2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Information Fusion for Cocaine Dependence Recognition using fMRI

Fabio A. Faria, Fábio A. Cappabianco

Group of Innovation Based on
Images and Signals
Federal University of São Paulo
S.J. dos Campos, SP 12231, Brazil
Email: {ffaria,cappabianco}@unifesp.br

Chiang-shan R. Li, Jaime S. Ide

Department of Psychiatry
Yale University School of Medicine
New Haven, CT 06519, USA
Email: {chiang-shan.li, jaime.ide}@yale.edu

***Abstract*—Cocaine dependence devastates millions of human lives. Despite of a variety of treatments, there is a very high rate of individual relapse to drug use. In the last decade, functional magnetic resonance imaging (fMRI) proved to be a powerful tool to diagnose and understand different pathologies. This work provides advances in the identification of cocaine dependence and in the relapse prediction based on fMRI classification. We improve the traditional methodology of the literature called multi-voxel pattern analysis (MVPA), which is used for feature extraction and classification. In addition, we propose new features that use specific functional connectivity measures. An extensive evaluation was conducted comparing our methodology with MVPA, as well as, several learning methods with distinct feature sets. We could identify the neural patterns that lead to improve classification accuracies and evaluate the advantages of employing an information fusion approach through an ensemble of classifiers. Experimental results show an improvement of final accuracy over the state-of-the-art methods.**

I. Introduction

Identifying neural phenotypes of cocaine dependence and relapse is of extreme importance to public health. According to the 2013 National Survey on Drug Use and Health, approximately 1.5 million Americans are currently addicted to cocaine. Despite treatment, individuals relapse to drug use at a very high rate, leading to prolonged, devastating impacts on human lives and the society. Understanding why cocaine addicts are unable to refrain from drug use is of critical importance in addiction neuroscience. For that purpose, the study of brain functions by means of functional magnetic resonance imaging (fMRI) has become increasingly popular during the last decade.

Two main approaches have being proposed for the analysis of fMRI data of the brain: the standard mass-univariate approach known as general linear modeling (GLM) [3] and the multi-voxel pattern analysis (MVPA) using pattern recognition techniques [7], [6]. GLM requires a priori task-design and is limited to observables or at least measurable task conditions [4]. On the other hand, MVPA has the ability to delineate complex associations between multiple voxel signals, stimuli, or mental states in a data-driven way. Unlike most previous work, our focus consists in predicting neuropathology, that is, identifing cocaine dependence and predicting relapse. This is a challenging task because of the commonly large inter-subject variability and small sample size of the data set [8].

Our experiments use a fMRI data set with seventy-five cocaine dependent and eighty-eight healthy individuals matched in age and gender. Imaging data were collected while subjects performed a cognitive paradigm known as the stop-signal task [16], [15]. More specifically, we examine the brain responses to conflict anticipation as altered in addiction [10], [9].

From each fMRI image, we extract mean-based patterns of brain responses as feature set, just as is commonly used by univariate GLM analyses and extant MVPA applications [7], [19], [6]. However, we also include a measure of signal complexity called power-spectrum scale invariance (PSSI) [18], [11] as well as functional connectivity [12], [22], since individual variability cannot be fully captured without considering more elaborate measures such as signal complexity and connectivity [21]. Finally, we classify samples using both single classifiers and a fusion method based on diversity measures [2]. The results are compared with PRONTO, a popular MVPA package [19].

In summary, we add to traditional MVPA methodology: (i) by using specific functional connectivity measures as new features; (ii) evaluating learning methods with distinct feature sets; and (iii) employing classifier selection and fusion for cocaine dependence recognition.

II. Cocaine dependence, cognitive control and Bayesian updating

Cognitive control is a critical executive function, defined as the ability to withhold or modify actions in response to a dynamically changing environment. With a variety of laboratory paradigms, numerous studies have characterized deficits in cognitive control in chronic cocaine users [5]. For example, in the stop signal task (SST) where ones respond quickly to an imperative go stimulus and must withhold response as instructed by an occasional stop signal, cocaine dependent individuals (CD) demonstrated diminished response inhibition and error processing, as compared to demographically matched healthy controls (HC) [17].

