Classifying Brain States for Cognitive Tasks: a Functional MRI Study in Children with Reading Impairments

C. Froehlich*, F. Meneguzzi*, A. R. Franco* and A. Buchweitz*

* PUCRS, Porto Alegre, Brazil
email: caroline.froehlich@acad.pucrs.br

Abstract: Dyslexia is a developmental reading disorder characterized by a persistent difficulty to learn how to read fluently, despite normal cognitive abilities. This study aims at investigating the neural underpinnings of this reading disorder in children and teenagers by applying existing machine learning techniques that identify cognitive states based solely on one's brain activation. The technique is used to identify whether participants were performing a reading task (identifying whether a word exists or not) or simply resting. The results show above 90% accuracy for classifying whether a participant was performing a reading task. The technique and results are discussed in turn.

Keywords: Feature Selection, Classifier, Dyslexia, Functional MRI, Words and Pseudowords

Introduction

The aim of the present study is to identify cognitive states in dyslexic readers based solely on their brain activation. The study is part of an umbrella project, named ACERTA, in Portuguese, which stands for Evaluation of Children at Risk for Reading Difficulties. The goal of the project is to understand the differences that underpin the inability of dyslexic children to learn to read fluently, in comparison to their normal reading peers. To achieve this goal, the project applies Functional Magnetic Resonance Imaging (fMRI) to obtain brain-imaging data from children with dyslexia.

Functional MRI (fMRI) is an imaging method that indirectly measures neural activity over time. Some neural activity patterns are known to indicate a person's cognitive state or the presence of a neuropsychological disorder. Cognitive disorders are known to have been successfully identified using fMRI data [1]. Moreover, it has been shown that fMRI data, in combination with machine learning techniques, can be used to predict the cognitive state of subjects [2]–[4]]. Additionally, given its non-intrusiveness to patients, fMRI is widely recognized as a powerful diagnostic tool for conditions with a neurological basis. In this sense, learning disabilities such as dyslexia may be investigated using fMRI to identify the differences in brain function that may underlie such developmental difficulty [5]–[7].

The diagnosis of dyslexia involves a complex, multidisciplinary evaluation of reading performance, cognitive abilities and intelligence, and school and medical history. It is necessary to wait for two years of regular schooling before a child is diagnosed with dyslexia (see DSM-5 criteria); in general, that would mean a child might be diagnosed between ages 8-9. However, the diagnosis is generally made at a much later time. Unpublished data from the Reading Clinic that evaluates children for project ACERTA show an average 10.5 years of age of children at diagnosis of dyslexia (with the range extending to 15 years). In this sense, identifying early indicators of children at risk for learning disabilities, such as dyslexia, may help understand early signs of reading impairment. . In this paper, we take the first steps towards our overarching objective, which is to use brain imaging data to identify specific cognitive states associated with reading impairment; as a first challenge, in this study we set out to identify whether participants were at rest or performing a word reading task.

fMRI data presents a methodological challenge for the unveiling of brain activation patters that may identify early signs of risk for dyslexia. For instance, fMRI data presents a significant amount of information that needs to be analyzed at once. An fMRI scan can generate more than 100,000 voxels of data. Furthermore, the difference between an active portion of the brain and a much less active portion is often minimal. Thus, there is significant potential for using machine learning algorithms to find complex brain activation patterns [5][8]. Machine learning techniques allow for the investigation of significant amounts of data about the brain, and they are sensitive to minimal changes in the fMRI test.

In this paper, we report on our initial efforts to use classification techniques on fMRI data as a diagnostic tool for dyslexia. The goal is to apply machine-learning to identify whether a participant is performing a word-reading task or not. We used a classification technique to predict whether a single patient is performing a task, or resting at any given time during the fMRI scan. Specifically, we trained a Support Vector Machine (SVM) learning algorithm with fMRI data from single patients whose data were partitioned into discrete time points of the fMRI scan. We show that the resulting classifier has an average of 92% accuracy on the final test partition of the dataset.

