BIOCYBERNETICS AND BIOMEDICAL ENGINEERING XXX (2018) XXX–XXX

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

# Fast statistical model-based classification of epileptic EEG signals

*Antonio Quintero-Rincón* [a,*], *Marcelo Pereyra* [b], *Carlos D'Giano* [c], *Marcelo Risk* [a,d], *Hadj Batatia* [e]

[a] *Department of Bioengineering, Instituto Tecnológico de Buenos Aires (ITBA), Av. Eduardo Madero 399, C1106ACD Buenos Aires, Argentina*
[b] *School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK*
[c] *Centro Integral de Epilepsia y Telemetría, Fundación Lucha contra las Enfermedades Neurológicas Infantiles (FLENI), Montañeses 2325, C1428AQK Buenos Aires, Argentina*
[d] *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina*
[e] *University of Toulouse, IRIT – INPT, Toulouse, 2 rue Charles Camichel, B.P. 7122, 31071 Cedex 7, France*

ARTICLE INFO

*Article history:*
Received 5 March 2018
Received in revised form
29 July 2018
Accepted 2 August 2018
Available online xxx

*Keywords:*
Generalized Gaussian distribution
Wavelet filter banks
EEG
Epilepsy
Kolmogorov–Smirnov
Cramer-von-Mises
Leave-one-out cross-validation
Linear classifier

ABSTRACT

This paper presents a supervised classification method to accurately detect epileptic brain activity in real-time from electroencephalography (EEG) data. The proposed method has three main strengths: it has low computational cost, making it suitable for real-time implementation in EEG devices; it performs detection separately for each brain rhythm or EEG spectral band, following the current medical practices; and it can be trained with small datasets, which is key in clinical problems where there is limited annotated data available. This is in sharp contrast with modern approaches based on machine learning techniques, which achieve very high sensitivity and specificity but require large training sets with expert annotations that may not be available. The proposed method proceeds by first separating EEG signals into their five brain rhythms by using a wavelet filter bank. Each brain rhythm signal is then mapped to a low-dimensional manifold by using a generalized Gaussian statistical model; this dimensionality reduction step is computationally straightforward and greatly improves supervised classification performance in problems with little training data available. Finally, this is followed by parallel linear classifications on the statistical manifold to detect if the signals exhibit healthy or abnormal brain activity in each spectral band. The good performance of the proposed method is demonstrated with an application to paediatric neurology using 39 EEG recordings from the Children's Hospital Boston database, where it achieves an average sensitivity of 98%, specificity of 88%, and detection latency of 4 s, performing similarly to the best approaches from the literature.



* *Corresponding author at*: Department of Bioengineering, Instituto Tecnológico de Buenos Aires (ITBA), Av. Eduardo Madero 399, C1106ACD Buenos Aires, Argentina.
E-mail address: aquinter@itba.edu.ar (A. Quintero-Rincón).

https://doi.org/10.1016/j.bbe.2018.08.002
0208-5216/

Please cite this article in press as: Quintero-Rincón A, et al. Fast statistical model-based classification of epileptic EEG signals. Biocybern Biomed Eng (2018), https://doi.org/10.1016/j.bbe.2018.08.002

## 1. Introduction

Q4 Epilepsy is a disease that produces brain activity disorders [1–3]. Its diagnosis relies strongly on the analysis of electroencephalography (EEG) data, a non-invasive and widely available biomedical modality that allows neurologists to monitor abnormal activity and characterize its nature. In particular, neurologists analyze characteristic waveforms to localize and quantify the epileptogenic zone. These brain activity disorders can lead to epileptic seizures that have a sudden onset, spread quickly, and are very brief.

Epileptic seizure detection methods based on EEG signals stem from the observation that EEG signal descriptors allow discriminating between normal and abnormal brain activity. This practice originated half a century ago with works by Viglione et al. [4], Liss et al. [5], Ktonas et al. [6] and Gotman et al. [7]; and continued with Iasemidis et al. [8,9] mainly in the medical literature and by using analogue EEG devices. Later, as EEG systems adopted digital signal processing capacity, this stimulated the development of pattern recognition methods to detect and analyse abnormal brain activity automatically. A main practical advantage of EEG technology is that it is very economically accessible. This has significantly contributed to the wide adoption of EEG in developing countries (whereas other more advanced modalities, such as magnetoencephalography (MEG), are expensive and have not been widely adopted as a result).

