Nonconvulsive Epileptic Seizure Detection

Yissel Rodríguez Aldana,
Borbála Hunyadi, *Member, IEEE,*
Enrique Marañón Reyes,
Valia Rodríguez Rodríguez,
and Sabine Van Huffel, *Fellow, IEEE*

***Abstract*—Nonconvulsive status epilepticus is a condition where the patient is exposed to abnormally prolonged epileptic seizures without evident physical symptoms. Since these continuous seizures may cause permanent brain damage, it constitutes a medical emergency. This paper proposes a method to detect nonconvulsive seizures for a further nonconvulsive status epilepticus diagnosis. To differentiate between the normal and seizure electroencephalogram (EEG), a K-Nearest Neighbor and a Radial Basis Support Vector Machine classifier are used. The classifier features are obtained from the Canonical Polyadic Decomposition (CPD) and Block Term Decomposition (BTD) of the EEG data represented as a third order tensor. To expand the EEG into a tensor, Wavelet or Hilbert-Huang transform are used. The algorithm is tested on a scalp EEG database of 139 seizures of different duration. The experimental results suggest that a Hilbert-Huang tensor representation and the CPD analysis provide the most suitable framework for nonconvulsive seizure detection. The K-Nearest Neighbor classifier shows the best performance with sensitivity , specificity, and accuracy values over 90%.**

***Index Terms*—Hilbert Huang Tranform, Multiway Data Analysis, Nonconvulsive epileptic seizures, Wavelet Transform.**

I. Introduction

STATUS Epilepticus (SE) is a condition where the patient is exposed to abnormally prolonged epileptic seizures. This condition could have long-term consequences as neuronal death, neuronal injury, and alteration of neuronal networks, depending on the seizure type and duration [1]. SE can be divided into two major categories, convulsive SE (CSE) and nonconvulsive SE (NCSE). CSE is characterized by muscular spasms typical for epileptic convulsive seizures (ECS). On the other hand, NCSE symptoms are too subtle to be noticed since this status is characterized by nonconvulsive seizures, which show persistent epileptic activity in the electroencephalogram (EEG) without evident physical symptoms.

Nonconvulsive epileptic seizures (NCES) can be found associated to coma/stupor or not (i.e. absence status, focal SE). Patients without coma/stupor are often awake and attend to the emergency services. Symptoms such as automatisms, facial twitching, eye deviation or jerking might help NCES diagnosis. On the other hand, patients with coma/stupor treated at intensive care units (ICU) are usually unconscious or unresponsive. Consequently, these NCES have no physical observable symptoms at all. Then, the diagnosis can only be accomplished by means of EEG monitoring. As explained before, these seizures are often associated with serious brain damage with very poor prognosis [2]. Therefore, it is crucial to monitor the EEG at the ICU, in order to allow physicians to detect a seizure when it is still possible to prevent it from causing permanent brain damage.

Due to the nature of long-term EEG recordings, its analysis is a time-consuming task. This analysis can be performed by doctors or a seizure detection software. The visual analysis performed by clinicians might lead to intra-observational inconsistencies in the results (the same expert can produce different outputs if (s)he reviews the data at different times)[3]. Software approaches are more time efficient, more consistent and objective.

Most of the available automated seizure detection software has been extensively trained on seizures obtained from patients with seizure disorders without brain injuries. Nevertheless, these seizures and those from comatose/brain-injured patients are different. The ictal activity from brain-injured/comatose patients has a longer duration, is less defined in time (unclear on-set /off-set), less organized and lower in maximum frequency [3].

The frequencies for NCSE range from $0.5 - 5Hz$. The occurrence of generalized or lateralized periodic discharges frequently below $2.5Hz$ (about $0.3Hz$) or rhythmic discharges greater than $0.5Hz$ (ranging from $0.2 - 3Hz$, usually at $0.5 - 2Hz$) are considered as activity preceding NCSE [4], [5], [6]. The occurrence of rhythmic delta-theta activity below $4Hz$ is also considered a possible sign of NCSE [5]. The NCSE without coma/stupor is characterized by an EEG pattern

This work has been supported by the Belgian foreign Affairs-Development Cooperation through VLIR-UOS (2013-2019) (Flemish Interuniversity Council-University Cooperation for Development) in the context of the Institutional University Cooperation program with Universidad de Oriente.

The research leading to these results has received funding from imec funds 2017 and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

Y. Rodríguez is with the Universidad de Oriente. Center of Neuroscience and Signals and Image Processing Santiago de Cuba, Cuba.(e-mail:yaldana@uo.edu.cu)

B. Hunyadi is with Stadius Center for Dynamical Systems, Signal Processing and Data Analytics and IMEC, Leuven, BE. Leuven, Belgium. (e-mail:borbala.hunyadi@esat.kuleuven.be)

E.J. Maraón Reyes are with the Universidad de Oriente. Center of Neuroscience and Signals and Image Processing Santiago de Cuba, Cuba. (e-mail:enriquem@uo.edu.cu)

V. Rodríguez Rodríguez was with ICU-Ameijeiras Clinic Surgical Hospital, La Habana, Cuba and Aston University, United Kingdom.(e-mail:v.rodriguez@aston.ac.uk)

S. Van Huffel is with Stadius Center for Dynamical Systems, Signal Processing and Data Analytics and IMEC, Leuven, BE. Leuven, Belgium. (e-mail:sabine.vanhuffel@esat.kuleuven.be)

of spike-and-wave complexes with frequencies from $3Hz$ to $3.5Hz$ [4]. This pattern can also be found in patients with coma/stupor at NCSE risk [5]. NCSE in a coma/stupor patient is diagnosed after the emergence of epileptiform discharges above $2.5Hz$ [4], [5].

A big number of published seizure detection methods address CES [7], [8] and only few of them address NCES [9]. The proposed methods for NCES detection published until 2015 are summarized in the work of Ansari and Sharma [9]. These methods can be split according to their principal aim: those that detect NCES without coma/stupor (i.e. absence seizures) [10], [11], [12], [13], [14] and the ones that perform NCES detection on ICU patients [15], [16], [17], [18], [19]. All databases used are private and very dissimilar in size and composition, thereby complicating objective comparison between the diverse methods.

The methodology used in these studies is also very diverse. Classifiers such as Support Vector Machines, Neural Networks, and Linear Discriminant functions are the most commonly used ones. For the data description, the most widespread features are Wavelet Transform (WT) scales [10], [11], [12], [18], Entropy [13], [14] and nonlinear parameters [11], [13], [15]. The major drawbacks of these methods are related to the arbitrary nature of the selected thresholds for performing the detection. Seizure duration and the number of channels affected by the seizure activity are the most popular criteria for thresholding [12], [16], [19], [18]. Seizures are not detected if they are too short in time or too localized (just a few channels affected by the seizure activity). NCES are heterogeneous across patients, the patterns present in the EEG will depend on the etiology of the seizure. Therefore, a threshold which works for one patient will not necessarily work for another. Also, there are more meaningful ways to describe the seizure's spatial localization than the number of channels; that is, the seizure topography.

This research aims to propose a method to detect NCES, which alleviates these shortcomings. In particular, this method attempts to identify the NCES based on their similarity to the first NCES detected by the physician at the EEG [20]. The explored features are obtained by means of a multiway analysis of the EEG signal represented as a third order tensor $X \in \mathbb{R}^{(F \times T \times Ch)}$ with axes $frequency \times time \times channels$. The tensors are computed by expanding the EEG segments using WT or Hilbert Huang Transform (HHT). WT is a popular tool for seizure EEG tensor construction [21], [22]. However, we propose here the HHT since it resolves $time \times frequency$ events with a finer resolution than WT and thanks to the adaptive nature of the empirical mode decomposition provides a more meaningful physical interpretation of the underlying EEG data. The tensor decomposition is performed with Canonical Polyadic Decomposition (CPD) and Block Term Decomposition (BTD). It has been shown that tensor decompositions of multiway EEG representations can extract seizure sources and accurately characterize the seizure pattern [21], [23], [22]. The multiway analysis exploits the EEG high dimensional structure by analyzing its spectral, temporal, and spatial properties simultaneously. As such, this approach represents a neat and practical alternative to analyze large, heterogeneous and multidimensional feature sets. Moreover, it eliminate the need for thresholding since the classifiers will implicitly evaluate the similarity of the signatures.

The paper is organized as follows: Section II provides some basic definitions and the description of the proposed method. In section III the results obtained are presented and discussed. Finally, in section IV our conclusions are summarized.

II. Material and Methods

A. EEG Data

EEG data were collected as part of patients' clinical assessment at the Epilepsy Unit of the Cuban Neurological Restoration Center (CIREN) and at the Intensive Care Units (ICU) of the Clinical Surgical Hospital Hermanos Ameijeiras, both in Havana City. Data were anonymized prior to its use in this study. The visual inspection and labeling were performed by trhee neurophysiologists (VRR). Since the patients come from different hospitals and areas, the acquisition protocol shows some differences; yet the sampling rate for all recordings is 200 Hz. The number of electrodes used for the recordings varies between 8 and 19. In all cases, the electrodes were placed according to the 10-20 montage system. The use of the 10-20 montage for all records guarantees that the electrodes' name and position do not change in the records.

The dataset comprises 14 adult patients ranged between 18 and 57 years old with various brain disorders that present NCES. A total of 139 seizures have been registered, all of them NCES according to the neurophysiologist' and neurologist' diagnosis (55 with coma/stupor). A more detailed description is given in Table I. All procedures were reviewed and approved by the Ethical Committees of the CIREN and Hermanos Ameijeiras Hospital respectively.

B. Tensor formulation

The EEG from each recording is divided into $3s$ duration segments (epochs). All epochs are expanded in the time-frequency domain using a WT or Hilbert-Huang Transform (HHT). A 3rd order tensor is built from every epoch with $frequency \times time \times channels$ axes.

1) Wavelet Transform: The Continous Wavelet Transform (CWT) performs a multi-resolution analysis of a signal $x(t)$ decomposing it into sub-band signals representing activity at different time scales. WT has been reported in the literature as one of the preferred techniques to perform epileptic seizure detection [8]. The signal decomposition is achieved by convolving scaled and shifted versions of a mother wavelet $\psi(n)$ with $x(t)$ and it is defined as follows,

$$x_w(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt \qquad (1)$$

Where $()^*$ indicates the complex conjugate, ψ is the analyzing wavelet, $a(>0)$ is the scale parameter and b is the shift parameter [24].

