Functional subtypes for prediction in schizophrenia

Christian Dansereau, Pierre Orban, Sebastian Urchs, Pierre Bellec

Abstract—Obtaining good prediction accuracy of schizophrenia status using functional connectivity is a difficult task and this is likely explained by the heterogeneity of symptoms observed in this mental disease. We propose an approach to tackle this heterogeneity by functionally subtyping the population. We also propose a two-stage classification framework to identify subtypes associated with good vs poor prediction accuracy. Using the degree of association of each subject functional maps to those subtypes we were able to 1) effectively compress complex spatio-temporal information while being significantly associated with the pathology 2)

Keywords—Schizophrenia, classification, multiscale, feature selection, SVM, margin optimization.

1 Introduction

SCHIZOPHRENIA has been defined as a disconnection syndrome, which is characterized by aberrant functional brain connectivity. Schizophrenia is also a very heterogeneous clinical entity, with marked variations in both the symptoms experienced by patients and their response to treatments. This variability is most likely associated with heterogeneous neurobiological aberrations. This heterogeneity can also be found in the so-called 'normal' population resulting in a lot of variations that is not accounted for with the current labels. We hypothesized that this neurobiological heterogeneity would be reflected in patterns of full brain functional connectivity (FC) in fMRI, and that specific subtypes of FC would be much more amenable than others to predict clinical symptoms.

Regarding the current clinical labels use in classification problems they are most likely poorly defined and therefore it is not possible to expect a reasonable prediction accuracy when training a classification model on those ill-defined labels.

• *Mr. Dansereau is with the Department of Computer science, University of Montreal, Montreal, CA.*
E-mail: christiandansereau@gmail.com

Manuscript received March 17, 2016

1.1 Objectives

Our goals for this study were therefore 1) to identify subtypes of functional connectivity, in a sample mixing patients diagnosed with schizophrenia and healthy controls, 2) Since prediction accuracy does not perform well with poorly defined labels, show that subtypes can predict more accurately a subgroup of the population by identifying easy cases, i.e. a subgroup of subjects that can be accurately diagnosed based on FC.

1.2 Public code and data

The code used in this experiment is available on a GitHub repository at the following URL[1]. A IPython Notebook is also provided with all of the figure generation scripts. The functional imaging data is publicly available on the NITRC web site[2].

2 Method

2.1 Dataset

The dataset used in this paper is the COBRE (The Center for Biomedical Research Excellence) dataset part of the INDI initiative. It consist of raw anatomical and functional MR data from 72 patients with Schizophrenia and

1. https://github.com/cdansereau/vision_or/tree/master/code_project2
2. http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html

75 healthy controls (ages ranging from 18 to 65 in each group). All subjects were screened and excluded if they had; history of neurological disorder, history of mental retardation, history of severe head trauma with more than 5 minutes loss of consciousness, history of substance abuse or dependence within the last 12 months. Diagnostic information was collected using the Structured Clinical Interview used for DSM Disorders (SCID).

A multi-echo MPRAGE (MEMPR) sequence was used with the following parameters: TR/TE/TI = 2530/[1.64, 3.5, 5.36, 7.22, 9.08]/900 ms, flip angle = 7, FOV = 256x256 mm, Slab thickness = 176 mm, Matrix = 256x256x176, Voxel size = 1x1x1 mm, Number of echos = 5, Pixel bandwidth = 650 Hz, Total scan time = 6 min. With 5 echoes, the TR, TI and time to encode partitions for the MEMPR are similar to that of a conventional MPRAGE, resulting in similar GM/WM/CSF contrast.

Resting state data was collected with single-shot full k-space echo-planar imaging (EPI) with ramp sampling correction using the intercomissural line (AC-PC) as a reference (TR: 2 s, TE: 29 ms, matrix size: 64x64, 32 slices, voxel size: 3x3x4 mm3).

We selected subjects matched for age, gender and motion for a total of 101 subjects (mean age 32.5, SD 10.3). 101 subjects were drawn from the COBRE study (resting-state fMRI, [1]), while 86 subjects were scanned in Montreal (emotional memory task fMRI, [2]). Individual fMRI scans were preprocessed using the NIAK pipeline [3]. We identified 12 functional brain networks using a Bootstrap Analysis of Stable Clusters (BASC [4]) on a separate dataset [5]. These networks were used as seeds to generate FC maps for each subject.

