Using AI to interpret BI: machine learning for decoding and characterization of brain activity patterns

Malin Björnsdotter*, Simon Beckmann, Erik Ziegler & Johan Wessberg

Abstract

The appealing properties of artificial intelligence (AI) methods are being increasingly acknowledged by the neuroimaging community, as evidenced by the recent surge of brain activity pattern recognition studies [19]. Supervised learning and classification, in particular, are appreciated tools for localizing and distinguishing intricate brain response patterns and making predictions about otherwise undetectable neural states. Our group refines and applies such methods in order to implement sensitive and dynamic tools for characterization of neurophysiological data. Specifically, we employ support vector machines (SVMs), particle swarm optimization (PSO), independent component analysis (ICA), as well as both genetic and memetic algorithms on functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) data. This paper provides a brief overview of our recent advances in the development and utilization of AI-based analysis, with the particular aspiration to characterize human brain activation patterns produced by touch.

1 Introduction

Machine learning approaches for mining neurophysiological data are rapidly gaining popularity, and justifiably so, as they provide a level of analysis not possible with conventional methods [19]. Commonly used systems for acquiring such brain activity data non-invasively include electroencephalography (EEG), which provides an estimate of electrical currents produced by the brain, and functional magnetic resonance imaging (fMRI), where local blood flow changes are quantified. Classically, the analysis of neurophysiological data relies on descriptive statistical measures where average activity changes in single data points are related to an experimental condition and particular data characteristic of interest (e.g. brain regions in fMRI and frequencies in EEG). Artificial intelligence techniques, in contrast, are capable of distinguishing complex and subtle data patterns, distributed across numerous measuring points, on a single-trial basis. The benefits of AI were acknowledged early for real-time classification of EEG signals in brain-computer interfaces (see e.g. [4] for a review), and more recently for clinical evaluation (see e.g. [9, 28, 24]), as well as a wide range of fMRI analyses (see e.g. [11] and [19]).

Generally, supervised learning techniques are used: a classifier is trained to recognize and decode subtle intrinsic signal patterns correlated to given brain states, such as the fMRI activity produced by a single touch stimulus [2]. Supervised learning methods not only enable real-time single-trial classification (such as brain state tracking [29] or intent decoding [30]), but by virtue of considering information encoded over multiple measuring points they also provide improved condition differentiation sensitivity [27, 19, 13].

In the following sections, we summarize our recent advances in the development and application of such AI-based analysis in both fMRI and EEG studies. In particular, we utilize these methods to characterize human brain activation patterns produced by touch.

2 Functional Magnetic Resonance Imaging

fMRI involves the estimation of local blood flow changes in measuring points, known as voxels,

*All authors are with the Department of Physiology, Institute of Neuroscience and Physiology, University of Gothenburg, Sweden. Email: malin.bjornsdotter@neuro.gu.se

which are evenly distributed across the brain volume. Conventional analysis is limited to localization of brain regions which are activated on average by a specific condition using general linear model (GLM) methods, where each voxel is treated independently of any other [16]. Classifier-based approaches, popularly termed multivoxel pattern analysis (MVPA), on the other hand, utilize multiple voxels simultaneously in a multivariate fashion, allowing identification of brain regions containing spatially distributed activity patterns. Numerous studies have demonstrated that by extracting such spatially-encoded information, otherwise non-discernible pattern differences can be identified (e.g. in lie detection [12], the decoding of single visual stimuli – visible [21], as well as invisible [18] – biofeedback [33], and various types of real-time fMRI analysis [23, 14]). In addition, the ability of MVPA to decode single-trial brain states provides an effective tool for observing brain-state changes in real time (somewhat equivocally termed "mind reading"; see e.g. [29, 30]).

