| | |
|---|---|
| **Citation/Reference** | Borbála Hunyadi, Simon Tousseyn, Patrick Dupont, Sabine Van Huffel, Wim Van Paesschen and Maarten De Vos (2014), **Automatic selection of epileptic independent fMRI components** Proc. of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2014), 26-30 August 2014, 3853-3856 |
| **Archived version** | Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher |
| **Published version** | http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6944464 |
| **Journal homepage** | http://embc.embs.org/2014/ |
| **Author contact** | borbala.hunyadi@esat.kuleuven.be *+ 32 (0)16 321799* |
| **IR** | https://lirias.kuleuven.be/handle/123456789/466667 |

*(article begins on next page)*

# Automatic selection of epileptic independent fMRI components

Borbála Hunyadi[1,2], Simon Tousseyn[3,4], Patrick Dupont[3,4,5],
Sabine Van Huffel[1,2], Wim Van Paesschen[3,4] and Maarten De Vos[6,7]

***Abstract*— EEG-correlated fMRI analysis has proven to be useful in localizing regions of BOLD activation related to epileptic activity. However, as EEG does not always provide reliable information, purely fMRI-based data-driven techniques are invaluable. Recently, we have shown that independent component analysis (ICA) can extract sources related to the epileptic network even in such EEG-negative cases [1]. Moreover, these sources were shown to be informative with respect to the seizure onset zone (SOZ). In order to utilize this concept in clinical practice in a prospective manner, this work aims at developing an automatic technique for selecting the epileptic sources. The proposed approach applies a cascade of two classifiers. In the first step artifact related sources are discarded. In the second step the sources are characterized by four discriminative features and epileptic sources are selected from among other BOLD-related components. Our technique reaches a promising 77% specificity and provides concordant sources with the EEG-correlated fMRI activation maps or with the SOZ in 71% of the cases.**

## I. INTRODUCTION

Epilepsy is a neurological disorder, affecting 1% of the worldwide population. The manifestation of this disease is the epileptic seizure, an abnormal, synchronous activity of the neurons in the brain. More than 30% of epilepsy patients continue to have seizures despite of medication, hence their quality of life is seriously compromised. Surgical resection of the epileptogenic focus might offer cure for these patients. The success of the surgery strongly depends on the accurate identification of the brain region responsible for the generation of seizures.

In recent years many research have shown the usefulness of EEG-correlated fMRI analysis for this purpose [2]. This approach heavily relies on the visual identification of interictal epileptic discharges (IEDs) in the EEG. However, EEG does not always provide accurate information, due to bad signal quality, or simply due to the absence of IEDs during the limited time of the recordings. Purely fMRI-based data-driven approaches are invaluable in such cases.

Relying on the assumption that the sources generating the measured blood oxygen level dependent (BOLD) signals are mutually statistically independent in space, independent component analysis (ICA) has been used in numerous fMRI studies. It has successfully extracted sources related to epileptic activity from fMRI time series recorded during ictal [3] and interictal [4] period. In the latter case ICs were automatically assigned to an artifact related or to a BOLD signal related group using the technique presented by [5]. Within the BOLD related class the epileptic ICs were identified based on the known timing of the epileptic events observed on EEG. However, the application of ICA is of particular interest in those cases where no epileptic events can be identified during the recordings. Recently, we have proven in a retrospective manner that ICA can extract sources related to the epileptic network [1] even in such EEG-negative cases. Moreover, these sources were shown to be informative with respect to the seizure onset zone (SOZ). In order to put this theoretical concept into practice, the epileptic source has to be selected blindly. Therefore, the goal of the current paper is to develop a prospective technique which can automatically select the epileptic fMRI source independently from EEG information.

## II. MATERIALS AND METHODS

### A. Data collection

For the purpose of this study patient data were included based on the following criteria: (1) consecutive adults who underwent a full presurgical evaluation for refractory focal epilepsy between August 2010 and November 2013 (2) concordant data pointing to one epileptic focus using all presurgical investigations, including EEG-correlated fMRI (3) identical recording parameters (TE = 33ms, TR = 2.5s, voxel size 2.6×3×2.6 mm). Finally, one recording session was selected arbitrarily from each patient. The final dataset consisted of the interictal fMRI time series of 10 patients.

In addition, fMRI data from 13 healthy individuals were included as well, in order to assess the behavior of the proposed method in the absence of pathological activity.