978-1-5090-4846-5/16/$31.00 ©2016 IEEE

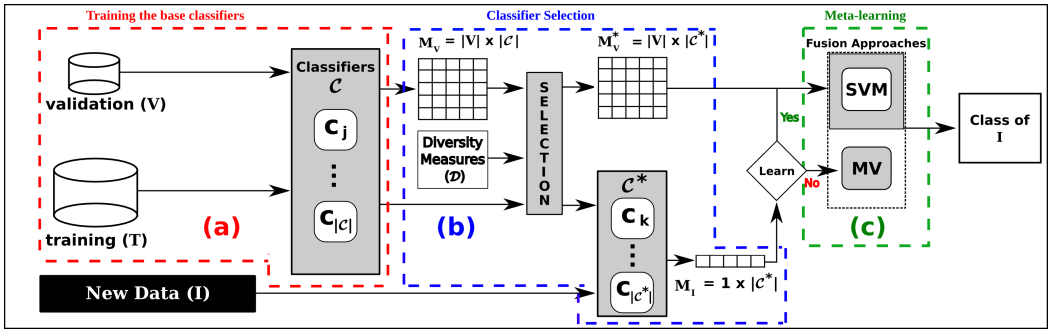


Fig. 1. Classifier selection and fusion framework is composed of three steps: (a) Training the classifier set $\mathcal{C}$; (b) Classifier selection approach reduces the amount of classifiers from $\mathcal{C}$ to $\mathcal{C}^*$; and (c) Meta-learning approach which combines the $\mathcal{C}^*$ classifier outcomes to achieve the final class [2].

Here, we employ a Bayes-optimal decision-making model to capture the adaptive nature of cognitive control in the SST, in which participants learn to anticipate an infrequent stop signal through trial by trial learning [20]. We hypothesize that participants choose to go or stop based on accumulating sensory evidence within a trial, as well as prior belief about the likelihood of a stop trial prior to stimulus onset. This rational strategy explains classic stopping behavior. In particular, by augmenting this decision-making model with trial-by-trial learning, we were able to account for the sequential effect in the SST [13]; that is, go reaction time slows down after a run of stop trials and speeds up after a preponderance of go trials. Further, we have shown that a core cognitive control deficit in cocaine addiction is impairment in learning from and adapting to changes in contextual information [10].

III. Classifier Selection and Fusion Framework

In order to improve the efficacy of classifiers, we employ a classifier fusion approach, building on a diversity measure as derived from the agreements/disagreements among distinct classifiers. For this purpose, a successful classifier selection and fusion framework originally proposed for multimedia recognition [2] is adopted. In this paper, a tuple containing a single learning method (e.g., k-Nearest Neighbors – kNN) and a description technique (e.g., a color channel from the RGB color channels) defines a classifier.

A. Classifier Fusion

Fig. 1 presents the framework as proposed in [2] for classifier selection and fusion. The classifiers first learn to identify patterns from a training set (T) of samples. Afterwards, classification models ($|\mathcal{C}|$) are generated and applied to a validation data set (V). The result of this process is a matrix M_V composed of classifiers outcomes, which is defined as $|M_V| = |V| \times |\mathcal{C}|$, where $|V|$ is the number of examples from a validation set V (see Fig. 1-(a)).

Then, diversity measures ($\mathcal{D}$) are computed using the contents of M_V, reflecting the degree of agreement and disagreement of all $|\mathcal{C}|$ available classifiers [14]. Based on $\mathcal{D}$, it is possible to determine the most appropriate classifiers that should be combined.

The objective of this process is to combine the set of most suitable classifiers in terms of the diversity and accuracy, given by $\mathcal{C}^* \subset \mathcal{C}$, as shown in Fig. 1 - (b)). During this step a new matrix $M_V^* \subset M_V$ is computed.

Given new data I, it is classified through a meta-learning approach embodied by a fusion technique (e.g., Support Vector Machines) over a new matrix M_V^* (Fig. 1 - (c)). Next section describes the selection process using diversity measures in more details.

B. Classifier Selection

The methodology to select classifiers on the basis of diversity measures as proposed in [2] is illustrated in Fig. 2. It comprises the following steps:

(a) Quantify the agreement among available classifiers in $\mathcal{C}$ by means of the set of diversity measures $\mathcal{D}$. It takes into account the computed values of M_V. Five different measures have been used here (*Correlation Coefficient* p, *Double-Fault Measure*, *Disagreement Measure*, *Interrater Agreement* k, and *Q-Statistic* [14]);
(b) Sort the pairs of classifiers according to their diversity score. Because distinct diversity metrics are used, this step generates a set of ranked lists $\mathcal{R}$;
(c) Compose another set of ranked lists $\mathcal{R}^t$ from the top t pairs of classifiers from each ranked list in $\mathcal{R}$;
(d) Generate a histogram $\mathcal{H}$ counting the occurrences of a classifier in all ranked lists of $\mathcal{R}^t$;
(e) Combine classifiers that are among the most frequent in $\mathcal{H}$, of which the accuracy is greater than a given threshold $\mathcal{T}$ in a fusion approach. $\mathcal{T}$ is given by the average accuracy among all classifiers used in the validation set V.