To expand the EEG into a three-way tensor expressed as $X \in \mathbb{R}^{(F \times T \times Ch)}$ the CWT is applied to each channel transforming it in a matrix, where the rows are the wavelet

TABLE I
EEG RECORDS DETAILED DESCRIPTION.

Patient data					Record		Protocol				
Patient	Gender	Age	Diagnosis	Type	Duration (h:min:s)	Seizures	Channels	EKG	EOG	EMG	Video
1	M	33	*Temporal lobe epilepsy.*	Ambulatory	*00:38:48*	*6*	*19*	o	o	o	o
2	M	*37*	*Focal Epilepsy.*	Ambulatory	*00:16:44*	*3*	*19*	o	o	o	o
3	M	*29*	*Temporal lobe epilepsy.*	Ambulatory	*00:48:29*	*13*	*19*	o	o	o	o
4	F	*37*	*Partial Temporal Lobe Epilepsy.*	Ambulatory	*00:14:05*	*5*	*19*	o	o	o	o
5	F	*18*	*Partial psychic seizures.*	Ambulatory	*00:16:05*	*2*	*19*	o	o	o	o
6	F	*18*	*Lenns-Gastaut. Atonic Neck.*	Ambulatory	*00:39:30*	*2*	*19*	o	o	o	o
7	F	*23*	*Head trauma.*	Ambulatory	*00:28:15*	*12*	*17*	x	x	x	o
8	F	*32*	*Spike Wave and Polyspike wave.*	Ambulatory	*08:05:58*	*6*	*17*	x	x	x	o
9	M	*31*	*Absence seizures.*	Ambulatory	*07:42:52*	*34*	*17*	x	x	x	o
10	F	*30*	*Leukocytosis,possible thrombocytopenic purpura.*	ICU	*19:36:38*	*40*	*19*	x	x	x	x
11	M	*48*	*Highly malignant astrocytoma (epileptic shock).*	ICU	*03:34:50*	*2*	*13*	x	o	x	x
12	F	*36*	*Unknown*	ICU	*21:12:46*	*6*	*8*	x	x	x	x
13	F	*57*	*Connective tissue mixed disease.*	ICU	*00:47:20*	*5*	*19*	x	x	x	x
14	F	*50*	*Cerebral gliomatotosis.*	ICU	*01:02:32*	*3*	*14*	x	o	x	x

scales (F), the columns time instants (T) and the elements are the coefficients. The Ch dimension represents the spatial distribution among the channels. Since the NCES activity is below $5Hz$ [6], five wavelet scales were computed, which correspond to $1-5Hz$ frequencies, with $1Hz$ resolution. As mother wavelet, a Mexican Hat is used as reported in [21], [22].

2) Hilbert-Huang Transform: Hilbert-Huang Transform (HHT) is a method that uses two steps to analyze the data. First, the data are decomposed into a number of components and a residual using the empirical mode decomposition (EMD). Second, a Hilbert transform is applied to those components to construct a $frequency \times time \times energy$ distribution, defined as the Hilbert spectrum. In this spectrum, all events in the time domain will be preserved by the instantaneous frequency computed by the Hilbert transform [25].

The main characteristic of EMD is to decompose a signal into so-called intrinsic mode functions (IMF) plus a residue which is conventionally defined as the temporal trend of the series. By definition, an IMF satisfies two conditions: (i) for the entire data set, the number of extrema and the number of zero crossings must be either equal or differ at most by one and (ii) the mean at any point of the contour defined by interpolating the local maxima (upper envelope) and the contour defined by interpolating the local minima (lower envelope) is zero [26]. For any one-dimensional discrete signal, the EMD can be expressed by the following equation,

$$x(t) = \sum_{k=1}^{n} imf_k(t) + r(t) \tag{2}$$

where imf_k is the kth IMF of the signal, and r is the residue [25].

Once the IMFs are computed each of them is transformed with the Hilbert Transform defined as,

$$y_{imf_k}(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{imf_k(t)}{t - t_0} dt \tag{3}$$

where P indicates the Cauchy principal value. $y_{imf_k}(t)$ and $imf_k(t)$ form a complex conjugate pair that describes an analytic signal $z_{imf_k}(t)$,

$$z_{imf_k}(t) = imf_k(t) + j y_{imf_k}(t) = a_{imf_k}(t) e^{j\theta_{imf_k}(t)} \tag{4}$$

where a_{imf_k} is the instantaneous amplitude and θ_{imf_k} is the instantaneous phase defined as,

$$a_{imf_k}(t) = \sqrt{[imf_k(t)]^2 + [y_{imf_k}(t)]^2} \tag{5}$$

and

$$\theta_{imf_k}(t) = \arctan\left(\frac{y_{imf_k}(t)}{imf_k(t)}\right) \tag{6}$$

respectively.

The instantaneous frequency is computed as

$$w_{imf_k}(t) = \frac{d\theta_{imf_k}(t)}{dt} \tag{7}$$

Then the $x_{HH}(w,t)$ matrices are constructed, these contain the energy at each time point t and at each instantaneous frequency w in the following way,

$$x_{HH}(w,t) = \sum_{\substack{k=1 \\ w=w_{imf_k}(t)}}^{n} s_{imf_k}(t) \tag{8}$$

where n is the number of IMFs and the energy s_{imf_k} is defined as

$$s_{imf_k}(t) = a_{imf_k}(t)^2 \tag{9}$$

As implemented with CWT, the HHT is applied to every EEG channel. The Hilbert spectrum is truncated at $5Hz$. The $x_{HH}(w,t)$ matrices are arranged as frontal slices of a tensor $X \in \mathbb{R}^{(F \times T \times Ch)}$, further on defined as HH_Tensor, where F indicates the $1-5Hz$ frequency values, T the number of time samples and Ch the number of EEG channels.

C. Tensor Decomposition

The tensor obtained by WT or HHT will be analyzed using CPD and BTD. CPD provides a trilinear description of the data, which is a very compact and interpretable model. This model assumes that the seizure pattern maintains the same frequency and topography within a certain period of time. CPD has been applied successfully in several studies that represent the EEG as a third-order tensor to recognize epileptic seizure activity [21], [27], [23], [22]. This fact makes CPD an attractive method for our study. On the other hand, BTD is more flexible, complex and facilitates a more accurate model when the data frequency or topography change in time. Some studies have demonstrated that BTD outperforms CPD for EEG applications where the pattern of interest shows evolution in frequency, morphology and/or topography [22], [28]. Epileptic seizures have been reported as having such kind of behavior. Therefore, BTD exploration is also interesting for this application.

1) Canonical Polyadic Decomposition (CPD): The Canonical Polyadic Decomposition (CPD) represents a third-order tensor $X \in \mathbb{R}^{(F\times T\times Ch)}$ as the outer product (represented as '$\circ$') of R rank-1 tensors in the following way,

$$X = \sum_{r=1}^{R} a_r \circ b_r \circ c_r \tag{10}$$

Where $a_r \in \mathbb{R}^F, b_r \in \mathbb{R}^T, c_r \in \mathbb{R}^{Ch}$ are nonzero vectors that define the mode-n signature, n=1,...,N (for $N = 3$). With $1 \leq r \leq R$.

The tensor rank of *X* is equal to the minimal number of components R that generates an 'exact' CPD of *X*, where 'exact' means that there is equality at (10) [29]. A tensor CPD is unique under mild conditions up to permutation and scaling [29]. Let $A = [a_1 \dots a_R]$, $B = [b_1 \dots b_R]$ and $C = [c_1 \dots c_R]$ be the factor matrices corresponding to each mode, and k_A, k_B, and k_C their respective k-rank. The k-rank k_A of a matrix A is the maximum value ensuring that any subset of k_A columns are linearly independent, with $k_A < r_A$. According to Kruskal's theorem [30] a sufficient condition for the uniqueness of a third-order tensor CPD up to permutation and scaling uncertainty is given by,

$$k_A + k_B + k_C \geq 2R + 2 \tag{11}$$

2) Block Term Decomposition: The (L,M,O)-BTD of a third-order tensor $X \in \mathbb{R}^{(F\times T\times Ch)}$ is written as,

$$X = \sum_{r=1}^{R} D_r \bullet_1 A_r \bullet_2 B_r \bullet_3 C_r \tag{12}$$

where $\bullet_n$ denote the mode-n product, $D_r \in \mathbb{R}^{(L\times M\times O)}$ are full rank-(L,M,O), $A_r \in \mathbb{R}^{(F\times L)}$ (with $F \geq L$), $B_r \in \mathbb{R}^{(T\times M)}$ (with $T \geq M$), and $C_r \in \mathbb{R}^{(Ch\times O)}$ (with $Ch \geq O$) are full column rank, $1 \leq r \leq R$. BTD is unique up to permutation and scaling uncertainty inherited from CPD under more restrictive conditions. Essential uniqueness conditions for the (L,M,O)-BTD can be found in [31].

The two obtained tensor sets, W_Tensor and HH_Tensor, are decomposed using CPD and BTD. The EEG epochs that contain seizure activity will be labeled as positives and the rest as negatives.

To differentiate the features sets, those obtained from a tensor built with HHT are denoted by 'h', 'w'is used for those obtained with WT. Space, Frequency and Time signatures obtained from the tensor decomposition, are referred as 'spac', 'freq', and 'time'. The training using all assembled signatures together is specified as 'asm'. To tag the decomposition method 'CP'and 'BT'will be used to refer to CPD and BTD respectively. As an example, the Space signature obtained by the CPD decomposition of a tensor built with WT will be denoted as: w_spac_CP .

