A Novel Framework for Identifying DTI-Based Brain Patterns of Schizophrenia

P. Wang[1], R. Gur[1], and R. Verma[1]

[1]Univ. of Pennsylvnia, Philadelphia, PA, United States

INTRODUCTION. Diffusion tensor imaging (DTI) provides rich information about brain tissue structure [1], especially white matter and has therefore gained attention in studying pathology of several diseases, such as schizophrenia, in a group-based analysis. In such research, it is important to accurately identify abnormal brain tissue that can classify patients and controls into two groups, for diagnosis, treatment and prognosis. However, conventional statistical methods are difficult to apply to DTI data due to its high dimensionality and nonlinearity. Most of the existing methods resort to a statistical t-test based on FA or kernel PCA features from tensors. However, the t-test usually assumes a single Gaussian distribution for each group, and a multivariate hypothesis testing is usually constrained by limited number of available samples and high-dimensional data. To address these issues, we present a novel group analysis framework in this paper. Rooted in the pattern classification theory, our method directly estimates the overlap between different groups using Bayes error rate. Unlike t-test, our method does not assume global Gaussian distribution for each group, but only considers data distribution to be locally smooth, thus facilitating separation of different groups when the data could have highly nonlinear structure, which is the case in DTI data. This framework also has the capability of combining multiple measurements that are extracted from tensor data, thus being able to handle multivariate testing. In this paper, the method is first used to identify abnormal brain regions that distinguish the patient group from control group, and is further validated by classifying patients from controls based on these regions. Experiments on 36 controls and 34 patients with schizophrenia show encouraging results.

METHOD. In our method, we normalize the DTI data of the subjects to a template space. Scalar features, such as Fractional Anisotropy (FA), Coefficient of linearity (Cl), Coefficient of planarity (Cp), diffusivity and/or their combinations, are then computed from tensor data. Due to limited space, we mainly present the results based on FA. The method is the same when using other features. Based on extracted features, a novel method is presented to estimate the overlap between different groups to identify abnormal regions in brain tissue. Our method is summarized as follows. First, pair-wised distances, e.g. d^k_{ij} between i- and j-th samples at k-th voxel, are computed using scalar features extracted at each voxel. In this method, L-2 normal distances are applied, whereas other appropriate distance measures are also applicable. Second, we apply a framework to estimate the Bayes error based on distance measures. This estimation framework is initially presented in [2]. This method is based on density estimation using Parzen window, to estimates the misclassification rate based on the Bayes rule. The error rate indicates how accurately group identities of samples can be predicted by local neighbors in a feature space [3]. The error rate at 50% means that two groups are not distinguishable, and lower error rate refers to better capability of differentiating groups. A significant merit of this method is that it does not assume any global Gaussian assumptions for each group, but only imposes local smooth constraints by using Parzen window. Also, since the estimation method only takes the distance measures as input, it can combine multiple features if appropriate distance metrics are defined. In this research, the method is applied to the group analysis of controls and patients with schizophrenia, to identify abnormal regions, and to further validate the regions by pattern classification experiments. To identify the abnormal regions, voxel-wise Bayes error is first estimated. At each voxel, the scalar features averaged over a 3×3×3 neighborhood are used to obtain a distance matrix, based on which the Bayes error is estimated. To obtain regions where abnormality exists, only those voxels where the error rates are below a threshold are preserved. Then connected component analysis is applied to remove isolated voxels that may be caused by random noise and registration error. The map of estimated error rates is shown in Fig.1 (b), and identified regions are shown in Fig 1 (c). Furthermore, to validate those identified regions, the features in the selected regions are applied in pattern classification to distinguish patients from controls. The accuracy and generalization of this method is validated through a 10-fold cross-validation, in which error estimation, region selection, and classifier training are only performed exclusively from the training sets. The averaged accuracy on testing sets indicates the performance of our method.