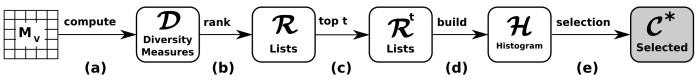


Fig. 2. The five steps for classifier selection are: (a) Computation of diversity measures; (b) Ranking of pairs of classifiers by their diversity measures scores; (c) Selection of the top t; (d) Computation of a histogram $\mathcal{H}$; (e) Select the most appropriate classifiers $|\mathcal{C}^*|$.

IV. Material and Methods

A. Subjects and behavioral task

Seventy-five cocaine dependent (CD) and eighty-eight healthy controls (HC) matched in age and gender participated in this study. A complete description of the experiment can be found in [11]. CD were recruited from the local, greater New Haven area in a prospective study [17] and met criteria for current cocaine dependence, as diagnosed by the Structured Clinical Interview for DSM-IV. They were drug-free while staying in an inpatient treatment unit The Human Investigation committee at Yale University School of Medicine approved the study, and all subjects signed an informed consent prior to participation.

We employed a simple reaction time task in this stop-signal paradigm [16], [15] (Fig. 3). There were two trial types: 'Go' (75%) and 'Stop' (25%), randomly intermixed, with an inter-trial interval of 2s. A small fixation dot appeared on the screen to engage attention at the beginning of the trials. After a randomized fore-period (FP) between 1 and 5 s, the dot turned into a circle (the go signal), prompting the subjects to quickly press a button. The circle disappeared at a button press or after 1 s had elapsed, and the trial terminated. The time between go signal and the button press is known as reaction time (RT). In a Stop trial, an additional X, the stop signal, replaced the go signal and instructed subjects to withhold their response. Failure to withhold the go response for the 1 s constituted a stop error (SE), and stop success (SS) otherwise. The stop signal reaction time (SSRT) is an index of motor response inhibition performance, and was estimated by subtracting the critical stop signal delay (the time interval between go and stop signals) from the median of Go trials RT.

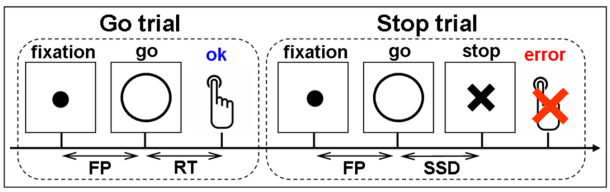


Fig. 3. Trial-types in the stop-signal task (SST) paradigm.

B. fMRI data acquisition and pre-processing

FMRI data were collected with 3T Siemens Trio scanner [15]. Each scan comprised four 10-min runs of the SST.

Functional blood oxygen level dependent (BOLD) signals were acquired with a single-shot gradient echo echo-planar imaging (EPI) sequence, with 32 axial slices parallel to the AC-PC line covering the whole brain [15]: TR=2000 ms, TE=25 ms, bandwidth=2004 Hz/pixel, flip angle=85, FOV=220×220 mm^2, matrix=66×64, slice thickness=4 mm and no gap. A high-resolution 3D structural image (MPRAGE; 1 mm resolution) was also obtained for anatomical co-registration. Three hundred images were acquired in each session.

Functional MRI data was pre-processed with standard pipeline using Statistical Parametric Mapping 12 (SPM12) (Wellcome Department of Imaging Neuroscience, University College London, U.K.).

C. Feature extraction

1) **Brain activation:** We constructed GLM's and localized brain regions responding to conflict anticipation (encoded by the posterior probability P(stop)) at the group level, as in our previous works [13], [9]. In short, four main conditions were defined according to trial type and outcome: go success, go error, stop success and stop error. P(stop), Go RT, and SSD were entered as parametric modulators. In the group level analysis including both CD and HC, the regions responding to P(stop) comprised the bilateral parietal cortex, the inferior frontal gyrus (IFG) and the right middle frontal gyrus (MFG) with peak MNI coordinates at [39,53,-1] and [42,23,38], respectively, in mm; and regions of motor slowing: bilateral insula ([-33, 17, 8] and [30, 20, 2]), the left precentral cortex (L.PC) ([-36,-13,56]), and the supplementary motor area (SMA) ([-9,-1,50]) (Fig. 4), consistent with our previous studies [13], [9]. These regions of interest (ROIs) were used as masks to extract the activation feature set, which included 1005 voxels.