All fMRI images were realigned, slice-time corrected, normalized to MNI space and spatially smoothed with an isotropic Gaussian kernel of 6 mm full width at half maximum. EEG-correlated fMRI analysis was performed within the general linear model (GLM) framework using the SPM8 toolbox (`http://www.fil.ion.ucl.ac.uk/spm/software/spm8`). The timing of the preponderant

[1]STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering (ESAT), KU Leuven, Belgium
[2]iMinds Medical IT
[3]Laboratory for Epilepsy Research, UZ Leuven & KU Leuven, Belgium
[4]Medical Imaging Research Center, UZ Leuven & KU Leuven, Belgium
[5]Laboratory for Cognitive Neurology, UZ Leuven & KU Leuven, Belgium
[6]Neuropsychology Lab, Department of Psychology, University of Oldenburg, Germany
[7]Hearing4all, University of Oldenburg, Germany
This work is supported by Research Council KUL: GOA/10/09 MaNet, CoE PFV/10/002 (OPTEC), PhD/Postdoc grants; FWO: G.0427.10N (Integrated EEG-fMRI), IWT: TBM 080658-MRI (EEG-fMRI); PhD/Postdoc grants; iMinds Medical Information Technologies SBO 2014; IUAP P7/19/ (DYSCO, 2012-2017); VLIR UOS programs; EU: ERC Advanced Grant: BIOTENSORS (nr. 339804)

IEDs convolved with the hemodynamic response function was used as a regressor of interest. The six rigid-body motion correction parameters, the fMRI signal averaged over the lateral ventricles, and the signal averaged over a region within the white matter were included as confounding covariates. The activation maps were thresholded at a significance level of $p < 0.05$ with family wise error correction.

### B. Blind selection method: overview

As a first step, ICA is performed on the fMRI time series. Consecutively, a cascade of classifiers is applied: after discarding artifact related ICs in the first classification step, the epileptic ICs are selected from the remaining reduced set of BOLD related ICs using the second classifier. Finally, localization information is retrieved from the spatial map corresponding to the epileptic ICs.

The ICA step and the discrimination between BOLD and artifact related ICs, together with the corresponding feature extraction was performed using the Fix plug-in of the FSL toolbox (`http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX`). Detailed explanation about the extracted features and the applied classification technique is available in [6]. The contribution of this paper is the development of the features and the classifier for the second discrimination step in the cascade. These steps are explained in detail below.

### C. Feature extraction

The aim of this study is to automatically select epileptic independent components. In [1] we argue that there might be multiple ICs corresponding to partially overlapping parts of the epileptic network, reflecting different aspects of epileptic activity. However, considering that the automatic recognition of an epileptic IC corresponding to the SOZ would be the clinically most relevant finding, we concentrate on characteristics which are peculiar to such ideal components. Therefore, the following features were extracted from the fMRI ICs.

*Number of clusters*. Suprathreshold voxels in the spatial map corresponding to an IC are spatially organized in one or more clusters. We only take into account clusters which comprise at least 30 suprathreshold voxels. The number of activation clusters in an epileptic IC is ideally 1, corresponding to the SOZ. In contrast, various resting state networks (RSNs) consist of multiple active regions.

*Activation asymmetry*. The SOZ of a unifocal epilepsy patient is restricted to a region in strictly one hemisphere, thus will show asymmetry. The activation asymmetry of an IC is assessed by the following formula:

$$AA = |\sum_{i=1}^{H} v_i^{(l)} - v_i^{(r)}|, \quad (1)$$

where $v_i^{(l)}$ denotes the $i^{th}$ voxel in the left hemisphere, $v_i^{(r)}$ denotes the corresponding contralateral voxel and $H$ is the total number of voxels in one hemisphere.

*Sparsity in activelet basis*. Activelets are a recently developed dictionary of wavelet basis functions [7], which has been applied to detect interictal epileptic activity from fMRI [8]. The activelet waveforms were specifically constructed to fit the BOLD signal in response to a sparse transient event. As such, the representation of a signal comprising sparse transient events, such as interictal epileptic spikes, will be sparse in the activelet basis. Practically, the neural activity $x(t),\ t = 1,...,T$ can be estimated by the linear combination of a few activelet basis vectors in the dictionary $\phi$ as $\hat{x} = \phi\beta_0$, where $\beta_0$ is the solution of the convex $l_1$ optimization problem:

$$\min_{\beta}(\frac{1}{2}\|x - \phi\beta\|_2^2 + \lambda\|\beta\|_1) \quad (2)$$

The regularization parameter $\lambda$ controls the trade-off between the sparsity of the solution and the reconstruction error, a higher value favoring a sparser solution. The value of $\lambda$ was set to 2.5 in this study. $\phi$ is the overcomplete dictionary matrix containing the basis functions of the undecimated activelet transform. It is a matrix of size $T \times P$, where T is the length of the time series and $P = 3 \cdot T$ as the number of wavelet decomposition scales was set to 3 [8].

