PsychNology Journal, 2008
Volume 6, Number 2, 189 – 201

Decoding Cognitive States from fMRI Data Using Support Vector Regression

Maria Grazia Di Bono*♦ and Marco Zorzi♥

♦Department of Developmental Psychology and Socialization, University of Padova (Italy)

♥Department of General Psychology and Center for Cognitive Science, University of Padova (Italy)

ABSTRACT

In this paper we describe a method based on Support Vector machines for Regression (SVR) to decode cognitive states from functional Magnetic Resonance Imaging (fMRI) data. In the context of the Pittsburgh Brain Activity Interpretation Competition (PBAIC, 2007), three participants were scanned during three runs of 20-minute immersion in a Virtual Reality Environment (VRE) where they played a game that engaged them in various search tasks. A set of objective feature ratings was automatically extracted from the VRE during the scanning session, whereas a set of subjective features was then derived from each individual experience. The aim of the present study was to explore the feasibility of the SVR approach in the case of an extremely complex regression problem, in which subjective experience of participants immersed in a VRE had to be predicted from their fMRI data. The proposed methodology was modeled as a multiphase process: a pre-processing phase, based on a filter approach, for fMRI image voxel selection, and a prediction phase, implemented by nonlinear SVR, for decoding subjective cognitive states from the selected voxel time series. Results highlight the generalization ability of nonlinear SVR, making this approach particularly interesting for real world application of Brain Computer Interface (BCI).

Keywords: *Brain Computer Interfaces, Signal Processing, fMRI Data, Multivariate Analysis, Support Vector Machine.*

Paper Received 04/09/2007; received in revised form 15/08/2008; accepted 20/08/2008.

1. Introduction

Recent advances in brain imaging and machine learning provide the foundations for the development of Brain-Computer Interfaces (BCI) based on functional Magnetic Resonance Imaging (fMRI) (Weiskopf et al., 2004). In the last decade, fMRI has

Cite as:
Di Bono, M. G., & Zorzi, M. (2008). Decoding Cognitive States from fMRI Data Using Support Vector Regression. *PsychNology Journal, 6*(2), 189 – 201. Retrieved [month] [day], [year], from www.psychnology.org.

* Corresponding Author:
Maria Grazia Di Bono
Department of Developmental Psychology and Socialization, University of Padova
Via Venezia 12, 35131 Padova, Italy
E-mail: mariagrazia.dibono@unipd.it

become the most widely used non-invasive technique for investigating human brain functions.

The main problem addressed by fMRI studies is to correlate neural signals temporally changing in different cortical areas to certain events (task conditions) opportunely encoded in the experimental paradigm. A BCI based on fMRI needs to interpret the relationships between neural signals from fMRI data and the subjective experience of scanned participants, in order to perform a sort of mind reading and to translate the predicted mental states into actions.

Conventional fMRI data analysis techniques are based on the statistical *Parametric Approaches* (e.g. General Linear Model – GLM) which correlate external regressors (task conditions) with the activity in specific brain regions, generating brain maps of localised activity (Friston et al., 1995). Traditional statistical methods measure the activity from many thousands of voxels in the brain images, analysing each voxel time series separately and comparing two or more task conditions at each voxel location. Recent methods, belonging to the class of Multivariate Analysis (*Non Parametric Approaches*), have the potential to improve our understanding of the complex pattern of brain activity measured by fMRI. These approaches are based on boosting the weak information available at each voxel location by a simultaneous analysis of hundreds or thousands of voxels to predict specific cognitive states.