2) **Brain complexity:** Dynamically dysregulated circuit can be characterized in a computationally efficient way using nonlinear complexity measures, such as PSSI. We showed significant complexity differences in frontoparietal networks between CD and HC, which may provide useful information in relapse prediction [11]. Using optimized methods [18], we estimated the PSSI-β from each FFT-transformed time series S(f) by linear fit of the decaying slope, within the frequency range 0.01-0.25 Hz. The whole-brain PSSI maps were masked using the same ROIs (Fig. 4).

3) **Functional connectivity:** We analyzed the frontoparietal circuit involved in Bayesian predictions and motor consequences using a standard temporal correlation analysis [22] and multivariate Granger causality analysis or mGCA [12]. In Fig. 5, we illustrate fifteen correlation coefficients derived from the six ROIs comprising the frontoparietal anticipation circuit for each individual CD and HC. Using mGCA, we observed that the connectivities from bilateral parietal to L.PC and SMA were disrupted in CD (Fig. 4). Interestingly, the connection strength from bilateral parietal to the L.PC was negatively correlated with earlier relapse (p=0.0067), and decreased connectivity from MFG to IFG was correlated with decrease in inhibition performance (i.e., prolonged SSRT, p=0.0038; Fig. 4). These findings offer circuit-level evidence of altered cognitive control in cocaine addiction [5].

The extracted features can be summarized as follows:

- **P(stop) 1005**: brain responses to conflict anticipation, comprised of 1005 values within the 6 ROIs.
- **PSSI 1005**: PSSI complexity, comprised of 1005 values within the 6 ROIs.
- **Connectivity measures**: 4 distinct metrics, 15 pairwise Pearson correlations between the fMRI time courses of the 6 ROIs (*cc15*), and mGCA connectivity weights, 30 F-values (*fs30*), 30 degrees of influence (*doi30*) and 30 F_{Geweke} (*fgeweke30*). All mGCA metrics indicated directional influences between ROIs [12].

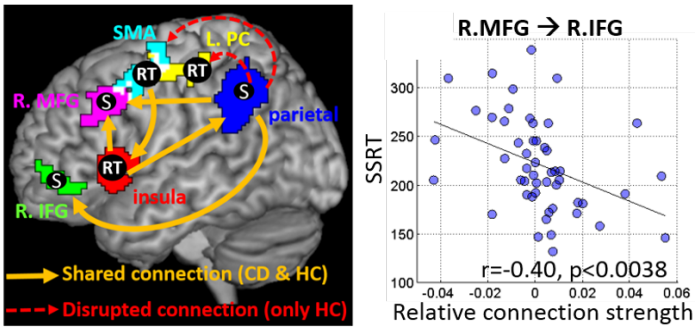


Fig. 4. Disrupted frontoparietal circuit in cocaine addicts. The frontoparietal circuit included six regions responding to Bayesian conflict anticipation ("S") and regions of motor slowing ("RT"). (a) CD and HC shared connections (orange arrows). (b) The relative connection strength or degree of influence (doi) between R.MFG and R.IFG is negatively correlated with SSRT;

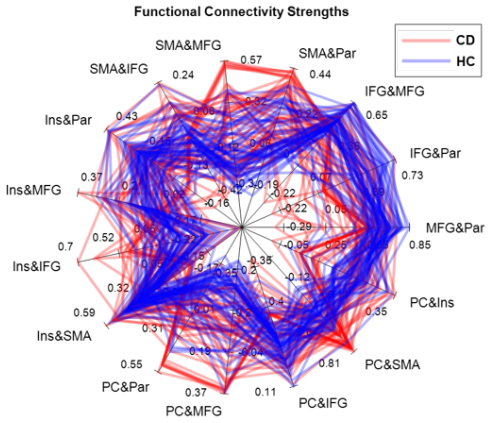


Fig. 5. Connectivity strengths between nodes in the frontoparietal circuit. We represent connectivity strengths between nodes for each individual subject in CD (red line) and HC (blue line) groups.

