Classification of epileptic activity through temporal and spatial characterization of intracranical recordings

Vanessa D'Amario[(1)], Annalisa Barla[(1)], Gabriele Arnulfo[(1)], Lino Nobili[(2)]

(1) DIBRIS, Università degli Studi di Genova, Genova, Italy,
(2) Ospedale Niguarda Ca' Granda, Milano, Italy

Abstract. Focal epilepsy is a chronic condition characterized by hyper-activity and abnormal synchronization of a specific brain region. For pharmacoresistant patients, the surgical resection of the critical area is considered a valid clinical solution, therefore, an accurate localization is crucial to minimize neurological damage. In current clinical routine the characterization of the epileptogenic zone (EZ) is performed using invasive methods, such as Stereo-Electroencephalography (SEEG). Medical experts perform the tag of neural electrophysiological recordings by visually inspecting the acquired data, a highly time consuming and subjective procedure. Here we show the results of an automatic multi-modal classification method for the evaluation of critical areas in focal epileptic patients. The proposed method represents an attempt in the characterization of brain areas which integrates the anatomical information on neural tissue, inferred using Magnetic Resonance Imaging (MRI) in combination with spectral features extracted from SEEG recordings.

1 Introduction

Epilepsy is a neurological disorder characterized by abnormal neural activity that leads to abrupt onset of seizures. Among world population, about 50 million people suffer from generalized epilepsy. For the majority of the affected patients, symptoms can be pharmacologically controlled. Unfortunately, 30% of patients are refractory to medication and, when diagnosed with focal onset, brain surgery can be considered as treatment. In these cases, complex and multimodal investigations are mandatory to accurately localize the Epileptogenic Zone (EZ), defined as the minimum amount of cortex that should be removed to produce seizure-free subjects. While the patient is hospitalized, multiple Magnetic Resonance Imaging tests (MRIs) as well as scalp electroencephalography data are acquired to define putative EZ. Nevertheless, this protocol shows clear evidence of malformations (e.g., tumors or dysplasia) only in a small percentage of patients candidate for surgery. Moreover, even in presence of positive results the border of the EZ or the localization of the onset zone might be elusive. In these cases, neurophysiologists require the acquisition of invasive intracerebral recordings such as Stereo-Electroencephalography (SEEG) [3]. It consists in the implantation in the brain tissue of depth filiform electrodes, whose number depends on the severity of the case, each endowed with several acquisition channels, that record local field potential at high sampling frequency.

Clinicians then perform a very time-consuming and highly subjective visual inspection on the signal acquired from each channel, looking for epileptic biomarkers, such as spike or spike-and-wave patterns and characterizing the relationship between brain regions by co- or lagged-occurrence of these pathological patterns [6]. This tagging procedure is a time- and resource-consuming task, with medical experts spending, on average, about two hours for the analysis of a 10 minutes neural activity recording. Moreover, even if SEEG is a highly precise acquisition method, surgical resection, does not lead to positive outcomes in a relevant portion of patients [17]. Among the possible

reasons of this unsatisfactory success rate there is the highly subjectivity due to the EZ identification procedure.

Therefore, it is clear that the definition of an automatic tool for the detection of the pathological tissue may prove as a great advancement in this context.

1.1   Related works

Most of the available works are not only restricted to the localization of the epileptogenic areas but also to seizure prediction.

Since the advent of new medical devices, developed to monitor the neural activity and to forecast critical events, high interest concerns the transition from *interictal* to *ictal* states, that is the change from a normal state to a seizure. Motivated by this, a great deal of recent literature focuses on automatic seizure detection tools. We mention, for instance the Kaggle challenge on the Melbourne dataset[1].

State of the art methods both for classification epileptogenic areas and seizure forecasting perform a feature extraction stage. The scientific community agrees in the importance of spectral quantities both in temporal and frequency domain as descriptors of the neurological signal. Relative amount of signal power in frequency bands of interests and temporal events characterized by critical amplitudes are indeed considered pathological biomarkers [19, 20, 1, 15]. In Truong [19] et al. the authors analyze the Melbourne dataset using standard spectral analysis methods as Fast Fourier Transform (FFT) to measure correlations across channels, in order to infer which are the mostly involved in the seizure generation. In the work of Vila et al. [20] some criteria for the localization of Seizure Onset Zone (SOZ) are established using spectral measures. The study is performed on seven patients, and characterizes the transition between interictal and ictal states. Mean activation measure, defined as the average of the instantaneous activity during the seizure epochs and relative time average power of every channel is shown to be significatly higher in SOZ across all patients in the α (8-13 Hz) and β (13-30 Hz) rhythms. Oscillations in the β and γ (30-70 Hz) range, rapid discharges, spectral and temporal aspects are taken into account to discriminate SOZ from physiological areas. The presence of rapid discharges in a given brain area immediately before the ictal state is also shown to be a good measure of epileptogenicity of the zone [1]. In this regard significant changes of the activity are evaluated using thresholds. The authors demonstrate a statistical correlation between duration of high energy phenomena and epileptogenicity of the area.

