Nonconvulsive Epileptic Seizure Detection in Scalp EEG Using Multiway Data Analysis

Yissel Rodríguez Aldana, Borbála Hunyadi, *Member, IEEE*, Enrique J. 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, a Radial Basis Support Vector Machine, and a Linear Discriminant Analysis classifier are used. The classifier features are obtained from the Canonical Polyadic Decomposition (CPD) and Block Term Decomposition of the EEG data represented as 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 Radial Basis Support Vector Machine classifier shows the best performance with sensitivity, specificity, and accuracy values over 98%. A rough comparison with other methods proposed in the literature shows the superior performance of the proposed method for nonconvulsive epileptic seizure detection.**

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

Manuscript received November 1, 2017; revised February 5, 2018 and March 6, 2018; accepted April 17, 2018. Date of publication April 27, 2018; date of current version March 6, 2019. This work was 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, in part by IMEC funds 2017, and in part by the European Research Council under the European Union's Seventh Framework Programme FP7/2007-2013/ERC Advanced Grant: BIOTENSORS 339804. *(Corresponding author: Yissel Rodríguez Aldana.)*

Y. R. Aldana and E. J. Marañón Reyes are with the Universidad de Oriente, Center of Neuroscience and Signals and Image Processing, Santiago de Cuba 90500, Cuba (e-mail: yaldana@uo.edu.cu; enriquem@uo.edu.cu).

B. Hunyadi and S. Van Huffel are with KU Leuven, Department of Electrical Engineering (ESAT), Stadius Center for Dynamical Systems, Signal Processing and Data Analytics and IMEC, Leuven 3001, Belgium (e-mail: borbala.hunyadi@esat.kuleuven.be; sabine.vanhuffel@esat.kuleuven.be).

V. R. Rodríguez is with the ICU-Almeijeiras Clinic Surgical Hospital, La Habana, Cuba, and also with the Aston University, Birmingham B4 7ET, U.K. (e-mail: v.rodriguez@aston.ac.uk).

Digital Object Identifier 10.1109/JBHI.2018.2829877

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 EEG without evident physical symptoms.

Nonconvulsive epileptic seizures (NCES) can be found associated with coma/stupor or not (i.e. absence status, focal SE). Patients without coma/stupor are often awake and are treated in 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 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].

A big number of published seizure detection methods address CES [4], [5] and only a few of them address NCES [6].

2168-2194 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

The proposed methods for NCES detection published until 2015 are summarized in the work of Ansari and Sharma [6]. These methods can be split according to their principal aim: those that detect NCES without coma/stupor (i.e. absence seizures) [7]–[11] and the ones that perform the NCES detection on ICU patients [12]–[16]. 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 [7]–[9], [16], Entropy [10], [11] and nonlinear parameters [8], [11], [12]. 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 [9], [14]–[16]. 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 [17]. 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 [18], [19]. However, we propose 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 (EMD) provides a more meaningful physical interpretation of the underlying EEG data. The tensor decomposition is performed with CPD and BTD. It has been shown that tensor decompositions of multiway EEG representations can extract seizure sources and accurately characterize the seizure pattern [18]–[20]. 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 eliminates the need for thresholding since the classifiers will implicitly evaluate the similarity of the signatures. Fig. 1 displays the general block diagram of the proposed method. Clinically, the goal of the algorithm is twofold. First, it indicates all seizure occurrences (sz flag). Besides, it sets an alarm at the very first seizure, indicating that the patient has a seizure disorder, and subsequently sets an alarm in case of continuous seizure activity, indicating that the patient might develop status epilepticus.

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 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 three neurophysiologists (including 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 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' 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 segments (epochs) of 3 seconds long. All epochs are expanded in the time-frequency domain using a WT or HHT. A 3rd order tensor is built from every epoch with $frequency \times time \times channels$ axes.

1) Wavelet Transform: The Continous WT (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 by the authors to perform the epileptic seizure detection [5]. 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 [21].

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 scales (F), the columns time instants (T) and the elements are coefficients. The Ch dimension represents the spatial distribution among the channels. Since the NCES activity is below 5 Hz [22], five wavelet scales were computed, which correspond to $1-5$ Hz frequencies, with 1 Hz resolution. As mother wavelet, a Mexican Hat is used as reported in [18], [19]. The tensors obtained with CWT will be referred further on as W_Tensor.