D. Multilinear Rank Analysis

Before performing CPD and BTD, the tensor rank must be determined. The multilinear rank (*ml_rank*) of a high-order tensor is the *n-tuple* of the *mode-n* ranks. Contrary to the matrix case, each *mode-n* rank may be different. In order to find the best low *ml_rank* approximation for the computed tensors, the multilinear singular value decomposition (MLSVD) is used [32]. The MLSVD is a multilinear generalization of the matrix SVD. The MLSVD of a third-order tensor $X \in \mathbb{R}^{(F\times T\times Ch)}$ can be written as,

$$X = Z \bullet_1 U^{(1)} \bullet_2 U^{(2)} \bullet_3 U^{(3)} \tag{13}$$

where $U^{(n)}$, $n = 1, \dots, N, N = 3$, are the orthonormal basis for the n different subspaces of the *mode-n* vectors and $Z \in \mathbb{R}^{(L\times M\times O)}$ an all-orthogonal and ordered tensor [33]. Essentially, the MLSVD considers a given tensor as N sets of *mode-n* vectors and computes the matrix SVD of these sets. The truncation of the MLSVD gives a suboptimal solution to the tensor best low *ml-rank* approximation problem which can be used as initialization to iterative algorithms such as the low multilinear rank approximation (LMLRA). The LMLRA can be formulated as MLSVD, the difference between these two algorithms lies in the computation and the optimality of the decomposition. LMLRA refines the MLSVD initial guess using an optimization based method (i.e. nonlinear unconstrained optimization, nonlinear least squares, adaptive cross approximation) [33].

To determine the tensor *ml_rank* , the MLSVD is performed on all tensors obtained from the first seizure and the seizure-free EEG prior to it for both tensor sets (for more information about the available data see Table II). Then, the obtained singular values (SV) are inspected in order to truncate the MLSVD core where the SV represented more than 95 % of the data variance. The truncated MLSVD core was used to initialize the LMLRA. The LMLRA method is then performed to find the multilinear rank that best approximates the tensor in the least squares sense [33].

As before, this procedure is performed for all tensors obtained from the EEG epochs corresponding to the first seizure and all available seizure-free epochs prior to it. Therefore, for every segment, we obtained a rank value in each of the 3 modes (i.e. frequency, time and channel). Afterwards,

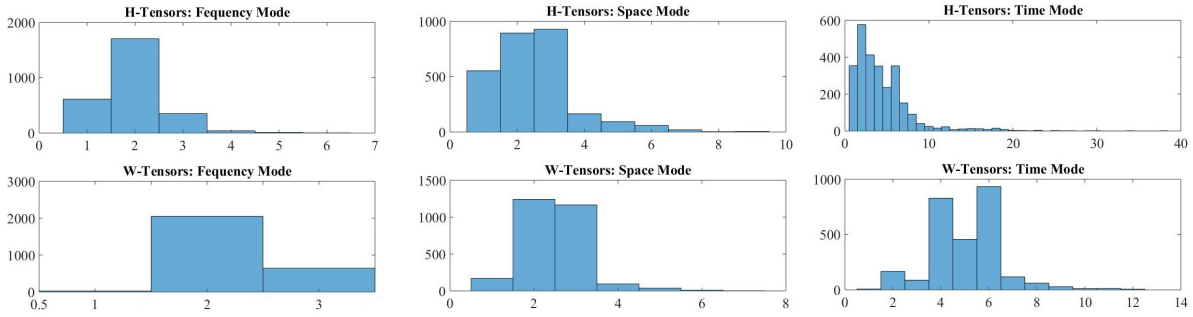
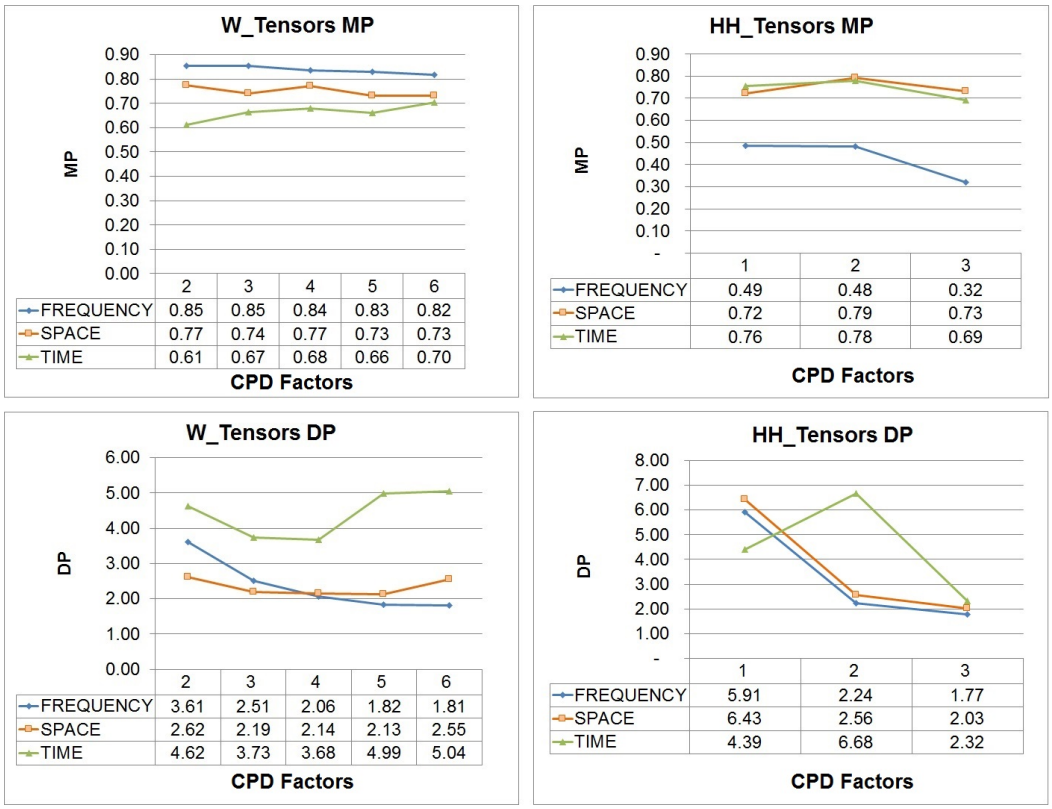


Fig. 1. MLRA analysis Frequency, Time and Space mode histograms. Axis X and Y represent the number of components and tensors respectively.



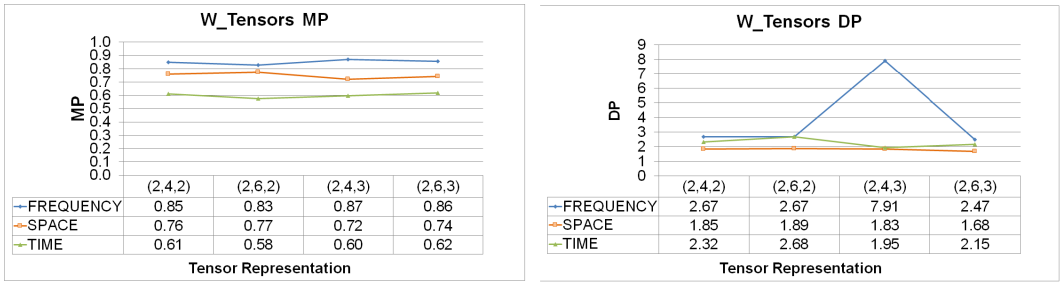
W_Tensors MP	2	3	4	5	6
FREQUENCY	0.85	0.85	0.84	0.83	0.82
SPACE	0.77	0.74	0.77	0.73	0.73
TIME	0.61	0.67	0.68	0.66	0.70

HH_Tensors MP	1	2	3
FREQUENCY	0.49	0.48	0.32
SPACE	0.72	0.79	0.73
TIME	0.76	0.78	0.69

W_Tensors DP	2	3	4	5	6
FREQUENCY	3.61	2.51	2.06	1.82	1.81
SPACE	2.62	2.19	2.14	2.13	2.55
TIME	4.62	3.73	3.68	4.99	5.04

HH_Tensors DP	1	2	3
FREQUENCY	5.91	2.24	1.77
SPACE	6.43	2.56	2.03
TIME	4.39	6.68	2.32

Fig. 2. Obtained MP and DP for HH_Tensors and W_Tensor CPD signatures computed for several number of factors. The displayed MP values are computed for the feature matrices of the first NCES. The MP and DP values displayed on the charts are averaged over the number of features of each mode.



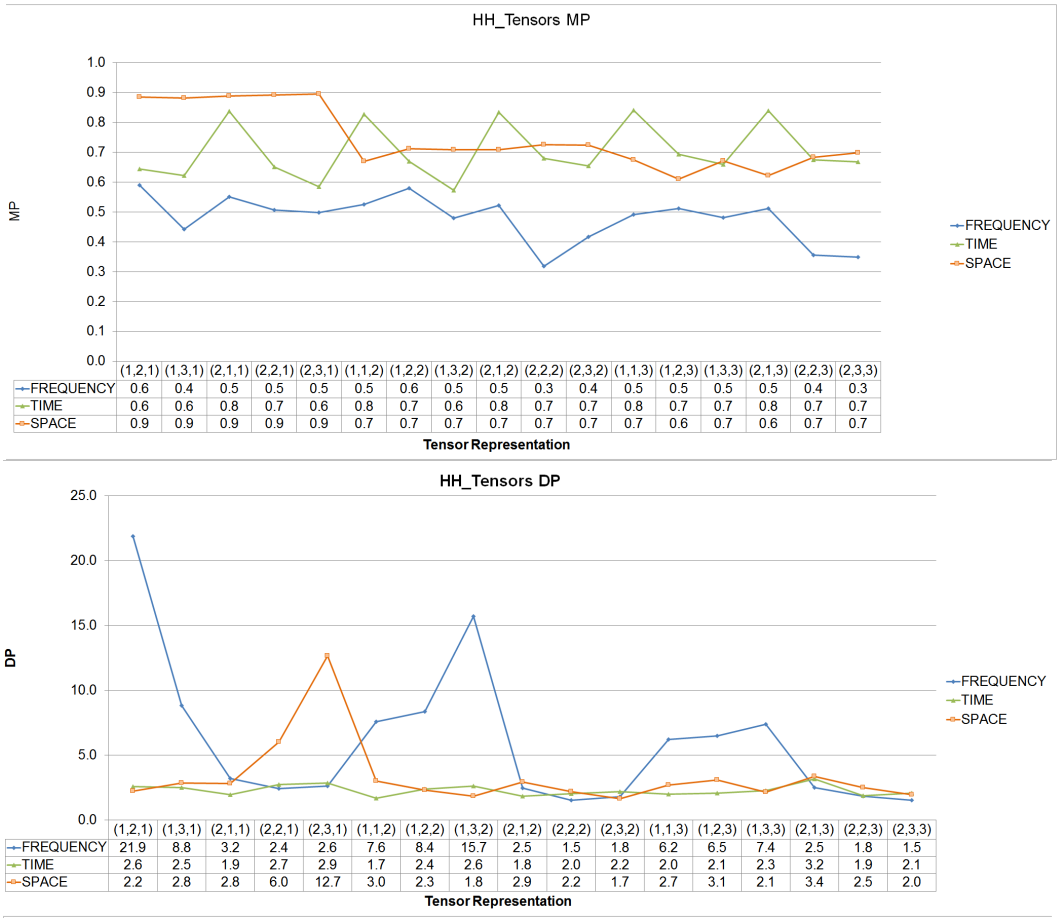
W_Tensors MP	(2,4,2)	(2,6,2)	(2,4,3)	(2,6,3)
FREQUENCY	0.85	0.83	0.87	0.86
SPACE	0.76	0.77	0.72	0.74
TIME	0.61	0.58	0.60	0.62

W_Tensors DP	(2,4,2)	(2,6,2)	(2,4,3)	(2,6,3)
FREQUENCY	2.67	2.67	7.91	2.47
SPACE	1.85	1.89	1.83	1.68
TIME	2.32	2.68	1.95	2.15

Fig. 3. MP and DP values computed for all tensor representations considered for BTD in the W_Tensor ml_rank analysis. he displayed MP values are computed for the feature matrices of the first NCES. The MP and DP values displayed on the charts are averaged over the number of features of each mode.

rank histograms are constructed by counting the number of tensors with a certain rank, as shown in Fig.1. Based on these histograms, the most frequently occurring ranks at each mode are selected as upper bound to perform CPD and BTD.