There are currently a wide range of EEG signal processing methods to detect brain seizures accurately. Most methods use classification techniques from the supervised machine learning literature, such as support vector machines [10–12] and discriminant analysis [13], and differ mainly in terms of their feature extraction methods and the features classification approaches. Many methods use time-frequency descriptors, either explicitly (e.g., short-term Fourier or wavelet representations) [14–17,11,18,19], empirical mode decomposition [20–22], or implicitly by learning neural networks [23–25] or by using component analysis or common spatial patterns (see for example [26–28]). Some also use statistical descriptors such as signal entropy [17,11,29–31] or fractal dimension [32,33].

The main approaches from the state of the art are summarised in Table 1, together with their detection performance on a test dataset. Observe that most modern methods perform remarkably well and achieve true positive rates (TPR) or sensitivities of the order of 95–99%, and true negative rates or specificities of the order of 85–95%, depending on the specific method and dataset considered. This good performance is achieved by using advanced signal processing techniques that are generally very computationally intensive. As a result, state-of-the-art detection methods cannot be incorporated into EEG devices to perform detection in real time. For example, the method [26] uses common spatial patterns that require estimating covariance matrices and performing singular value decompositions at each detection step. This limitation is motivating the development of detection methods that use cloud computing technology to perform detection on a high performance computing server that is accessed remotely (see for example [28]). This strategy is potentially very interesting in some settings, but it would be difficult to implement in developing countries where many hospitals still have limited Internet access and poor IT infrastructure.

Another limitation of state-of-the-art methods is that they pull information from all spectral bands to improve detection performance [26]. While beneficial in terms of classification accuracy, this can be problematic in many clinical applications where the current practice is to detect seizures independently in each physiological spectral band or *brain rhythm* (these bands are specified in Section 2). Finally, state-of-the-art methods also rely increasingly on large training datasets, which is a drawback in clinical applications where there is limited annotated data available. Also, many existing methods use feature-based classification techniques, with a significant number of features in order to handle the inherent variability of such features.

This paper seeks to address these limitations of the existing methodology by developing an automatic EEG detection technique that has low computational cost, that performs detection independently in each brain rhythm following current clinical practice, and that can be trained with small datasets, with a detection performance that is similar to that of state-of-the-art algorithms. In contrast with existing methods, the proposed method adopts a model-based classification approach. Model-based classification has been used in various applications [34–36]. The idea is to capture the statistical properties of the signal using the parameters of a probabilistic model. This approach is interesting compared to feature-based classification, especially when features are numerous or exhibit large variability. It can be viewed as an interesting dimensionality reduction technique facing the curse of dimensionality and leading to low computational cost classification. Despite its interest, this approach has not been widely investigated in EEG signal processing. Precisely, our classification method is driven by a parametric statistical model that captures the statistical properties of the signals and their evolution in time, with the model parameters acting as classification features. This approach is an interesting alternative to the non-parametric features (e.g., signal power spectrum, variance, entropy, etc.) commonly used in the literature because the parametric structure of the model acts as a dimensionality reduction mechanism that regularizes the classification problem and consequently improves the stability and robustness of the classification, and which at the same time significantly reduces the associated computational cost. Despite its advantages, to the best of our knowledge this promising approach has not been investigated for EEG signal classification. Note however that statistical approaches have been successfully applied to other challenging EEG processing problems (see for example [37,38]).

The remainder of the paper is structured as follows. Section 2 introduces notation and specifies the detection problem considered. Section 3 presents the proposed method, with its three main steps detailed in Sections 3.1–3.3. Section 4 presents a range of experimental results with EEG recordings from the Children's Hospital Boston database and reports detection performance in terms of sensitivity, specificity, and latency. Advantages and limitations, Conclusions and perspectives for future work are finally reported in Sections 5 and 6 respectively.



Please cite this article in press as: Quintero-Rincón A, et al. Fast statistical model-based classification of epileptic EEG signals. Biocybern Biomed Eng (2018), https://doi.org/10.1016/j.bbe.2018.08.002