Multivariate Bayesian classification of epilepsy EEG signals

Antonio Quintero Rincon, Jorge Prendes, Marcelo Alejandro Pereyra, Hadj Batatia, Marcelo Risk

MULTIVARIATE BAYESIAN CLASSIFICATION OF EPILEPSY EEG SIGNALS

Antonio Quintero-Rincón[1], *Jorge Prendes*[2], *Marcelo Pereyra*[3], *Hadj Batatia*[2], *Marcelo Risk*[1,4]

[1]Deparment of Bioengineering, Instituto Tecnológico de Buenos Aires, Buenos Aires, Argentina.
[2]University of Toulouse, IRIT-INP-ENSEEIHT, Toulouse, France.
[3]Department of Mathematics, University of Bristol, University Walk, Bristol, U.K.
[4]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

ABSTRACT

The classification of epileptic seizure events in EEG signals is an important problem in biomedical engineering. In this paper we propose a Bayesian classification method for multivariate EEG signals. The method is based on a multilevel 2D wavelet decomposition that captures the distribution of energy across the different brain rhythms and regions, coupled with a generalised Gaussian statistical representation and a multivariate Bayesian classification scheme. The proposed approach is demonstrated on a challenging paediatric dataset containing both epileptic events and normal brain function signals, where it outperforms a state-of-the-art method both in terms of classification sensitivity and specificity.

Index Terms— Bayesian classifiers, Epilepsy, Multilevel 2D wavelet , Generalized Gaussian distribution, EEG.

1. INTRODUCTION

Seizures and epilepsy are clinical phenomena resulting from the hyperexcitability of neurons [1]. The electroencephalogram (EEG) is the predominant modality to study abnormal cerebral activity and diagnose epilepsy and inform its treatment. In particular, it is the main modality to classify epilepsy conditions, analyze epilepsy syndromes, and perform seizure onset detection and propagation analysis.

The methods to analyse epileptic seizure signals can be classified into univariate or multivariate approaches. Univariate approaches analyze the state of a single brain region, while multivariate analyse many regions simultaneously as well as their interactions [2]. EEG signal classifiers play a particularly important role in EEG signal processing. Classification is based on features extracted from single channels, multiple channels or a combination of these [3].

Different feature sets have been investigated in the literature [4–7]. The predominant approach is to use ad-hoc features that describe the time-frequency properties of the EEG signals, as this is related to the brain activity at the different brain rhythms of frequency bands. Recent works have formalised this approach by using a wavelet representation to decompose the univariate EEG signal into the different brain rhythms, followed by statistical modelling of the wavelet coefficients [8, 18]. The works [8, 18] consider different statistical models and conclude that the generalised Gaussian provides the better model fit to data. In this paper we extend the approach by using a 2D wavelet representation coupled with a Bayesian classification scheme [5–7, 9] to operate with multivariate EEG signals so as to analyse several brain regions simultaneously.

The remainder of the paper is organised as follows. Section 2 describes the proposed methodology that combines a 2D wavelet representation, a generalised Gaussian statistical model, and a Bayesian classification scheme for multivariate EEG data. In Section 3 the proposed methodology is applied to real EEG signals from patients suffering from epileptic seizures. Conclusions and perspectives for future work are finally reported in Section 4.

2. METHODOLOGY

Let $\boldsymbol{X} \in \mathbb{R}^{N\times M}$ denote the matrix gathering M EEG signals $\boldsymbol{x}_m \in \mathbb{R}^{N\times 1}$ measured simultaneously on different channels and at N discrete time instants. We use the representation [10]

$$\boldsymbol{X} = \boldsymbol{K}\,\boldsymbol{J} + \boldsymbol{\nu} \tag{1}$$

where $\boldsymbol{J}$ is a matrix representing the sources, $\boldsymbol{K}$ is the so called lead field or gain matrix, and $\boldsymbol{\nu}$ is an additive noise.