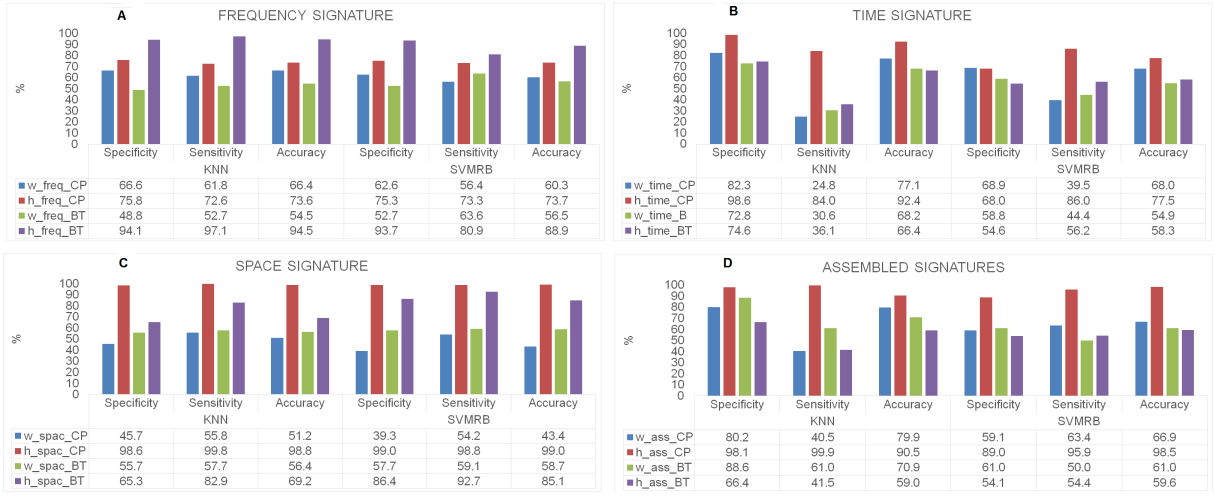
The best tensor model for the classification task is selected by computing the modeling power (MP) and the discrimination



	(1,2,1)	(1,3,1)	(2,1,1)	(2,2,1)	(2,3,1)	(1,1,2)	(1,2,2)	(1,3,2)	(2,1,2)	(2,2,2)	(2,3,2)	(1,1,3)	(1,2,3)	(1,3,3)	(2,1,3)	(2,2,3)	(2,3,3)
FREQUENCY	0.6	0.4	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.3	0.4	0.5	0.5	0.5	0.5	0.4	0.3
TIME	0.6	0.6	0.8	0.7	0.6	0.8	0.7	0.6	0.8	0.7	0.7	0.8	0.7	0.7	0.8	0.7	0.7
SPACE	0.9	0.9	0.9	0.9	0.9	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.7	0.7

	(1,2,1)	(1,3,1)	(2,1,1)	(2,2,1)	(2,3,1)	(1,1,2)	(1,2,2)	(1,3,2)	(2,1,2)	(2,2,2)	(2,3,2)	(1,1,3)	(1,2,3)	(1,3,3)	(2,1,3)	(2,2,3)	(2,3,3)
FREQUENCY	21.9	8.8	3.2	2.4	2.6	7.6	8.4	15.7	2.5	1.5	1.8	6.2	6.5	7.4	2.5	1.8	1.5
TIME	2.6	2.5	1.9	2.7	2.9	1.7	2.4	2.6	1.8	2.0	2.2	2.0	2.1	2.3	3.2	1.9	2.1
SPACE	2.2	2.8	2.8	6.0	12.7	3.0	2.3	1.8	2.9	2.2	1.7	2.7	3.1	2.1	3.4	2.5	2.0

Fig. 4. MP and DP values computed for all tensor representations considered for BTD in the H_Tensor ml_rank analysis. he displayed MP values are computed for the feature matrices of the first NCES. The MP and DP values displayed on the charts are averaged over the number of features of each mode.



	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
	KNN			SVMRB		
w_freq_CP	66.6	61.8	66.4	62.6	56.4	60.3
h_freq_CP	75.8	72.6	73.6	75.3	73.3	73.7
w_freq_BT	48.8	52.7	54.5	52.7	63.6	56.5
h_freq_BT	94.1	97.1	94.5	93.7	80.9	88.9

	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
	KNN			SVMRB		
w_time_CP	82.3	24.8	77.1	68.9	39.5	68.0
h_time_CP	98.6	84.0	92.4	68.0	86.0	77.5
w_time_B	72.8	30.6	68.2	58.8	44.4	54.9
h_time_BT	74.6	36.1	66.4	54.6	56.2	58.3

	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
	KNN			SVMRB		
w_spac_CP	45.7	55.8	51.2	39.3	54.2	43.4
h_spac_CP	98.6	99.8	98.8	99.0	98.8	99.0
w_spac_BT	55.7	57.7	56.4	57.7	59.1	58.7
h_spac_BT	65.3	82.9	69.2	86.4	92.7	85.1

	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
	KNN			SVMRB		
w_ass_CP	80.2	40.5	79.9	59.1	63.4	66.9
h_ass_CP	98.1	99.9	90.5	89.0	95.9	98.5
w_ass_BT	88.6	61.0	70.9	61.0	50.0	61.0
h_ass_BT	66.4	41.5	59.0	54.1	54.4	59.6

Fig. 5. Sensitivity , Specificity and Accuracy scores obtained in the classification for KNN and SVMRB.

power (DP) of the signatures obtained from the CPD and BTD with a certain rank. The MP describes the variable relevance for the model and the DP expresses how well a variable discriminates between different classes. The MP and DP were proposed by Wold et. al. in [34] for feature selection in chemical applications. Subsequently, these measures have been used for feature selection in diverse applications as image processing [35], DNA samples selecction [36], and biomedical problems modeling [37], [38]. The MP and DP analysis was performed for the signatures obtained from the tensors computed for the EEG epochs corresponding to the first seizure and all available seizure-free epochs prior to it.



Fig. 6. FPR values displayed by the KNN and SVMRB classifiers during the different trainings.

Before introducing the DP and MP computation, let us define the feature matrices $H^A \in \mathbb{R}^{I_A \times F}$, $H^B \in \mathbb{R}^{I_B \times T}$and $H^C \in \mathbb{R}^{I_C \times Ch}$ corresponding to the Frequency, Time and Space modes respectively. $I_A = E_p \cdot L$, $I_B = E_p \cdot M$, and $I_C = E_p \cdot O$ where, E_p is the number of EEG epochs analyzed, L, M, and O are the ml-rank. $A_j \in \mathbb{R}^{F \times L}$, $B_j \in \mathbb{R}^{T \times M}$ and $C_j \in \mathbb{R}^{Ch \times O}$ are the factor matrices of the Frequency, Time and Space modes respectively, of the jth epoch, $j = 1, \cdots, E_p$ with

$$H^A = \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_{Ep}^T \end{bmatrix}$$

and H^B and H^C similarly defined.

To compute the MP and the DP, at first stage, principal component analysis (PCA) is carried out on the feature matrices H^A, H^B and H^C separately. The number of principal components (PC) λ^{H^A}, λ^{H^B} and λ^{H^C} that represents 95% of the variance of the feature matrices H^A, H^B and H^C respectively are retained.

Let us take the H^A feature matrix to exemplify the MP and DP computation. For more clarity in the notation, let H^A be replaced by H and substitute the number of frequency scales F by the index J as the number of column of H^A. Then, the MP of the jth column of H (H_j) is defined as [34]:

$$MP_j^H = 1 - \frac{\sqrt{\sum_{i=1}^{n^H} \frac{(e_{ij}^H)^2}{(n^H - \lambda^H - 1)}}}{\sqrt{\sum_{i=1}^{n^H} \frac{(H_{ij} - \mu_j^H)^2}{n^H - 1}}} \tag{14}$$

Where μ_j^H is the mean of H_j, n^H is the number of elements in H_j, e_{ij}^H is the $(i,j)^{th}$ entry of the residual matrix E^H obtained from H as follows:

$$E^H = H - P_{\lambda^H}^H {Q_{\lambda^H}^H}^T \tag{15}$$

$P_{\lambda^H}^H$ are the λ^H retained PC and $Q_{\lambda^H}^H$ are the retained PC loadings respectively after truncated PCA.

This measure takes values between 0 and 1. An MP value close to 1 implies that the variable is relevant for the class description. If this value is close to 0, then it can be concluded that the variable in terms of modeling is less significant [34].