The time course of an epilepsy related IC is expected to have a sparser representation in the activelet basis compared to non-related ICs. The sparsity of the representation in the activelet basis was quantified with the Gini index [9], which measures the statistical dispersion of the magnitude of the coefficients.

*Sparsity in sine dictionary*. In contrast, the time course of resting state networks is characterized by low-frequency (0.01-0.1 Hz) fluctuations [10]. As such, they have a sparse representation in a sine dictionary restricted to this frequency band. A matching pursuit algorithm was used to retrieve the coefficients corresponding to the best nonlinear approximation of the fMRI IC. Again, the sparsity was quantified using the Gini index.

Kruskal-Wallis tests were performed in order to assess whether the extracted features differentiate between epileptic and non-epileptic ICs (see below). The distribution of the feature values are shown in Figure 1. The statistical significances are indicated in brackets. The difference of the number of clusters was marginally insignificant, while the other features were significantly different between the two groups.

### D. Classification

A least-squares support vector machine (LS-SVM) classifier was applied to learn an optimal classifier based on the above extracted features. A linear kernel was chosen, and the kernel parameter was tuned using leave-one-component-out crossvalidation on the training data. Positive training examples, i.e. the class of epileptic ICs, consisted of the ICs showing the largest overlap with the cluster containing the maximally activated voxel in the GLM-based fMRI activation map in each patient. In addition, ICs which showed at least 10% overlap with the same cluster and significantly correlated with the timing of the IEDs, were also included in the epileptic class. All other ICs were included as negative training examples, i.e. in the class of non-epileptic ICs.

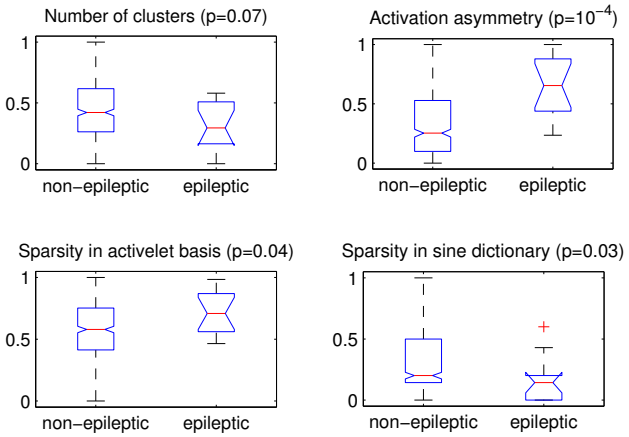



Fig. 1: The feature values significantly differ between the epileptic and non-epileptic ICs, except, the difference in number of clusters is marginally insignificant.

Note that our final goal is to select an IC which overlaps with the SOZ, nevertheless, we prefer to define the training examples based on the GLM-based activation maps for the following reason. While the SOZ reflect ictal phenomena, in this study interictal processes were recorded in the fMRI. We do expect that the fMRI-based maps will contain the SOZ, but perfect overlap is unlikely. The GLM-based fMRI maps provide a more reliable image of how the epileptic ICs should look like.

Recall that LS-SVM takes decisions according to:

$$y(\mathbf{x}) = \text{sign}(\mathbf{w}^T \varphi(\mathbf{x}) + b), \tag{3}$$

with $y(\mathbf{x}) \in \{-1, 1\}$ are the class labels, $\mathbf{x}$ is the input pattern, $\mathbf{w}$ is a weighting vector, $b$ is a bias term and $\varphi$ is a feature mapping. A linear kernel, i.e. a linear mapping was used in this work. Note that this formula might assign multiple ICs to the epileptic class, however, we are interested in selecting exactly one epileptic IC in each patient. Therefore, we modify the decision function as follows:

$$y(x_i) = \begin{cases} 1 & \text{if } \arg\max_{\mathbf{x}}(\mathbf{w}^T \varphi(\mathbf{x})) = i \text{ and } \mathbf{w}^T \varphi(\mathbf{x}) + b > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

This way at most one IC is selected in each patient. The values $\mathbf{w}^T \varphi(\mathbf{x})$ determine a ranking of the ICs, the highest value corresponds to the IC which resembles most the epileptic ICs in the training data. In case this value exceeds the threshold $-b$, the first ranked IC is selected as epileptic. Otherwise, if no IC shows enough resemblance to the training epileptic ICs, no selection is made.

