

Exploiting Temporal Information in Functional Magnetic Resonance Imaging Brain Data*

Lei Zhang^{1,2}, Dimitris Samaras¹, Dardo Tomasi², Nelly Alia-Klein²,
Lisa Cottone², Andreana Leskovjan², Nora Volkow², and Rita Goldstein²

¹ Department of Computer Science, SUNY at Stony Brook, NY, USA

² Medical Department, Brookhaven National Laboratory, NY, USA

Abstract. Functional Magnetic Resonance Imaging (fMRI) has enabled scientists to look into the active human brain, leading to a flood of new data, thus encouraging the development of new data analysis methods. In this paper, we contribute a comprehensive framework for spatial and temporal exploration of fMRI data, and apply it to a challenging case study: separating drug addicted subjects from healthy non-drug-using controls. To our knowledge, this is the first time that learning on fMRI data is performed explicitly on temporal information for classification in such applications. Experimental results demonstrate that, by selecting discriminative features, group classification can be successfully performed on our case study although training data are exceptionally high dimensional, sparse and noisy fMRI sequences. The classification performance can be significantly improved by incorporating temporal information into machine learning. Both statistical and neuroscientific validation of the method's generalization ability are provided. We demonstrate that incorporation of computer science principles into functional neuroimaging clinical studies, facilitates deduction about the behavioral probes from the brain activation data, thus providing a valid tool that incorporates objective brain imaging data into clinical classification of psychopathologies and identification of genetic vulnerabilities.

1 Introduction

Functional Magnetic Resonance Imaging (fMRI) has enabled scientists to look into the active human brain by providing sequences of 3D brain images. This has revealed exciting insights into the spatial and temporal changes underlying a broad range of brain functions, including basic functions such as how we see, feel,

* We thank Steve Berry, B.A., for help with preliminary data analyses, F. Telang, E.C. Caparelli, L. Chang, T. Ernst and N.K. Squires for helpful discussions; This study was supported by grants from the National Institute on Drug Abuse (to NDV: DA06891-06; and to RZG: 1K23 DA15517-01), Laboratory Directed Research and Development from U.S. Department of Energy (OBER), NARSAD Young Investigator Award, SB/BNL seed grant (79/1025459), National Institute on Alcohol Abuse and Alcoholism (AA/ODO9481-04), ONDCP, and General Clinical Research Center (5-MO1-RR-10710).

move, understand. Concomitantly, this new instrumentation has led to a flood of new data: a twenty-minute fMRI session with a single human subject produces a series of 3D brain images each containing approximately 150,000 voxels, collected once a second (or two), yielding tens of millions of data observations. Thus, developing appropriate data analysis methods is needed for truly comprehensive exploration of this ample volume of data. We suggest that through incorporation of computer data analysis principles into functional neuroimaging studies we will be able to identify unique patterns of variability in brain states and deduce about the behavioral probes from the brain activation data (in contrast to the reverse: deducing about brain activation data from behavioral probes). We further propose that this interscientific incorporation may provide a valid tool where objective brain imaging data are used for clinical purpose of classification of psychopathologies and identification of genetic vulnerabilities.

Functional Magnetic Resonance Imaging (fMRI) [1][2] is based on the increase in blood flow to the local vasculature that accompanies neural activity in the brain, so that human cortical functions can be observed without the use of exogenous contrast agents. To date, the analyses and interpretation of fMRI data that are most commonly employed by neuroscientists depend on the cognitive-behavioral probes that are developed to tap regional brain function. Thus, brain responses are a-priori labeled based on the putative underlying task condition (e.g., regions involved in reward vs. regions involved in punishment) and are then used to separate a priori defined groups of subjects. A variety of machine learning methods have also been used for exploratory analysis of fMRI data[3][4][5]. In recent research[6][7][8][9], machine learning methods have been applied for fMRI data analysis but only statistical maps [8][9] or the mean of each fMRI time interval[7] are used while temporal information has yet to be fully employed. Discarding temporal information results in more manageable data sizes; however, the cost of such information loss is still unclear.