2.2 Preprocessing and feature extraction

Each fMRI dataset was corrected for slice timing; a rigid-body motion was then estimated for each time frame, both within and between runs, as well as between one fMRI run and the T1 scan for each subject [1]. The T1 scan was itself non-linearly co-registered to the Montreal Neurological Institute (MNI) ICBM152 stereotaxic symmetric template [2], using the CIVET pipeline [3]. The rigid-body, fMRI-to-T1 and T1-to-stereotaxic transformations were all combined to resample the fMRI in MNI space at a 3 mm isotropic resolution. To minimize artifacts due to excessive motion, all time frames showing a displacement greater than 0.5 mm were removed [4]. A minimum of 50 unscrubbed volumes per run was required for further analysis (13 subjects were rejected). The following nuisance covariates were regressed out from fMRI time series: slow time drifts (basis of discrete cosines with a 0.01 Hz highpass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first principal components (accounting for 95% variance) of the six rigid-body motion parameters and their squares [5], [6]. The fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring kernel. A more detailed description of the pipeline can be found on the NIAK website[3] and Github[4].

2.3 functional connectivity maps

functional connectomes were obtained using seed-based FC maps from an atlas of 12 networks and resolution of the maps were computed at the region-level (444 regions covering all the brain) instead of the voxel-level. Resulting in connectomes of size 12x444. The two atlases were obtained using a functional template based on an independent dataset of ~ 200 subjects form the 1000 functional connectome project (Cambridge dataset) REF.

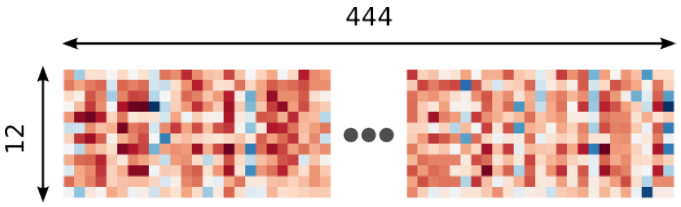

Fig. 1. Caption

2.4 Cross-validation

A nested cross-validation was performed for accuracy estimation and parameters optimization using leave-one out cross validation in

3. http://niak.simexp-lab.org/pipe_preprocessing.html
4. https://github.com/SIMEXP/multisite

order to estimate the generalizability of the model.

2.5 Subtypes

Instead of performing our classification on a correlation matrix we introduce a new metric called subtype weights that is computed as follow: For each network, we identified 10 subgroups by grouping subjects with similar FC maps through hierarchical agglomerative clustering resulting in subtypes template. We then derived subtype weights for each individual using the spatial correlation of the individual FC map and each subtype template. This dimensionality reduction show the degree of association for an individual connectivity map with the corresponding subtypes templates.

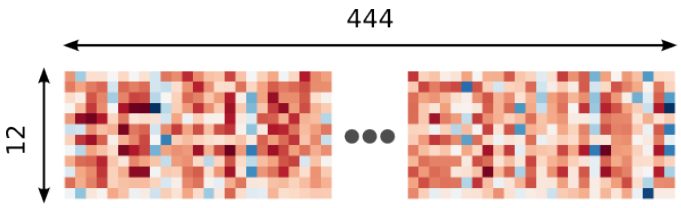

Fig. 2. Caption

2.6 Feature selection

To identify the subtype weights significantly associated with the pathology, we used a GLM including age, gender and frame displacement (FD) as covariates (corrected to have a zero mean across subjects). A p-value was generated for each subtype to quantify the probability that the difference between the control subjects and the patient with schizophrenia was significant [7]. The number of false discovery was also controlled ($\alpha = 0.01$) using a BenjaminiHochberg false discovery rate (FDR) procedure [8].

2.7 Two-steps classification

Using the training set from the previously generated subtype weights, we first trained a logistic regression (LR) model [9] for the classification of the clinical labels (LRC). Second, an other logistic regression model (LRHM) was trained to predict the hit (correct classification) and miss (incorrect classification) labels from the LRC, using the subtype weights as an input. We then used the test set to evaluate the prediction accuracy of the two subgroups (easy and hard to classify subjects) by deriving the accuracy score separately for the easy and difficult cases.