The proposed methodology is composed of three stages. The first stage splits the original signal $\boldsymbol{X}$ in to set of non-overlapping 2 seconds segments using a rectangular sliding window so that

$$\boldsymbol{X}^{(i)} = \boldsymbol{\Omega}^{(i)}\boldsymbol{X} \tag{2}$$

$$\boldsymbol{\Omega}^{(i)} = \left[\mathbf{0}^{L\times iL}, \boldsymbol{I}^{L\times L}, \mathbf{0}^{L\times N-iL-L}\right] \tag{3}$$

where $\mathbf{0}^{N\times M} \in \mathbb{R}^{N\times M}$ is the null matrix, $\boldsymbol{I}^{N\times N} \in \mathbb{R}^{N\times N}$ is the identity matrix and L is the number of measurement obtained in 2 seconds. The second stage consists of representing each segment $\boldsymbol{X}^{(i)}$ using a time-frequency Daubechies 2D wavelet decomposition [11] with 6 scales. The purpose of this decomposition is to evaluate the energy distribution

Part of this work was funded by a) the *SuSTaIN* program, EPSRC grant EP/D063485/1 at the Department of Mathematics, University of Bristol and b) *ITBACyT* grant DP/No.557/2015, ScientificTechnical Activities at Research Department, Instituto Tecnológico de Buenos Aires. Marcelo Pereyra holds a Marie Curie Intra-European Fellowship for Career Development.

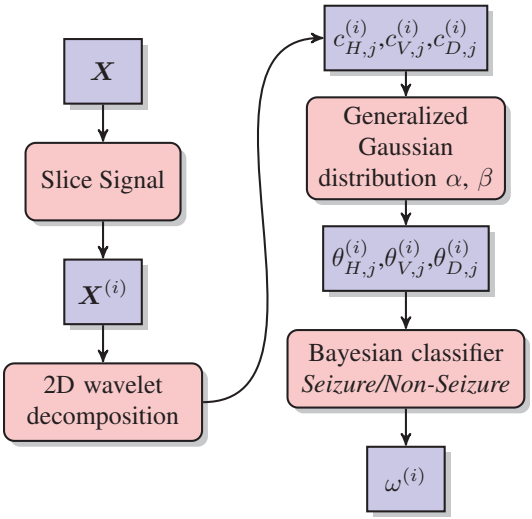


Fig. 1. Algorithm used in the methodology proposed.

throughout the neurological frequency spectrum or brain rhythm (namely the *delta, theta, alpha, beta* and *gamma* bands [12, 13]). Finally, in the third stage, the statistical distribution of the wavelet coefficients is represented by using a zero-mean generalized Gaussian distribution. Precisely, each scale of the wavelet decomposition is summarised by estimating the statistical parameters α and β of the generalized Gaussian distribution [14] to obtain the proposed feature set $\boldsymbol{\theta}^{(i)}$ collecting the parameters associated with all wavelet scales for a 2-second segment.

Finally, the feature vector associated with each time segment is classified by using a Bayesian classifier as *Seizure/Non-Seizure*. Here we use a multivariate Gaussian classifier [15] which provides a robust second order approximation to more general Bayesian classification methods and which has the important advantage or requiring little training data. This 3-stage methodology is summarised in Figure 1.

We now introduce the 2D wavelet decomposition, the generalised Gaussian statistical model, and the Bayesian classifier used in this paper.

2.1. Multilevel 2D Wavelet Decomposition

Wavelets are localized waves, where instead of oscillating forever, they drop to zero; they come from the iteration of filters with scaling [16]. They are obtained from a single prototype "mother" wavelet $\psi(t)$ by rescaling and shifting, i.e.,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\Big(\frac{t-b}{a}\Big) \tag{4}$$

where a is the scaling parameter and b is the shifting parameter. The 1D wavelet transform is given by

$$W_f(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}(t)\mathrm{dt} \tag{5}$$

