Unsupervised fMRI Analysis

David R. Hardoon
Department of Computer Science
University College London
London WC1E 6BT
United Kingdom
D.Hardoon@cs.ucl.ac.uk

Janaina Mourão-Miranda
Brain Image Analysis Unit
Centre for Neuroimaging Sciences
Institute of Psychiatry, De Crespigny Park
London SE5 8AF
Janaina.Mourao-Miranda

Michael Brammer
Brain Image Analysis Unit
Centre for Neuroimaging Sciences
Institute of Psychiatry, De Crespigny Park
London SE5 8AF
Michael.Brammer@iop.kcl.ac.uk

John Shawe-Taylor
Department of Computer Science
University College London
London WC1E 6BT
United Kingdom
jst@cs.ucl.ac.uk

Abstract

Recently machine learning methodology has been used increasing to analyze the relationship between stimulus categories and fMRI responses [2, 14, 15, 11, 13, 8, 9, 1, 12, 7]. Here, we introduce a new unsupervised machine learning approach to fMRI analysis approach, in which the simple categorical description of stimulus type (e.g. type of task) is replaced by a more informative vector of stimulus features. We compared this new approach with a standard Support Vector Machine (SVM) analysis of fMRI data using a categorical description of stimulus type.

The following study differs from conventional unsupervised approaches in that we make use of the stimulus characteristics. We use kernel Canonical Correlation Analysis (KCCA) to learn the correlation between the fMRI volume and the corresponding stimulus features presented at a particular time point. CCA can be seen as the problem of finding basis vectors for two sets of variables such that the correlation of the projections of the variables onto these basis vectors are mutually maximised. KCCA first projects the data into a higher dimensional feature space before performing CCA in the new feature space.

To classify a new stimulus we compute the kernel between the test and the training samples and project that kernel onto the learned semantic space. The output gives a score which can be thresholded to allocate a category to each test example.

CCA [3, 4] and KCCA [5] have been used in previous work for fMRI analysis using only conventional categorical stimulus descriptions and did not explore the possibility of using the complex characteristics of the stimuli as the base for the feature selection in the fMRI data.

The fMRI data used is from an experiment in which we studied the responses to stimuli designed to evoke different types of emotional responses (pleasant/unpleasant). The pleasant images consisted of women in swimsuits while the unpleasant images comprised image of

skin diseases. For experiment details see [7]. Each stimulus image was represented using Scale Invariant Feature Transformation (SIFT) [10] features, which are used in a similar fashion to the bag-of-words (word-frequency) models for text documents.

The complete dataset consisted of 1344 fMRI volumes and their corresponding image-stimuli. From this dataset, we used 670 fMRI volumes ($517,845$ voxels per volume) and their corresponding image-stimuli for training and the remaining 674 for testing. In both training and testing datasets half the stimuli were *Pleasant* and half *Unpleasant*, hence a random classifier would achieve 50% accuracy. The experiments are repeated ten times, each with a random permutation of the training-testing set. We use linear, centralised, kernels in our KCCA framework.

The KCCA regularisation parameter [6] is computed using $2-$fold cross validation on the training data. KCCA achieves an accuracy of $81.34\% \pm 3.29\%$ while the supervised SVM attains an accuracy of $90.28\% \pm 0.93\%$. These classifications accuracy's are significantly different than of chance expectation *t*-test $p << 0.0001$. The unthresholded weights vectors for both methods are displayed in 1. The similarities in the patterns of the weight vectors obtained by the two classification methods provides reassurance of the validity of the KCCA analysis, which is effectively obtaining the categorical stimulus type directly from the data.

