EEG-based Classification of Epileptic and Non-epileptic Events using Multi-array Decomposition

Evangelia Pippa, Vasileios G. Kanas, Evangelia I. Zacharaki, Vasiliki Tsirka, Michael Koutroumanidis, Vasileios Megalooikonomou

EEG-based Classification of Epileptic and Non-epileptic Events using Multi-array Decomposition

[*] **Evangelia Pippa[1], Vasileios G. Kanas[1,2], Evangelia I. Zacharaki[3], Vasiliki Tsirka[4] Michael Koutroumanidis[4] and Vasileios Megalooikonomou[1]**

[1]Multidimensional Data Analysis and Knowledge Management Laboratory

Dept. of Computer Engineering and Informatics, University of Patras

26500 Rion-Patras, Greece

[2]Dept. of Electrical and Computer Engineering, University of Patras, Greece

[3]CVC, Department of Applied Mathematics, Centrale Supélec - Equipe GALEN, INRIA Saclay, France

[4]Dept. of Clinical Neurophysiology and Epilepsies

Guy's & St. Thomas' and Evelina Hospital for Children, NHS Foundation Trust/ King's College, London, United Kingdom

[*]Corresponding Author's email: pippa@ceid.upatras.gr

Abstract: In this paper, the classification of epileptic and non-epileptic events from multi-channel EEG data is investigated based on temporal and spectral analysis and two different schemes for the formulation of the training set. Although matrix representation which treats EEG features as concatenated vectors allows capturing dependencies across EEG channels, it leads to significant increase of feature vector dimensionality and lacks a means of modeling dependencies between features. Thus in this paper, we compare the commonly used matrix representation in which features are concatenated from all channels in order to capture the total spatiotemporal context with a tensor-based scheme which extracts signature features to feed the classification models. TUCKER decomposition is applied to learn the essence of original, high-dimensional domain of feature space and extract a multi-linear discriminative subspace. In contrast to relevant studies found in the literature, in this study, the non-epileptic class consists of two types of paroxysmal episodes of loss of consciousness, namely the psychogenic non-epileptic seizure (PNES) and the vasovagal syncope (VVS). The classification schemes were evaluated on EEG epochs from 11 subjects in an inter-subject cross-validation setting. The proposed tensor scheme achieved an accuracy of 97,7% which is better

compared to the spatiotemporal model even after trying to improve the latter by dimensionality reduction through principal component analysis (PCA) and feature selection by feature ranking.

Keywords: electroencephalography, seizure-like events, tensors, multi-array decomposition, multi-linear data structures

1. Introduction

One of the most challenging medical cases a clinician usually faces in everyday practice is that of patients reporting episodes of transient loss of consciousness (TLoC or blackout), altered awareness, abnormal limb movements or incontinence. The common causes of such episodes are mainly that of epiletpic seizures, posssible psychogenic non-epileptic seizures (PNES) and vasovagal syncopal attacks (VVS) [1, 2]. The similar seizure-like reactions of both epileptic and non-epileptic events make their diagnosis a difficult task. In clinical practice, the diagnosis is based on historical information assisted by specific tests [3]. However, since patients may have limited or no recall of the event and a witness report might not be available clinical information can be either missing or fragmented.

Diagnostic uncertainty may has costs in terms of patients' distress, unnecessary lifestyle changes, social exclusion and financial deprivation associated with hospitalization and repeated investigations [4]. In the worst case scenario, a misdiagnosis of epilepsy can result in mistreatment, with potentially important side effects from the use of antiepileptic drugs and also may have significant medical implications if a serious condition remains undiagnosed or untreated. Furthermore, the financial burden on health services accompanied by an incorrect diagnosis is significant. Taking into account the estimated proportion of the worldwide population with active epilepsy (according to WHO, it is estimated between 4 to 10 per 1000 people) and the unnecessary treatment costs, the estimated annual cost of epilepsy misdiagnosis only in England is around £189 million [5].

The most common diagnostic issue that medical experts routinely deal with, is the differentiation between an epileptic seizure commonly manifested by generalized spike wave discharges (GSW), a psychogenic non-epileptic seizure (PNES) [6] and a vasovagal/ vasodepressor syncope (VVS) [7]. Figures 1 to 3 show examples of the different epileptic and non-epileptic events investigated in our study.

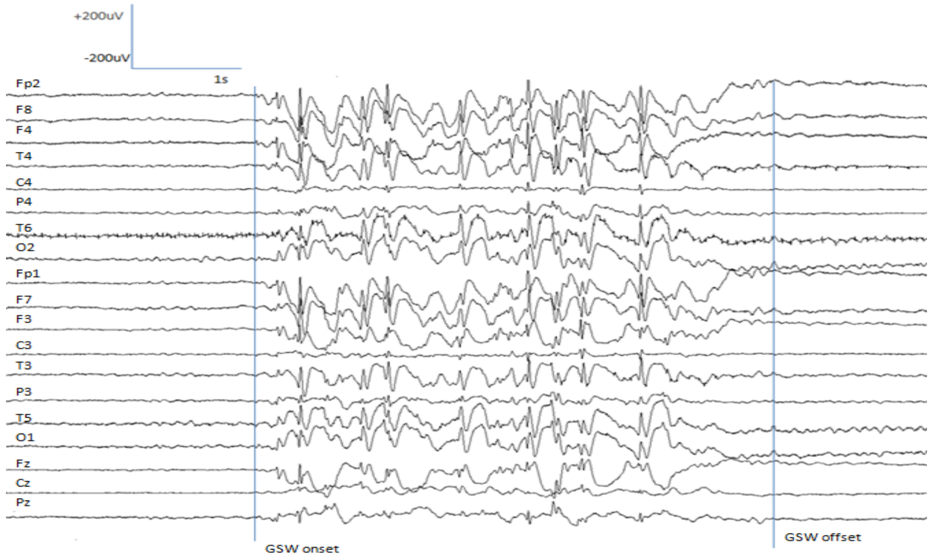

Fig. 1 Generalized Spike Wave (GSW) example. The first marker indicates the beginning of the GSW event and the second marker its end.