The discrete wavelet transform (DWT) transforms a discrete time signal to a discrete wavelet representation. It converts an input series $\boldsymbol{x} = [x_0, \ldots, x_{L-1}]^T$ of length L, into one high-pass ($\boldsymbol{h}$) wavelet coefficient series and one low-pass ($\boldsymbol{l}$) wavelet coefficient series, each one of length $\frac{L}{2}$, given by

$$h_j = \sum_{k=0}^{K-1} x_{2j-k}\, s_k, \;\; l_j = \sum_{k=0}^{K-1} x_{2j-k}\, t_k \;\; \forall\, 0 \le j < \frac{L}{2} \tag{6}$$

where $\boldsymbol{s} = [s_0, \ldots, s_{K-1}]^T$ and $\boldsymbol{t} = [t_0, \ldots, t_{K-1}]^T$ are called the wavelet filters.

Recalling that $X^{(i)}$ represents a multichannel signal where each column contains a different channel and each row represents the temporal evolution of the EEG signal, the multilevel 2D wavelet transform decomposes the matrix $\boldsymbol{X}^{(i)}$ using (6) into four component matrices, namely $\boldsymbol{LL}^{(i)}_j$, $\boldsymbol{LH}^{(i)}_j$, $\boldsymbol{HL}^{(i)}_j$ and $\boldsymbol{HH}^{(i)}_j$, where the first letter corresponds to applying either a low-pass (L) or high-pass (H) frequency operation to the temporal component (rows) of $\boldsymbol{X}^{(i)}$ and the second letter refers to the filter applied to the channel component (columns) of $\boldsymbol{X}^{(i)}$, each one according the scale j. The lowest frequency sub-band $\mathbf{LL}^{(i)}_j$ is the approximation coefficients of the original signal $\boldsymbol{X}^{(i)}$. The remaining three frequency sub-bands are the detail parts of the signal and give the vertical high ($\boldsymbol{LH}^{(i)}_j$), horizontal high ($\boldsymbol{HL}^{(i)}_j$) and diagonal high ($\boldsymbol{HH}^{(i)}_j$) coefficients. This process is repeated recursively replacing the input signal $X^{(i)}$ with the last approximation series $\boldsymbol{LL}^{(i)}_j$ until the desired number of scales $j = [1, 2, \ldots, J]^T$ is obtained. We refer the reader to [16, 17] for a comprehensive treatment of the mathematical properties of wavelets and filter banks.

Table 1. Frequencies of the different scales of the DWT.

Decomposed Signal	Frequency range (Hz)	Band
D3 - H3 - V3	32-64	Gamma
D4 - H4 - V4	16-32	Beta
D5 - H5 - V5	8-16	Alpha
D6 - H6 - V6	4-8	Theta
A6	0-4	Delta

Table 1 presents frequencies corresponding to different levels of decomposition for the Daubechies wavelets of order 4 with a sampling frequency of 256 Hz, where H, V and D refer to horizontal, vertical and diagonal details respectively and the number is the scale [12, 13]. The rest of approximations and details are discarded because they are outside of the brain rhythms.

2.2. Generalized Gaussian distribution

The univariate generalized Gaussian distribution (GGD) is a flexible statistical model for one-dimensional signals that has found numerous applications in science and engineering. Their application to epilepsy signal has been studied in [8,18]. Since the wavelet detail coefficients arise from high-pass filtering a zero-mean EEG signal matrix, it can be safely assumed that they also have mean value of zero [15]. Con-

sequently, the distribution of the wavelet coefficients $\boldsymbol{C}_j^{(i)}$ (where $\boldsymbol{C}$ can be one of $\boldsymbol{LH}$, $\boldsymbol{HL}$ or $\boldsymbol{HH}$) can be represented by using a zero-mean GGD statistical model [19] with probability density function (PDF) given by

$$f_{\text{GGD}}(x;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(\beta^{-1})}\exp\left(-\left|\frac{x}{\alpha}\right|^{\beta}\right) \tag{7}$$

where $\alpha \in \mathbb{R}^+$ is a scale parameter and $\beta \in \mathbb{R}^+$ is a shape parameter that controls the shape of the density tail and $\Gamma(\cdot)$ is the Gamma function. Note that the GGD parametric distribution family includes many popular distributions that are commonly used in biomedical signal processing. For example, setting $\beta = 1$ leads to a Laplacian or double-exponential distribution, $\beta = 2$ leads to Gaussian or normal distribution, and $\beta \to \infty$ leads to a uniform distribution (we refer the reader to [20] for a comprehensive treatment of the mathematical properties of the GGD).

