DEFORMATION ANALYSIS OF DIFFUSION TENSOR DATA USING RANDOM FORESTS

Neda Sadeghi[1], M. Okan Irfanoglu[1,2], Amritha Nayak[1,2], Cibu Thomas[1,3], Carlo Pierpaoli[1]

[1]Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, [2]Henry M. Jackson Foundation, [3]Center for Neuroscience and Regenerative Medicine

The standard practice for analyzing neuroimaging data (e.g. diffusion tensor imaging) is to register the individual images to a common coordinate space and perform a voxel-wise comparison of diffusion derived metrics between groups. Neuroimaging data is inherently high dimensional and massive multiple comparisons correction is necessary in the voxel-wise analysis; the choice of the method used for correction can produce differing results. Another important method, that is less common in diffusion tensor imaging, is to analyze the deformation fields that map individual images from their native space to a common template. In this work, we use the log of the Jacobian of the deformation fields in regions of interest as input to random forests. Random forests are a set of data driven classification algorithms that highlight important features. A random forest is iteratively fit to the data, at each iteration the least important variables are eliminated until the current out of bag (OOB) error becomes larger than the previous OOB rate + OOB standard error. This results in regions that achieve high classification rate and are important in characterizing the disease. We present the methodology applied to a group of patients that have brain atrophy.