A major challenge in MVPA analysis is to identify which voxels (out of hundreds of thousands) are in fact relevant to the classification task. To this end, we have proposed a number of combinatorial optimization techniques based on evolutionary approaches combined with a classifier for fitness evaluation [20]. First we implemented a simple genetic algorithm (GA) to show that only a handful of voxels where sufficient to successfully decode gentle brushing of the forearm (subject mean of 74.3% correct with a linear support vector machine classifier) and that the performance was significantly improved compared to a univariate GLM-based voxel selection scheme (see [1]; Figure 1).

While producing high classification rates by identifying few, representative voxels which were *sufficiently* informative, the simple GA did, however, yield spatially sparse and distributed brain maps which were of limited use for a physiological examination of underlying neural processes. The algorithm was therefore refined to include elements of voxel clustering, with the goal of identifying larger brain regions containing *useful* voxels [8].

In contrast to the simple GA, the clustering algorithm proved to be highly useful in revealing physiologically relevant brain regions. We could, for example, detect which regions of the so-called insular cortex were differentially activated by gentle touch

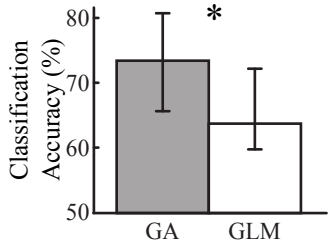

Figure 1: Classification results on fMRI data from six individuals demonstrating that voxels selected by a simple genetic algorithm are highly effective in predicting whether or not the individuals were sensing a tactile stimulation (soft brush) on their forearm. Importantly, a significantly higher classification rate was obtained on the voxels selected by the genetic algorithm in comparison with the conventional general linear model (GLM) approach. Chance classification is 50%.

of the forearm and thigh, where conventional GLM methods failed (see [6]; Figure 2).

The clustering GA proved, however, to be erratic when attempting to identify a more complex distribution of multiple voxel clusters. This issue is currently being resolved by incorporating a local search element according to a memetic algorithm (MA) framework, and the improved algorithm is successful in detecting multiple clusters of voxels distributed across the brain volume. Illustratively, the memetic algorithm contrasts the simple GA by being substantially better at detecting *all* useful voxels as opposed to those *sufficient* for good performance (Figure 3). On real data, the memetic algorithm was highly successful in detecting brain regions which could decode finger movement brain patterns learnt from seven subjects and applied to an eighth individual (Figure 4).

The proposed evolutionary algorithms are relatively complex to implement. As a simpler alternative, we are currently exploring particle swarm optimization (PSO). Our first attempt included an implementation which was similar to the simple GA, with no explicit spatial clustering of the voxels [26]. Again, this algorithm provided high brain-state classification results, but yielded maps which were difficult to interpret (Figure 5).

In a second implementation, clustering was forced such that each particle, instead of coding a number of distributed voxels, represented a sin-

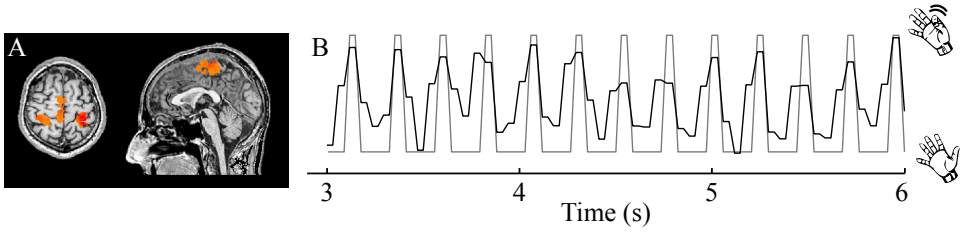

Figure 4: Decoding finger movements from fMRI data in brain regions selected by the memetic algorithm. A) The brain regions identified by the algorithm. B) A temporal trace of the classification performance (thick line). The true brain state is indicated by the thin gray line.