To compute DP, let us define NCES and non-NCES as two classes $sz = NCES$ and $nsz = non - NCES$ with feature matrices H_{sz} and H_{nsz} for certain tensor representation. H_{sz} and H_{nsz} are analyzed separately with PCA and the $\lambda^{H_{sz}}$ and $\lambda^{H_{nsz}}$ PCs that represent 95% of the features variance (for each class are) retained. Then, two PCA models are obtained with $\lambda^{H_{sz}}$ and $\lambda^{H_{nsz}}$ PCs. The DP between the two classes, sz and nsz, of the feature j is defined as,

$$DP_j^{H_{sz},H_{nsz}} = \sqrt{\frac{S_j^{H_{nsz}} (H_{sz})^2 + S_j^{H_{sz}} (H_{nsz})^2}{S_j^{H_{nsz}} (H_{nsz})^2 + S_j^{H_{sz}} (H_{sz})^2}} - 1 \tag{16}$$

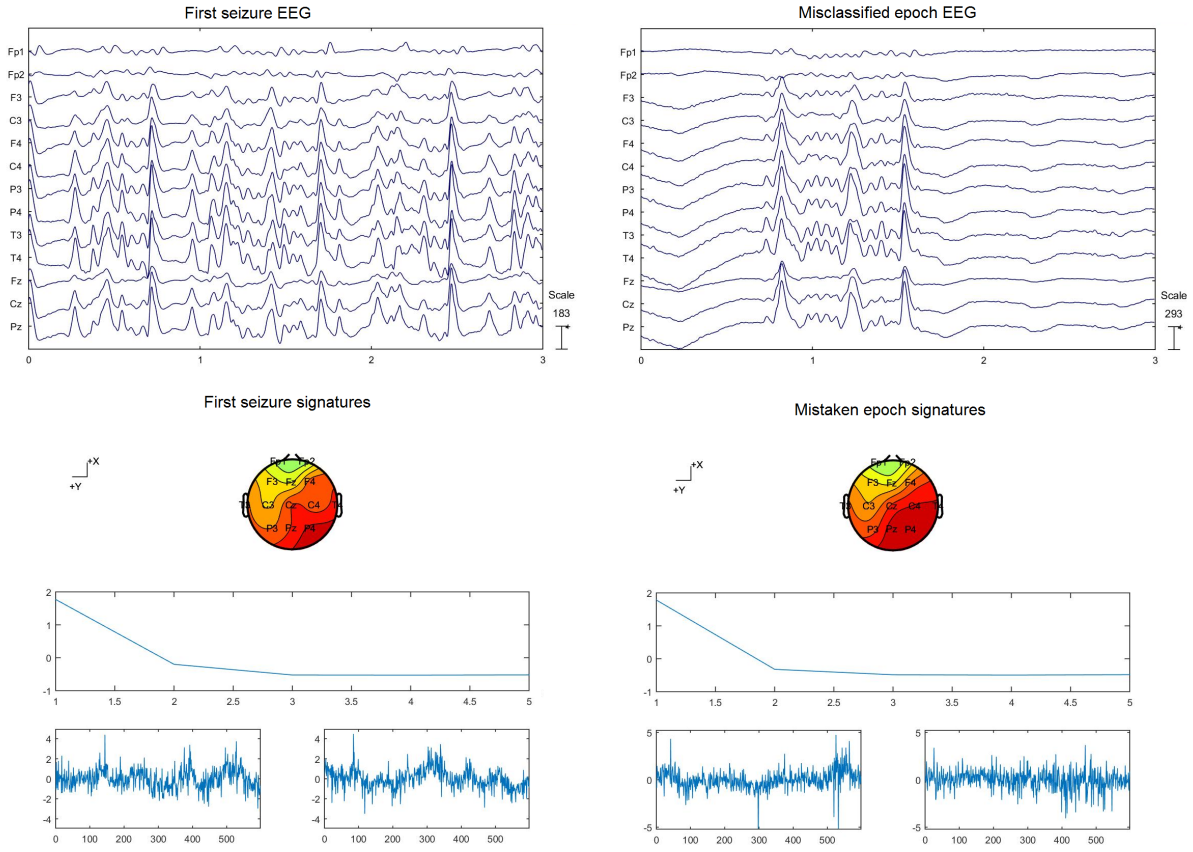


Fig. 7. H_Tensors BTD false positive classification example. Signatures are represented in the fallowing order: Space, Frequency and Time.

Where $S_j^{H_{nsz}}(H_{nsz})^2$ and $S_j^{H_{sz}}(H_{sz})^2$ are the standard deviation computed for the residual matrices $E^{H_{nsz}}$ and $E^{H_{sz}}$ obtained for H_{nsz} and H_{sz} PCA models with (15), and are defined as

$$S_j^{H_{nsz}}(H_{nsz})^2 = \sum_{i=1}^{n^{H_{nsz}}} \frac{(e_{ij}^{H_{nsz}})^2}{(n^{H_{nsz}} - \lambda^{H_{nsz}} - 1)} \tag{17}$$

and

$$S_j^{H_{sz}}(H_{sz})^2 = \sum_{i=1}^{n^{H_{sz}}} \frac{(e_{ij}^{H_{sz}})^2}{(n^{H_{sz}} - \lambda^{H_{sz}} - 1)} \tag{18}$$

were $\lambda^{H_{nsz}}$ and $\lambda^{H_{sz}}$ are the number of retained PCs for H_{nsz} and H_{sz} PCA models respectively. The $S_j^{H_{nsz}}(H_{sz})^2$ and $S_j^{H_{sz}}(H_{nsz})^2$ terms are the residual standard deviation of H_{nsz} and H_{sz} when fitting samples from the class H_{sz} onto the PCA model obtained for class H_{nsz}, and when fitting samples from the class H_{nsz} onto the PCA model obtained for class H_{sz} respectively [34]. These terms are defined as,

$$S_j^{H_{sz}}(H_{nsz})^2 = \frac{J}{n^{H_{sz}}(J - \lambda^{H_{nsz}})} \sum_{i=1}^{n^{H_{sz}}} (e_{ij}^{H_{nsz}})^2 \tag{19}$$

and

$$S_j^{H_{nsz}}(H_{sz})^2 = \frac{J}{n^{H_{nsz}}(J - \lambda^{H_{sz}})} \sum_{i=1}^{n^{H_{nsz}}} (e_{ij}^{H_{sz}})^2 \tag{20}$$

Where J is the number of columns of H_{nsz} and H_{sz}, $n^{H_{sz}}$ and $n^{H_{nsz}}$ are the number of rows of H_{nsz} and H_{sz} respectively.

If the DP value is close to 0, low discriminatory power is observed. If it is much higher than 1, a good discriminatory power is observed [34].

Fig. 2 shows the computed MP and DP values for the signatures obtained from the HH_Tensor and W_Tensor after applying the CPD. Based on the LMLRA (see Fig. 1) the CPD is performed with $R = 1, 2, 3$ for HH_Tensor and $R = 2, \ldots, 6$ for W_Tensor. BTD is applied to all rank combinations of $F = 1, 2$, $T = 1, 2, 3$ and $S = 1, 2, 3$ for HH_Tensor and $F = 2$, $T = 4, 6$ and $S = 2, 3$ for W_Tensor (where F, T and S are the possible rank values of the Frequency, Time and Space modes). This results in a total of 17 different tensor representations for HH_Tensor (excluding the rank-(1,1,1) since it is equivalent to a $R = 1$ CPD, see Fig. 4) and 4 tensor representations for W_Tensor, see Fig. 3.

The average MP value is computed by averaging the values obtained by the three signatures for each tensor representation. The same procedure is performed for the DP. Then, the tensor representation with the average MP closest to 1 and the largest

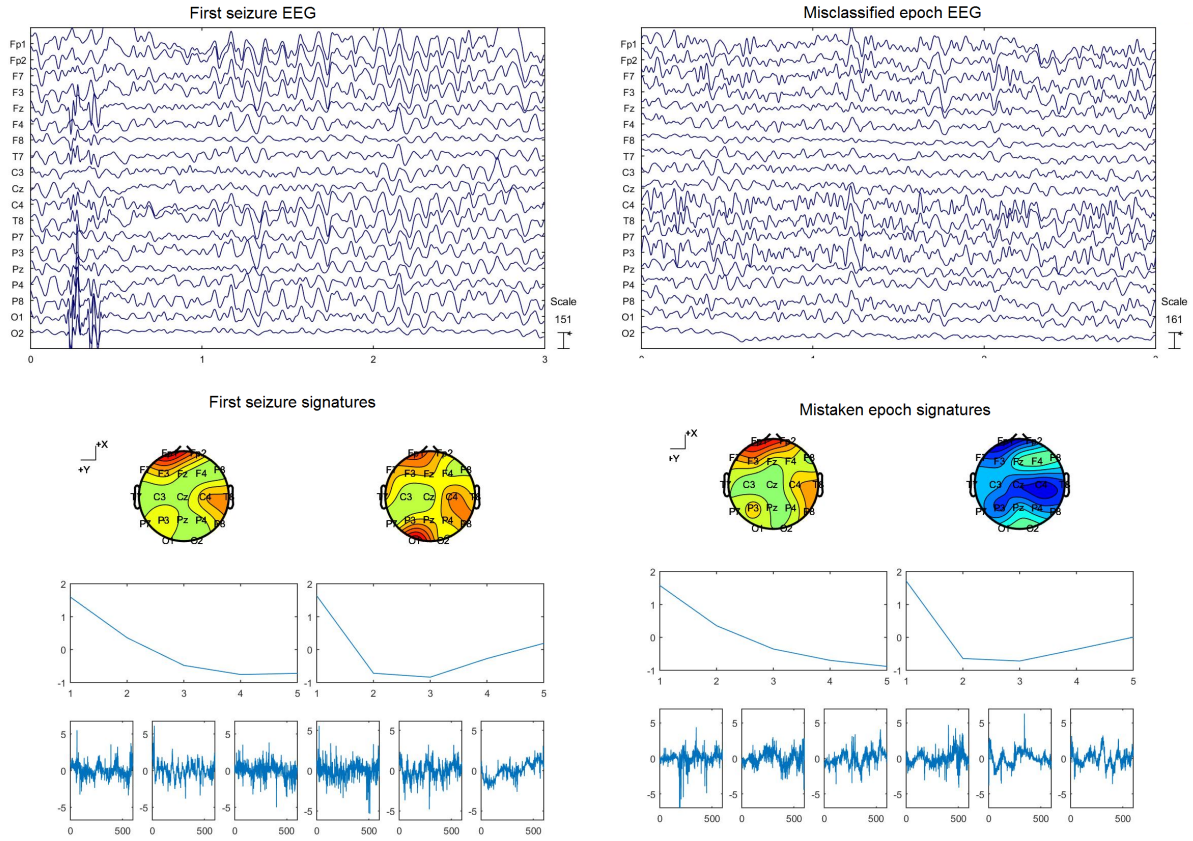


Fig. 8. W_Tensors BTD false positive classification example. Signatures are represented in the following order: Space, Frequency and Time.