Many pattern recognition methods have been employed as multivariate techniques for fMRI data analysis. Machine learning techniques based on artificial neural networks (Chuang, Chiu, Lin, & Chen, 1999; Voultsidou, Dodel, & Herrmann, 2005) or different clustering algorithms (Meyer & Chinrungrueng, 2003; Liao, 2005; Heller, Stanley, Yekutieli, Rubin, & Beniamini 2006; Chen, Bouman, & Lowe, 2004; Chen, H., Yuan, Yao, Chen, L., & Chen, W., 2006) have been employed for time series data analysis in different domain applications, among which fMRI data analysis. Other methodologies, such as independent component analysis (ICA), have also been used for processing fMRI data (Hu et al., 2005; Meyer-Baese, Wismueller, & Lange, 2004). One of the most widely used Machine Learning techniques for fMRI data analysis are Support Vector Machines (SVM), which are kernel-based methods designed to find functions of the input data that enable both classification and regression (Vapnik, 1995). In particular, SVMs classify data with different class labels by determining a set of support vectors, which are members of the training set, outlining a hyperplane in the feature space. SVM provides a mechanism to fit the hyperplane surface to the training data using a specific kernel function. SMV classifiers are well known for their very good generalization ability and have been used in recent studies of fMRI data analysis (Cox & Savoy, 2003; Mitchell et al., 2004; Kamitani & Tong, 2005; Haynes & Rees, 2005; La Conte et al., 2005, Mourão-Miranda, Friston, & Brammer, 2007). However, most of the previous studies have focused on classification problems. In the case of regression problems, the goal is to find a functional shape for the function that can correctly predict new cases that the SVM has not been presented with before. This latter method is usually referred to as Support Vector Regression (SVR; see Smola & Schölkopf, 2004, for a review). Thus, our goal was to explore different techniques of feature selection and the feasibility of SVR in the case of an extremely complex regression problem whereby the subjective experience of participants immersed in a virtual reality environment (VRE) must be predicted from a set of fMRI data.

2. Materials and Methods

The VRE experiment was organized in the context of the Pittsburgh Brain Activity Interpretation Competition (PBAIC, 2007). The purpose of this competition was to infer subjective experience of the participants experiencing an immersion in a virtual reality environment, from a contextually gathered set of fMRI data.

The VRE experiment, the gathered fMRI data and the proposed fMRI decoding method are described in the next sections.

2.1 Participants

Fifteen subjects participated in this study. In particular, after a selection procedure made by the PBAIC staff (see PBAIC, 2007 for details), the only data available for the competition were relative to three subjects (age range: 20-26 years).

2.2 Procedure

Participants were instructed to play a game in a virtual world during three runs of fMRI data acquisition. In the game they were paid by an anthropology department grant to gather information about urban people. In particular, they were visiting several times the virtual reality environment, outside and inside some specified places, and were instructed to collect as much as possible samples of toy weapons and fruits, in a predefined order; moreover, they had to take pictures of people with piercings and avoid an aggressive dog. Participants were also informed and asked to keep in mind that any money obtained in the game corresponded to an earning in real life.

The study was completed over a period of four days. During the first day, participants watched a 13-minute video for a first phase of familiarization with the VRE and completed a battery of questionnaires, implicit association tests for assessing the ingroup/outgroup and canines perception, the level of anxiety and sickness, the sense of direction, the computer familiarity scale (see PBAIC, 2007 for details). During the second day, participants played the virtual reality game outside the scanner and every 2 minutes they were asked to rate their level of sickness. At the end of the session they completed three questionnaires for assessing the level of simulator sickness and the level of presence and comfort during the navigation. In the third day, subjects were asked to perform search tasks during three 20 minute runs of the game inside the scanner. As in the previous day, participants were asked to rate every 2 minutes their level of sickness.

During the navigation of the virtual world, a set of target feature ratings was gathered for a total of 13 required and 10 optional features. In particular, some features were obtained through software loggings of subject actions in the virtual world, soundtrack analysis and eye-tracking based analysis of video from each run, and were referred to as *objective features* (e.g., *Hits*: whenever subjects correctly picked up fruit or weapon or took pictures of persons with piercings; *Instructions*: whenever task instructions were presented; *Search people; Search weapons; Search fruit; Faces*: whenever subjects looked at faces of pierced or unpierced people).

The other features were referred to as *subjective features*, such as the level of arousal or valence (how positive or negative is the environment), and they were assigned by each participant on the last day of the study while watching a playback of the own actions in the virtual world.

Figure 1 shows a screenshot of the virtual world and the behavioural time vector ratings of multiple categories representing what participants perceived/experienced during the navigation of the virtual world. For the first two runs, videos of the subject's path through the virtual reality environment along with 20 minutes of continuous fMRI data and target feature ratings were provided, whereas for the third run the ratings were not provided.