Methods

Participants – The study included 4 participants (1 female) who were diagnosed with dyslexia or were identified as having reading difficulties (poor readers). Participants age ranged from 9 to 13 years (Mean = 10.5; SD = 1.9); they had to be regularly enrolled in elementary school, aged eight to 16 years, and have formal complaints from parents and/or the school that indicate persistent below-average reading performance. In other words, participants had to have at least two years of formal elementary schooling during which they showed persistent below-average performance. Participants who met the inclusion criteria were evaluated at a Reading Clinic (RC), a pro-bono service set up by the ACERTA project. The RC evaluation and research protocol included evaluation with a medical doctor (medical history; exclusion criteria: history of psychiatric illness), a psychologist (nonverbal IQ; WISC - Wechsler Intelligence Scale III; exclusion criteria: IQ < 80), and a speech therapist (standard reading and writing tasks for the Brazilian population). After diagnosis, participants were scanned at our facility according to the protocols described below. This study was approved by the PUCRS Research and Ethics Committee (process #3629513.0.0000.5336), and each participant's parent or guardian signed an informed consent form approved by the Committee.

Study Design – An event-related experiment was conducted using a word and pseudoword reading task. Stimuli were selected from Salles et al. [9] single word/pseudoword reading task. The set of stimuli is controlled for regularity of letter-sound association, word length (long and short words), and frequency (frequent and infrequent). The reading tasks consist of 20 regular words, 20 irregular words, and 20 pseudowords. The 60 stimuli were divided into two 30-item presentation sets. The division was made to give participants a break in the middle of the task. Words and pseudowords were presented in separate trials on the screen, for seven seconds; a question was presented to participants together with each word (is this a real word?) to which they had to select "Yes" or "No" by pressing response buttons. Mapping of left -hand button, for "Yes," and right-hand button, for "No" matched left-right presentation of "Yes" and "No" on the screen.

Stimulus presentation was offset by jittered intervals; i.e. variable interstimulus interval. The jitter ranged from 1 to 3 seconds. The jitter was inserted after each trial. After 10 trials (10 words) either a baseline condition or rest was inserted in the study. The baseline condition consisted of presentation of a crosshair in the middle of the screen; each baseline lasted 30s. There were two 30-s baseline conditions. The rest period also consisted of presentation of a crosshair, but lasted only seven seconds. Rest was not a condition and was not explicitly modeled in the analysis; the goal of inserting rest between a set of trials is to give participants a short break after 10 trials. A six-second dummy scan was inserted at the beginning of each 30-word set of tasks to ensure T1 magnetization reaches an equilibrium state.

Data Collection – All data was collected on a GE HDxT 3.0T MRI scanner with an 8-channel head coil. Initially, patients underwent a T1 structural scan (TR/TE = 6.16/2.18ms, isotropic 1mm^3 voxels). Subsequently a two 5min 26sec functional FMRI EPI sequences were performed with the following parameters: TR = 2000ms, TE = 30ms, 29 interleaved slices, slice thickness = 3.6mm matrix size = 64x64, FOV = 216x216mm, voxel size = 3.4x3.4x3.6mm^3.

Processing – Initially, functional data was upsampled to have a TR=1 time resolution. The first 6 seconds of each functional run was discarded to eliminate T1 equilibrium effects and subsequently concatenated. Data was then despiked, slice-time and motion corrected, blurred with a 6mm full-width-half-max Gaussian kernel, and aligned to a standard space (MNI152) using the T1 structural volume for improving the registration. Finally, in order to further remove noise from the data, a general linear model was calculated using the motion estimation parameters as nuisance variables. All preprocessing was performed using the AFNI software[1].

Feature Selection – The data of each patient is composed of over 100,000 voxels, most of which are not involved in the neural activity regarding the pseudo word task. Since such irrelevant voxels have activation values, which can interfere with the training of the classifier, we want to remove these voxels from the dataset before training the classifier to generate cleaner results. Thus, instead of using all of the voxels, we follow Buchweitz's [8] approach to select a fixed number of most stable voxels. By stable voxel we mean a voxel that has a minimal standard deviation value for its activation over the times when patients are seeing words within the time series, i.e. that these voxels are consistently activated throughout the tasks. Furthermore, we want the selected voxels to be more or less evenly distributed throughout the brain instead of being clustered in just a few brain locations (otherwise, activation tends to cluster around the occipital lobe, due to the nature of a visual task). Therefore, we partition the brain into 4 lobes (frontal, temporal, occipital, parietal) and find the 200 most stable voxels in each lobe, resulting in 800 most voxels distributed over the brain which will be used as features for the classification algorithm.