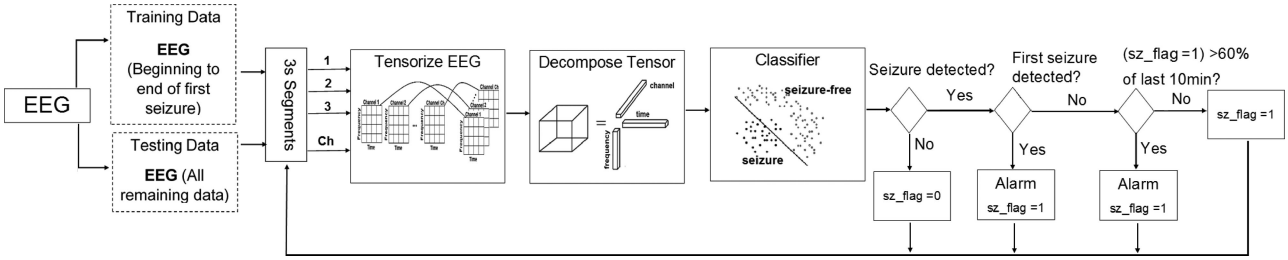


Fig. 1. General block diagram of the proposed method.

TABLE I
DETAILED DESCRIPTION OF THE EEG DATABASE

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

2) Hilbert-Huang Transform: The 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 EMD. Second, the Hilbert transform is applied to those components to construct a $frequency \times time \times energy$ distribution, designated 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 [23].

The main characteristic of EMD is to decompose a signal into a so-called Intrinsic Mode Function (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 [24]. For any one-dimensional discrete signal, EMD can be presented with 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 [23].

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

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

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

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

where $a_{\mathrm{imf_k}}$ is the instantaneous amplitude and $\theta_{\mathrm{imf_k}}$ is the instantaneous phase defined as,

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

and

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

respectively.

Then the instantaneous frequency is computed as

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

The $x_{\mathrm{HH}}(w,t)$ matrices are now constructed, representing the energy at each time point t and at each instantaneous frequency w the in following way,

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

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

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

As implemented with CWT, the HHT is applied to every EEG channel. The Hilbert spectrum is truncated at 5 Hz. The $x_{\mathrm{HH}}(w,t)$ 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 $1-5$ Hz 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 [18]–[20], [25]. This fact makes CPD an attractive method for our study. On the other hand, BTD is more flexible, and facilitates a more accurate model when the data frequency or topography change over 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 [19], [26]. 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 CPD represents a third-order tensor $X \in \mathbb{R}^{(F\times T\times Ch)}$ as the outer product (represented as '$\circ$') of rank-1 tensors in the following way,

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

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

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) [27]. A tensor CPD is unique under mild conditions up to permutation and scaling [27]. Let $A = [a_1 \ldots a_R]$, $B = [b_1 \ldots b_R]$ and $C = [c_1 \ldots c_R]$ be the factor matrices corresponding to each *mode-n*, 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 $k_A < r_A$ (r_A rank of the factor mattrix A). According to Kruskal's theorem [28] a sufficient condition for the uniqueness of a third-order tensor CPD up to permutation and scaling uncertainty is given by,

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

2) Block Term Decomposition (BTD): 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 $F \ge L$), $B_r \in \mathbb{R}^{(T\times M)}$ (with $T \ge M$), and $C_r \in \mathbb{R}^{(Ch\times O)}$ (with $Ch \ge O$) are full column rank, $1 \le r \le R$. BTD is unique up to permutation and scaling inherited from CPD under more restrictive conditions. Essential uniqueness conditions for the (L,M,O)-BTD can be found in [29].

The two tensor sets obtained, 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 [30]. 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 $U^{(n)}$, $n = 1,\ldots,N$, $N = 3$, are the orthonormal basis for the n different subspaces of the *mode-n* vectors and $Z \in \mathbb{R}^{(L\times M\times O)}$ is an all-orthogonal and ordered tensor [31]. 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) [31].

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 [31].

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

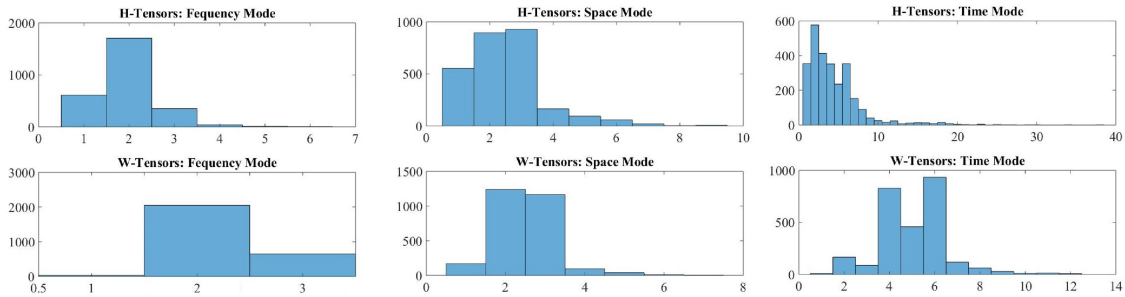


Fig. 2. LMLRA analysis Frequency, Time and Space modes histograms. Axis X and Y represent the number of components and tensors, respectively.