D. Experimental Methodology

1) Learning Methods: We used seven learning methods in the framework: Naïve Bayes (NB), Decision Tree (DT), Naïve Bayes Tree (NBT), k-Nearest Neighborhood (kNN) using $k = 1$, $k = 3$, and $k = 5$, and Support Vector Machine (SVM). We used the implementation of learning methods as available in the WEKA[1] data mining library. All learning techniques were used with default parameters.

2) Evaluation Measure: In our experiments, we used evaluation measure from the confusion matrix. The balanced accuracy can be calculated according to the expression in Equation 1.

3) Cross-Validation: We used the 10-fold cross-validation protocol and all results are reported in terms of Average Balanced Accuracy.

$$Balanced\ Accuracy = \frac{Specificity + Sensitivity}{2} \qquad (1)$$

[1]http://www.cs.waikato.ac.nz/~ml/weka (As of 04/15/2016).

V. Results

This section shows results from two different experiments. In Section V-A, we compare different learning methods for each extracted feature. In Section V-B, we employ classifiers fusion to explore the use of complementary information provided by each extracted features.

A. Comparative Study: Learning Methods and Features

In this experiment, we used six feature sets derived from three different sources: brain activation features $P(stop)1005$; brain complexity features $PSSI1005$; and four functional connectivity features: $cc15$, $fs30$, $doi30$, and $fgeweke30$. The numeric suffix in front of the feature name indexes the number of dimensions. Additional ten new feature sets were created through combination of any 2 or 3, or all connectivity features ($2-Combined$, $3-Combined$, $4-Combined$ or ALL). All results are reported in terms of average balanced accuracy from the 10-folds cross-validation protocol.

*1) **Cocaine Dependence vs. Healthy Control**:* Table I shows effectiveness results of the seven learning methods for each extracted feature for cocaine dependence dataset. As shown, for single features, the NB and NBT using $cc15$ feature achieved the best results than any other learning method and single feature, with an average balanced accuracy of 70.94% and 69.28%, respectively. Also, the functional connectivity features (single or combined) achieved the best results for all seven learning methods used in this experiments (in green and blue). This suggests that the functional connectivity features best describe the data.

Furthermore, the best result among all classifiers was achieved by kNN1 using $2-Combined$ features ($cc15+doi30$ and $cc15+fS30$) with average balanced accuracy of 71.91%. Note that all best combinations contain $cc15$, the best feature to describe cocaine dependence (blue, Table I).

*2) **Relapse Prediction**:* Since the relapse dataset is composed of unbalanced classes (63 non-relapsors and 13 relapsors), we created 400% synthetic instances to smaller class with the well-known SMOTE approach [1]. Therefore, after dividing the dataset on 10-folds cross validation, we applied SMOTE approach only on the training sets to equalize both classes.

Table II shows effectiveness results achieved with seven learning methods for each extracted feature for relapse prediction. For single features, the kNN1 and kNN5 using $fgeweke$ feature achieved the best results as compared to any other learning method and single feature, with an average balanced accuracy of 69.88% and 76.43%, respectively. Moreover, the latter is the best effectiveness result among all classifiers.

*3) **Comparison with state of the art**:* We executed the same previous experiments with the popular fMRI classification package PRONTO [19], using brain activation and complexity as feature sets. For cocaine dependent dataset, PRONTO achieved best balanced accuracies of 55.57% for the feature set P(stop) 1005, and of 66.46% for the feature set PSSI 1005. For cocaine relapse dataset, PRONTO achieved best balanced accuracies of 49.63% for the feature set P(stop) 1005, and of

TABLE I

EFFECTIVENESS RESULTS ACHIEVED OF SEVEN LEARNING METHODS USING SIXTEEN DIFFERENT DESCRIPTORS FOR COCAINE DEPENDENCE DATASET.