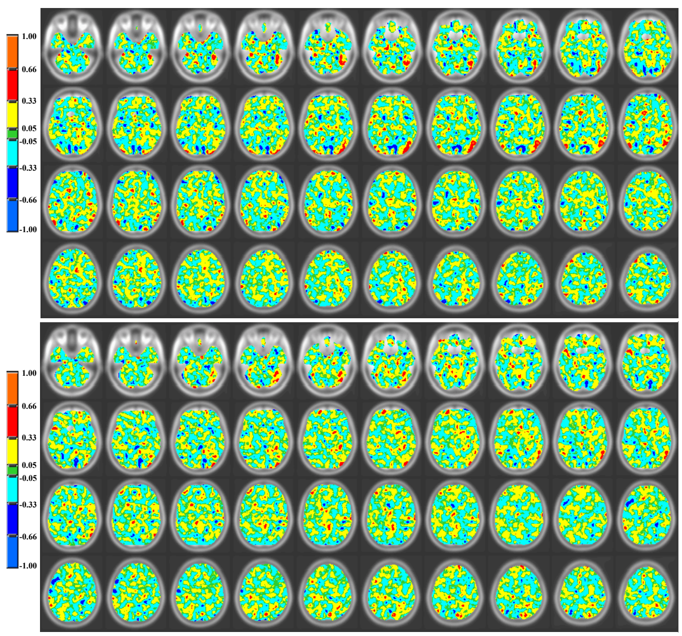

Figure 1: The unthresholded weight values showing the contrast between viewing *Pleasant* vs. *Unpleasant*. We use the blue scale for negative (*Unpleasant*) values and the red scale for the positive values (*Pleasant*). Top figure: The discrimination analysis on the training data was performed with labels $(+1/-1)$. Bottom figure: The discrimination analysis on the training data was performed without labels. The class discrimination is automatically extracted from the analysis.

Acknowledgements

The authors would like to thank Larry M. Manevitz from the *HIACS* Reseach Centre, University of Haifa for allowing them to use the centre's computation facility *eagle*. This work was supported by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778. This publication only reflects the authors views.

References

[1] Davatzikos C., Ruparel K., Fan Y., Shen D.G., Acharyya M., Loughead J.W., Gur R.C., and Langleben D.D. Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, 28:663–668, 2005.

[2] Cox D.D. and Savoy R.L. Functional magnetic resonance imaging (fmri) 'brain reading': detecting and classifying distributed patterns of fmri activity in human visual cortex. *Neuroimage*, 19:261–270, 2003.

[3] O. Friman, M. Borga, P. Lundberg, and H. Knutsson. Adaptive analysis of fMRI data. *NeuroImage*, 2003. In Press.

[4] O. Friman, J. Carlsson, P. Lundberg, M. Borga, and H. Knutsson. Detection of neural activity in functional MRI using canonical correlation analysis. *Magnetic Resonance in Medicine*, 45(2):323–330, February 2001.

[5] David R. Hardoon, John Shawe-Taylor, and Ola Friman. KCCA for fMRI Analysis. In *Proceedings of Medical Image Understanding and Analysis*, London, UK, 2004.

[6] David R. Hardoon, Sandor Szedmak, and John Shawe-Taylor. Canonical correlation analysis: an overview with application to learning methods. *Neural Computation*, 16:2639–2664, 2004.

[7] Mourao-Miranda J., Reynaud E., McGlone F., Calvert G., and Brammer M. The impact of temporal compression and space selection on svm analysis of single-subject and multi-subject fmri data. *NeuroImage (accepted)*, 2006.

[8] Mourao-Miranda J., Bokde A. L.W., Born C., Hampel H., and Stetter S. Classifying brain states and determining the discriminating activation patterns: support vector machine on functional mri data. *NeuroImage*, 28:980–995, 2005.

[9] Haynes J.D. and Rees G. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8:686–691, 2005.

[10] D.G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the 7th IEEE International Conference on Computer vision*, pages 1150–1157, Kerkyra Greece, 1999.

[11] T.M. Mitchell, R. Hutchinson, R.S. Niculescu, F. Pereira, X. Wang, M. Just, and S. Newman. Learning to decode cognitive states from brain images. *Machine Learning*, 1-2:145–175, 2004.

[12] Kriegeskorte N., Goebel R., and Bandettini P. Information-based functional brain mapping. *PANAS*, 103:3863–3868, 2006.

[13] LaConte S., Strother S., Anderson J. Cherkassky V., and Hu X. Support vector machines for temporal classification of block design fmri data. *NeuroImage*, 26:317–329, 2005.

[14] Carlson T.A., Schrater P., and He S. Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience*, 15:704–717, 2003.

[15] Wang X, Hutchinson R., and Mitchell T.M. Training fmri classifiers to detect cognitive states across multiple human subjects. In *In Proceedings of the 2003 Conference on Neural Information Processing Systems*, 2003.