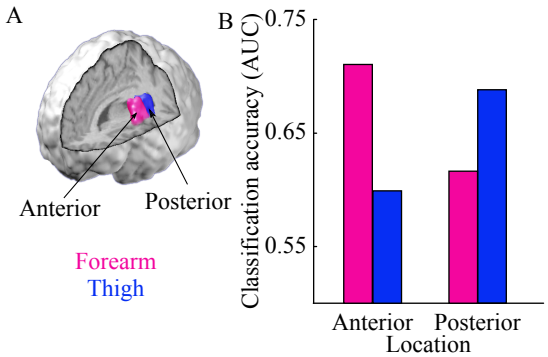

Figure 2: A) Brain regions identified for forearm (purple) and thigh (blue) gentle brushing by the clustering genetic algorithm. B) The brain state decoding accuracies (measured in area under the receiver operating characteristic curve, AUC) obtained when attempting classification of the forearm data in the forearm brain region and vice versa for all permutations of data and regions. The results demonstrate that this patch of the brain, called the insular cortex, is organized in a somatotopic fashion with forearm stimulation projecting anterior to thigh.

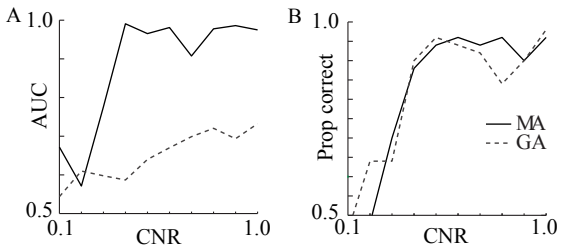

Figure 3: The memetic algorithm (MA) is substantially better than the simple genetic algorithm (GA) in detecting all useful fMRI voxels (A: performance measured in area under the receiver operating characteristic curve, AUC), although both algorithms detect useful voxels sufficient for good classification (B). The methods were evaluated on simulated data of varying contrast-to-noise ratio (CNR) containing useful voxels of known location.

gle spherical cluster of voxels [7]. This approach was inspired by the highly popular "searchlight" method, where such search spheres are sequentially centered on each voxel in the brain [22]. The proposed PSO method was successful in detecting brain areas where the sensation of a soft brush stimulus could be accurately decoded (Figure 6), and was substantially faster than the searchlight (e.g. 6.7 minutes compared to 9 hours to map a whole brain).

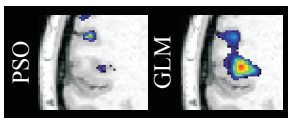

Figure 5: Voxel selection frequency using a simple particle swarm optimization (PSO) algorithm to identify brain areas which are activated by gentle touch, compared to a standard general linear model (GLM) map. The values are scaled to reflect minimum (blue) to maximum (red) map values.

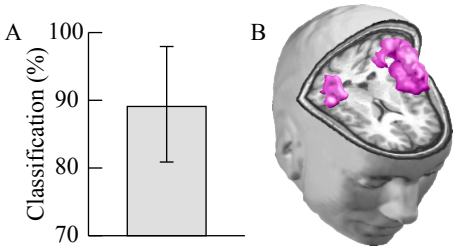

Figure 6: The average classification score (A) and corresponding brain areas (B) detected by the clustering particle swarm optimization algorithm (PSO).

3 Electroencephalography

EEG involves the registration of electrical brain activity at a temporal resolution of milliseconds using electrodes attached to the scalp [15]. The signals are believed to mainly reflect post-synaptic currents in nerve cells, and synchronous activity of thousands of cells is required to produce a measurable potential. With recent advances in the complex problem of localizing signal sources within the brain, EEG has emerged as an excellent tool for both characterizing temporal dynamics and localizing cognitive processing.

Standard EEG analysis involves averaging over numerous events to produce event-related potentials (ERPs) which are subsequently analyzed based on the latency and amplitude of characteristic deflections. The averaging procedure, however, may mask or even eliminate relevant information potentially concealed in complex signal patterns. We therefore explored an EEG analysis approach based on independent component analysis (ICA) combined with supervised learning. ICA blindly decomposes multi-channel EEG data into maximally independent component processes (ICs) that typically express either particularly brain generated EEG activities or some type of non-brain artifacts (e.g. line noise or muscle activity [10]).

