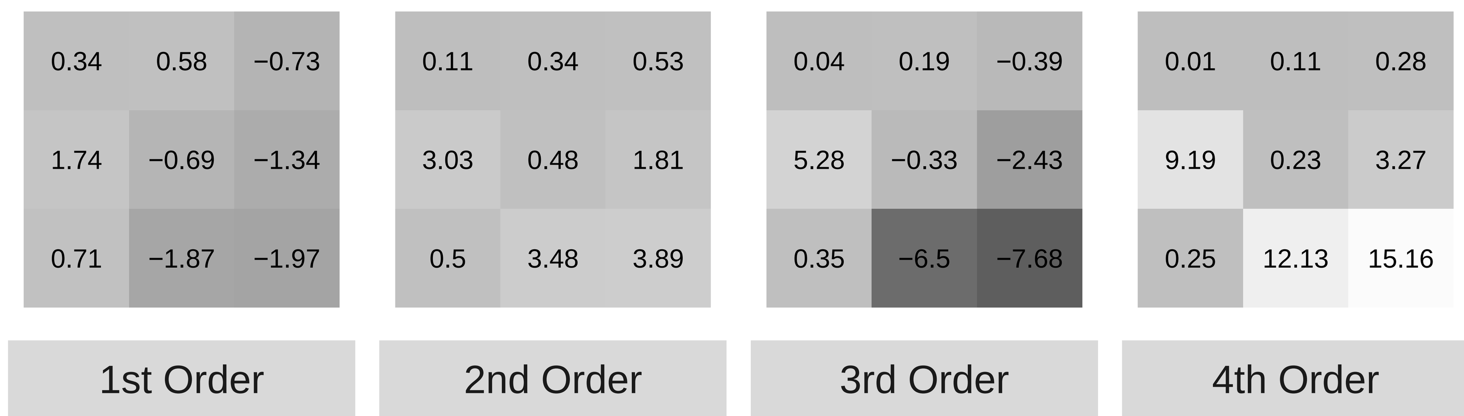

Figure 2: 2-D mock example of moments of local neighbors

$$
\begin{aligned}
X_{ij} = &(0.34, 0.58, -0.73, 1.74, -0.69, -1.34, 0.71, -1.87, -1.97,\\
&0.11, 0.34, 0.53, 3.03, 0.48, 1.81, 0.50, 3.48, 3.89,\\
&0.04, 0.19, -0.39, 5.28, -0.33, -2.43, 0.35, -6.50, -7.68,\\
&0.01, 0.11, 0.28, 9.19, 0.23, 3.27, 0.25, 12.13, 15.16).
\end{aligned}
$$

This gives the jth row for subject i. If we are interested in considering additional imaging modalities, we simply add those as additional columns in the matrix. We preform this operation for each voxel for each subject and stack these vectors into a (potentially large) matrix $\mathbf{X}$.

Once the matrix $\mathbf{X}$ is built, we center and scale each column to ensure that the observed information is not dominated by the scale of higher moments and conduct PCA. This is equivalent to conducting PCA on the correlation matrix between the columns of $\mathbf{X}$. The size of the matrix $\mathbf{X}$ is large and there is a need to use this matrix without loading it in the computer memory.

Finally, we use the scores from the principal component analysis to fit a principal component regression model where an expert manual segmentation is the outcome. Data are separated into a training and testing set to assess performance.

Computational Challenges

Since the size of the matrix $\mathbf{X}$ is large, we need to decompose this matrix without loading it into memory. To overcome this challenge, we iteratively read in each subject and calculate sufficient statistics for the principal component analysis of the correlation matrix between the columns of $\mathbf{X}$. This is implemented in the R package MEDALS (Memory Efficient Decomposition for Analysis of Local neighborhood moments for Segmentation) and is available at https://github.com/JMMaronge/medals.

Results

Below we show partial ROC (pROC) curves for each subject, as well as overall training and overall testing results. We also show an example of how we performed on a particular subject.