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 space, $(R_f, R_t, R_{\text{ch}})$). Afterwards, rank histograms are constructed by counting the number of tensors with a certain rank, as shown in Fig. 2 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 power (DP) of the signatures obtained from the CPD and BTD with a certain rank. The MP describes the relevance of the variable 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 [32] for feature selection in chemical applications. Subsequently, these measures have been used for feature selection in diverse applications as image processing [33], DNA samples selection [34], and biomedical problems modeling [35], [36]. The MP and DP analysis were performed for the signatures obtained from the tensors computed from the EEG epochs corresponding to the first seizure and all available seizure-free epochs prior to it.

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. $I_A = Ep \cdot L$, $I_B = Ep \cdot M$, and $I_C = Ep \cdot O$ where, Ep is the number of EEG epochs analyzed, L, M, and O are the ml-rank and $A_j \in \mathbb{R}^{F \times L}$, $B_j \in \mathbb{R}^{T \times M}$ 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, $j = 1, \ldots, Ep$ with

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

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 columns of H^A. Then, the MP of the jth column of H (H_j) is defined as [32] [34]:

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

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:

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

where $P_{\lambda^H}^H$ 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 [32].

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 $\lambda^{H_{\mathrm{sz}}}$ and $\lambda^{H_{\mathrm{nsz}}}$ PCs that represent 95% of the features variance (for each class) are retained. Then, two PCA models are obtained with $\lambda^{H_{\mathrm{sz}}}$ and $\lambda^{H_{\mathrm{nsz}}}$ PCs. The DP between the two classes, sz and nsz, of the feature j is defined as,

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

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

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

and

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

where $\lambda^{H_{\mathrm{nsz}}}$ and $\lambda^{H_{\mathrm{sz}}}$ are the number of retained PCs for H_{nsz} and H_{sz} PCA models respectively. The $S_j^{H_{\mathrm{nsz}}}(H_{\mathrm{sz}})^2$ and $S_j^{H_{\mathrm{sz}}}(H_{\mathrm{nsz}})^2$ 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 [32], these terms are defined as,

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

and

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

where J is the number of columns of H_{nsz} and H_{sz}, $n^{H_{\mathrm{sz}}}$ and $n^{H_{\mathrm{nsz}}}$ 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 [32].

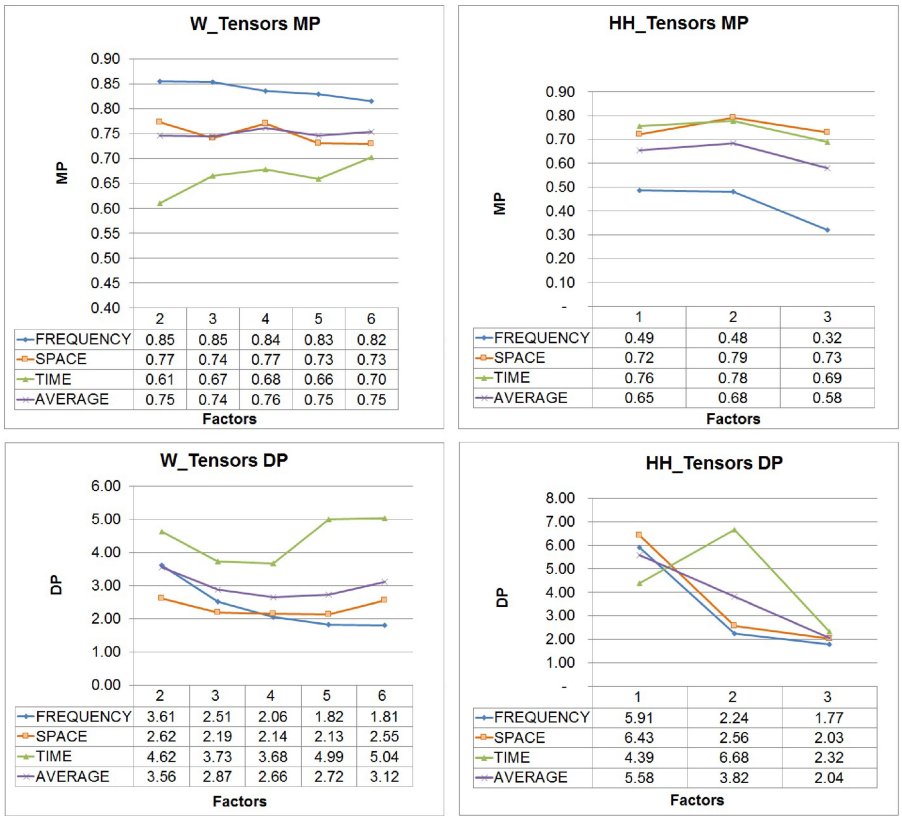
Fig. 3 shows the computed MP and DP values for the signatures obtained from HH_Tensor and W_Tensor after applying CPD. Based on the LMLRA (see Fig. 2) the element of the *ml-rank* triplet $(R_f, R_t, R_{\mathrm{ch}})$ most commonly occurring at each mode is inspected. Then, the largest rank among the three modes, denoted as R, is selected as an upper bound to perform CPD. This led to performing CPD with $R = 1, 2, 3$ for HH_Tensor and $R = 2, \ldots, 6$ for W_Tensor. BTD is applied to all rank combinations of $R_f = 1, 2$, $R_t = 1, 2, 3$ and $R_{\mathrm{ch}} = 1, 2, 3$ for HH_Tensor and $R_f = 2$, $R_t = 4, 6$ and $R_{ch} = 2, 3$ for W_Tensor (where R_f, R_t, and R_{ch} are the possible rank values of the Frequency, Time and Space modes respectively). Combinations violating the multilinear rank property rules $R_f < R_t \times R_{\mathrm{ch}}$, $R_t < R_f \times R_{\mathrm{ch}}$ and $R_{\mathrm{ch}} < R_f \times R_t$ were removed [29], [30], [31]. This results in a total of 8 different tensor representations for HH_Tensor (excluding the *ml_rank*-$(1,1,1)$ since it is equivalent to an $R = 1$ CPD) and 3 tensor representations for W_Tensor.