In this work we define a feature extraction pipeline which leverages on spectral features in line with the ones defined above. In particular we are interested in (i) implementing the feature extraction in such a way to obtain a characterization of the neural signal independent from the specific patient, so to realize an automatic classification method across patients, (ii) including anatomical knowledge derived from imaging test (MRI) and (iii) measuring the classification performances of our pipeline in the analysis of interictal signal.

The paper is organized as follows: Section 2 provides a description of the SEEG and MRI dataset, the feature extraction pipeline and the machine learning methods used for the analysis. Section 3 regards the statistics and the obtained results, in Section 4 we conclude by describing the ongoing work on epileptic signal classification through network analysis.

2   Materials and methods

We acquired the dataset at the Hospital Niguarda Neurology Unit (Milan, Italy). Patients provided written consent for the analysis of the data. The dataset consists of SEEG and MRI data for forty patients and SEEG data only for another set of twenty

---

[1]`https://www.kaggle.com/c/melbourne-university-seizure-prediction`

patients.

We registered local field potential with common reference in white matter, using platinum-iridium, multi-lead electrodes. The number of contacts for each electrodes varies from 8 to 15, each is 2 mm long, 0.8 mm of thickness and have distance of 1.5 mm from its neighbours (DIXI medical, Besancon, France). We acquired 10 minutes of spontaneous resting state activity, at a sampling frequency of 1kHz, with eyes closed, using a 192-channels SEEG amplifier system (NIHON-KOHDEN NEUROFAX-110). We automatically ascertained the position of each recording contact using a dedicated segmentation software [14].

We fused MRI-pre with CT-post (Computed Tomography) using affine rigid-body coregistration, [3]. After the coregistration phase, the algorithm automatically segments each contact contained in the multi-lead electrodes by searching its center of mass.

The total amount of channels for this set of patients is 5315, only 1342 have been marked as pathological by an equipe of medical experts. On average, the total number of channels per patient is 140 ± 20 (mean $\pm$ std), of which 34 ± 21 (mean $\pm$ std) are epileptogenic or characterized by critical activity.

For what concerns the SEEG acquisition, the extraction of relevant spectral features is first preceded by a preprocessing stage which consists of two steps: i) local reference of potential, ii) removal of power line effects. We proceed in the former case to the computation of the potential difference of neighbour channels on the same electrode. This local reference of potential is shown to decrease the correlation of spurious electrical activity which propagates through fibers in the white tissue [12] and for this reason it is preferable than average reference of the potential or other settings. We remove the power line effects through notch filters peaked at 50 Hz and harmonics (Butterworth, 2nd order).

2.1 Time-frequency features

In order to get effective descriptors of the average activity in the interictal stage, we extracted spectral features on temporal windows of 300 seconds. First, we measured basic features as the first moments: *variance*, *skewness* and *kurtosis*.

Then, to capture the variability across patients, we measure the *mean energy* values [15] for the frequency bands of clinical interest δ 1-4 Hz, θ 4-8 Hz, α 8-13 Hz, β 13-30 Hz, γ 30-70 Hz, high-γ 70-90 Hz, using a 2nd order Butterworth filter and dividing every measure at a specific band by the sum of energy across all frequencies, up to the Nyquist frequency (500 Hz). Recent results [11, 5, 7] show a relevant contribution of high frequency patterns in determining the pathological state. To this aim, we include in the analysis mean energy values at frequencies higher than high-γ band. Band pass filters of width 50 Hz has been used, spanning the frequency space from high-γ up to Nyquist frequencies.

We also evaluate the *wavelet entropy* measure, which has been shown to be a discriminative quantity in the evaluation of signal coherence in neurophysiology [16], especially in pathological activity detection [13]. We use an orthogonal discrete mother wavelet (Daubechies, 2nd order) to perform the decomposition of the signal.