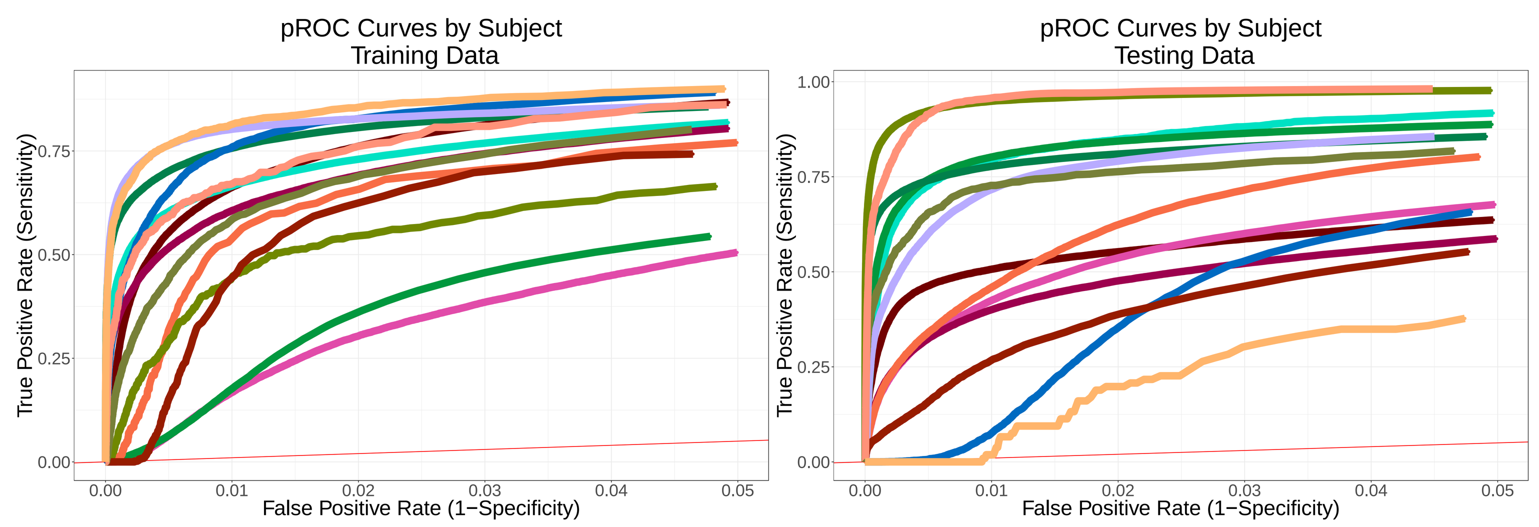

Figure 3: Subject level pROC curves for training set and testing set

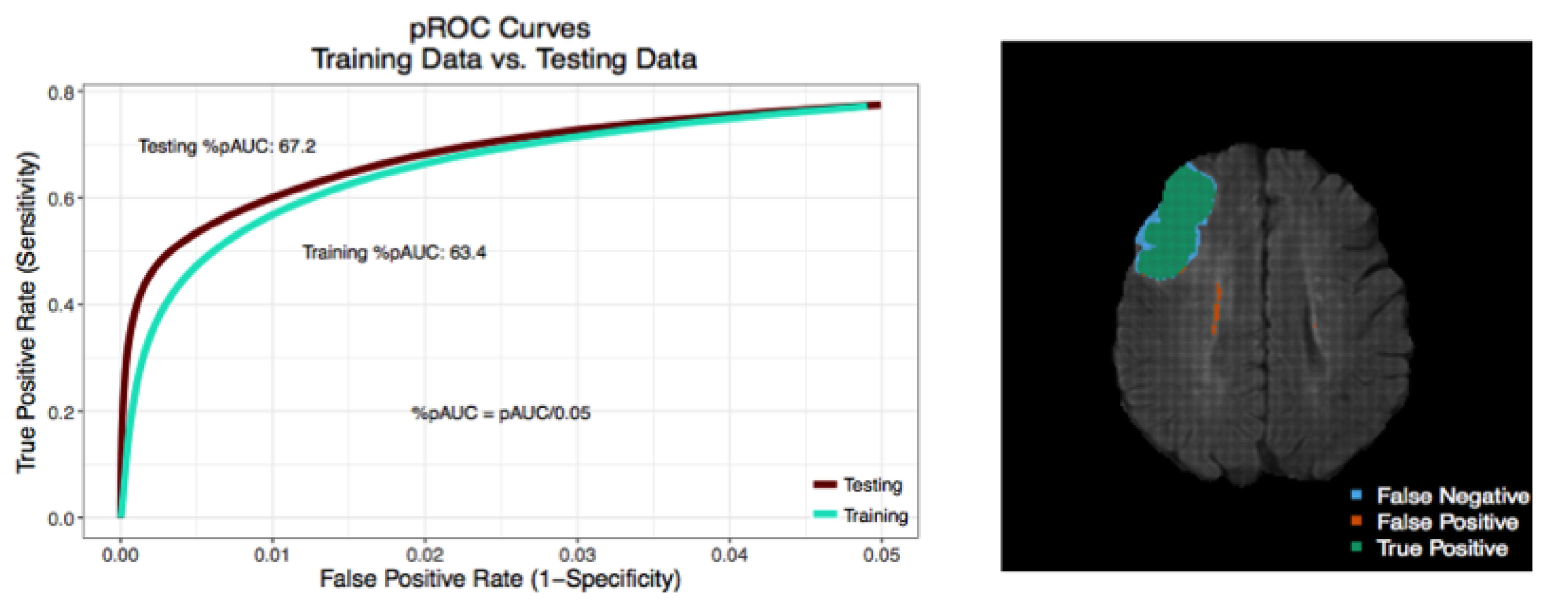

Figure 4: (left) Overall training results and overall testing results (right) Example of performance on one subject

Conclusions

In conclusion, carefully leveraging the information contained within neighborhoods appears to lead to interesting results in data applications. Future work should address how these methods compare to strictly vectorized approaches.