We used this approach to identify and characterize differential temporal patterns in the EEG responses to stimulation of the fingerpad with surfaces of varying roughness. Our classifier (a linear support vector machine, SVM) was successfully trained to differentiate between the temporal patterns evoked by the two different textures

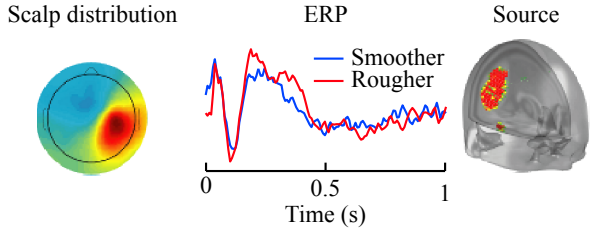

Figure 7: Example of a contralateral somatosensory component where the support vector machine classifier could differentiate two surfaces of different textures. The scalp distribution plot indicates the component's scalp map projection. The component's event-related potentials (ERPs) are plotted for each of the textures (rough spatial period: 1920 μm, smooth: 520 μm) and a source image was constructed to indicate the component's origin.

in some ICs but not in others [3]. For example, an EEG component generated in the contralateral somatosensory cortex with activation peaks at 100 ms after onset (P100) of stimulation significantly differentiated the textures (Figure 7).

4 Concluding remarks

AI-based approaches for analyzing brain activity provide a highly appealing complement to conventional statistical methods, enable a deeper understanding of brain function, and promote the development of novel medical techniques. Whereas we primarily use AI for the characterization of brain activation patterns, other brain signal decoding applications include brain-computer interfaces (BCIs), biofeedback [32, 31, 33], real-time signal analysis [14], disease diagnosis [25, 17], enabling prosthesis control, and opening communication channels with locked-in patients [5]. The application of AI concepts in neuroimaging is in its infancy, and further refinement of these algorithms will undoubtedly facilitate our understanding of the human brain.

Acknowledgments

These studies were supported by the Swedish Research Council (grant K2007-63X-3548), the Sahlgrenska University Hospital (grant ALFGBG 3161) and NANOBIOTACT (`nanobiotact.org`,

EU-FP6-NMP-033287). We are grateful to Karin Göthner and Linda Lundblad for assistance with data acquisition.

References

[1] M. Åberg and J. Wessberg. Evolutionary optimization of classifiers and features for single trial EEG discrimination. *BioMedical Engineering Online*, 6(32), 2007.

[2] M. S. Beauchamp, S. LaConte, and N. Yasar. Distributed representation of single touches in somatosensory and visual cortex. *Human Brain Mapping*, (in press), 2009.

[3] S. Beckmann, M. Björnsdotter, H. Backlund-Wasling, and J. Wessberg. Brain decoding of texture processing using independent component analysis and support vector machines. pages 287 –292, aug. 2009.

[4] N. Birbaumer. Breaking the silence: brain-computer interfaces (bci) for communication and motor control. *Psychophysiology*, 43(6):517–32, 2006.

[5] N. Birbaumer and L. G. Cohen. Brain-computer interfaces: communication and restoration of movement in paralysis. *Journal of Physiology*, 579(Pt 3):621–636, Mar 2007.

[6] M. Björnsdotter, L. S. Löken, H. Olausson, Å. B. Vallbo, and J. Wessberg. Somatotopic organization of gentle touch processing in the posterior insular cortex. *Journal of Neuroscience*, 29(29):9314–9320, 2009.

[7] M. Björnsdotter and J. Wessberg. Identification and localization of human brain activity patterns using particle swarm optimization. In *Proceedings of the 2nd International Conference on Computer and Automation Engineering*, 2010.

[8] M. Björnsdotter Åberg and J. Wessberg. An evolutionary approach to the identification of informative voxel clusters for brain state discrimination. *IEEE Journal of Selected Topics in Signal Processing*, 2(6):919–928, 2008.