DP (over 1) is chosen. From this analysis it is decided to decompose the W_Tensor set with $ml_rank - (2, 6, 2)$ and $R = 1$ for BTD and $R = 2$ for CPD. The HH_Tensor analysis is performed with $ml_rank - (1, 2, 1)$ for BTD and $R = 1$ for both BTD and CPD.

Note that in the case of HH_Tensor, the maximum average MP and DP values of the signatures for CPD are achieved for different tensor representations. The maximum average MP is obtained for $R = 2$ with $MP = 0.68$ and $DP = 3.82$ and the maximum average DP is obtained for $R = 1$ with $MP = 0.66$ and $DP = 5.58$. As can be appreciated the difference in MP values for these tensor representations is small compared to the difference in DP values. Since the task to be solved is a classification task, we do not look for an exact approximation of the data but a sufficiently good approximation that allows separating between seizure and non-seizure data. Therefore, the R with higher average DP value is selected.

E. Classification

Two classifiers are used to perform the NCES detection: A K-Nearest Neighbor classifier (KNN) with K=3 and a Radial Basis SVM (SVMRB). The implementation used for the SVMRB estimates automatically the optimal radial basis kernel and the regularization parameter. The kernel optimization is performed by quadratic programming using the QP solver provided by MatLab. The regularization parameter is estimated by the leave-one-out cross-validation. The classifier selection is based on a previous study [39] where it was concluded that these are the best classifiers when using tensor-based signatures as features. The KNN and SVMRB implementations are available in the MATLAB toolbox PRTools [40].

The classification is performed for each patient individually. The training set for each case is assembled with the epochs from the first seizure (= positive class) and the same number of non-seizure epochs from the EEG prior to the first seizure (=negative class). Details about the training sets are given in Table II. It was impossible to balance the training sets for the cases 2 and 3 since the duration of the EEG prior to the first seizure was too short to meet the number of epochs from the seizure(the first seizure started after a few seconds of recording). Yet, this does not seem to affect the classifier performance. The classification was performed in two ways: (1) with the signatures from the different modes (frequency, time,channel) separately, (2) with all mode signatures assembled.

The classifiers' performance is evaluated in terms of sensitivity, specificity, accuracy and false negative rate. Where the sensitivity is calculated as,

$$Sen = \frac{TP}{TP + FN} \tag{21}$$

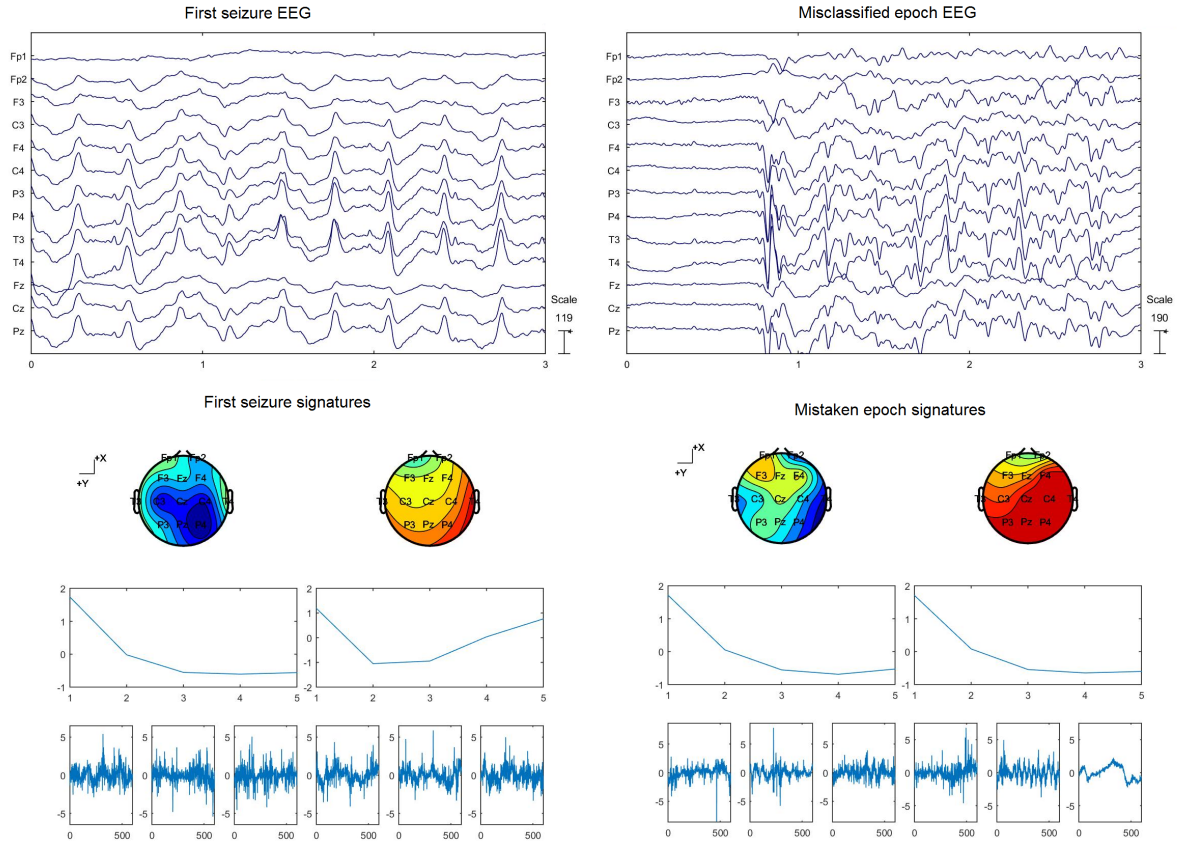


Fig. 9. W_Tensors BTD false negative classification example. Signatures are represented in the fallowing order: Space, Frequency and Time.

The specificity is defined as,

$$Spec = \frac{TN}{TN + FP} \tag{22}$$

The accuracy is termed as,

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{23}$$

The false positive rate denotes the Type I error ratio (false epochs classified as positives) and is defined as

$$FPR = \frac{FP}{FP + TN} \tag{24}$$

Where TP is the number of seizure epochs correctly detected, FP is the number of false alarms, TN is the number of seizure free epochs correctly detected and FN is the number of missed seizure epochs.

III. Results and Discussion

The classification results are presented for the epoch based analysis since the duration of the epochs is the minimum duration of an epileptic event in the EEG. This means that an alarm should be displayed after each positive detection. This is not a practical implementation in real life dynamical setting. For the real life setting, an alarm will be set at the seizure activity onset. Then, subsequent alarms are suppressed and a new alarm is allowed after 10 minutes (low bound established for NCSE diagnosis [1]) of continuous seizure activity.

The classifiers' performance for the frequency features is shown in Fig. 5 A. The best classification results are obtained with a KNN classifier for h_freq_BT signatures with 94.1, 97.1 and 94.5% of specificity, sensitivity and accuracy values respectively,and with a false detection rate per hour of $1.6 * 10^{-5}$ for h_freq_BT. According to the analysis of MP values, frequency signatures offer the best model of the data when they are obtained from the decomposition of W_Tensor set either with CPD or BTD. On the other hand, the highest DP is obtained for h_freq_BT signatures followed by the h_freq_CP signatures. This analysis revealed that even if the signatures allow an accurate prediction model of the original data, as is the case for W_Tensors signatures, discrimination between classes is not guaranteed. The rest of the signatures achieved sensitivity and accuracy values under 80%.

Regarding the Time feature, the highest DP value is obtained with h_time_CP signatures (see Fig. 5 B). Consequently, the best classification results for this feature are achieved by h_time_CP signatures with KNN classifier. The classifier reached 98.6, 84.0 and 92.4% for specificity, sensitivity and accuracy values respectively. This feature achieved a false positive detection rate of $3.5 * 10^{-6}$.

The false positive detection number for the Time feature increases for h_time_BT, w_time_CP, and w_time_BT signatures. This could be explained by the increase in the number of factors used for these signatures. The greater the number of

TABLE II
AVAILABLE DATA FOR NCES AND SEIZURE FREE (N-NCES) CLASS AND THE RESULTING TRAINING SETS FOR EACH EEG RECORD GIVEN IN EPOCHS OF 3S DURATION

	Epochs							
	Available Data		First seizure and prior EEG				Test Set	
			Modeling Data		Training Set			
Record	NCES	n-NCES	NCES	n-NCES	NCES	n-NCES	NCES	n-NCES
1	58	345	4	177	4	4	54	168
2	39	127	26	5	26	5	13	122
3	140	424	12	1	12	1	128	423
4	73	112	8	30	8	8	65	82
5	10	18	3	7	3	3	7	11
6	84	269	13	18	13	13	77	251
7	6	237	3	143	3	3	3	94
8	13	4827	3	18	3	3	7	4809
9	163	4558	3	24	3	3	160	4534
10	1074	3346	19	125	19	19	1055	3221
11	16	1841	12	1381	12	12	4	460
12	875	1036	87	152	87	87	788	884
13	90	860	16	33	16	16	74	827
14	44	472	13	59	13	13	31	413
Total	2685	18472	222	2173	222	190	2466	16299

factors, the more uninteresting patterns are modeled by the signatures, which increases uncertainty in the classification task.

The combination of the signatures obtained with CPD and BTD is unique. Yet, the computed components (signatures) on their own are not. When more than one component is computed, it is expected that some of them fit the pattern of interest (hopefully most of them), and others model the noise in the data or other underlying patterns that may appear. The order in which these components are obtained is not unique and may vary from epoch by epoch. It is not possible to determine which component contains NCES information based on it. Then, the features computed from different epochs may not correspond each to other, thereby hampering the classification based only on the signatures.

We further discuss the Space signature performance. As low DP values were obtained for W_Tensors CPD and BTD space signatures, the classifier performance for these signatures is lower than that obtained for the HH_Tensors signatures, Fig.5 C. The highest specificity, sensitivity and accuracy values for the Space feature are achieved by SVMRB with 99.0%, 98.8% and 99.0% respectively for the h_spac_CP signatures. The KNN classifier obtained similar results for these signatures with a specificity of 99.8%, a sensitivity of 98.6%, and accuracy of 98.2%. The h_spac_CP signatures combined with SVMRB achieved a $3.4 * 10^{-6}$ as false positive detection rate value.