Type	Descriptors	Learning Methods						
		NB	DT	NBT	kNN1	kNN3	kNN5	SVM
Brain Activation	$P(stop)1005$	53.88	48.58	50.67	50.90	45.81	45.36	57.53
Brain Complexity	$PSSI1005$	62.17	53.57	55.41	54.09	64.41	60.50	61.04
Functional connectivity	$cc15$	**70.94**	67.86	**69.28**	64.03	65.94	67.21	69.23
	$doi30$	58.57	62.05	62.07	66.75	63.50	64.21	60.80
	$fgeweke30$	65.99	59.95	60.52	58.37	61.04	63.04	68.70
	$fs30$	65.99	59.95	60.52	58.37	61.04	63.04	68.70
2-Combined	$cc15 + doi30$	66.74	67.43	65.95	69.94	**69.37**	**68.71**	**71.11**
	$cc15 + fgeweke30$	67.19	**70.80**	66.20	**71.91**	68.49	65.96	68.97
	$cc15 + fs30$	67.19	**70.80**	66.20	**71.91**	68.49	65.96	69.04
	$doi30 + fgeweke30$	62.86	59.95	62.52	60.78	59.41	61.38	68.39
	$doi30 + fs30$	62.86	59.95	62.52	60.78	59.41	61.38	67.68
	$fgeweke30 + fs30$	65.99	59.95	55.16	58.37	61.04	63.04	65.37
3-Combined	$cc15 + doi30 + fgeweke30$	63.88	67.60	61.22	67.20	68.28	66.43	71.02
	$cc15 + doi30 + fs30$	63.88	67.60	61.22	67.20	68.28	66.43	71.02
	$doi30 + fgweke30 + fs30$	63.97	60.74	64.19	60.71	61.38	61.23	64.51
4-Combined	ALL	64.60	67.60	64.31	65.97	68.28	67.34	70.08

TABLE II

EFFECTIVENESS RESULTS ACHIEVED OF SEVEN LEARNING METHODS USING SIXTEEN DIFFERENT FEATURES FOR COCAINE RELAPSE DATASET.

Type	Descriptors	Learning Methods						
		NB	DT	NBT	kNN1	kNN3	kNN5	SVM
Brain Activation	$P(stop)1005$	42.74	43.45	55.36	45.83	45.12	53.69	52.86
Brain Complexity	$PSSI1005$	59.40	58.45	53.69	63.57	68.10	66.55	60.24
Functional connectivity	$cc15$	57.02	50.48	46.19	59.40	48.69	52.86	50.71
	$doi30$	58.81	57.02	62.14	68.81	68.93	69.64	60.60
	$fgeweke30$	45.36	51.07	49.76	**69.88**	70.60	**76.43**	67.26
	$fs30$	42.02	44.40	56.19	54.64	62.26	63.10	62.14
2-Combined	$cc15 + doi30$	**61.90**	54.29	**63.81**	60.71	61.43	67.26	61.31
	$cc15 + fgeweke30$	49.29	39.40	52.74	67.26	**75.60**	71.55	61.07
	$cc15 + fs30$	47.74	51.07	57.02	59.64	54.29	58.45	53.57
	$doi30 + fgeweke30$	41.90	44.05	53.69	65.48	67.14	73.81	64.52
	$doi30 + fs30$	49.52	41.31	44.76	57.14	62.14	67.14	66.07
	$fgeweke30 + fs30$	38.57	**62.74**	39.76	62.02	59.52	66.43	60.24
3-Combined	$cc15 + doi30 + fgweke30$	46.07	60.24	57.26	66.31	66.43	75.60	65.24
	$cc15 + doi30 + fs30$	46.90	52.02	58.21	56.31	66.31	64.88	58.81
	$doi30 + fgweke30 + fs30$	38.69	51.19	53.69	60.48	57.86	67.86	**70.24**
4-Combined	ALL	47.62	48.81	55.12	67.14	67.98	69.64	57.02

62.78% for the feature set PSSI 1005. These results have been obtained with PRONTO using SVM, and they are consistent with our SVM implementation results (Tables I and II, SVM).

Together, these findings support the hypothesis that functional connectivity features are best in identifying cocaine dependence.

B. Selection and Classifier Fusion

We adopt a classifier selection and fusion framework (FSVM) to combine the most suitable classifiers using all available functional connectivity features [2].

In this experiment, four different fusion approaches were used: (1) MV is a majority voting technique; (2) FSVM-POLY is a SVM technique with polynomial kernel; (3) FSVM-NORM is a SVM technique with normalized polynomial kernel; (4) FSVM-RBF is a SVM technique with RBF kernel. FSVM refers to the selection and fusion framework, as implemented through SVM in the fusion step (meta-learning).