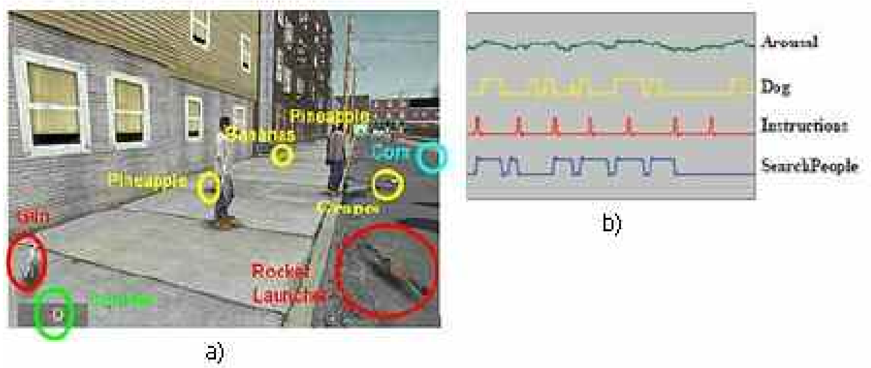

Figure 1. Screenshot of the virtual world (a); illustration of behavioral time vector ratings of multiple categories (b).

The purpose of the competition was to predict feature rating vectors for the third segment.

2.3 fMRI Dataset

3T EPI fMRI data from three subjects in three runs were downloadable from the Pittsburgh Brain Activity Interpretation Competition web site (PBAIC 2007). Twenty minutes of continuous functional data, consisting of 704 of 64x64x34 image volumes, were available for each participant in each run. The acquired images were motion corrected, slice time corrected, detrended and spatially normalized.

The fMRI data of the first two runs were used for the learning phase, whereas the last run was used as test set for the prediction of the related ratings.

2.4 fMRI Decoding Method

The proposed fMRI decoding method is modelled as a multiphase process (pre-processing phase, prediction phase) as shown in Figure 2.

In the *pre-processing phase* we first extracted only those voxels belonging to the brain cortex, by using the respective masks available for each subject. Then, for each subject, all brain voxel time series were normalized to the same mean intensity and temporally filtered. We then performed the voxel subset selection based on a filter approach. For each feature rating, convolved with the canonical Hemodynamic Response Function (HRF), we computed the correlation with each voxel time series in the image volumes, separately for each subject and run. Then we selected only those voxels showing a correlation that was significant at the 0.05 level ($r > 0.45$, $p < 0.05$). The subsets extracted for the first two runs, used as training set, were merged to form the final set of voxel time series.

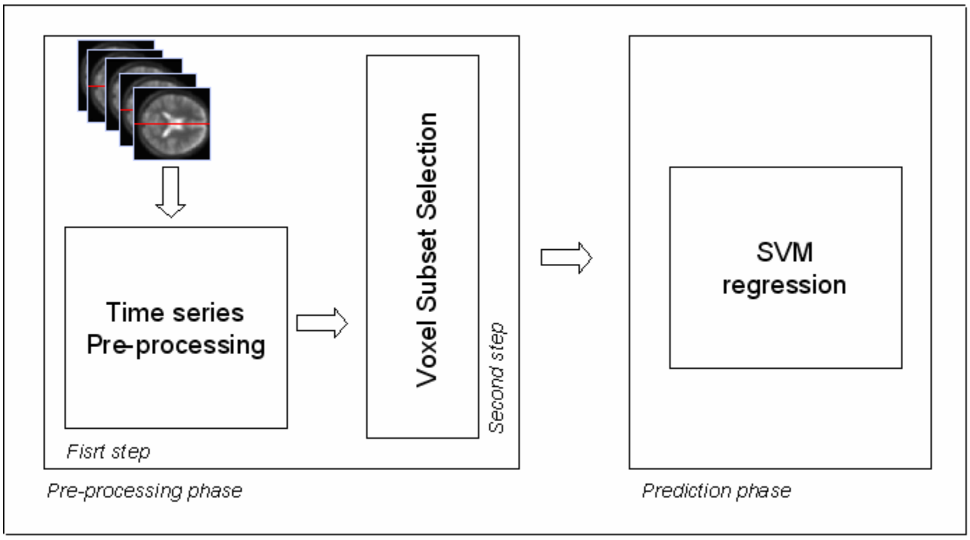

Figure 2. Architecture of the system.