From (7), the statistical properties of a wavelet coefficient matrix $\boldsymbol{C}_j^{(i)}$ can be summarized with the maximum likelihood parameter vector $\boldsymbol{\theta}_{\boldsymbol{C}_j^{(i)}}$ such that

$$\boldsymbol{\theta}_{\boldsymbol{C}_j^{(i)}} = \left[\alpha_j^{(i)}, \beta_j^{(i)}\right]^T \tag{8}$$

$$= \arg\max_{[\alpha,\beta]^T} f_{\text{GGD}}(\boldsymbol{C}_j^{(i)};\alpha,\beta). \tag{9}$$

For a detailed explanation on the estimation of the GGD parameters we refer the reader to [14, 19]. Finally, a feature vector summarizing the statistical properties of the wavelet details for each brain rhythm (delta, theta, alpha, beta and gamma [12, 13]) can be obtained as

$$\boldsymbol{\theta}^{(i)} = \left[\boldsymbol{\theta}_3^{(i)}, \boldsymbol{\theta}_4^{(i)}, \boldsymbol{\theta}_5^{(i)}, \boldsymbol{\theta}_6^{(i)}, \boldsymbol{\theta}_{\boldsymbol{LL}_6^{(i)}}^T\right]^T \tag{10}$$

$$\boldsymbol{\theta}_j^{(i)} = \left[\boldsymbol{\theta}_{\boldsymbol{LH}_j^{(i)}}^T, \boldsymbol{\theta}_{\boldsymbol{HL}_j^{(i)}}^T, \boldsymbol{\theta}_{\boldsymbol{HH}_j^{(i)}}^T\right] \tag{11}$$

where $\boldsymbol{\theta}^{(i)}$ is a 13 dimensional vector.

2.3. Multivariate Normal Bayesian Classifier

Consider a classification into J possible classes $\omega_1, \ldots, \omega_J$. For a feature vector $\boldsymbol{\theta}^{(i)}$ belonging to the class ω_j, we assume that $\boldsymbol{\theta}^{(i)}$ has a multivariate normal distribution with mean value $\boldsymbol{\mu}_j$ and covariance matrix $\boldsymbol{\Sigma}_j$, i.e.,

$$\rho\left(\boldsymbol{\theta}^{(i)}\middle|\omega_j\right) = \frac{\exp\left[-\frac{1}{2}(\boldsymbol{\theta}^{(i)}-\boldsymbol{\mu}_j)^T\boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\theta}^{(i)}-\boldsymbol{\mu}_j)\right]}{(2\pi)^{K/2}|\Sigma_j|^{1/2}} \tag{12}$$

where $\rho(\cdot)$ is the probability of a particular event, and K is the size of the vector $\boldsymbol{\theta}^{(i)}$.

The Bayes decision rule states that the estimated class $\hat{\omega}^{(i)}$ corresponding to $\boldsymbol{\theta}^{(i)}$ is

$$\hat{\omega}^{(i)} = \arg\max_j \rho\left(\boldsymbol{\theta}^{(i)}\middle|\omega_j\right)\rho(\omega_j) \tag{13}$$

or equivalently using the logarithmic likelihood we obtain the equivalent rule

$$g_j\left(\boldsymbol{\theta}^{(i)}\right) = \log\rho\left(\boldsymbol{\theta}^{(i)}\middle|\omega_j\right) + \log\rho(\omega_j) \tag{14}$$

$$\hat{\omega}^{(i)} = \arg\max_j g_j\left(\boldsymbol{\theta}^{(i)}\right) \tag{15}$$

where $g_j(\cdot)$ is the so called discriminant function.