The KNN classifier achieved sensitivity and accuracy values below 85% for h_spac_BT. It is expected to obtain lower sensitivity and accuracy values for h_spac_BT signatures given that the DP for these is 3.9 compared to 6.4 computed for h_spac_CP. Yet, the SVMRB classifier obtained accuracy and sensitivity values over 85% outperforming KNN. Even when lower performance scores are expected for h_spac_BT based on the DP values, these should not be as low as those obtained for KNN, since a 3.9 is not a negligible DP value. Hence, a higher performance could be expected from KNN. Inpection of the h_spac_BT signatures showed that scatter distribution of these rsulted in more overlap between the classes, explaining KNN poor performance.

The training set using the assembled signatures achieved similar results to those achieved by the space signature. The best classifier for the assembled signatures training is KNN for HH_Tensor CPD signatures with a specificity value of 98.1%, sensitivity value of 99.9% and accuracy of 90.5%. The average DP computed for these signatures is 5.5, it is lower than the 7.5 obtained for h_ass_BT signatures. Yet, if the DP values are analyzed individually for each signature, it can be noticed that h_freq_BT displays a large DP of 21.9 while the DP values of the remaining h_time_BT and h_spac_BT are low, 2.2 and 2.6 respectively. The DP values for HH_Tensor CPD are 5.9, 4.4 and 6.4 for h_freq_CP, h_time_CP and h_spac_CP respectively. Thus, this implies that the assembled training results will depend directly on the individual signatures quality. The false detection rate for the assembled HH_Tensor CPD training are the same as achieved by the h_spac_CP signatures training.

The performed experiments showed that the tensor built with HHT separates better the NCES data, thereby outperforming the tensor obtained with the WT. Note that we also tested datasets created using the energy of the wavelets coefficients and other mother wavelets. None of these data sets obtained better results in the training than those obtained using the wavelet coefficients computed with the Mexican Hat mother wavelet or HHT.

The HH_Tensor signatures obtained with CPD outperform those computed with BTD in most of the training sets. This is somewhat unexpected, as NCES is known to evolve along the EEG. A low-rank BTD is a suitable model for analyzing such evolving patterns [22]. The NCES can evolve along the EEG

record in frequency, morphology (time), or location (space). The Salzburg Consensus Criteria for Non-Convulsive Status Epilepticus proposed in [6] by adopting the American Clinical Neurophysiology Society (ACNS) criterion provided more detailed, unambiguous guidelines about the characteristics of this "evolution" in each of these dimensions which could clarify this unexpected result. According to the ACNS, evolution in frequency is defined as at least 2 consecutive changes in the same direction by at least $0.5Hz$, e.g. from 2 to 2.5 to $3Hz$, or from 3 to 2 to $1.5Hz$; Evolution in morphology implies at least 2 consecutive changes to a novel morphology; Evolution in location is defined as sequentially spreading into or sequentially out of at least two different standard 10-20 electrode locations. In order to qualify as "present", a single frequency or location must persist at least $3s$ [6]. This implies that the changes usually occur slowly, and a certain pattern tends to persist 3s or longer. The CPD model assumes that the source data preserve the same frequency and location within the observed epoch [22]. Due to the analyzed epochs duration ($3s$), it is therefore unlikely to observe a lot of evolution in frequency, time or space occurring within the same epoch. Then, the NCES data analyzed in these small epochs fulfill the trilinear criteria justifying CPD and explaining why BTD overfits the model.

Both classifiers, KNN and SVMBR, showed similar performance for classification for the different training sets. However, different combinations of signatures and classifiers performed better for the different training scenarios(e.g. h_spac_BT combined with SVMRB outperform the combination of h_spac_BT and KNN, h_freq_BT combined with KNN outperform h_freq_BT combined with SVMRB). Therefore, is not possible to establish a clear superiority of one over the other given the obtained results.

As the achieved false positive rate values suggest, Fig.6, the number of non-seizure events classified as seizures is low. The majority of these misclassifications are due to the appearance of long duration paroxysms (Fig 7) or a large number of successive spikes in the EEG. These events are long enough to resemble a seizure but too brief to be classified as seizure according to the physician's criteria. False positive detections also occur at epochs with pre-ictal or post-ictal EEG activity as can be seen in Fig. 8.

False negative detections are found in records longer than 10 hours. These are related to changes in the EEG patterns of the most distant seizures in time since the first detected seizure (Fig. 9). These changes may be associated with the patient's disease evolution. Including new EEG patterns from subsequent seizures to the initial training set could help to solve this.

Some misclassifications are caused by high-frequency artifacts probably related to poor electrode contact. These are only found at the epochs where the artifact affected more than one electrode at the same time instant. All signatures proved to be robust to artifacts affecting isolated electrodes.

Table III presents a rough comparison of our results with the results achieved by other methods for NCES detection, in terms of sensitivity and specificity. As can be noticed, the methods from Sharma et. al. [18] and Fatma et. al. [41] obtained superior results in terms of sensitivity than those achieved by the method proposed here. The database used in these publications consists of 13 EEG records, 6 of which have only one seizure. The parameters extracted from epochs of the same seizure tend to display a high similarity. Hence, any classification task made with parameters from epochs of the same seizure will reach high sensitivity values.

TABLE III
PROPOSED METHOD PERFORMANCE COMPARISON WITH OTHER EXISTING METHODS IN TERMS OF SENSITIVITY AND SPECIFICITY.

	Sensitivity (%)	Specificity (%)
Proposed method	98.8	99
Jacquin et.al. [11]	83	96
Minasyan et. al.[15]	71	99
Khan et. al. [16]	86.8	96.9
Sharma et. al. [18]	100	93.3
Fatma et. al.[41]	100	88.02

TABLE IV
PROPOSED METHOD PERFORMANCE COMPARISON WITH OTHER EXISTING METHODS IN TERMS OF ACCURACY.

	Accuracy (%)
Proposed Method	99
Liang et. al. [14]	97.5
Xanthopoulos et. al.[12]	95
Kollialil et. al. [13]	99.6
Fatma et. al.[41]	88.2

In terms of accuracy, Table IV shows that our method is only outperformed by Kollialil et. al. [13]. The method proposed by Kollialil uses an SVM to classify between normal, interictal and seizure EEG. In this publication, three different labeled datasets are used. The classification was not performed in a continuous recording. It should be noted that with such test data, the classifier does not face issues as artifacts, EEG pattern changes or paroxysms as done in our study.

Regarding the execution time, which is a major concern for detection algorithms, our method displays a runtime from $15.2s$ to $3.3min$ for the training step. All tests are performed in a computer with an Intel Core-i3 processor at 1.70 GHz with 8 GB of RAM. The training results are obtained for the EEG records with the shortest (9s) and the largest (4.35min) first seizure respectively. It should be noted, that the training will only be performed once at the beginning of the monitoring. The training step includes the tensorization and analysis of all first seizure epochs and the classifier training. Then, the execution time will depend on the first seizure duration. It could be decided to use only part of the seizure for training if this is very long. Yet, it should be kept in mind that doing so fewer training patterns will be included. In this sense, there is a compromise between the number of positive epochs to include in the training set and the quality of the classifier.

The classification runtime ranged from $2.4s$ to $3s$ which means that our method can run in real-time. These results are obtained by classifying 100 epochs from different EEG records. The duration of the classification step will depend on the pattern complexity. As the pattern becomes more complex, more time is required to perform the EMD.

IV. Conclusion

The presented paper introduced a NCES detection method that exploits techniques such as multiway data analysis and HHT. It is the first time that HHT is used for seizure detection, and the first algorithm to apply tensor methodology on this type of seizures. The HHT provides an accurate definition of particular events in the $time \times frequency$ space capturing the NCES data in a tensor close to R=1. This turns all signatures obtained from HH Tensor decomposition in relevant features to perform the NCES detection.

The best results at the classification were achieved for HH_Tensor Space and Assembled features computed with CPD combined with the SVMRB classifier, closely followed by HH_Tensor Frequency features computed with BTD combined with KNN.

According to the obtained results, the proposed method proved to be one of the best NCES methods so far. The method was validated in continuous clinical data in a realistic training setting proving that can run in real-time.

The number of false positives for the training with HH_Tensor Space and Assembled signatures obtained with CPD was low. Yet, the appearance of EEG activity resembling the seizure affected the performance of these signatures. The false negatives found were related to EEG pattern changes at the seizures distant in time from the first recorded seizure. Due to this seizure EEG pattern evolution caused by the underlying disease progression, adaptive training strategies must be implemented to improve this.

References

[1] E. Trinka, H. Cock, D. Hesdorffer, A. O. Rossetti, I. E. Scheffer, S. Shinnar, S. Shorvon, and D. H. Lowenstein, "A definition and classification of status epilepticus–report of the ilae task force on classification of status epilepticus," *Epilepsia*, vol. 56, no. 10, pp. 1515–1523, 2015.

[2] P. M. Vespa, V. Nenov, and M. R. Nuwer, "Continuous eeg monitoring in the intensive care unit: early findings and clinical efficacy," *Journal of Clinical Neurophysiology*, vol. 16, no. 1, pp. 1–13, 1999.

[3] D. L. Schomer and F. L. Da Silva, *Niedermeyer's electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2012.

[4] P. W. Kaplan, "Eeg criteria for nonconvulsive status epilepticus," *Epilepsia*, vol. 48, no. s8, pp. 39–41, 2007.

[5] E. Trinka and M. Leitinger, "Which eeg patterns in coma are nonconvulsive status epilepticus?" *Epilepsy & Behavior*, vol. 49, pp. 203–222, 2015.

[6] M. Leitinger, S. Beniczky, A. Rohracher, E. Gardella, G. Kalss, E. Qerama, J. Höfler, A. H. Lindberg-Larsen, G. Kuchukhidze, J. Dobesberger *et al.*, "Salzburg consensus criteria for non-convulsive status epilepticus–approach to clinical application," *Epilepsy & Behavior*, vol. 49, pp. 158–163, 2015.

[7] A. T. Tzallas, D. G. Tsalikakis, E. C. Karvounis, L. Astrakas, M. Tzaphlidou, M. G. Tsipouras, and S. Konitsiotis, *Automated epileptic seizure detection methods: a review study*. Citeseer, 2012.