In order to validate the methodology, we initially used only the first run for training and the second one for testing the method. In particular we developed two different approaches for the pre-processing phase. In the first approach we used the subset of voxels extracted by the correlation filter directly as input to SVM. In the second approach we clustered the selected voxels using hierarchical clustering and k-means algorithms into 100 clusters, then we extracted the centroid of each defined cluster and used it like a super-voxel as input to SVM for the prediction phase.

After this initial step, dedicated to the validation of the methodology, we selected the pre-processing approach that provided the best results with the first two runs. Thus, in the *prediction phase*, we used run 1 and run 2 of the same subject as training data to predict, in the test phase, the feature ratings for the third run. For each subject each feature was predicted separately.

2.5 SVM Regression (SVR)

Support Vector Machines (SVM) were developed by Vapnik (1995) to solve classification problems, but recently SVM have been successfully extended to regression and density estimation problems (Smola & Schölkopf, 2004, for a review).

Suppose we have training data $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \subset \Re^M \times \Re$, in the ε-insensitive Support Vector Regression (SVR) technique (Vapnik, 1995) the goal is to find the function $f(x)$ that has at most ε deviation from the actually obtained target y_i for all the vectors of observation x_i in the training data, and at the same time is as flat as possible. In nonlinear SVR, the input vector x is first mapped onto a high dimensional feature space using some fixed nonlinear mapping, and then a linear model is constructed in this feature space as:

$$f(x) = \langle w, \Phi(x) \rangle + b \ , \ w \in \Re^M, b \in \Re \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product in $\Re^M$, w is the weight vector, $\Phi(x)$ is the nonlinear mapping function, and b is the bias.

However, one also want to allow for some errors, thus, analogously to the soft margin loss function adopted by Vapnik (1995), one can introduce (non-negative) slack variables $\xi_i, \xi_i^*, i = 1,..., N$ to measure the deviation of training samples outside the ε-insensitive zone. Thus the problem can be modelled as a convex optimization problem that, in the nonlinear case, can be formulated as:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \text{ subject to } \begin{cases} y_i - \langle w, \Phi(x_i) \rangle - b \le \varepsilon + \xi_i^* \\ \langle w, \Phi(x_i) \rangle + b - y_i \le \varepsilon + \xi_i \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{2}$$

In most cases, problem (2) can be easily solved in its dual formulation, that provides the key to extend SVR to nonlinear cases. The solution of the dual convex problem can be expressed as:

$$f(x) = \sum_{i=1}^{N}\left(\alpha_i - \alpha_i^*\right)K(x_i, x) + b \tag{3}$$

where α_i, α_i^* are the so called Lagrange multipliers and $K(x, x')$ is the kernel function. SVM algorithm only depends on dot products between patterns x_i, hence it is sufficient to know the kernel function $K(x, x') = \langle \Phi(x), \Phi(x') \rangle$ rather then Φ explicitly.

From the optimality constraints that are behind the dual problem definition, it is possible to derive w as a linear combination of the training patterns:

$$w = \sum_{i=1}^{N}\left(\alpha_i - \alpha_i^*\right)\Phi(x_i) \tag{4}$$

The difference from the linear case is that w can no longer given explicitly, whereas the flattest function $f(x)$, that has to be found in the feature space (not in the input space), can be expressed through the trick of the kernel function.

Several functions, such as polynomial functions, radial basis functions (RBF), splines, hyperbolic tangent functions, can be used as kernel in SVR (Burges, 1998; Smola & Schölkopf, 2004). These functions have to satisfy the conditions of the Mercer's theorem (Mercer, 1909), that is the conditions under which it is possible to write $k(x, x')$

as a dot product in some feature space. In particular, translation invariant kernels $k(x, x') = k(x - x')$, that are proved be admissible kernels, are widespread, among which one of the most widely used is the RBF kernel that can be written as:

$$k(x, x') = e^{-\frac{\|x - x'\|^2}{2\sigma^2}} \tag{5}$$

It is well known that SVM generalization performance (estimation accuracy) depends on a good setting of meta-parameters *C*, ε and the kernel parameters (Burges, 1998; Smola & Schölkopf, 2004). The parameter *C* determines the trade off between the model complexity and the degree to which deviations larger than ε are tolerated in optimization formulation. Parameter ε controls the width of the ε-insensitive zone, used to fit the training data. The value of ε can affect the number of support vectors used to construct the regression function. The bigger ε, the fewer support vectors are selected. Hence, both *C* and ε-values affect model complexity, but in a different way.