From (12) and (14) the discriminant functions becomes

$$g_j\left(\boldsymbol{\theta}^{(i)}\right) = -\frac{1}{2}\left(\boldsymbol{\theta}^{(i)}-\boldsymbol{\mu}_j\right)^T\boldsymbol{\Sigma}_j^{-1}\left(\boldsymbol{\theta}^{(i)}-\boldsymbol{\mu}_j\right) - \frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}_j| + \log\rho(\omega_j) \tag{16}$$

We refer the reader to [21, 22] for a comprehensive treatment of the mathematical properties of Bayesian classifier for multivariate normal distributions.

3. RESULTS

In this section we evaluate the proposed methodology using the Children Hospital Boston database. This dataset consists of 36 bipolar 256Hz EEG recordings from paediatric subjects suffering from intractable seizures [4, 23]. In this work we have used 18 recordings from 11 different subjects. Each recording contains a seizure event, whose onset time has been labeled by an expert neurologist. Here we used the expert annotations to extract a short epoch from each recording such that it is focused on the seizure and that it contains both seizure and non-seizure signals (the epochs have a duration of the order of 5 minutes). The goal is to use these data to train and subsequently test the capacity of our classification scheme to identify seizure and non-seizure signals. Moreover, for comparison we use the state-of-the-art classification method of Shoeb et al. [4], which is also based on energy features of a space-frequency representation of the EEG signal array (here we use an implementation with a 2D wavelet transform to make the comparison fair).

Table 2 reports the performance of each classification method assessed by using a leave-one-out- cross validation approach to calculate the method classification confusion matrix [24] (the results for the proposed method are denoted by Q and those of the method [23] by S). These performance matrices are composed of the following measures that characterise the different aspects of the classifiers: the sensitivity or true positives rate (TPR); the false positive rate (FPR); the sensitivity or true negative rate (TNR); and the overall classification accuracy (ACC), calculated as the total number of correct classifications out of 36 events (18 Seizure and 18 Non-Seizure). Notice that the classification results are performed and reported separately for each brain rhythm or frequency band because this information is relevant to neurologists and allows discriminating clinical events of different nature. To simplify the visual interpretation we highlight in red the method that achieves the highest sensitivity, specificity, and overall accuracy for each frequency band. We observe from Table 2 that the proposed method clearly outperforms the competing approach [4] in terms of overall

Table 2. Comparison between the proposed epilepsy classification method (Q) and the state-of-the-art method [4] (S) using 36 events (18 seizure and 18 non-seizure), for each brain rhythm or frequency band, and using the following performance metrics: sensitivity or true positives rate (TPR); the false positive rate (FPR); the specificity or true negative rate (TNR); and overall classification accuracy (ACC).

	TPR		FPR		TNR		ACC	
Bands	Q	S	Q	S	Q	S	Q	S
Delta	0.83	0.80	0.17	0.27	0.83	0.72	30	27
Theta	0.75	0.76	0.19	0.27	0.80	0.81	28	28
Alpha	0.83	0.75	0.20	0.50	0.79	0.75	29	27
Beta	0.88	0.76	0.14	0.48	0.86	0.76	31	27
Gamma	0.84	0.84	0.15	0.17	0.85	0.82	30	29

accuracy (ACC), and achieves a superior sensitivity (TPR) and specificity (TNR) for most frequency bands.