As in Bartolomei et al. [1] we also measure the *hyperactivity* of each channel, defined as abnormal signal amplitudes, with respect to the baseline activity. To distinguish between baseline and hyperactivity, we set different threshold values on the filtered signal at different bands. We compute the absolute value for each outcome in the temporal domain and estimate the length of hyperactivity periods. The thresholds were learned on the 20 subjects with only the SEEG data available, we excluded this subset of patients from the classification pipeline, in order to prevent from potential overfit issues. From these recordings, we consider the 1968 physiological channels to estimate an adaptive signal baseline. We first filter these recordings in the frequency bands defined above,

band f	$\langle \sigma_f \rangle$ [μV]
δ	13.19
θ	15.54
α	12.45
β	11.09
γ	4.58
high-γ	1.19
90-140 Hz	1.07
140-190 Hz	$5.78 \cdot 10^{-1}$
190-240 Hz	$3.91 \cdot 10^{-1}$
240-290 Hz	$2.98 \cdot 10^{-1}$
290-340 Hz	$2.44 \cdot 10^{-1}$
340-390 Hz	$2.03 \cdot 10^{-1}$
390-440 Hz	$1.83 \cdot 10^{-1}$
440-490 Hz	$1.90 \cdot 10^{-1}$

Table 1: Mean values of $\langle \sigma_f \rangle$ for different energy bands f evaluated on 1968 physiological channels for which the MR images where not available. These values can be considered as a physiological standard activity across patients

and consider the standard deviation of the activity at each band. We denote by $\langle \sigma_f \rangle$ the mean standard deviation for the band f for this subset of patients. We then computed $\langle \sigma_f \rangle$ over the channels. The values relative to each band are shown in Table 1. Then we measure the time spent over a variable threshold $a \cdot \sigma_f$, where a assumes discrete values in the range $[2, 7]$, by first filtering the time series in the band f.

2.2 Spatial feature

To the best of our knowledge there has been no attempt to integrate anatomical quantities together with spectral features in automatic learning pipelines for detection of critical areas.

With this regard, the recent work of Mercier [12] gives a broad insight in the role played by white matter in the signal propagation through the brain. He points out the improvement in the analysis of brain potential obtained by computing local difference of potential for neighboring channels, in order to decouple spurious activity. His work also shows the relevance of quantifying the anatomical nature of the brain tissue in the area of acquisition of the signal.

The characterization of brain regions is based on the differentiation of gray and white matter evaluated through *FreeSurfer* [8], a software tool which parcellates cortical and subcortical regions from MRI acquisition. By defining each voxel imaging as 1 mm^3, the Partial Tissue Density (PTD) index is defined as follows

$$\text{PTD} = \frac{\text{Vox Gray} - \text{Vox White}}{\text{Vox Gray} + \text{Vox White}}$$

where Vox Gray and Vox White correspond respectively to the number of gray and white voxels contained in a volume of $3 \times 3 \times 3$ mm^3 centered around the electrode position [12]. Indeed in the same work, the electrode position is proved to be crucial for clinical evaluations, as there is a high correlation between signal power and PTD index, and signal amplitude is greater in the gray matter than in the white matter regions. It is well known in the clinical routine that the amplitude of the signal at low frequencies is a pathological biomarker of the epileptic activity, which is originated by the excitation of neural population, localized in the gray matter only. The PTD quantifies the proximity between the gray matter and white matter and assumes continuous values between [-1, 1].

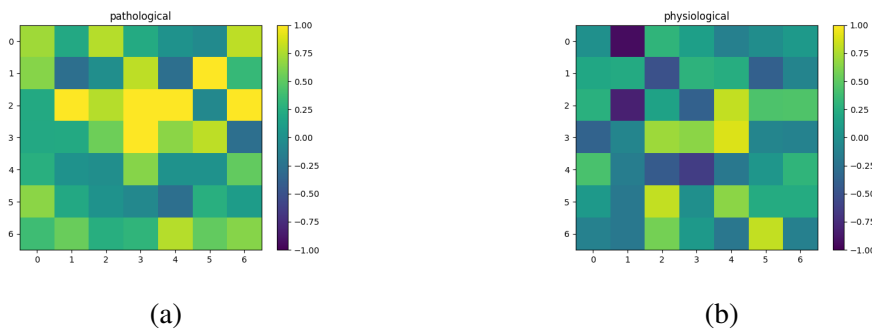(a)   (b)

Figure 1: Plot of PTD values for a single patient, with a common number of epileptic and non epileptic channels. The colormap identifies the PTD index. In (a), channels tagged as pathological, in (b) there are instead PTD values relative to physiological channels. The matrix of non epileptic channels contains also low values of PTD, that can be easily excluded from the analysis, given the absence of gray matter in those areas.