RESULTS. In this study, there are 34 patients diagnosed with schizophrenia compared with 36 healthy controls. The data is acquired at dimensions of 128×128×40 and a voxel size of 1.72×1.72×3.0mm. For each subject, DTI images were registered to a common space for analysis. Fig. 1 (a) show FA slices for a control and Fig. 1 (b) shows distributions of estimated Bayes error. The red color shows regions with lower overlap, i.e., higher group separation, and blue color represents regions with lower group separation. Fig. 1 (c) shows the extracted abnormal regions, obtained from estimated error below a threshold (set as 0.48 here). The affected regions are in the internal capsule, corpus callossum, corona radiata and cingulum bundle. Although these regions have been shown to be affected in schizophrenia, we further validated the importance of these regions for group separation by using the FA values in the abnormal regions for classification. We test two types of classifiers: SVM and KNN, in our method. The classification performance of using FA feature and KNN classifier has achieved a 22.86% error rate on all the 70 subjects. The ROC curve of the cross validation is depicted in Fig.2.

DISCUSSION. We have presented a novel method for identifying abnormal regions in brain tissue using DTI data, and for selecting region features for classification. We have shown that the method is able to detect nonlinear group differences without assuming any parametric distributions of data. The error rate maps provide rich information on tissue changes, and provide normalized and predictive value that could potentially aid in disease diagnosis. This framework is general and applicable to other high dimensional, multi-modality datasets.

(a)

(b)

(c)

Fig.1. Identification of abnormal regions in DTI. (a) FA slices of a control. (b) Maps of estimated Bayes error across all the subjects from both patient and control group. (c) Identified abnormal regions (threshold set as 0.48)

References.
1. D. LeBihan, J. F. Mangin, C. Poupon, C. A. Clark, S. Papata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: concepts and applications," *Journal of Magnetic Resonance Imaging*, vol. 13, pp. 534-546, 2001.
2. Peng Wang, Christian Kohler, and Ragini Verma, Estimating Cluster Overlap on Manifolds and its Application to Neuropsychiatric Disorders, *IEEE Workshop on Component Analysis Methods*, 2007
3. Duda RO, Hart PE, and Stork DG, *Pattern Classification*: John Wiley and Sons, Inc., 2001
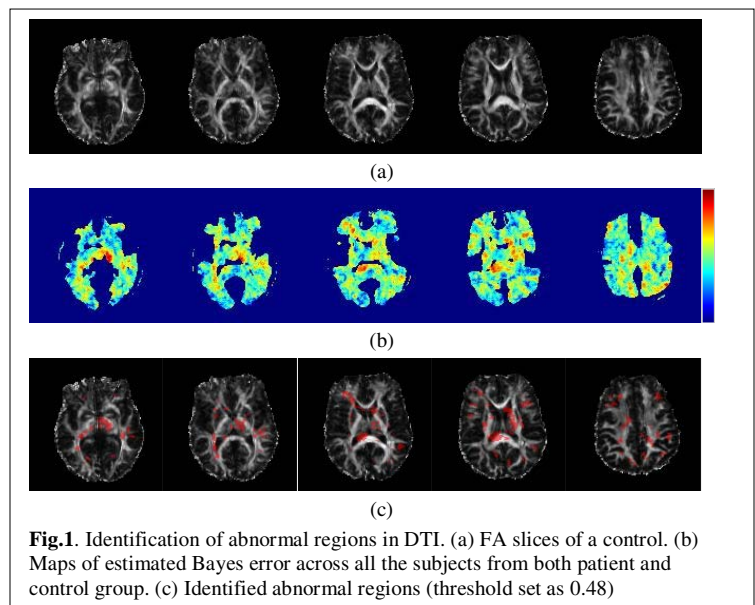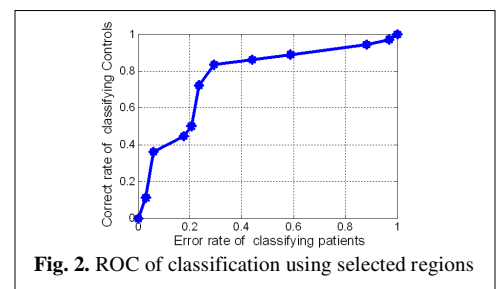
Fig. 2. ROC of classification using selected regions