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 DP (over 1) is chosen. From this analysis, it is decided to decompose the W_Tensor set with *ml-rank* $L = (2, 4, 2)$ and $R = 1$ for BTD and $R = 2$ for CPD. For HH_Tensor BTD analysis is performed with *ml-rank-l* $= (2, 1, 2)$ and $R = 1$ for BTD, $R = 1$ for 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

Three classifiers are used to perform the NCES detection: A K-Nearest Neighbor classifier (KNN) with $K = 3$, a Radial Basis Support Vector Machine (SVMRB), and a Linear Discriminant Analysis (LDA). 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 [37] where it was concluded



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
AVERAGE	0.75	0.74	0.76	0.75	0.75

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
AVERAGE	0.65	0.68	0.58

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
AVERAGE	3.56	2.87	2.66	2.72	3.12

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
AVERAGE	5.58	3.82	2.04

Fig. 3. Obtained MP and DP for HH_Tensors and W_Tensor CPD signatures computed for several numbers of factors. The displayed MP values are computed for the features matrices of the first NCES. The MP and DP values displayed on the charts are averaged over the number of features of each mode.

that these are the best classifiers when using tensor-based signatures as features. The KNN and SVMRB implementation are available at the MATLAB toolbox PRTools [38].

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, Space) 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,

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

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 seizure-free epochs classified as seizure, 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 (lower bound established for NCSE diagnosis [1]) of continuous seizure activity.

The classifiers' performance for the Frequency signatures is shown in Fig. 4(a). The best classification results are obtained

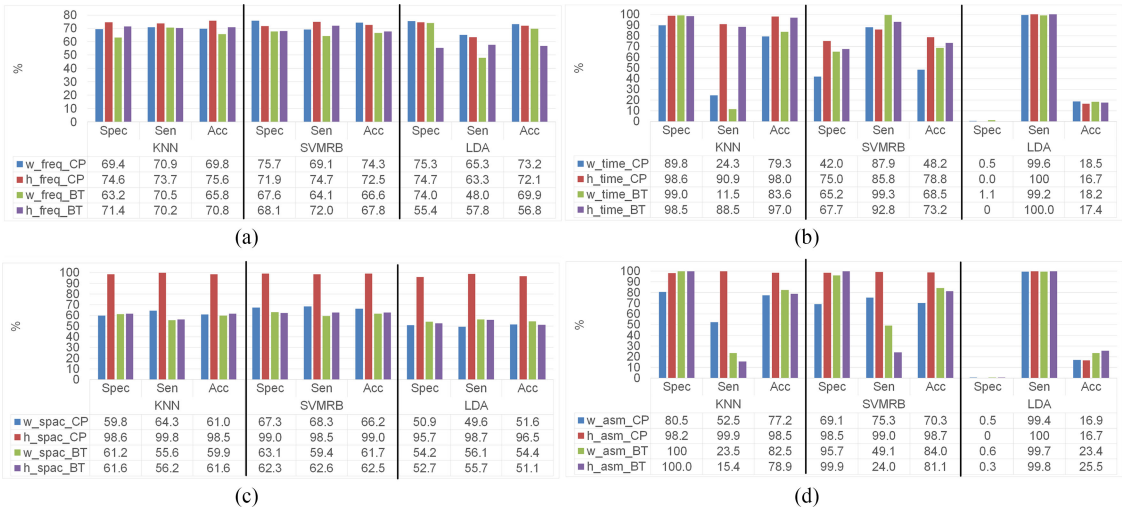


Fig. 4. Specificity, Sensitivity, and Accuracy scores obtained in the classification of the test set for KNN, SVMRB and LDA. The results for the different signatures are displayed in the following order: (a) Frequency, (b) time, (c) space, and (d) assembled signatures. The columns in the graph for each classifier stand for specificity, sensitivity, and accuracy respectively.

by a KNN classifier for h_freq_CP signatures with a 74.6, 73.7, and 75.6% of specificity, sensitivity, and accuracy values respectively, and with a false detection rate of 0.385. According to the analysis of MP values, Frequency signatures offer the best model of the data when they are obtained from the decomposition of the W_Tensor set either with CPD or BTD. The performance for the frequency signatures for all decompositions is very similar. 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 Frequency signatures achieved sensitivity and accuracy values under 80%.