3. Results and Discussion

In developing the decoding method, we tested and tuned the SVR using run 1 for all subjects as training and run 2 as test set. We then applied our method for predicting the feature ratings of the third run after training on the voxels and ratings of the first two runs. All the algorithms used here were developed in *Matlab 7.0.1 (R14)*, by using the *SVM toolbox* (Gunn, 2007) for developing regression algorithms and the tools of *NIfTI (ANALYZE) MR image* (Shen, 2005) for fMRI volume visualization.

In the *pre-processing* phase, all brain voxel time series were normalized to the same mean intensity (subtracting their mean and dividing by their standard deviation) and temporally filtered, by using a running average filter with window size = 7 (Figure 3).

As mentioned in section 2.4, we tested two different approaches of SVR input preparation, one based on a correlation filter for voxel selection and the other based on voxel selection followed by feature extraction. In particular, in the first approach we extracted for each subject the set of voxels showing a correlation with each feature rating (convolved with the canonical HRF) of run 1 that was significant at the 0.05 level ($r > 0.45$, $p < 0.05$) and merged them to obtain the final set. We then used the coordinates of the extracted voxels for selecting the same set of voxels from run 2. SVR was finally employed for predicting the feature ratings of the second run using the voxel time series and the target ratings of the first run as training set. In the second approach, we applied a clustering step in the pre-processing phase based on *hierarchical clustering* and *k-means* algorithm, but we did not find any improvement in run 2 feature rating prediction. In contrast, we observed a general degradation of the prediction performance. We suggest that this deterioration could be due to a loss of the original distributed information which was compressed in the clustering phase. Thus, for the prediction of the target feature ratings of the third run we did not apply any clustering procedure or feature extraction.

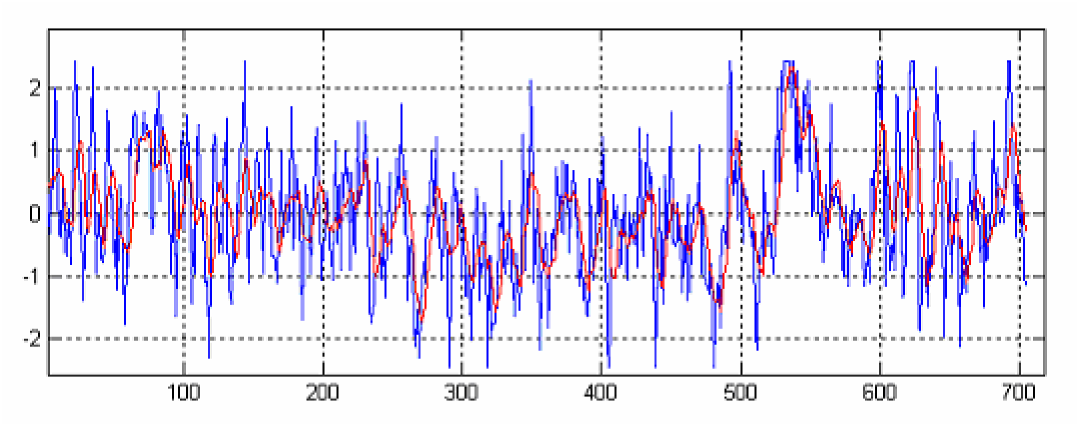

Figure 3. Example of voxel time series after normalization (in blue) and temporal filtering (in red).