Classification – Before applying a classification algorithm, it is necessary to extract the discrete training examples from the time series of voxel activation values extracted from the MRI scan. The study included 4 patients who performed a task during a scan session; the task had 4 conditions (regular word, irregular word, pseudoword and baseline). Patient data is analyzed individually: the training and testing of the classifier is carried out using data from the same patient. Thus, we can separate a patient's data between the task condition (when patients are seeing words) and when they are resting (baseline), comprising the task and rest examples. For the task examples, the following procedure was observed for selection of which portion of the brain imaging data would be used for classification of the task and of rest: (1) Task: each word was remained on the screen for seven seconds. The brain imaging data used included the average activation of a voxels for images collected two seconds after the beginning of the presentation of a word and the images in the following four

[1] http://afni.nimh.nih.gov/

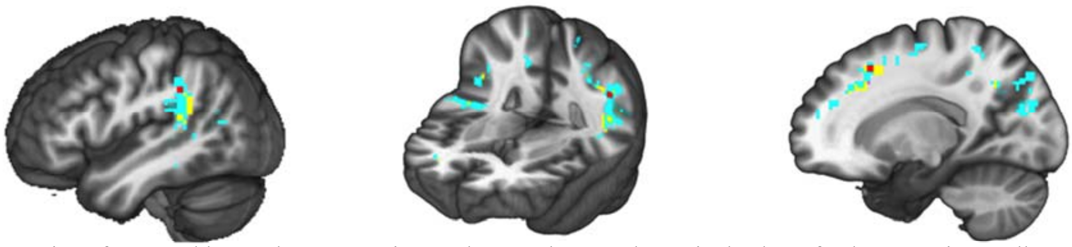

Figure 1: Location of most stable voxels across patients. Blue voxels were chosen in the data of only one patient, yellow voxels were chosen in the data of 2 patients, red voxels were chosen in the data of 3 patients, and pink voxels were chosen in the data of all 4 patients. Notice that only one voxel was chosen in the data of 4 patients, shown in the right image. The regions where pixels clustered are listed in Table 1.

seconds; (2) Rest: the rest/baseline condition lasted 30 seconds; images for the training of the classifier included those four seconds after the baseline block began, with seven four-second examples in sequence for the remaining 26-second block; where each example is the average brain activation for four seconds. The result of this data processing is 20 examples of regular words, 20 of irregular words, 20 of pseudowords (one example per word), and 14 examples of baseline (7 examples per block). After the examples are generated, the final step in the preparation for the classifier algorithm is to manipulate the examples, by performing feature scaling using the scikit-learn [10] package.

Cortical location	Coordinates (x,y,z)	Number of Voxels
Left Superior temporal Gyrus	-47, -41, 25	473
Right Superior Frontal Gyrus	18, 24, 38	379
Left Cuneus	-9, -80, 25	263
Right Cuneus	14, -79, 24	160
Left Fusiform Gyrus	-24, -79, -4	129
Right Superior Temporal Gyrus	51, -28, 5	121

Table1: Regions where the most stable voxels across subjects, considering only clusters that contain more than 100 voxels. Results shown in Figure 1.