In this paper, we consider a different classification problem: separating different groups of human subjects based on the observed fMRI time sequences. We contribute a comprehensive framework of spatially and temporally exploring fMRI data, and apply it to a challenging case study: separating drug addicted subjects from healthy non-drug-using controls based on their observed fMRI time sequences. This learning problem is challenging for a number of reasons: 1) oversized dimensionality of the fMRI BOLD sequences; 2) undersized data space; 3) increased inter-subject variability and intra-subject variability: even for the same person, activations are different from trial to trial due to brain-behavior complexity; and 4) decreased between group experimental heterogeneity. Fig. 1 shows time sequences of one voxel in three human brains.

To our knowledge, this is the first time that machine learning is performed on the temporal information of fMRI data for classification purposes. In this work, we aim to answer the following questions: 1) given the difficulties inherent in this type of the data, what are the most discriminative features for this classification problem? 2) Will temporal/function information help us with classification? We explore fMRI data in two different representations: 3D brain

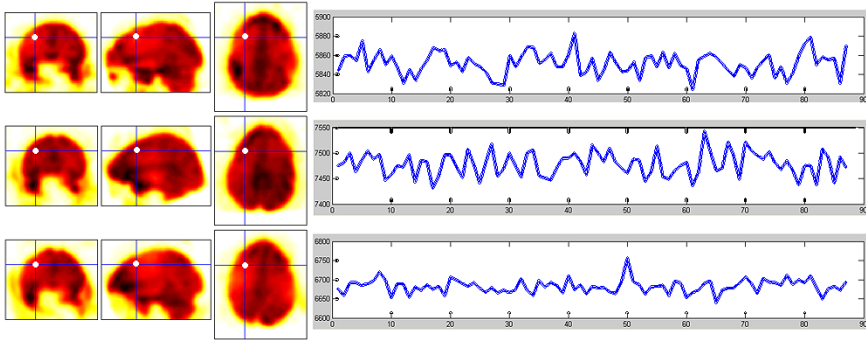


Fig. 1. In each row, the left three images show three slides of the 3D fMRI scan in different views with a white point representing the location of one voxel, followed by a time sequence of this voxel across the 87 task-sequence time points

“video” and a set of voxel-specific time series. Dimensionality reduction and feature selection techniques are proposed and examined in each representation. Experimental results demonstrate that group classification is improved by selecting discriminative features and incorporating fMRI temporal information into a machine learning framework. The automatically selected features prove reliable, i.e., stable in cross-task validation. Furthermore, these blindly selected features prove valid, i.e., they include voxels in regions previously implicated in the successful performance of the core cognitive-behavioral task. Thus, our approach is generalizable, as tested statistically and integrated in a wider neuroscience context. For example, the anterior cingulate gyrus has been assumed to be essential for error detection, conflict resolution, and in emotional regulation. The emergence of anterior cingulate voxels, blindly selected through our analyses, confirms its their central role in underlying the core characteristics of addiction and its relevance to performing this specific sustained attention task.

Drug addiction is a complex disorder characterized by compromised inhibitory control. Individuals with compromised mechanisms of control are difficult to identify unless they are directly subjected to challenging conditions. Solving this problem is essential because patterns of variability in brain states may be unique to a certain psychopathology and could be therefore used for improving diagnosis and prevention efforts (e.g. diagnosis of drug addiction, prevention of relapse or craving). In addition, the development of this “*clinical machine learning framework*” can be applied to further our understanding of other human disorders and states such as those impacting insight and awareness, that similarly to drug addiction are currently identified based mostly on subjective criteria and self-report.