Regarding the Time feature, the highest DP value is obtained with h_time_CP signatures. Consequently, the best classification results for this feature are achieved by h_time_CP signatures with a KNN classifier [see Fig. 4(b)]. The classifier reached 98.6, 90.9 and 98.0% of specificity, sensitivity and accuracy values respectively. This feature achieved a false positive detection rate of 0.010.

The false positive detection number for the Time signatures increases for 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 factors, the more uninteresting patterns are modeled by the signatures, which increases the uncertainty in the classification task.

The combination of the signatures obtained with CPD and BTD is unique. Yet, the components computed (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 to epoch. It is not possible to determine which component contains NCES information based on it. Hence, the features computed from different epochs may not correspond to each 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. 4(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 98.6%, a sensitivity of 99.8%, and accuracy of 98.5%. The h_spac_CP signatures combined with SVMRB achieved a 0.014 as false positive detection rate value per hour.

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 SVMRB for HH_Tensor CPD signatures with a specificity value of 98.5%, a sensitivity value of 99.0% and accuracy of 98.7%.

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 [19]. The NCES can evolve along the EEG record in frequency, morphology (time), or location (space). The Salzburg Consensus Criteria for NCSE proposed in [22] 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 0.5 Hz,

e.g. from 2 to 2.5 to 3 Hz, or from 3 to 2 to 1.5 Hz; 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 3 s [22]. 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 [19]. 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. Hence, the NCES data analyzed in these small epochs fulfill the trilinear criteria justifying CPD and explaining why BTD overfits the model.

Although HHT and WT are used for this proposal, different tensorizations would also be possible, e.g. empirical wavelet transform and tunable-Q wavelet transform [39], [40], which were successfully used to detect epileptic seizures. Such methods could be tried on NCES data in order demonstrate if these algorithms are capable of capturing this seizure activity as well.

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

The LDA classifier showed to be able to classify NCES data with the Frequency and Space signatures. This classifier obtained values of specificity, sensitivity, and accuracy over 95% for the h_spac_CP. However, for high dimensional features such as Time and Assembled signatures LDA showed to be unable to identify the seizure from seizure-free data. For the Time and Assembled signatures, LDA displayed high sensitivity but low values of specificity and accuracy.

As the achieved false positive rates values suggest, Fig. 5, 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. 6) or a large number of successive spikes in the EEG. These events, are long enough to resemble a seizure but are too brief to be classified as a seizure according to the physician's criteria. False positives detections also occur at epochs with pre-ictal or post-ictal EEG activity.

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 compared to the first detected seizure (Fig. 7). 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. This will make the method proposed here robust to the changes in the seizure pattern. There is no need to stop the classification to retrain the classifier, this can be done in parallel. This is a strength comparing our proposal to threshold-based methods. The threshold based methods are unable to respond to seizure changes and background EEG pattern changes. This lack of adaptability makes the threshold based methods less accurate for long-term EEG recordings.

TABLE III
ROUGH COMPARISON OF THE PROPOSED METHOD PERFORMANCE WITH OTHER EXISTING METHODS IN TERMS OF ACCURACY, SENSITIVITY AND SPECIFICITY

Method	Accuracy	Sensitivity	Specificity
Proposed Method	99.0%	98.8%	99.0%
Liang et. al. [14]	97.5%	N/R	N/R
Xanthopoulos et. al. [12]	95.0%	N/R	N/R
Kollialil et. al. [13]	99.6%	N/R	N/R
Fatma et. al. [41]	88.2%	N/R	N/R
Jacquin et. al. [11]	N/R	83.0%	96.0%
Minasyan et. al. [15]	N/R	71.0%	99.0%
Khan et. al. [16]	N/R	86.8%	96.9%
Sharma et. al. [18]	N/R	100.0%	93.3%
Fatma et. al. [41]	N/R	100.0%	88.0%

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.

Based on the performed analysis, the best combination for NCES detection is provided by the Space signatures as features and the SVMRB as classifier. The Space signature must be obtained with HHT and CPD as tensorization and tensor decomposition method respectively.

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. [16] 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.