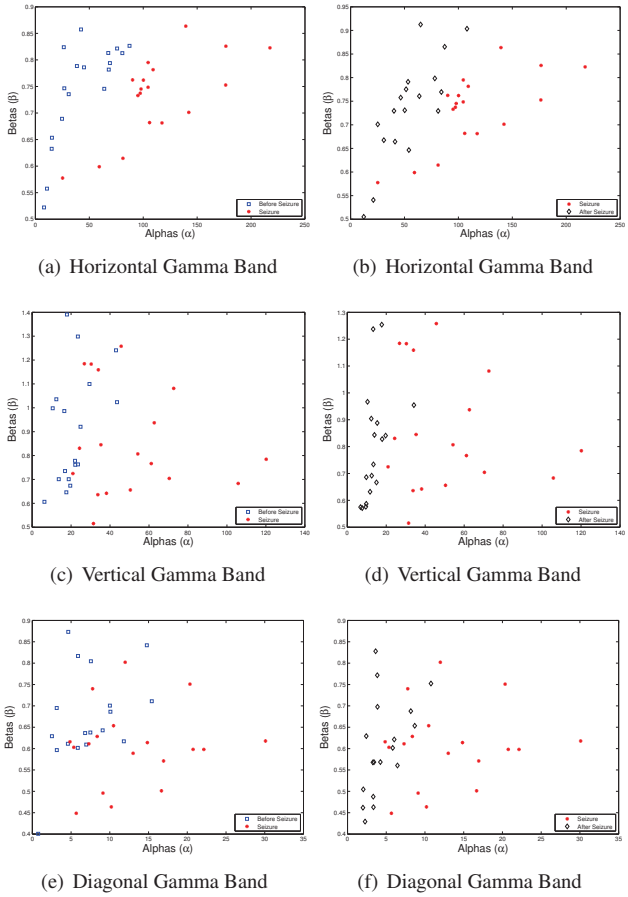
Finally, to develop an intuition for the good performance of the proposed classification scheme, Fig. 2 shows scatter plots of the generalised Gaussian parameters α and β for seizure events (red circles) and non-seizure events (blue squares and black diamonds) observed through the *Gamma* frequency band. We observe that the proposed representation, based on a generalised Gaussian model for the wavelet coefficients, leads to a very clear discrimination of seizure and non-seizure events. In particular, notice that by using this representation it is possible to discriminate events with separating line or hyper-plane, which is essentially what is achieved by using the multivariate Gaussian classifier.

4. CONCLUSION

This work presented a new multivariate Bayesian classification method to detect epileptic seizure events in EEG signals. The method is based on a multilevel 2D wavelet decomposition that captures the distribution of energy across the different brain rhythms and brain regions, coupled with a generalised Gaussian statistical model that summarises this information, and a multivariate Bayesian classification scheme that discriminates seizure events from normal brain function. The proposed methodology was demonstrated on a real dataset containing 36 multivariate EEG recordings related to both seizure and non-seizure events, and by performing comparisons with the state-of-the-art classification method [4]. Future work will focus on an extensive evaluation of the proposed approach and on deriving instances of the method that are tailored for specific medical applications, such as time and location of epilepsy onset detection.

5. ACKNOWLEDGEMENTS

We are grateful to Dr Carlos D'Giano and Dr Gabriela Ugarnes from Fundación contra las Enfermedades Neurológicas Infantiles (FLENI), Argentina, for useful discussion and for providing expert annotations for the data.



(a) Horizontal Gamma Band (b) Horizontal Gamma Band

(c) Vertical Gamma Band (d) Vertical Gamma Band

(e) Diagonal Gamma Band (f) Diagonal Gamma Band

Fig. 2. Scatter plots for the generalised Gaussian parameters α and β for seizure events (red circles) and non-seizure events (blue squares and black diamonds) observed through the *Gamma* frequency band, showing the good linear discrimination power of the proposed approach.

6. REFERENCES

[1] Richard Appleton, Andrew Nicolson, David Smith, David Chadwick, and James MacKenzie, *Atlas of epilepsy*, CRC Press, 2007.

[2] Bjorn Schelter, Jens Timmer, and Andreas Schulze-Bonhage, *Seizure Prediction in Epilepsy: From Basic Mechanisms to Clinical Applications*, Wiley-VCH, 2008.

[3] Andrea Varsavsky, Iven Mareels, and Mark Cook, *Epileptic seizures and the EEG Measurement, Models, Detection and Prediction*, CRC Press, Taylor and Francis Group, 2011.

[4] Ali Shoeb, Herman Edwards, Jack Connolly, Blaise Bourgeois, S. Ted Treves, and John Guttagf, "Patient-specific seizure onset detection," *Epilepsy and Behavior*, vol. 5, pp. 483–498, 2004.