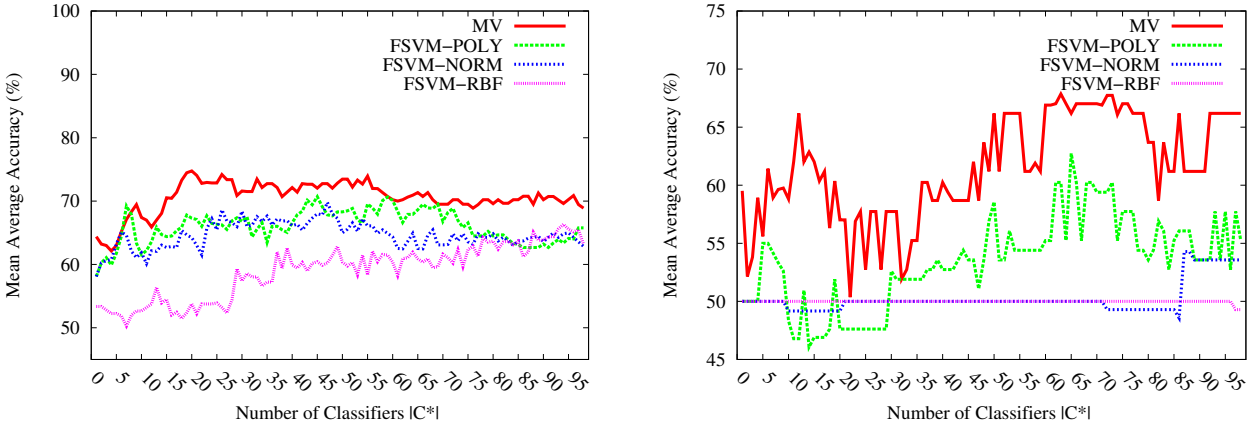
Figs. 6 (a) and (b) show four curves, MV, FSVM-POLY, FSVM-NORM, and FSVM-RBF, describing a behavioral analysis of each fusion approach using different number of classifiers $|C| = \{1, \ldots, 98\}$ (14 types of features $\times$ 7 learning methods $= 98$ classifiers).

As shown in Fig. 6 (a), the MV approach achieved better results in the range of $\{8, \ldots, 98\}$ classifiers. However, FSVM-POLY obtained the best results when less than 8 classifiers were considered. The best result, in identifying CD from HC, was achieved by MV using 20 classifiers with average balanced accuracy of 74.75%.

In Fig. 6 (b), the same behavior seen previously could not be observed for the cocaine relapse dataset. This occurs because of the challenging learning task with very few instances in the training set (small training task). In this scenario, the adopted framework has not produced good results. However, it is a promising approach to be explored in the future as more data are collected.

VI. CONCLUSION

In this work, we proposed the use of specific functional connectivity measures for cocaine dependence recognition task. We showed that connectivity measures achieved better effectiveness results than traditional MVPA features. Furthermore, a comparative study of seven learning methods and sixteen different features was performed. The $2 - Combined$ features ($cc15 + fgeweke30$ and $cc15 + fs30$) obtained the best average balanced accuracy in identifying cocaine dependence with 71.91% against 66.46% by tradicional package PRONTO. In



(a) Cocaine Dependent Dataset. (b) Cocaine Relapse Dataset.

Fig. 6. In (a) are effectiveness results for dependents dataset. In (b) are results for relapse dataset.

relapse prediction, the connectivity feature $fgeweke30$ also achieved better result than PRONTO with 76.43% and 62.78%, respectively. Thus, regional directional connectivity is a critical neural phenotype of cocaine dependence. In addition, the classifier selection and fusion framework has improved classification effectiveness results of cocaine dependence with 74.75% against 66.46% obtained by package PRONTO. However, the same improvement was not observed in relapse prediction. This latter may relate to the imbalance of sample size between classes that hinders the learning of classifiers.

ACKNOWLEDGMENT

This study was supported by NIH grant DA023248 (Li) and CNPq grant #486988/2013-9 (Cappabianco). We thank the support of scientific funding agency, grants #2010/14910-0 and #2016/17064-0, São Paulo Research Foundation (FAPESP).

REFERENCES

[1] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *J. Artif. Int. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[2] F. A. Faria, J. A. dos Santos, A. Rocha, and R. da S. Torres, "A framework for selection and fusion of pattern classifiers in multimedia recognition," *Pattern Recognition Letters*, vol. 39, no. 0, pp. 52 – 64, 2014.

[3] K. J. Friston, A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Human Brain Mapping*, vol. 2, no. 4, pp. 189–210, 1995.

[4] K. J. Friston, "Modalities, modes, and models in functional neuroimaging," *Science*, vol. 326, no. 5951, pp. 399–403, 2009.

