Support Vector Machine Classification of Stroke Using Resting State Functional Connectivity

Svyatoslav Vergun[1], Veena Nair[2], Matthew Jensen[3], Marcus Chacon[3], Justin Sattin[3], and Vivek Prabhakaran[2]

[1]*Medical Physics, UW-Madison, Madison, WI, United States,* [2]*Radiology, UW-Madison, Madison, WI, United States,* [3]*Neurology, UW-Madison, Madison, WI, United States*

Target Audience

This work will be of interest to scientists studying functional connectivity of healthy and pathological aging populations, and in particular to those using multivariate pattern analysis tools to investigate these groups' differences.

Purpose

Resting-state functional MRI (rs-fMRI) is a powerful technique for studying whole brain neural connectivity, and allows us to examine the dynamics of activity within large scale networks potentially affected by stroke. In this study we seek to accurately classify normal and stroke subjects based on single rs-fMRI scans, examine whole brain functional connectivity differences between the groups and extract underlying connections that drive the classifications.

Methods

50 resting state functional MRI (rs-fMRI) scans from 24 healthy subjects (11 female, mean age = 47.4 years) and 26 acute stroke subjects (11 female, mean age = 58.6 years) were acquired on two GE 750 3T scanners with a gradient echo EPI sequence (40 slices, 231 volumes, 2.6ms TR, 3.5×3.5×3.5 mm). Data were preprocessed using scripts adapted from the 1000 Functional Connectome Project, [1] which included slice-timing correction, motion correction, band-pass filtering (0.005 – 0.1 Hz), linear and quadratic detrending, transformation into MNI space (3×3×3 mm), spatial smoothing (6 mm FWHM), and regression of white matter, CSF and global signal. Time series were then extracted from 160 previously defined regions of interest (ROIs) that were generated from meta-analyses focused on error-processing, default-mode, memory, language and sensorimotor functions. [2] Correlations from every ROI pair were fed into a linear support vector machine (SVM) classifier as features for each subject. Classification was performed with the Spider Machine Learning Toolbox [3] as well as custom scripts implemented in MATLAB.

Results

A linear kernel SVM classifier discriminated between stroke and normal subjects with 80% accuracy using leave-one-out cross-validation (p-value < 1×10^{-5}, sensitivity = 81%, specificity = 79%). Table 1 lists the classification of subjects. Figure 1 shows the top 10 features or connections that drive the classifier, and Figure 2 the top 10 ROIs. There was a significant age difference between the two groups (p-value = 0.0144), but, when age was added as a feature, the accuracy and subject classification were not affected.

Discussion

The classifier was able to predict both groups with high accuracy and was slightly more accurate in predicting stroke subjects. Classification was most influenced by the differences in connectivity of the cingulo-opercular (23% relative weight) and sensorimotor (41% relative weight) networks, with the sensorimotor network containing 7 of the top 10 ROIs.

Conclusion

Multivariate pattern analysis techniques have been successful in predicting healthy and disease brain states, and here we show that they can be used to accurately classify stroke and normal subjects based on functional connectivity. An important aspect of a linear SVM is its ability to extract features that drive the classification, allowing insight into pathological aging and healthy aging subject connectivity.

References

1. http://fcon_1000.projects.nitrc.org/ . Accessed November, 2012.
2. Dosenbach, N. U., et al. Prediction of individual brain maturity using fMRI. Science 2010; 329:1358–1361.
3. Weston, J., et al. The spider machine learning toolbox. http://people.kyb.tuebingen.mpg.de/spider/

Table 1. Classification of subjects

	Predicted Normal	Predicted Stroke
Normal	19	5
Stroke	5	21

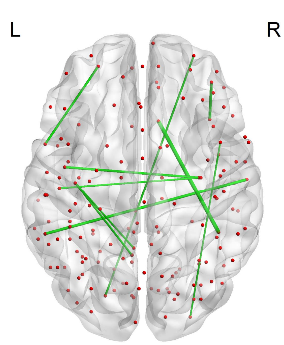

Figure 1. Top 10 features driving the classifier. Connection width represents its importance/weight.

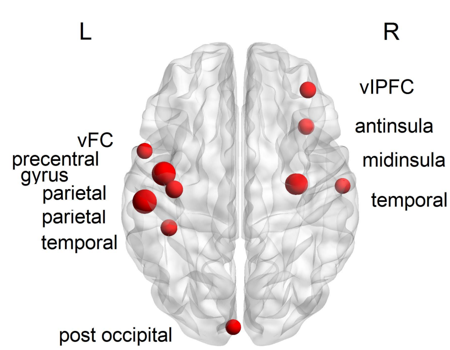

Figure 2. Top 10 ROIs and their labels, with ROI size indicating its importance. ROI weights are the sum of the weights of connections to and from that ROI.