In terms of accuracy, Table III shows that our method is only outperformed by Kollialil et. al. [11]. 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 6.04 s to 3.3 min for the training step. All tests are performed on 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 (9 s) and the largest (4.35 min) 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. Hence, the execution time will depend on the first seizure duration. It could be decided

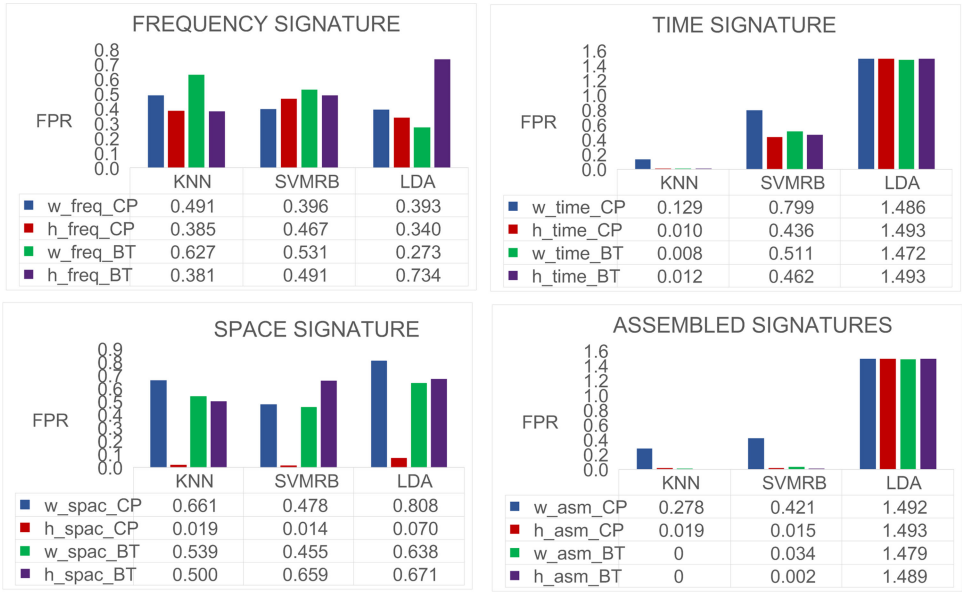


Fig. 5. FPR values displayed by the KNN, SVMRB and LDA classifiers during the different trainings.