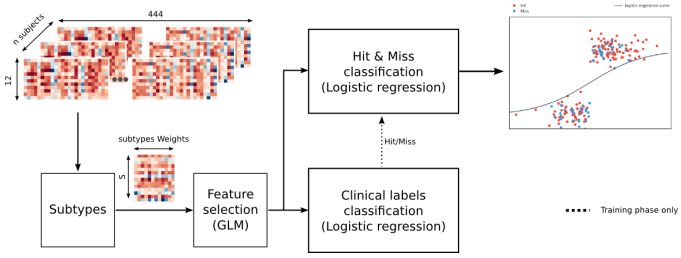

Fig. 3. Caption

3 Results

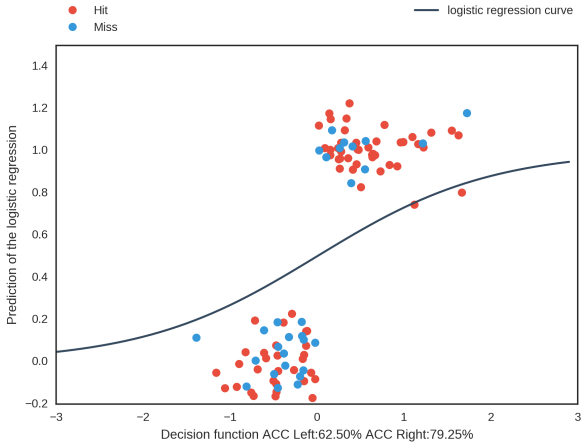

Fig. 4. Caption

If we combine the confound regression with the normalization (unit variance and zero mean for the connectivity of each subject) and we search for the optimal parameter C for the linear classifier we obtain 69.89% and an AUC of 0.80 see Figure **??**.

4 Discussion and Conclusion

The use of subtypes weight is a better way to reduce the dimension of the problem and for coding for complex structure in a compact representation.

Difficult cases: factors, Noise in the data, cognitive state (sleep etc...) Easy case: factors

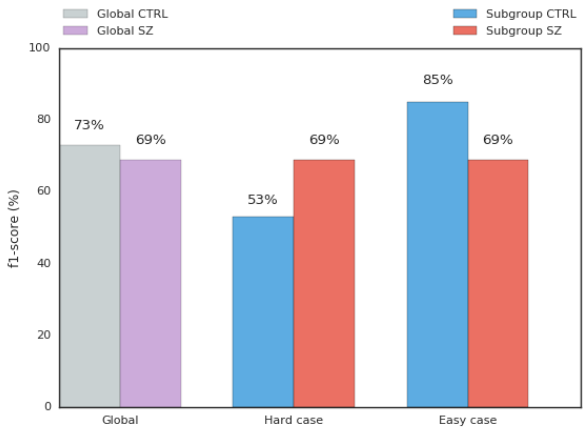

Fig. 5. Caption

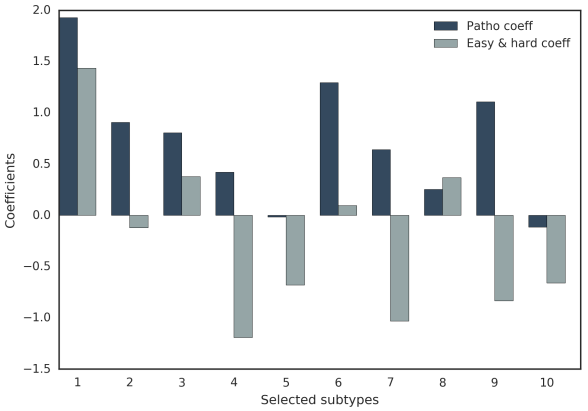

Fig. 6. Caption

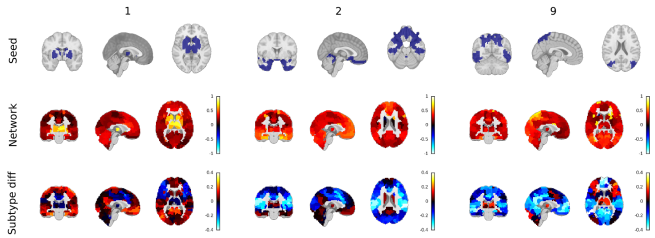

Fig. 7. Caption

TABLE 1
Summary results with no split. The accuracy is 71.3%

Class	*precision*	*recall*	*f1-score*	*support*
0	78%	69%	73%	58
1	64%	74%	69%	43
avg/total	72%	71%	71%	101