[9] C. Cao, R.L. Tutwiler, and S. Slobounov. Automatic classification of athletes with residual functional deficits following concussion by means of eeg signal using support vector machine. *IEEE Trans Neural Syst Rehabil Eng.*, 16(4):327–35, 2008.

[10] Pierre Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287 – 314, 1994. Higher Order Statistics.

[11] D. D. Cox and R. L. Savoy. Functional magnetic resonance imaging (fMRI) 'brain reading': detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, 19(2 Pt 1):261–270, June 2003.

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

[13] F. De Martino, G. Valente, N. Staeren, J. Ashburner, R. Goebel, and E. Formisano. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *NeuroImage*, 43(1):44–58, 2008.

[14] R. C. deCharms. Applications of real-time fMRI. *Nature Reviews Neuroscience*, 9(9):720–729, Sep 2008.

[15] Bruce J. Fisch. *Fisch & Spehlmann's EEG Primer; Basic Principles of Digital and Analog EEG*. Elsevier Science, third edition edition, 1999.

[16] 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*, 2(4):189–210, 1994.

[17] C. H. Y. Fu, J. Mourão-Miranda, S. G. Costafreda, A. Khanna, A. F. Marquand, S. C. R. Williams, and M. J. Brammer. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. *Biological Psychiatry*, 63(7):656–662, Apr 2008.

[18] J.-D. Haynes and G. Rees. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8(5):686–691, April 2005.

[19] J.-D. Haynes and G. Rees. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7):523–534, July 2006.

[20] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. The MIT Press, April 1992.

[21] Y. Kamitani and F. Tong. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5):679–685, May 2005.

[22] N. Kriegeskorte, R. Goebel, and P. Bandettini. Information-based functional brain mapping. *Proceedings of the National Academy of Science*, 103:3863–3868, 2006.

[23] S. M. LaConte, S. J. Peltier, and X. P. Hu. Real-time fMRI using brain-state classification. *Human Brain Mapping*, 28(10):1033–1044, Oct 2007.

[24] J. Löfhede, N. Löfgren, M. Thordstein, A. Flisberg, I. Kjellmer, and K. Lindecrantz. Classification of burst and suppression in the neonatal electroencephalogram. *J Neural Eng.*, 5(4):402–10, 2008.

[25] A. F. Marquand, J. Mourão-Miranda, M. J. Brammer, A. J. Cleare, and C. H. Y. Fu. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport*, 19(15):1507–1511, Oct 2008.

[26] T. Niiniskorpi, M. Björnsdotter Åberg, and J. Wessberg. Particle swarm feature selection for fMRI pattern classification. In P. Encarnação and A. Veloso, editors, *Proceedings of the 2nd BIOSIGNALS*, pages 279–284. INSTICC, 2009.

[27] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby. Beyond mind-reading: multivoxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9):424–430, 2006.

[28] L.M. Patnaik and O.K. Manyam. Epileptic eeg detection using neural networks and post-classification. *Comput Methods Programs Biomed*, 91(4):100–9., 2008.

[29] S. M. Polyn, V. S. Natu, J. D. Cohen, and K. A. Norman. Category-specific cortical activity precedes retrieval during memory search. *Science*, 310(5756):1963–1966, 2005.

[30] Chun Siong Soon, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes. Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11(5):543–545, May 2008.

[31] N. Weiskopf, F. Scharnowski, R. Veit, R. Goebel, N. Birbaumer, and K. Mathiak. Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI). *Journal of Physiology (Paris)*, 98(4-6):357–373, 2004.

[32] N. Weiskopf, R. Veit, M. Erb, K. Mathiak, Grodd W., Goebel R., and Birbaumer N. Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *NeuroImage*, 19:577–586, 2003.

[33] S. S. Yoo, H. M. O'Leary, T. Fairneny, N. K. Chen, L. P. Panych, H. Park, and F. A. Jolesz. Increasing cortical activity in auditory areas through neurofeedback functional magnetic resonance imaging. *Neuroreport*, 17(12):1273–1278, 2006.