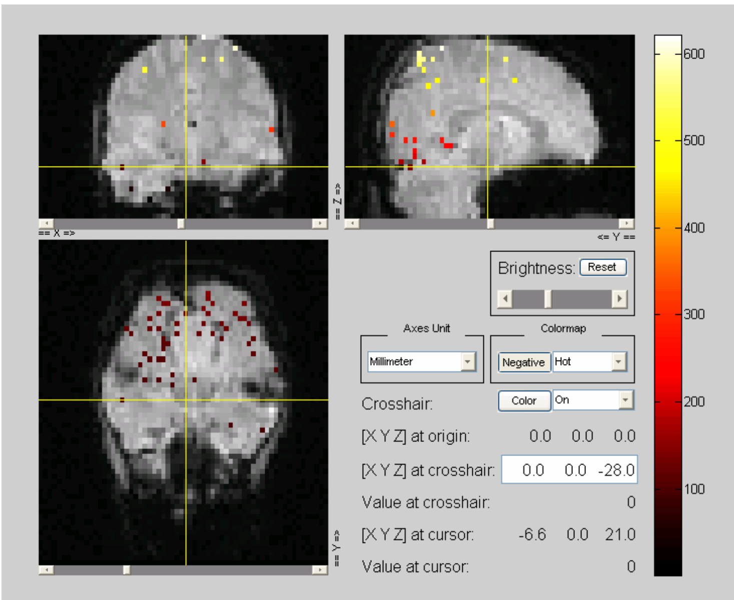

Figure 4. Selected voxels (subject 14) after computing the correlation with the convolved feature ratings (all features) on run 1 and run 2 and selecting the intersection between the two runs.

After the validation of the method, we used it for predicting target feature ratings of the third run. We therefore extracted two different subset of voxel time series, applying the correlation filter for both run 1 and run 2 with their respective feature ratings, and considered only the intersection between them as the final set to be used for selecting the voxels in the third run (Figure 4).

SVR was then employed to predict the feature ratings for run 3, using voxels and ratings of the first and second run as training set. As explained in section 2.4.1, in the prediction phase the choice of the regularization constant C, the kernel function and its parameters was a fundamental key for obtaining good generalization performance. We explored different sets of critical parameters, empirically obtaining the best results using *C=2* and the *Exponential Radial Basis function* (with $\sigma=6$) as nonlinear kernel.

Table 1 shows the results obtained with the first pre-processing approach for the prediction of the target feature ratings relative to the third run. The prediction scores are expressed in terms of standardised correlation coefficients, and, at least for the third run, were computed directly by the Experience Based Cognition (EBC) Project staff that had the relative target feature ratings. In particular, the scoring algorithm adopted by the EBC team was based on two steps. In the first step, the Pearson correlation coefficient r between the predicted feature and the observed subject rating was computed for each feature. In the second step, the Fisher transformation, that makes the scores normally distributed, was applied to each correlation calculated in the previous step, according to:

$$z' = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right) \tag{12}$$

The obtained predictions, at least for a subset of features, reached a good correlation with the target ones. A consistency across subjects can be noted with respect to the features that are more reliably predicted.

	Subject 1	**Subject 13**	**Subject 14**
Body	0.2421	0.4178	0.3438
Velocity	**0.4554**	**0.5197**	**0.7126**
Hits	**0.3035**	**0.4339**	0.3959
Instruction	**0.5453**	**0.7957**	**0.7842**
Faces	0.2611	0.3595	**0.5871**

Table 1: The best feature rating predictions, expressed in terms of correlations, achieved on run 3 for all subjects. The best three correlations for each subject are shown in bold font.

In conclusion, the aim of this study was to explore the feasibility of SVR in the case of an extremely complex regression problem, in which subjective experience of participants immersed in a VRE had to be predicted from their fMRI data. We used a decoding method modeled as a multiphase process: a pre-processing phase, based on a filter approach, for fMRI image voxel selection, and a prediction phase, implemented by nonlinear SVR, for decoding subjective cognitive states from the selected voxel time series. Results are quite good, at least for a subset of feature ratings. The emphasis of SVM/SVR on generalization ability makes this approach particularly interesting for real-world applications of BCI (Weiskopf et al., 2004; Piccione et al, 2006; Sitaram et al., 2007; Piccione et al., 2008), in particular when the amount of training data is limited and the input space has a high dimension (as in the case of fMRI data).

Planned extensions to this work include the evaluation of different feature extraction techniques to combine with SVM/SVR or the use of embedded methods, in the context of nonlinear kernels, for voxel selection and ranking, in order to extract a more compact and informative set of voxels and to further increase the prediction accuracy.

4. Acknowledgements

We thank the Experience Based Cognition (EBC) Project team for providing the fMRI data of the PBAIC 2007. Finally, we would like to thank three anonymous reviewers for their helpful suggestions. This study was supported by a grant from the Cariparo Foundation.

5. References

Chen, S., Bouman, C. A., & Lowe, M. J. (2004). Clustered components analysis for functional MRI. *IEEE Transactions on Medical Imaging, 23*, 85-98.