TABLE 2
Summary results with split on the right. The accuracy is 79.2%

Class	*precision*	*recall*	*f1-score*	*support*
0	81%	88%	**85%**	34
1	75%	63%	69%	19
avg/total	79%	79%	79%	53

TABLE 3
Summary results with split on the left. The accuracy is 62.5%

Class	*precision*	*recall*	*f1-score*	*support*
0	71%	42%	53%	24
1	59%	83%	69%	24
avg/total	65%	60%	61%	48

subgroup of subject were we have enough examples to reach a conclusion physiologically.

we found subtypes associated with the pathology and we have been able to learn from the prediction to identify a subgroup of subject that can be accurately classified.

By leveraging subtypes of FC, we were able to identify a subgroup of easy cases where it is possible to accurately (80%) predict clinical diagnosis based on individual FC maps. The prediction model was trained completely independently of our validation set, and the accuracy estimate are unbiased. The 80% accuracy is very satisfactory considering that both our training and validation sets comprised subjects scanned at different sites, using different imaging and behavioral protocols. Our findings suggest that the clinical labels may be ill-defined in terms of brain connectivity patterns, and that a subgroup population exists who presents a more definitive functional signature of clinical symptoms. We plan to investigate next the spatial organization of brain connectivity subtypes associated with good vs poor prediction accuracy.

A common mistake made when trying to tackle this kind of problem is to do the feature selection before cross-validation using the labels. This may lead to an overestimate of our accuracy, it may give very good result that will not generalize well. As an example we have performed the same experiment with the feature selection procedure (I-Relief) in and out of the cross-validation loop resulting in improved performance when selecting the features outside the cross-validation loop.

REFERENCES

[1] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, "Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space." *Journal of computer assisted tomography*, vol. 18, no. 2, pp. 192–205, 1994. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/8126267

[2] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry, D. L. Collins, and Brain Development Cooperative Group, "Unbiased average age-appropriate atlases for pediatric studies." *NeuroImage*, vol. 54, no. 1, pp. 313–327, Jan. 2011. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2010.07.033

[3] Y. Ad-Dab'bagh, D. Einarson, O. Lyttelton, J. S. Muehlboeck, K. Mok, O. Ivanov, R. D. Vincent, C. Lepage, J. Lerch, E. Fombonne, and A. C. Evans, "The CIVET Image-Processing environment: A fully automated comprehensive pipeline for anatomical neuroimaging research," in *Proceedings of the 12th Annual Meeting of the Human Brain Mapping Organization*, M. Corbetta, Ed. Florence, Italy: Neuroimage, 2006.

[4] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion." *NeuroImage*, vol. 59, no. 3, pp. 2142–2154, Feb. 2012. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2011.10.018

[5] F. Giove, T. Gili, V. Iacovella, E. Macaluso, and B. Maraviglia, "Images-based suppression of unwanted global signals in resting-state functional connectivity studies." *Magnetic resonance imaging*, vol. 27, no. 8, pp. 1058–1064, Oct. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.mri.2009.06.004

[6] T. E. Lund, K. H. Madsen, K. Sidaros, W.-L. Luo, and T. E. Nichols, "Non-white noise in fMRI: does modelling have an impact?" *NeuroImage*, vol. 29, no. 1, pp. 54–66, Jan. 2006. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2005.07.005

[7] K. J. Worsley and K. J. Friston, "Analysis of fMRI Time-Series Revisited—Again," *NeuroImage*, vol. 2, no. 3, pp. 173–181, Sep. 1995. [Online]. Available: http://dx.doi.org/10.1006/nimg.1995.1023

[8] Y. Benjamini and Y. Hochberg, "Controlling the false-discovery rate: a practical and powerful approach to multiple testing." *J. Roy. Statist. Soc. Ser. B*, vol. 57, pp. 289–300, 1995.

[9] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *The Journal of Machine Learning Research*, vol. 9, pp. 1871–1874, 2008.