2 Methodology

Acquisition of fMRI data: In our experiments, data were collected to study the neuropsychological problem of loss of sensitivity to the relative value of

money in cocaine users[10]. MRI studies were performed on a 4T Varian scanner and all stimuli were presented using LCD-goggles connected to a PC. The human participants pressed a button or refrained from pressing based on a picture shown to them. They received a monetary reward if they performed correctly. Specifically, three runs were repeated twice (T1, T2, T3; and T1R, T2R, T3R) and in each run, there were three monetary conditions (high money, low money, no money) and a baseline condition where a fixation cross was shown on the screen; the order of monetary conditions was pseudo-randomized and identical for all participants. Participants were informed about the monetary condition by a 3-sec instruction slide, which visually presented the stimuli: or \$0.45, \$0.01 or \$0.00. The feedback for correct responses in each condition consisted of the respective numeral designating the amount of money the subject has earned if correct or the symbol (X) otherwise. To simulate real-life motivational salience, subjects could gain up to \$50 depending on their performance on this task. 16 cocaine dependent individuals, 18-55 years of age, in good health, were matched with 12 non-drug-using controls on sex, race, education and general intellectual functioning.

In this work, we use Statistical Parametric Mapping (SPM)[11] to preprocess (realignment, normalization/registration and smoothing) the fMRI sequences.

Methodology: In this paper, we aim to separate the drug-addicted subjects from controls by applying machine learning methods to observed fMRI sequences. Following [7], the classification problem can be represented: $f : \langle fMRI - sequence(t1, t2) \rangle \rightarrow [DrugAddicted|Control]$ where fMRI-sequence($t1, t2$) is the sequence of fMRI images collected during the contiguous time interval $[t1, t2]$. The input data is an extremely high dimensional feature vector, consisting of hundreds of thousands of features (87 scans per sequence and $53 \times 63 \times 46$ voxels per scan). In order to evaluate the discriminative power of the extracted features, we first perform classification using a simple Euclidean method that uses the Euclidean distance on the original fMRI sequences. This is computationally expensive due to the high dimensionality and yields inferior results due to the large number of non-discriminative features. Feature selection is essential to achieve accurate classification when only a small number of data are available [12], hence we explore a variety of approaches for dimensionality reduction and feature selection. We group our approaches into two categories based on the different views of the fMRI sequences:

2.1 fMRI: 3D Brain “Video”

The 4D input fMRI data ($53 \times 63 \times 46 \times 87$) can be treated as sequences of 3D images. Previous work[13] demonstrated that, by traversing the 3D space of fMRI images using the Hilbert space-filling curve[14], a 3D image can be linearly mapped into 1D space(153594×1). A space filling curve defines a continuous path in a multidimensional grid, visiting each point exactly once and never crossing itself. Hilbert space-filling curve has been proven optimal in preserving the locality and clustering properties of data[15]. The Hilbert space-filling proceeds

in a recursive manner, following the same rotation and reflection pattern at each vertex of the basic curve (details in [16]). By using this technique, the original input fMRI sequences can be mapped into 2D spaces (153594×87) where each 3D fMRI scan is linearly mapped into 1D space. Hence, two groups of dimensionality reduction and feature selection techniques applied in the signal analysis domain can be applied: those based on global information, such as Singular Value Decomposition (SVD)[17]; and those based on local information, such as the Discrete Fourier Transform (DFT)[18]. Due to the dimensionality of the input data, we chose to use local dimensionality reduction techniques.

Dimensionality Reduction and Feature Selection: After mapping each of the 3D scans into 1D space, the input fMRI sequences are transformed into 2D matrices with one spatial dimension and one temporal dimension.

We applied two sets of signal analysis approaches for dimensionality reduction:

1. *1D Transformation:* by performing 1D DFT or DWT on the spatial dimension, the original input data can be represented by a small number of time series.
2. *2D Transformation:* by performing 2D DFT or DWT on the 2D matrices directly, the original input data can be represented by a small 2D feature matrix.

After dimensionality reduction, we can perform classification either by: i) employing the whole reduced feature space for classification or, ii) learning the most discriminative features for classification. As expected, experimental results in Sec. 3 demonstrate the need for learning discriminative features.