The performance of the proposed method was estimated in a leave-one-patient-out scheme: individual classifiers were trained for each patient using data from all other patients.

### *E. Evaluation measures*

The voxel values in the spatial maps of the selected ICs were converted to z-score and thresholded at 5 in order to obtain highly specific activation maps. Afterwards, the overlap between the IC activation maps and the SOZ as well as the GLM-based EEG-fMRI activation map were assessed. If the cluster containing the maximally activated voxel overlapped with the SOZ or the GLM-based map, the IC was considered concordant.

On one hand, an overlap with the GLM-based EEG-fMRI map indicates that the selected IC represents some aspect of the epileptic network. On the other hand, a method selecting an IC which overlaps with the SOZ would have real clinical significance. Following the considerations made in [11], sensitivity and specificity of the proposed method for localizing the SOZ were defined as follows.

Activation maps of the selected ICs overlapping with the SOZ were considered true positive cases. Patients where no IC was selected or no overlap was found were considered false negative cases. In order to define specificity, the data from healthy controls were analyzed. Controls where no IC was selected or the selected IC contained no suprathreshold voxels were considered true negatives, and others false positives. Finally, sensitivity and specificity was calculated with the standard formula:

$$\text{Sensitivity} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \tag{5}$$

$$\text{Specificity} = \frac{\text{true negative}}{\text{true negative} + \text{false positive}} \tag{6}$$

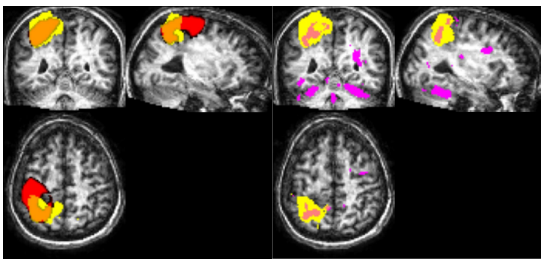


Fig. 2: Patient 6. The selected IC is indicated in yellow, while the SOZ (i.e. focal cortical displasia) and the GLM-based activation map are indicated in red and violet, respectively.

## III. Results

An IC was selected in 7 out of 10 patients. As an example, the selected IC, together with the SOZ and the GLM-based activation map of patient 6 are shown in Figure 2.

In 6 out of the 7 cases where a selection was made, the selected IC overlapped either with the SOZ or with the GLM-based EEG-fMRI activation map. Moreover, the cluster containing the maximally activated voxel overlapped either with the SOZ or with the GLM-based map in 5 out of 7 cases. These results indicate that the method is very selective for components related to the epileptic network.

However, an IC overlapping with the SOZ was selected only in 4 cases, i.e. the proposed method has a sensitivity of 40% for localizing the SOZ. Looking at the cluster with the maximally activated voxel, the sensitivity drops to 30 %.

Nevertheless, the method selected an IC in only 3 controls out of 13, corresponding to a specificity of 77%. This further supports that the ICs selected in patients are truly related to epilepsy.

The low sensitivity for selecting the SOZ is partly explained by the fact that we insist on selecting one IC per patient. However, there is strong evidence that in some patients more than one epileptic IC is present. Patients 3, 4, 5 and 9 showed more than one type of IEDs in the simultaneously recorded EEG. While the GLM activation maps based on the preponderant spike type overlaps with the SOZ in most cases [11], BOLD activations corresponding to other spike types may be remote from the SOZ. In fact, the IC activation maps selected in patients 3 and 5 correspond to the field of the secondary spike type. The ICs of the epileptic class, i.e. the ones concordant with the GLM maps based on the preponderant spike type, were ranked on average 3rd out of 39 across all patients. This high average ranking suggests that ICs reflecting different aspects of epileptic activity compete for the first ranked position.

## IV. Discussion

In this paper an automatic method was developed which selects the epileptic source among fMRI ICs. To this end, artifact related ICs were first rejected using a recently introduced and online available technique [6]. Subsequently, the remaining BOLD related ICs were characterized with four features, which were fed to a LS-SVM classifier. The proposed technique was evaluated on a dataset including fMRI recordings of 10 focal epilepsy patients and 13 healthy controls. It reached 77% specificity, indicating that the proposed technique reliably selects ICs related to epileptic activity. Indeed, in the vast majority of patients where a selection was made, the selected IC overlapped either with the SOZ or the GLM-based map.