[5] H. Garavan and R. Hester, "The role of cognitive control in cocaine dependence," *Neuropsychol Rev*, vol. 17, no. 3, pp. 337–45, 2007.

[6] J. V. Haxby, A. C. Connolly, and J. S. Guntupalli, "Decoding neural representational spaces using multivariate pattern analysis," *Annual Review of Neuroscience*, vol. 37, no. 1, pp. 435–456, 2014.

[7] J. D. Haynes and G. Rees, "Decoding mental states from brain activity in humans," *Nat Rev Neurosci*, vol. 7, no. 7, pp. 523–34, 2006.

[8] J. Honorio, "Classification on brain functional magnetic resonance imaging:: dimensionality, sample size, subject variability and noise," in *Frontiers of Medical Imaging*, C. Chen, Ed. World Scientific Publishing Company Pte Limited, 2014, ch. 14, pp. 266–290.

[9] S. Hu, J. S. Ide, S. Zhang, and C. S. Li, "Anticipating conflict: Neural correlates of a bayesian belief and its motor consequence," *Neuroimage*, vol. 119, pp. 286–95, 2015.

[10] J. S. Ide, S. Hu, S. Zhang, A. J. Yu, and C. S. Li, "Impaired bayesian learning for cognitive control in cocaine dependence," *Drug Alcohol Depend*, 2015.

[11] J. S. Ide, S. Hu, S. Zhang, L. R. Mujica-Parodi, and C. shan R. Li, "Power spectrum scale invariance as a neural marker of cocaine misuse and altered cognitive control," *NeuroImage: Clinical*, vol. 11, pp. 349 – 356, 2016.

[12] J. S. Ide and C. S. Li, "A cerebellar thalamic cortical circuit for error-related cognitive control," *NeuroImage*, vol. 54, no. 1, pp. 455–464, 2011.

[13] J. Ide, P. Shenoy, A. Yu, and C. Li, "Bayesian prediction and evaluation in the anterior cingulate cortex," *J Neurosci*, vol. 33, no. 5, pp. 2039–47, 2013.

[14] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.

[15] C. S. Li, C. Huang, R. T. Constable, and R. Sinha, "Imaging response inhibition in a stop-signal task: neural correlates independent of signal monitoring and post-response processing," *J Neurosci*, vol. 26, no. 1, pp. 186–92, 2006.

[16] G. D. Logan, W. B. Cowan, and K. A. Davis, "On the ability to inhibit simple and choice reaction time responses: a model and a method," *J Exp Psychol Hum Percept Perform*, vol. 10, no. 2, pp. 276–91, 1984.

[17] X. Luo, S. Zhang, S. Hu, S. R. Bednarski, E. Erdman, O. M. Farr, K.-I. Hong, R. Sinha, C. M. Mazure, and C. shan R. Li, "Error processing and gender-shared and -specific neural predictors of relapse in cocaine dependence," *Brain*, vol. 136, no. 4, pp. 1231–44, 2013.

[18] D. Rubin, T. Fekete, and L. R. Mujica-Parodi, "Optimizing complexity measures for fmri data: algorithm, artifact, and sensitivity," *PLoS One*, vol. 8, no. 5, 2013.

[19] J. Schrouff, M. J. Rosa, J. M. Rondina, A. F. Marquand, C. Chu, J. Ashburner, C. Phillips, J. Richiardi, and J. Mourao-Miranda, "Pronto: pattern recognition for neuroimaging toolbox," *Neuroinformatics*, vol. 11, no. 3, pp. 319–37, 2013.

[20] P. Shenoy, R. Rao, and A. Yu, "A rational decision making framework for inhibitory control." in *Advances in Neural Information Processing Systems (NIPS)*, vol. 23. MIT Press, 2011, pp. 2146–2154.

[21] S. M. Smith, T. E. Nichols, D. Vidaurre, A. M. Winkler, T. E. J. Behrens, M. F. Glasser, K. Ugurbil, D. M. Barch, D. C. Van Essen, and K. L. Miller, "A positive-negative mode of population covariation links brain connectivity, demographics and behavior," *Nat Neurosci*, vol. 18, no. 11, pp. 1565–1567, 2015.

[22] S. Zhang, J. S. Ide, and C. S. Li, "Resting-state functional connectivity of the medial superior frontal cortex," *Cereb Cortex*, vol. 22, no. 1, pp. 99–111, 2012.