2.2 fMRI: A Set of Voxel-Specific Time Series

In Sec. 2.1, we treat each fMRI sequence as a time series of 3D images. Alternatively, each fMRI sequence can also be thought of as a combination of N time series where N is the number of the voxels. In this section, we examine voxel-based feature selection methods for group classification. The most common approach for feature selection[19] is to select those features that best discriminate the target classes: given the goal of learning a target classification function, one common approach to feature selection is to rank the features by their mutual information with respect to the class variable, then to select the n highest scoring features. In our experiments, the information of each voxel is represented by a time sequence of intensities representing BOLD brain activations. Distance (dissimilarity) between voxel intensities is straightforward to define. However, there are different ways to measure this distance when the features are voxel time series. We examine various time series analysis techniques. More specifically, given training data Q and C , for the i th voxel, we propose to compute and examine the distance between two time series $q_{1..T}$ and $c_{1..T}$ by:

1. *Raw_Dist:* Euclidean Distance Metric directly on the two time series of the voxel: $D(q, c) = \sum_{t=1}^T (q_t - c_t)^2$;
2. *Norm_Dist:* Euclidean Distance Metric on normalized time series, where $X' = (X - \bar{X})/\sigma(X)$ with \bar{X} and $\sigma(X)$ the mean and standard deviation;

- 3. *Mean_Dist*: Distance of the mean values of the two series: $D(q, c) = \|\bar{q} - \bar{c}\|$;
- 4. *Var_Dist*: Distance of the variances of the two series: $D(q, c) = \|\sigma^2(q) - \sigma^2(c)\|$;

In this paper, we choose to use a simple kth-Nearest Neighbor(KNN)[19] retrieval scheme to evaluate classification using different sets of selected features.

3 Experiments and Results

In our data collection, there are totally 6 runs: T1, T2, T3, T1R, T2R and T3R with the three latter repeating the three former, grouped into 3 data sets *T1, T2, T3*. The first set of experiments is to evaluate classification performance in a "leave-one-out" cross validation procedure. Each of the *K* human subjects was used as a test subject while training on the remaining *K* - 1 subjects, and the mean classification rate of the *K* experiments is reported.

For comparison purposes, in Table 1, we report the classification performance of the plain Euclidean method(Sec.2) which uses the computationally expensive Euclidean distance on the original fMRI sequences and yields classification results that are close to random. Table 1 also presents the classification performance with the first 30 Fourier coefficients by applying DFT to the mapped matrices using Hilbert space-filling curve. The results of DWT(double the number of features) are similar to the results reported in Table 1. After dimensionality reduction, most of the features are still non-discriminative and a feature selection step can significantly improve classification although standard global dimensionality reduction techniques such as SVD, which also perform feature selection, cannot be applied.

Table 2 reports classification rates for voxel-specific time series analysis in each data set individually. It also reports cross validation using the features selected from voxel-based methods. Contrast map based classification rates in [9] are listed in the last row for comparison since the creation of contrast maps can be thought of a dimensionality reduction process using statistical inferences along

Table 1. 3D Brain "Video" Analysis: classification rates of the simple Euclidean method on the original fMRI sequences are close to random due to the large number of non-discriminative features. Classification using features computed from signal-based methods (DFT). The "ALL" column shows classification rates using all the first 30 Fourier coefficients and the "Discrim" column reports rates using 5 selected discriminative coefficients only. Experimental results demonstrate that in the locally reduced feature space, appropriate features can significantly improve classification.

| | Euclidean Distance | | 1D DFT | | 2D DFT | |
|----|--------------------|--------|--------|---------|--------|---------|
| | 1NN | 5NN | ALL | Discrim | ALL | Discrim |
| T1 | 51.92% | 53.84% | 65.38% | 84.62% | 63.46% | 84.62% |
| T2 | 51.92% | 51.92% | 61.54% | 80.77% | 63.46% | 82.69% |
| T3 | 50% | 52.08% | 60.42% | 72.92% | 62.50% | 68.75% |

Table 2. Voxel-Specific Time Series Analysis: classification rates using features selected from voxel-based methods. Experimental results in each individual data set demonstrate that classification performance can be significantly improved by incorporating temporal information into learning. Classification rates of using Contrast maps *Contrasts* are the best reported in [9] that use the kNN classifier. Experimental results across data sets demonstrate that the features selected using temporal information (*Raw_Dist*) stably perform better across data sets.