Chen, H., Yuan, H., Yao, D., Chen, L., & Chen, W. (2006). An integrated neighbourhood correlation and hierarchical clustering approach of functional MRI. *IEEE Transactions on Biomedical Engineering, 53*, 452-458.

Chuang, K. H, Chiu, M. J., Lin, C. C., & Chen, J. H. (1999). Model-free functional MRI analysis using Kohonen clustering neural network and fuzzy C-means. *IEEE Trans Med Imaging, 18* (12), 1117-1128.

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) ”brain reading”: detecting and classifying distributed patterns of fMRI activity in the human visual cortex. *Neuroimage, 19*, 261-270.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*(2), 121-167.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping, 2* (4), 189-210.

Gunn, S. R. (2007). Support Vector Machines for Classification and Regression. SVM toolbox [Computer software]. Retrieved August 27, 2007, from http://www.isis.ecs.soton.ac.uk/isystems/kernel/svm.zip.

Haynes, J. & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in primary visual cortex. *Nature Neuroscience, 8*, 686-691.

Heller, R., Stanley, D., Yekutieli, D., Rubin, N., & Beniamini, Y. (2006). Cluster-based analysis of FMRI data. *NeuroImage, 33*, 599-608.

Hu, D., Yan, L., Liu, Y., Zhou, Z., Friston, K. J., Tan, C., & Wu, D. (2005). Unified SPM–ICA for fMRI analysis. *NeuroImage, 25*, 746-755.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8*, 679-685.

LaConte, S., Strother, S., Cherkassky, V., Anderson, J., & Hu, X. (2005). Support vector machines for temporal classification of block design of fMRI data. *Neuroimage, 26*, 317-329.

Liao, T. W. (2005). Clustering of Time Series Data - A Survey. *Pattern Recognition, 38* (11), 1857-1874.

Meyer, F. G., & Chinrungrueng, J. (2003). Analysis of event-related fMRI data using best clustering bases. *IEEE Transactions on Medical Imaging, 22*, 933–939.

Meyer-Baese, A., Wismueller, A., & Lange, O. (2004). Comparison of two exploratory data analysis methods for fMRI: unsupervised clustering versus independent component analysis. *IEEE Transactions on Information Technology in Biomedicine, 8*, 387-398.

Mercer, J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society, London, A209*, 415-446.

Mitchell, T., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M., & Newman, S. (2004). Learning to Decode Cognitive States from Brain Images, *Machine Learning, 57*, 145-175.

Mourão-Miranda, J., Friston, K. J., & Brammer, M. (2007). Dynamic discrimination analysis: a spatial-temporal SVM. *Neuroimage, 36*(1), 88-99.

PBAIC (2007). Retrieved August 27, 2007, from http://www.ebc.pitt.edu/2007/competition.html.

Piccione, F., Giorgi, F., Tonin, P., Priftis, K., Giove, S., Silvoni, S., Palmas, G., & Beverina, F. (2006). P300-based brain computer interface: Reliability and performance in healthy and paralised participants. *Clin Neurophysiol*, *117*(3), 531-537.

Piccione, F., Priftis, K., Tonin, P., Vidale, D., Furlan, R., Cabinato, M., Merico, A., & Piron, L. (2008). Task and Stimulation Paradigm Effects in a P300 Brain Computer Interface Exploitable in a Virtual Environment: A Pilot Study. *PsychNology Journal, 6*(1), 99-108.

Shen, J. (2005). NIfTI (ANALYZE) MR image tool [Computer software]. Retrieved August 27, 2007 from http://www.mathworks.com/matlabcentral/.

Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., Shimizu, K., & Birbaumer, N. (2007). Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *Neuroimage, 34*(4), 1416-1427.

Smola, A., & Schölkopf. B. (2004). A tutorial on support vector regression. *Statistics and Computing, 14*, 199-222.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

Voultsidou, M., Dodel, S. & Herrmann, J., M. (2005). Neural Networks Approach to Clustering of Activity in fMRI Data. *IEEE Transactions on Medical Imaging, 24*, 987–996.

Weiskopf, N., Mathiak, K., Bock, S. W., Scharnowski, F., Veit, R., Grodd, W., Goebel, R., & Birbaumer, N. (2004). Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transactions on Biomedical Engineering, 51* (6), 966-970.