[5] M.E. Saab and J. Gotman, "A system to detect the onset

of epileptic seizures in scalp eeg," *Clinical Neurophysiology*, vol. 116, pp. 427–442, 2005.

[6] P. LeVan, E. Urrestarazu, and J. Gotman, "A system for automatic artifact removal in ictal scalp eeg based on independent component analysis and bayesian classification," *Cinical Neurophysiology*, vol. 117, no. 4, pp. 912–927, 2006.

[7] J.J.Halford, R.J. Schalkoff, J. Zhou, S.R. Benbadis, W.O. Tatum, R.P. Turner, S.R. Sinha, N.B. Fountain, A. Arain, P.B. Pritchard, E. Kutluay, G. Martz, J.C. Edwards, C. Waters, and B.C. Dean, "Standardized database development for eeg epileptiform transient detection: Eegnet scoring system and machine learning analysis," *The Journal of the Acoustical Society of America*, vol. 212, no. 2, pp. 308–316, 2014.

[8] Antonio Quintero-Rincón, Hadj Batatia, Marcelo Pereyra, and Marcelo Risk, "Detection of onset in epilepsy signals using generalized gaussian distribution," *Fifth International Conference on Advances in New Technologies, Interactive Interfaces and Communicability*, vol. 2014, 2014, ISBN:978.88.96.471.37.1, DOI: 10.978.8896471/371.

[9] P. Valenti, E. Cazamajou, M. Scarpettini, A. Aizemberg, W. Silva, and S. Kochen, "Automatic detection of interictal spikes using data mining models," *Journal of Neuroscience Methods*, vol. 150, no. 1, pp. 105–110, 2006.

[10] Matti Hamalainen, Riitta Hari, Risto J. Ilmoniemi, Jukka Knuutila, and Olli V. Lounasmaa, "Magnetoencephalography: Theory, instrumentation, and applications to noninvasive studies of the working human brain," *Review of Modern Physics*, vol. 65, no. 2, pp. 413–459, 1993.

[11] Ingrid Daubechies, *Ten Lectures On Wavelets Daubechies*, SIAM, 1992.

[12] Vincent J. Samar, Ajit Bopardikar, Raghuveer Rao, and Kenneth Swartz, "Wavelet analysis of neuroelectric waveforms a conceptual tutorial," *Brain and Language 66*, vol. 66, pp. 7–60, 1999.

[13] Hojjat Adeli, Ziqin Zhou, and Nahid Dadmehr, "Analysis of eeg records in an epileptic patient using wavelet transform," *Journal of Neuroscience Methods*, vol. 123, pp. 69–87, 2002.

[14] Frédéric Pascal, Lionel Bombrun, Jean-Yves Tourneret, and Yannick Berthoumieu, "Parameter estimation for multivariate generalized gaussian distributions," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5960–5971, 2013.

[15] Ernst Niedermeyer and Fernando Lopes da Silva, *Electroencephalography Basic Principles, Clinical Applications, and Related Fields*, Lippincott Williams and Wilkins, 2010.

[16] Gilbert Strang and Truong Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[17] Stephane Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2008.

[18] Antonio Quintero-Rincón, Marcelo Pereyra, Carlos D'Giano, Hadj Batatia, and Marcelo Risk, "A new algorithm for epilepsy seizure onset detection and spread estimation from eeg signals," *Journal of Physics: Conference Series*, vol. 705, no. 1, pp. 012032, 2016.

[19] Minh N. Do and Martin Vetterli, "Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146–158, 2002.

[20] Saralees Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, no. 7, pp. 685–694, 2005.

[21] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Journal of Clinical Neurophysiology*, vol. 29, pp. 131–163, 1997.

[22] Chris Solomon and Toby Breckon, *Fundamentals of Digital Image Processing*, Wiley Blackwell, 2011.

[23] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.Ch. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C-K. Peng, and H.E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation 101(23):*, vol. 101, no. 23, pp. e215–e220, 2000.

[24] Christopher M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.