| Training Testing | T1 | | | T2 | | | T3 | | |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | T1 | T2 | T3 | T1 | T2 | T3 | T1 | T2 | T3 |
| <i>Raw_Dist</i> | 94.2% | 90.4% | 87.5% | 92.3% | 96.2% | 91.7% | 86.5% | 88.5% | 93.8% |
| <i>Norm_Dist</i> | 84.6% | 75.0% | 70.8% | 78.9% | 90.4% | 72.9% | 71.2% | 75.0% | 85.4% |
| <i>Mean_Dist</i> | 88.5% | 76.9% | 70.8% | 80.8% | 90.4% | 75.0% | 75.0% | 76.9% | 85.4% |
| <i>Var_Dist</i> | 86.5% | 82.7% | 72.9% | 80.8% | 86.5% | 75.0% | 78.9% | 78.9% | 83.3% |
| <i>Contrasts</i> | 88.5% | N/A | N/A | N/A | 86.8% | N/A | N/A | N/A | 85.7% |

the temporal axis. Experimental results on each individual data set demonstrate that classification performance can be significantly improved by incorporating temporal information into learning. Classification rates of *Raw_Dist* outperformed other methods. In time series analysis, temporal distortion is a common problem. Dynamic Time Warping (DTW) has been successfully applied in many time series analysis problems [20] but is inapplicable here due to dimensionality. Luckily the nature of the data prevents significant temporal distortion (since each 87 point sequence can be divided into 6 short sub-sequences). For validation purposes, we performed a set of DTW experiments on 5 selected features. Experimental results demonstrate that DTW gives similar classification results as the Euclidean distance method.

In the three data sets, subjects are performing exactly the same task whereas the sequence of the monetary conditions is different. We address the generalization question by examining classification performance when selected features from one data set are applied onto other data sets. Table 2 reports the cross validation results. The features that are automatically selected using temporal information (*Raw_Dist*), prove very stable in cross-task validation. Hence, by selecting the most discriminative features using temporal information, group classification can be successfully performed.

We also examined the selected features under a neuro-scientific context: the selected most discriminative voxels cluster into two prefrontal brain regions: the middle frontal gyrus (dorsolateral prefrontal cortex) and the anterior cingulate gyrus (ACG). Those regions are known to be involved in sustained attention/working memory and in the processing of salient stimuli/inhibitory control, respectively. It is very intriguing that these blindly selected voxels represent the two core functions of the delayed forced-choice task used in this study. Further, their location in the prefrontal cortex lends support to our working hypothesis that this region is crucially involved in the underlying core characteristics of drug addiction. Specifically, the involvement of the ACG, a region which has been previously implicated in drug intoxication and craving [21], lends support

to the dysfunction in drug addiction of this corticolimbic region (see [22] where a hypofunctionality of the ACG to a GO/NO-GO task was reported in cocaine users compared to controls).

4 Conclusions and Future Work

We have demonstrated that, by selecting discriminative features, group classification can be successfully performed on a challenging case study although training data are fMRI sequences that are exceptionally high dimensional, sparse and noisy. We have also shown that classification rates can be significantly improved by incorporating temporal information into machine learning analysis of such data. To our knowledge, this is the first time that the temporal/functional information of the fMRI data is explicitly explored for machine learning classification purposes. This comprehensive framework of exploring spatial and temporal information of fMRI data for classification problems can be extended to many other fMRI analysis applications. Our analyses provide an additional method for validation of a regional-functionality specificity, however external validation using a lesioned sample would still be necessary to confirm a particular region's role in a specific function (i.e., loss of a certain function in individuals lacking a specific region).