[8] L. Orosco, A. G. Correa, and E. Laciar, "Review: a survey of performance and techniques for automatic epilepsy detection," *Journal of Medical and Biological Engineering*, vol. 33, no. 6, pp. 526–537, 2013.

[9] A. Q. Ansari and P. Sharma, "A review on automated detection of nonconvulsive seizures using eeg," in *Computational Intelligence & Communication Technology (CICT), 2016 Second International Conference on*. IEEE, 2016, pp. 283–286.

[10] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of eeg records in an epileptic patient using wavelet transform," *Journal of neuroscience methods*, vol. 123, no. 1, pp. 69–87, 2003.

[11] A. Jacquin, E. Causevic, and E. R. John, "Automatic identification of spike-wave events and non-convulsive seizures with a reduced set of electrodes," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 1928–1931.

[12] P. Xanthopoulos, S. Rebennack, C.-C. Liu, J. Zhang, G. L. Holmes, B. M. Uthman, and P. M. Pardalos, "A novel wavelet based algorithm for spike and wave detection in absence epilepsy," in *BioInformatics and BioEngineering (BIBE), 2010 IEEE International Conference on*. IEEE, 2010, pp. 14–19.

[13] E. S. Kollialil, G. K. Gopan, A. Harsha, and L. A. Joseph, "Single feature-based non-convulsive epileptic seizure detection using multi-class svm," in *Emerging Trends in Communication, Control, Signal Processing & Computing Applications (C2SPCA), 2013 International Conference on*. IEEE, 2013, pp. 1–6.

[14] S.-F. Liang, W.-L. Chang, and H. Chiueh, "Eeg-based absence seizure detection methods," in *Neural Networks (IJCNN), The 2010 International Joint Conference on*. IEEE, 2010, pp. 1–4.

[15] G. R. Minasyan, J. B. Chatten, and R. N. Harner, "Detection of epileptiform activity in unresponsive patients using ann," in *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*. IEEE, 2009, pp. 2117–2124.

[16] Y. U. Khan, O. Farooq, M. Tripathi, P. Sharma, and P. Alam, "Automatic detection of non-convulsive seizures using ar modeling," in *Power, Control and Embedded Systems (ICPCES), 2012 2nd International Conference on*. IEEE, 2012, pp. 1–4.

[17] J. C. Sackellares, D.-S. Shiau, J. J. Halford, S. M. LaRoche, and K. M. Kelly, "Quantitative eeg analysis for automated detection of nonconvulsive seizures in intensive care units," *Epilepsy & Behavior*, vol. 22, pp. S69–S73, 2011.

[18] P. Sharma, Y. U. Khan, O. Farooq, M. Tripathi, and H. Adeli, "A wavelet-statistical features approach for nonconvulsive seizure detection," *Clinical EEG and neuroscience*, vol. 45, no. 4, pp. 274–284, 2014.

[19] S. Varshney, Y. U. Khan, O. Farooq, P. Sharma, and M. Tripathi, "Latency study of non-convulsive seizures," in *Multimedia, Signal Processing and Communication Technologies (IMPACT), 2013 International Conference on*. IEEE, 2013, pp. 108–111.

[20] H. Qu and J. Gotman, "A seizure warning system for long-term epilepsy monitoring," *Neurology*, vol. 45, no. 12, pp. 2250–2254, 1995.

[21] E. Acar, C. Aykut-Bingol, H. Bingol, R. Bro, and B. Yener, "Multiway analysis of epilepsy tensors," *Bioinformatics*, vol. 23, no. 13, pp. 10–18, 2007.

[22] B. Hunyadi, D. Camps, L. Sorber, W. Van Paesschen, M. De Vos, S. Van Huffel, and L. De Lathauwer, "Block term decomposition for modelling epileptic seizures," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 139, 2014.

[23] W. Deburchgraeve, P. J. Cherian, M. De Vos, R. M. Swarte, J. H. Blok, G. H. Visser, P. Govaert, and S. Van Huffel, "Neonatal seizure localization using parafac decomposition," *Clinical Neurophysiology*, vol. 120, no. 10, pp. 1787–1796, 2009.

[24] C. Guerrero-Mosquera, M. Verleysen, and A. N. Vazquez, "Eeg feature selection using mutual information and support vector machine: A comparative analysis," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010, pp. 4946–4949.

[25] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.

[26] N.-F. Chang, T.-C. Chen, C.-Y. Chiang, and L.-G. Chen, "On-line empirical mode decomposition biomedical microprocessor for hilbert huang transform," in *Biomedical Circuits and Systems Conference (BioCAS), 2011 IEEE*. IEEE, 2011, pp. 420–423.

[27] E. A. Atamon, *Understanding epilepsy seizure structure using tensor analysis*. Rensselaer Polytechnic Institute, 2008.

[28] R. Zink, B. Hunyadi, S. Van Huffel, and M. De Vos, "Tensor-based classification of an auditory mobile bci without a subject-specific calibration phase," *Journal of neural engineering*, vol. 13, no. 2, p. 026005, 2016.

[29] M. Sørensen and L. De Lathauwer, "Tensor decompositions with block-toeplitz structure and applications in signal processing," in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*. IEEE, 2011, pp. 454–458.

[30] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of n-way arrays," *Journal of chemometrics*, vol. 14, no. 3, pp. 229–239, 2000.

[31] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms–part iii: Alternating least squares algorithms," *SIAM journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067–1083, 2008.

[32] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank-(r 1, r 2,..., rn) approximation of higher-order tensors," *SIAM journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.

[33] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," *available online, URL: www. tensorlab. net*, 2016.

[34] S. Wold and M. SJÖSTRÖM, "Simca: A method for analyzing chemical data in terms of similarity and analogy," in *Chemometrics: theory and application*. ACS Publications, 1977, pp. 243–282.

[35] S. Dabbaghchian, M. P. Ghaemmaghami, and A. Aghagolzadeh, "Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology," *Pattern Recognition*, vol. 43, no. 4, pp. 1431–1440, 2010.

[36] I. T. Bustamante, F. S. Mata, N. H. González, R. G. Gazapo, J. Palau, and M. M. C. Ferreira, "Application of chemometric tools for automatic classification and profile extraction of dna samples in forensic tasks," *Analytica chimica acta*, vol. 595, no. 1, pp. 43–50, 2007.

[37] S. Bicciato, A. Luchini, and C. Di Bello, "Pca disjoint models for multiclass cancer analysis using gene expression data," *Bioinformatics*, vol. 19, no. 5, pp. 571–578, 2003.

[38] E. A. Kanık, G. O. Temel, S. Erdoğan, and İ. E. Kaya, "Affected states soft independent modeling by class analogy from the relation between independent variables, number of independent variables and sample size," *Balkan medical journal*, vol. 30, no. 1, p. 28, 2013.

[39] Y. Rodríguez Aldana, B. Hunyadi, J. E. Marañón Reyes, V. Rodríguez Rodríguez, and S. Van Huffel, "Nonconvulsive epileptic seizure dtection using multiway data analysis," in *EUSIPCO 2017*. IEEE, 2017, p. 5.

[40] R. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. De Ridder, D. Tax, and S. Verzakov, "Prtools 4.1," *A Matlab Toolbox for Pattern Recognition, Software and Documentation downloaded May*, 2010.

[41] T. Fatma, O. Farooq, Y. U. Khan, M. Tripathi, and P. Sharma, "Automatic detection of non-convulsive seizures: A reduced complexity approach," *Journal of King Saud University-Computer and Information Sciences*, vol. 28, no. 4, pp. 407–415, 2016.



Yissel Rodríguez Aldana received the degree in informatic sciences engineering at the University of Informatics Sciences, Havana, Cuba and the M.Sc. degree on biomedical engineering from the Universidad de Oriente, Santiago de Cuba, Cuba in 2008 and 2013 respectively. She is currently working towards the Ph.D. degree in both, Universidad de Oriente, Cuba and KU Leuven, Belgium. Her research interests include machine learning and tensor methods applied to biomedical signal processing.



Borbála Hunyadi received a M.Sc. degree in computer and electrical engineering from the Pzmny Pter Catholic University, Hungary in 2009. Afterwards she joined STADIUS, Department of Electical Engineering (ESAT), KU Leuven as a PhD student in close collaboration with the Laboratory for Epilepsy Research, KU Leuven and obtained her doctorate in electrical engineering in 2014. Since then she has been working as a postdoctoral researcher on the ERC Advanced Grant project 'BIOTENSORS: Biomedical Data Fusion using Tensor based Blind Source Separation'. Her current research interests include developing signal processing and machine learning solutions for pattern recognition analysis, blind source separation and fusion of EEG and fMRI signals.



Enrique J. Maraón Reyes received the BS in electrical engineering in telecomunications in 1964 at the Universidad de la Habana, Havana, Cuba. Obtained the Ph.D. in Signals and Systems in 1985 at the Technical University of Praha. He is full professor at the department of Telecommunications and Electronics and Emeritus Professor of the Universidad de Oriente, Cuba. Is one of the founders 2004 and first Director of the Study Center of Neuroscience and Image and Signal Processing, Universidad de Oriente, Cuba. Where he currently works.



Valia Rodríguez Rodríguez received the MD in medical sciences in 1989 and the Certificate of specialist training in Clinical Neurophysiology in 1993 from University of Medical Sciences of Havana. Obtained the Ph.D. in Cognitive Neuroscience from the National Center for Scientific Research in Havana in 2003. She is a Professor and Senior Research Scientist at the Cuban Neuroscience Center, and a Consultant Clinical neurophysiologist in intensive care unit of the Clinical-Surgical Hospital 'Hnos Ameijeiras', Havana, Cuba. Recently she has joined as Lecturer in Neurophysiology, at the Aston University, Birmingham, UK.



Sabine VAN HUFFEL received the MD in computer science engineering in June 1981, the MD in Biomedical engineering in July 1985 and the Ph.D in electrical engineering in June 1987, all from KU Leuven, Belgium. She is full professor at the department of Electrical Engineering, KU Leuven, Belgium, and also affiliated with imec Leuven. In April 2013 she received an honorary doctorate from Eindhoven University of Technology, together with an appointment as a Distinguished professor.

She is heading the Biomedical Data Processing Research Group with focus on the development of numerically reliable and robust algorithms for improving medical diagnostics.