Considering the small size of the patient group, these results are only preliminary, however, promising. The performance of our technique is in line with the results obtained with another fully blind approach [8] using an activelet based representation and spatiotemporal clustering of the fMRI voxel time series. This approach performed very well on fMRI runs with only a few epileptic events as these data matched the sparsity assumption of the method. However, in a group of runs containing more than five events, 69% of the obtained activation maps were concordant with GLM activation maps. Our technique, tested on patients who showed plenty of IEDs, provided maps concordant with the GLM activation maps or with the SOZ in 5 out of 7 cases (71%).

Although we emphasized in [1] that multiple epileptic ICs may exist in each patient, we insist on selecting exactly one. The goal is to pinpoint an epileptic IC which can help identifying the SOZ without any further visual inspection or prior information from other modalities. An inherent limitation of our algorithm in its current stage is that it cannot differentiate between initial and propagated activity or sources corresponding to different IED types. Allowing the selection of more ICs per patient presumably increases the sensitivity of our method. However, manual intervention would be necessary to determine the one corresponding to the SOZ based on some prior knowledge, e.g. the spike field of the preponderant spike type. In order to overcome this limitation, future work will aim at characterizing the differences between ICs representing initial (preponderant) and propagated (secondary) activity.

Note that the dataset used in this study consisted of patients where IEDs were recorded in the EEG. However, our methodology is especially relevant in cases where no epileptic activity is present, hence, GLM-based analysis cannot be carried out. Therefore, the proposed blind selection method should be tested on a dataset of EEG-negative cases as well. We expect that the proposed technique will pinpoint the SOZ in at least a few such cases, which would yield a considerable improvement over the state of the art.

## References


[1] B. Hunyadi, S. Tousseyn, B. Mijović, P. Dupont, S. Van Huffel, W. Van Paesschen, and M. De Vos, "Ica extracts epileptic sources from fmri in eeg-negative patients: A retrospective validation study," *PLoS ONE*, vol. 8, no. 11, p. e78796, 11 2013.

[2] M. Zijlmans, G. Huiskamp, M. Hersevoort, J.-H. Seppenwoolde, A. C. van Huffelen, and F. S. S. Leijten, "Eeg–fmri in the preoperative work–up for epilepsy surgery," *Brain*, vol. 130, no. 9, pp. 2343–2353, 2007.

[3] P. LeVan, L. Tyvaert, F. Moeller, and J. Gotman, "Independent component analysis reveals dynamic ictal bold responses in eeg-fmri data from focal epilepsy patients," *NeuroImage*, vol. 49, no. 1, pp. 366 – 378, 2010.

[4] R. Rodionov, F. De Martino, H. Laufs, D. Carmichael, E. Formisano, M. Walker, J. Duncan, and L. Lemieux, "Independent component analysis of interictal fmri in focal epilepsy: Comparison with general linear model–based eeg–correlated fmri," *NeuroImage*, vol. 38, no. 3, pp. 488 – 500, 2007.

[5] F. De Martino, F. Gentile, F. Esposito, M. Balsi, F. Di Salle, R. Goebel, and E. Formisano, "Classification of fmri independent components using ic–fingerprints and support vector machine classifiers," *NeuroImage*, vol. 34, no. 1, pp. 177 – 194, 2007.

[6] G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti, and S. M. Smith, "Automatic denoising of functional mri data: Combining independent component analysis and hierarchical fusion of classifiers," *NeuroImage*, 2014.

[7] I. Khalidov, J. Fadili, F. Lazeyras, D. Van De Ville, and M. Unser, "Activelets: Wavelets for sparse representation of hemodynamic responses," *Signal Processing*, vol. 91, no. 12, pp. 2810 – 2821, 2011.

[8] R. Lopes, J. Lina, F. Fahoum, and J. Gotman, "Detection of epileptic activity in fmri without recording the eeg," *NeuroImage*, vol. 60, no. 3, pp. 1867 – 1879, 2012.

[9] N. Hurley and S. Rickard, "Comparing measures of sparsity," in *Machine Learning for Signal Processing, 2008. MLSP 2008. IEEE Workshop on*, oct. 2008, pp. 55 –60.

[10] D. Cordes, V. Haughton, K. Arfanakis, J. Carew, P. Turski, C. Moritz, M. Quigley, and M. Meyerand, "Frequencies contributing to functional connectivity in the cerebral cortex in resting–state data," *American Journal of Neuroradiology*, vol. 22, no. 7, pp. 1326–1333, 2001.

[11] S. Tousseyn, P. Dupont, S. Sunaert, and W. Van Paesschen, "Evaluation of interictal eeg–fmri sensitivity and specificity for detection of the ictal onset zone in refractory focal epilepsy," in *American Epilepsy Society Annual Meeting, Abstract No. 1.196. www.aesnet.org*, 2012.