2.3   Data representation

Each sample collects the features computed for a fixed channel position for a window of 5 minutes of activity. By applying the feature extraction pipeline we get the data matrix used for classification, which contains 10630 samples, each described by 156 time-spatial features.

2.4   Machine learning methods

We consider several machine learning techniques for classification, both linear and non-linear. In particular, we use sparse Logistic Regression (LR) [10], Support Vector Machines with linear kernel (SVMs) [4], Random Forest (RF) [2] and Gradient Boosting (GB) [9]. All the learning methods require the tuning of hyperparameters, which is performed through cross validation.

We used different metrics for the evaluation of results, performed on the test set. From the confusion matrix we extracted the number of false positive (FP), false negative (FN), true positive (TP), true negative (TN) which are needed to compute precision, recall, balanced accuracy and F1 score.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{balanced accuracy} = \frac{1}{2} \cdot \left(\frac{\text{TP}}{\text{TP} + \text{FN}} + \frac{\text{TN}}{\text{TN} + \text{FP}} \right)$$

$$\text{F1 score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

3   Results

The proportion of epileptic and non epileptic channels is unbalanced in favor of non epileptic channels, with random guess corresponding to 0.74%. For this reason, we computed the performance of our methods using metric scores that take into account the unbalancedness of the dataset. We split the dataset in 80% samples for training and 20% for test, using automatic *scikit-learn* procedures that split the dataset with respect to the unbalancedness of the original problem. The choice of the optimal hyperparameters for all the algorithms was performed by using three-fold cross-validation in the learning procedure.

In LR we imposed sparsity through the L_1 norm on the regularization term, with the regularization constant C varying in a logarithmically spaced range of twenty values between $(10^{-2}, 10^2)$.

For SVM, we fixed a linear kernel and let the cross validation choose the best values of C, in the same range of LR.

For what concerns RF we fixed the number of estimators to 10^3 where the tunable parameters were the percentage of maximum features with respect to the total, in the range $(0.1, 0.2, 0.3, 0.4, 0.5, 0.6)$.

In GB we fixed the learning rate to 10^{-3}, the tunable parameters were the max depth of trees, free to vary in a linearly space interval between $(3, 31)$ and the number of estimators, chosen between three linearly spaced values in $(100, 500)$.

The learning and testing procedures were repeated 50 times in order to get a statistically reliable outcome for the four classifiers.

classifier	precision	recall
LR	0.68 ± 0.03	0.34 ± 0.03
SVM	0.57 ± 0.04	0.15 ± 0.03
RF	$\mathbf{0.88 \pm 0.02}$	$\mathbf{0.57 \pm 0.02}$
GB	0.80 ± 0.07	0.35 ± 0.06
	balanced accuracy	F1 score
LR	0.64 ± 0.01	0.45 ± 0.03
SVM	0.55 ± 0.01	0.23 ± 0.05
RF	$\mathbf{0.77 \pm 0.01}$	$\mathbf{0.69 \pm 0.01}$
GB	0.66 ± 0.02	0.48 ± 0.05

Table 2: Scores of different learning algorithms trained on the extracted features. The values corresponds to mean $\pm$ std obtained from 50 repetition of the learning pipeline. Random forest classifier gives the best performances for all the score metrics used.

The results obtained from the testing phase of the machine learning algorithms introduced in previous section are shown in Table 2. Random Forest performs best for all the considered metrics. The values of balanced accuracy and F1 score are highlighted. Both metrics show a performance which is highly above chance level and is promising in the discrimination of epileptic areas. The precision value for RF indicates that the number of false positives is relatively low. This can help in the discrimination of epileptic areas, because areas classified as pathological have high probability to belong to that class.

4 Conclusions and future work

In this work we defined a pipeline for the analysis of brain activity in focal epileptic patients during the interictal period, with the aim of localizing critical areas, involved in seizure generation and propagation. By considering relative measures based on single patient, several machine learning methods have been trained across patients, all of them with results highly above chance. In particular random forest classifiers has been shown to achieve the best performances (balanced accuracy $= 0.77 \pm 0.01$, F_1 score $= 0.69 \pm 0.01$). To the best of our knowledge, the integration of spectral features with anatomical characteristics of the recorded areas represents a first attempt to merge multiple tests results, fixing a set of features which are both functional and structural descriptors of the epileptic brain. In the analysis we have considered long interictal period, segments of 5 minutes at high sampling frequency and the features extracted represents an average behavior of each area.