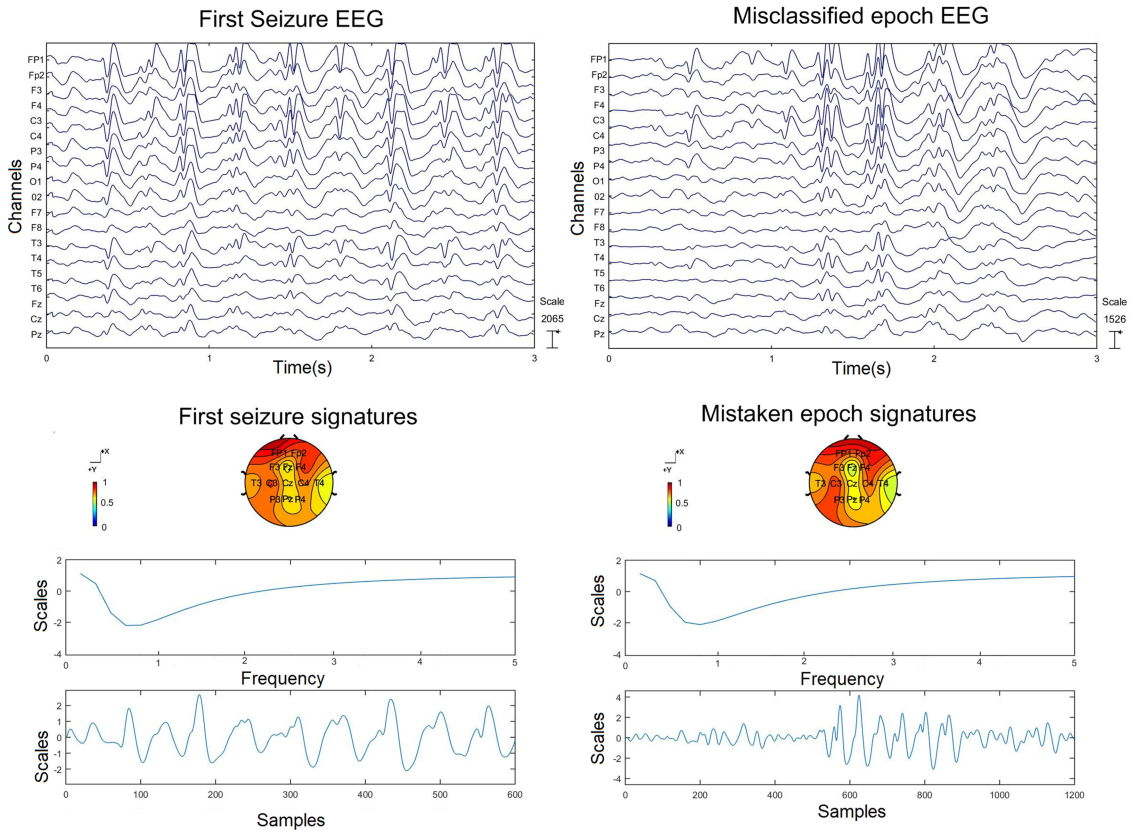


Fig. 6. HH_Tensors CPD: false positive classification example. Misclassified seizure-free epoch shows a short paroxysm of approximately 1 s of seizure-like activity in the EEG. This activity is captured by highly similar signatures as the first seizure's time and frequency signatures. This leads to the detection of the paroxysm as a seizure. Signatures are represented in the following order: Space, Frequency and Time.

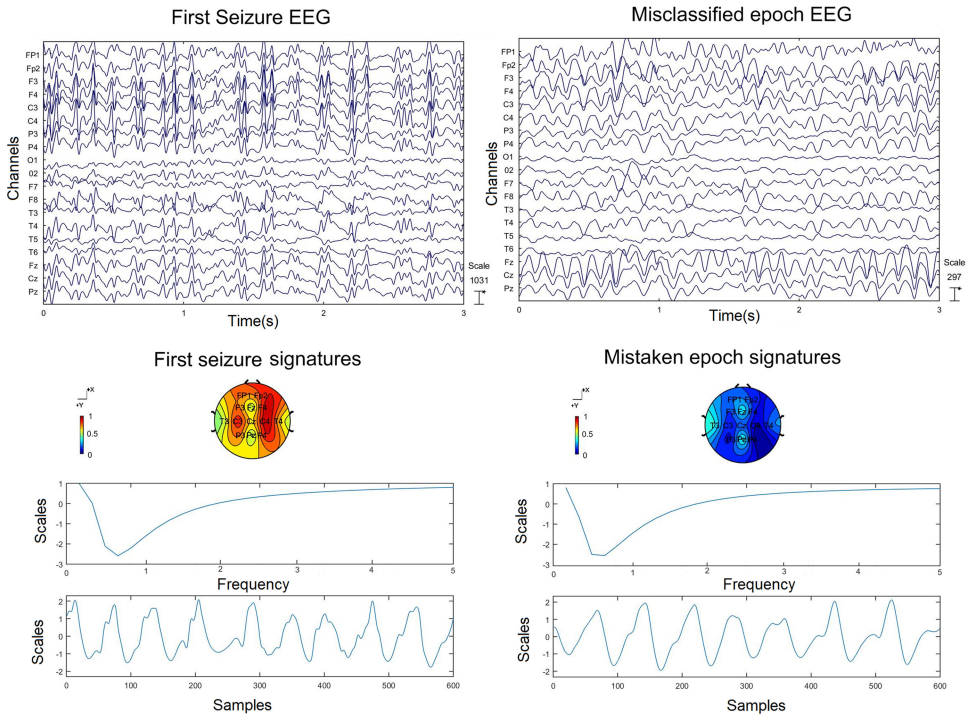


Fig. 7. HH_Tensors CPD: false negative classification example. Seizure epoch classified as a non-seizure. The EEG seizure patterns from the misclassified epoch are dissimilar to the patterns from the first seizure. The space signature from the misclassified epoch shows an inverted activation pattern compared to the one shown by the first seizure. This inverted pattern could be interpreted as a lighter seizure or a non-seizure EEG segment. The frequency component from the misclassified epoch is steeper than the first seizure frequency component. The time component from the misclassified epoch is smoother than the one from the first seizure. However, the time component from both epochs are quite similar. The dissimilarities found between the computed components could explain the classifier failure. Signatures are represented in the following order: Space, Frequency and Time.

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 0.37 to 3 s which means that our method can run in real-time. These results are obtained by classifying all epochs from the test set. 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. The average detection delay of the algorihtm is 5.4 s.

IV. Conclusion

The presented paper introduced a NCES detection method that exploits techniques such as multiway data analysis and HHT. It is the first algorithm to apply HHT to model seizure data with tensors. This is the first algorithm that uses this combination for NCES detection. 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 Space features computed with CPD 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 on continuous clinical data in a realistic training setting proving that the method 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 the seizure EEG pattern evolution caused by the underlying disease progression, adaptive training strategies must be implemented to improve this.

Acknowledgment

This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

References

[1] E. Trinka *et al.*, "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," *J. 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*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2012.

[4] A. T. Tzallas *et al.*, "Automated epileptic seizure detection methods: A review study," in *Epilepsy-Histological, Electroencephalographic and Psychological Aspects*, D. Stevanovic, Ed. Rijeka, Croatia:, InTech, 2012.