SVM [11] is a classification algorithm that represents the input space as a cartesian plane, mapping each example into a point in this space. It splits the state space in two using a hyperplane: one side contains the positive examples, that we label as 1, and the other side contains the negative examples, that we label as −1. Consequently, the SVM algorithm focus on finding the hyperplane that better splits the state space. One of its key parameters include a kernel, which is the function that creates this hyperplane, an example of which is the Radial Basis Function (RBF) kernel. From all classification algorithms, Although Buchweitz et al. [8] uses a Gaussian Naive Bayes along with the feature selection described in this paper, we opted for using SVM, as it seems promising for fMRI studies, where it has been successfully used elsewhere[5]. Particularly, SVM with an RBF kernel is interesting because previous work shows more accurate results than other algorithms when working with a small number of examples and a relatively large number of features [12]. Consequently, because our data contains only a few examples of each class (20 examples of each word type task and 14 examples of baseline), we use the libsvm [12] implementation with the RBF kernel.

For the classification, we partitioned the examples in the training set and the test set. In the training set, we use 16 from the 20 examples of each type of word, and 11 from the 14 baseline examples. The remaining 4 word examples and 3 baseline examples are in the test examples. When acquiring the results, first we perform a cross validation grid search with the training data in order to find the best parameters for the SVM (C and gamma). Second, we use these parameters to train the classifier with the training data. Finally, we test the classifier with the training set. Notice that there is no across-subject experiment.

Results – Our first experimental result of note is the selection of the most stable voxels by the feature selection algorithm for each patient (and an analysis of the overlap of voxels among all patients), which we illustrate in the three pictures of Figure 1.

We generated multiple classifiers for each patient, each classifier is trained with the examples of pairs of

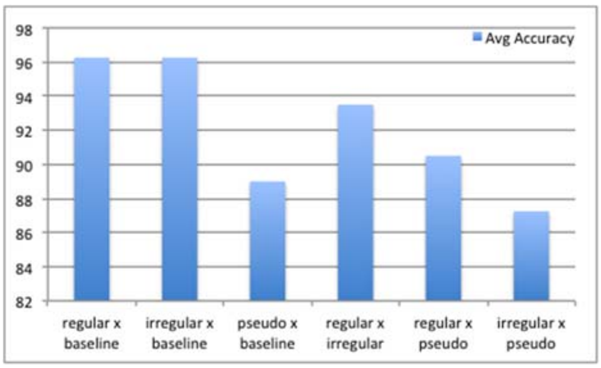

Figure 2: Average classifier accuracy in the 6 experiments. Each bar represents result the average accuracy of 4 patients for each experiment.

conditions, each pair consisting of 6 combinations of 4 classes of condition (see Study Design): regular word (regular), irregular word (irregular), pseudo word (pseudo) and baseline. Thus, for each patient, we train 6 classifiers to recognize the 6 possible condition combinations we selected: regular x baseline; irregular x baseline; pseudo x baseline; regular x irregular; regular x pseudo; and irregular x pseudo). The set of classifiers for each patient yielded different results for the classification tasks. These results are summarized in Figure 2, which shows the average classification accuracy for the 6 experiments using data from 4 patients.

Discussion – The results show that the classifiers successfully identify the cognitive state (performing a task versus resting) in all 6 experiments; they also show that the classifiers generalize the examples correctly with 92% average accuracy in all experiments. The lowest accuracy was 75% in irregular x pseudo experiment using data from patient 4 (well above chance). It is expected that identifying baseline from task would be more accurate than identifying 2 different task types. Alternatively, identifying brain activation patterns for irregular words versus pseudo words shows the worst results, as irregular words and pseudowords may present additional reading difficulty for dyslexic readers [13].

The locations of the most stable voxels include a voxel cluster in the right superior frontal gyrus. Activation of frontal-lobe networks, including right-hemisphere areas, is a characteristic of dyslexic readers; it is hypothesized that dyslexics may activate more areas of the frontal lobe than normal readers to compensate for their decoding and reading fluency difficulties [13]. Two left-hemisphere clusters identified are also part of the traditional language network of the brain: the left fusiform gyrus and the left superior temporal gyrus [14]; even though dyslexics are known to underactivate an area in the vicinity of the left superior temporal gyrus (i.e. the angular gyrus) [5][7], in the classification of reading versus rest it is expected that these two areas would be reasonably stable. Interestingly, there is also a significant cluster of voxels in the right-hemisphere homologue (right superior temporal gyrus), which may suggest a spillover activation mechanism; that is, as a task shows a certain level of difficulty, one of the mechanisms of brain adaptation is recruitment of cortical tissue in homologue areas [15]. The question remains of whether a control group would have more stable voxels in this network, or in the right-hemisphere as well.