In the future, we plan to improve this multi-modal approach in order to enrich the description of the brain areas for the classification task. We will characterize the brain dynamics more in detail, by considering chuncks of smaller length. For this scope, we will resort to graphical methods [18], which we again will combine to structural features extracted from MRI tests.

References

[1] Fabrice Bartolomei, Patrick Chauvel, and Fabrice Wendling. Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral eeg. *Brain*, 131(7):1818–1830, 2008.

[2] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[3] Francesco Cardinale, Massimo Cossu, Laura Castana, Giuseppe Casaceli, Marco Paolo Schiariti, Anna Miserocchi, Dalila Fuschillo, Alessio Moscato, Chiara Caborni, Gabriele Arnulfo, et al. Stereoelectroencephalography: surgical methodology, safety, and stereotactic application accuracy in 500 procedures. *Neurosurgery*, 72(3):353–366, 2012.

[4] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[5] Benoît Crépon, Vincent Navarro, Dominique Hasboun, Stéphane Clemenceau, Jacques Martinerie, Michel Baulac, Claude Adam, and Michel Le Van Quyen. Mapping interictal oscillations greater than 200 hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*, 133(1):33–45, 2009.

[6] Liset Menendez de la Prida, Richard J Staba, and Joshua A Dian. Conundrums of high-frequency oscillations (80-800 hz) in the epileptic brain. *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society*, 32(3):207, 2015.

[7] Matthias Dümpelmann, Julia Jacobs, Karolin Kerber, and Andreas Schulze-Bonhage. Automatic 80–250hz ripple high frequency oscillation detection in invasive subdural grid and strip recordings in epilepsy by a radial basis function neural network. *Clinical Neurophysiology*, 123(9):1721–1731, 2012.

[8] Bruce Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.

[9] Jerome H Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

[10] Trevor Hastie, Robert Tibshirani, and Martin Wainwright. *Statistical learning with sparsity: the lasso and generalizations*. CRC press, 2015.

[11] Jean-Philippe Lachaux, Nikolai Axmacher, Florian Mormann, Eric Halgren, and Nathan E Crone. High-frequency neural activity and human cognition: past, present and possible future of intracranial eeg research. *Progress in neurobiology*, 98(3):279–301, 2012.

[12] Manuel R Mercier, Stephan Bickel, Pierre Megevand, David M Groppe, Charles E Schroeder, Ashesh D Mehta, and Fred A Lado. Evaluation of cortical local field potential diffusion in stereotactic electro-encephalography recordings: A glimpse on white matter signal. *NeuroImage*, 147:219–232, 2017.

[13] Anne H Mooij, Birgit Frauscher, Mina Amiri, Willem M Otte, and Jean Gotman. Differentiating epileptic from non-epileptic high frequency intracerebral eeg signals with measures of wavelet entropy. *Clinical Neurophysiology*, 127(12):3529–3536, 2016.

[14] Massimo Narizzano, Gabriele Arnulfo, Serena Ricci, Benedetta Toselli, Martin Tisdall, Andrea Canessa, Marco Massimo Fato, and Francesco Cardinale. Seeg assistant: a 3dslicer extension to support epilepsy surgery. *BMC bioinformatics*, 18(1):124, 2017.

[15] Ibrahim Omerhodzic, Samir Avdakovic, Amir Nuhanovic, and Kemal Dizdarevic. Energy distribution of eeg signals: Eeg signal wavelet-neural network classifier. *arXiv preprint arXiv:1307.7897*, 2013.

[16] Osvaldo A Rosso, Susana Blanco, Juliana Yordanova, Vasil Kolev, Alejandra Figliola, Martin Schürmann, and Erol Başar. *Journal of neuroscience methods*, 105(1):65–75, 2001.

[17] Susan Spencer and Linda Huh. Outcomes of epilepsy surgery in adults and children. *The Lancet Neurology*, 7(6):525–537, 2008.

[18] F. Tomasi, V. Tozzo, S. Salzo, and A. Verri. Latent variable time-varying network inference. *ArXiv e-prints*, February 2018.

[19] Nhan Duy Truong, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Andrew Faulks, and Omid Kavehei. Supervised learning in automatic channel selection for epileptic seizure detection. *Expert Systems with Applications*, 86:199–207, 2017.

[20] Manel Vila-Vidal, Alessandro Principe, Miguel Ley, Gustavo Deco, Adrià Tauste Campo, and Rodrigo Rocamora. Detection of recurrent activation patterns across focal seizures: Application to seizure onset zone identification. *Clinical Neurophysiology*, 128(6):977–985, 2017.