[5] L. Orosco, A. G. Correa, and E. Laciar, "Review: A survey of performance and techniques for automatic epilepsy detection," *J. Med. Biol. Eng.*, vol. 33, no. 6, pp. 526–537, 2013.

[6] A. Q. Ansari and P. Sharma, "A review on automated detection of nonconvulsive seizures using EEG," in *Proc. IEEE 2nd Int. Conf. Comput. Intell. Commun. Technol.*, 2016, pp. 283–286.

[7] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *J. Neurosci. Methods*, vol. 123, no. 1, pp. 69–87, 2003.

[8] 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 *Proc. IEEE 29th Annu Int. Conf. Eng. Med. Biol. Soc.,* 2007, pp. 1928–1931.

[9] P. Xanthopoulos *et al.*, "A novel wavelet based algorithm for spike and wave detection in absence epilepsy," in *Proc. IEEE Int. Conf. Bioinformat. Bioeng.,* 2010, pp. 14–19.

[10] S.-F. Liang, W.-L. Chang, and H. Chiueh, "EEG-based absence seizure detection methods," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2010, pp. 1–4.

[11] 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 *Proc. IEEE Int. Conf. Emerg. Trends Commun., Control, Signal Process. Comput. Appl.*, 2013, pp. 1–6.

[12] G. R. Minasyan, J. B. Chatten, and R. N. Harner, "Detection of epileptiform activity in unresponsive patients using ann," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2009, pp. 2117–2124.

[13] 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.

[14] Y. U. Khan, O. Farooq, M. Tripathi, P. Sharma, and P. Alam, "Automatic detection of non-convulsive seizures using ar modeling," in *Proc. IEEE 2nd Int. Conf. Power, Control Embedded Syst.*, 2012, pp. 1–4.

[15] S. Varshney, Y. U. Khan, O. Farooq, P. Sharma, and M. Tripathi, "Latency study of non-convulsive seizures," in *Proc. IEEE Int. Conf. Multimedia, Signal Process. Commun. Technol.*, 2013, pp. 108–111.

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

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

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

[19] B. Hunyadi *et al.*, "Block term decomposition for modelling epileptic seizures," *EURASIP J. Adv. Signal Process.*, no. 139, pp. 1–19, 2014.

[20] W. Deburchgraeve *et al.*, "Neonatal seizure localization using PARAFAC decomposition," *Clinical Neurophysiology*, vol. 120, no. 10, pp. 1787–1796, 2009.

[21] C. Guerrero-Mosquera, M. Verleysen, and A. N. Vazquez, "EEG feature selection using mutual information and support vector machine: A comparative analysis," in *Proc. IEEE Annu Int. Conf. Eng. Med. Biol. Soc.*, 2010, pp. 4946–4949.

[22] M. Leitinger *et al.*, "Salzburg consensus criteria for non-convulsive status epilepticus–approach to clinical application," *Epilepsy Behavior*, vol. 49, pp. 158–163, 2015.

[23] N. E. Huang *et al.*, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, 1988, vol. 454, no. 1971, pp. 903–995.

[24] 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 *Proc. IEEE Biomed. Circuits Syst. Conf.*, 2011, pp. 420–423.

[25] E. A. Atamon, "Understanding epilepsy seizure structure using tensor analysis," Rensselaer Polytechnic Inst., Troy, NY, USA, 2008.

[26] 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," *J. Neural Eng.*, vol. 13, no. 2, 2016, Art no. 026005.

[27] M. Sørensen and L. De Lathauwer, "Tensor decompositions with block-toeplitz structure and applications in signal processing," in *Proc. IEEE 45th Asilomar Conf. Signals, Syst. Comput.*, 2011, pp. 454–458.

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

[29] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1067–1083, 2008.

[30] 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 J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1324–1342, 2000.

[31] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

[32] S. Wold and M. SJÖ STRÖM, "*Simca: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy.*" ACS Publications, 1977. InChemometrics, Theory and Application, ACS Symposium Series, vol. 52, pp. 243–282, 1977.

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

[34] 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.

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

[36] 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 Med. J.*, vol. 30, no. 1, pp. 28–32, 2013.

[37] Y. R. Aldana, B. Hunyadi, E. M. Reyes, V. R. Rodriguez, and S. V. Huffel, "Nonconvulsive epileptic seizures detection using multiway data analysis," in *Proc. 25th Eur. Signal Process. Conf.*, 2017, pp. 1–5.

[38] P. D. Robert, P. Juszczak and P. Paclik, "PRTools4, a matlab toolbox for pattern recognition. version 4.1," Delft University of Technology, Delft, Netherlands, 2007.

[39] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2003–2015, Sep. 2017.

[40] A. Bhattacharyya, R. B. Pachori, A. Upadhyay, and U. R. Acharya, "Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals," *Appl. Sci.*, vol. 7, no. 4, p. 385, 2017.

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