Since feature selection is the key for pattern recognition problems, especially when only a small number of data are available, as in most human subject research, one of our future research directions is to explore efficient global dimensionality reduction techniques[23] that can be applied on extremely high dimensional training data and examine more sophisticated classifiers. Another future research direction is to apply grouping/clustering as a preprocessing step to reduce dimensionality of the raw data and smooth noise. Finally, connectivity and interactivity information play important roles in brain activation patterns. For example, in the voxel-based feature selection part, voxels are considered to be independent, however, this is not accurate for the human brain. After further validation with other data sets (additional subjects with addiction or other psychopathology), we aim to explore the connectivity and interactivity between voxels to reveal more discriminative brain activation patterns.

References

1. Kwonget et al., K.: Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA* (1992) 5675–5679
2. Bandettini, P., Jesmanowicz, A., Wong, E., Hyde, J.: Processing strategies for time-course data sets in functional mri of the human brain. (*Magn Reson Med*)
3. Goutte, C., e.a.: On clustering fmri time series. (Technical Report IMM-REP-1998-11)
4. Penny, W.: Mixture models with adaptive spatial priors. In: *Concepts and Methods in NeuroImaging workshop, NIPS*. (2001)
5. LaConte, S., Strother, S., Cherkassky, V., Hu, X.: Predicting motor tasks in fmri data using support vector machines. In: *ISMRM*. (2003)

6. Wang, X., Hutchinson, R., Mitchell, T.: Training fmri classifiers to detect cognitive states across multiple human subjects. In: NIPS03. (2003)
7. Mitchell, T., Hutchinson, R., Niculescu, R., Pereira, F., Wang, X., Just, M., Newman, S.: Learning to decode cognitive states from brain images. (Machine Learning)
8. Ford, J., Farid, H., Makedon, F., Flashman, L., McAllister, T., Megalooikonomou, V., Saykin, A.: Patient classification of fmri activation maps. In: MICCAI. (2003)
9. Zhang, L., Samaras, D., Tomasi, D., Volkow, N., Goldstein, R.: Machine learning for clinical diagnosis from functional magnetic resonance imaging. In: CVPR. (2005)
10. Goldstein et al., R.: A modified role for the orbitofrontal cortex in attribution of salience to monetary reward in cocaine addiction: an fmri study at 4t. In: Human Brain Mapping Conference. (2004)
11. Friston, K., Holmes, A., Worsley, K., et al.: Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* (1995) 2:189–210
12. Levi, K., Weiss, Y.: Learning object detection from a small number of examples: The importance of good features. In: CVPR. (2004)
13. Wang, Q., Kontos, D., Li, G., Megalooikonomou, V.: Application of time series techniques to data mining and analysis of spatial patterns in 3d images. In: ICASSP. (2004) 525–528
14. Hilbert, D.: Uber die stetige abbildung einer linie auf flachenstuck. *Math. Annln.* (1891) 459–460
15. Moon, B., Jagadish, H., Faloutsos, C., Saltz, J.: Analysis of the clustering properties of the hilbert space-filling curve. *IEEE Transactions on Knowledge and Data Engineering* (2001) 124–141
16. J.G.Griffiths: An algorithm for displaying a class of space-filling curves. *Software Practice and Experience* **16** (1986) 403–411
17. Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S.: Dimensionality reduction for fast similarity search in large time series databases. *Journal of Knowledge and Information Systems* (2000)
18. Agrawal, R., Faloutsos, C., Swami, A.: Efficient similarity search in sequence databases. In: Proc. of the 4thConference on Foundations of Data Organization and Algorithms. (1993)
19. Mitchell, T.: *Machine Learning*. McGraw-Hill (1997)
20. Keogh, E.: Data mining and machine learning in time series databases. In: Tutorial in ICML. (2004)
21. Goldstein, R., Volkow, N.: Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. (*American Journal of Psychiatry*) 1642–1652
22. J.N.Kaufman, T.J.Ross, E.A.Stein, H.Garavan: Cingulate hypoactivity in cocaine users during a go-nogo task as revealed by event-related functional magnetic resonance imaging. *The Journal of Neuroscience* (2003)
23. Ye, J.: Generalized low rank approximations of matrices. In: ICML. (2004) 887–894