From the feature selection results shown in Figure 1, we observe that the most stable voxels chosen for each patient are sparse. Although most of the chosen voxels cluster when seeing each patient individually, they do not cluster across subject. Figure 1 shows a much larger number of stable voxels selected for a single patient than voxels, which were selected across multiple patients. Finally, only one voxel was chosen in the data of 4 all patients. Thus, as the features to train the classifier are sparse, and the feature selection process picked completely different voxel locations for each patient, we believe it will be difficult for a classifier algorithm to generalize across multiple patients.

Acknowledgements

The study was funded by CAPES/OBEDUC 0898/2013 project 23038.002530/2013-93 (ACERTA).

References

[1] R. Cabeza and L. Nyberg, "Imaging cognition II: An empirical review of 275 PET and fMRI studies.," *J. Cogn. Neurosci.*, vol. 12, pp. 1–47, 2000.

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

[3] T. M. Mitchell, S. V Shinkareva, A. Carlson, K.-M. Chang, V. L. Malave, R. a Mason, and M. A. Just, "Predicting human brain activity associated with the meanings of nouns.," *Science*, vol. 320, no. 5880, pp. 1191–5, May 2008.

[4] S. V Shinkareva, R. a Mason, V. L. Malave, W. Wang, T. M. Mitchell, and M. A. Just, "Using FMRI brain activation to identify cognitive states associated with perception of tools and dwellings.," *PLoS One*, vol. 3, no. 1, p. e1394, Jan. 2008.

[5] F. Hoeft, B. D. McCandliss, J. M. Black, A. Gantman, N. Zakerani, C. Hulme, H. Lyytinen, S. Whitfield-Gabrieli, G. H. Glover, A. L. Reiss, and J. D. E. Gabrieli, "Neural systems predicting long-term outcome in dyslexia.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 1, pp. 361–6, Jan. 2011.

[6] N. M. Raschle, M. Chang, and N. Gaab, "Structural brain alterations associated with dyslexia predate reading onset.," *Neuroimage*, vol. 57, no. 3, pp. 742–9, Aug. 2011.

[7] B. a Shaywitz, S. E. Shaywitz, K. R. Pugh, W. E. Mencl, R. K. Fulbright, P. Skudlarski, R. T. Constable, K. E. Marchione, J. M. Fletcher, G. R. Lyon, and J. C. Gore, "Disruption of posterior brain systems for reading in children with developmental dyslexia.," *Biol. Psychiatry*, vol. 52, no. 2, pp. 101–10, Jul. 2002.

[8] A. Buchweitz, S. V. Shinkareva, R. A. Mason, T. M. Mitchell, and M. A. Just, "Identifying bilingual semantic neural representations across languages," *Brain Lang.*, vol. 120, pp. 282–289, 2012.

[9] J. Salles, L. R. Piccolo, R. S. Zamo, and R. Toazza, "Normas de desempenho em tarefa de leitura de palavras / pseudopalavras isoladas (LPI) para crianças de 1º ano a 7º ano," *Estud. e Pesqui. em Psicol.*, vol. 13, no. 2, pp. 397–419, 2013.

[10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2012.

[11] V. N. Vapnik, *The Nature of Statistical Learning Theory*, vol. 8. 1995, p. 188.

[12] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 27:1–27:27, 2011.

[13] S. Shaywitz, *Overcoming Dyslexia: A New and Complete Science-Based Program for Reading Problems at Any Level (Vintage)*. Vintage, 2008.

[14] C. J. Price, "The anatomy of language : a review of 100 fMRI studies published in 2009," vol. 1191, pp. 62–88, 2010.

[15] C. S. Prat, R. A. Mason, and M. A. Just, "Individual differences in the neural basis of causal inferencing.," *Brain Lang.*, vol. 116, no. 1, pp. 1–13, Jan. 2011.