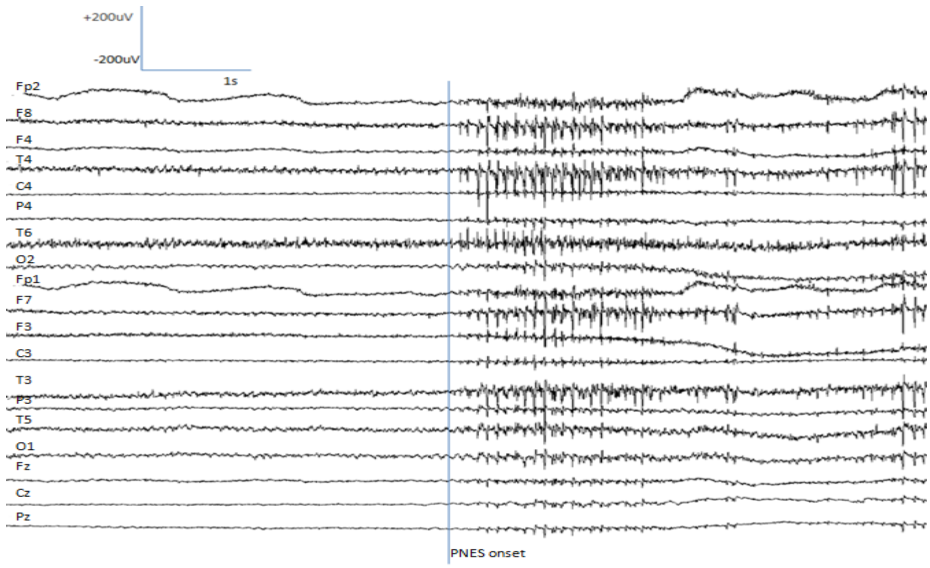

Fig. 2 Psychogeninc Non Epileptic Seizure (PNES) example. The marker indicates the beginning of the PNES event.

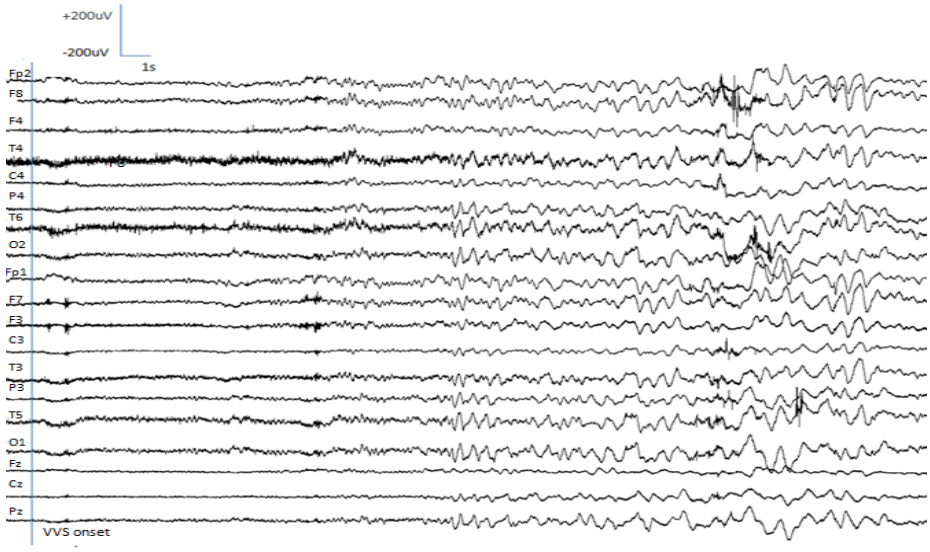

Fig. 3 Vasovagal Syncopal Event (VVS) example. The marker indicates the beginning of the VVS event.

Epileptic seizures are brief episodes of abnormal excessive or synchronous neuronal brain activity [8], characterized by typical ictal neurophysiological patterns and postictal and/or interictal abnormalities. Pshychogenic non-epileptic seizures (PNES) are sudden paroxysmal changes in behavior or consciousness, that resemble epilepsy but are not accompanied by the electrophysiological changes that characterize an epileptic seizure [9]. Vasovagal or vasodepressor syncope is a common type of syncope and various mechanisms have been postulated for explaining the characteristic association of hypotension and bradycardia. The term "vasovagal" indicates that both blood vessels and heart were implicated and since atropine reversed the bradycardia but not the hypotension he considered vasodilatation as the primary responsible factor. As such, PNES and VVS are generally considered to be physical symptoms of an underlying psychological disturbance, triggered by extreme stress-related or emotional events. Clinical characteristics, such as stable ictal heart rate, pelvic thrusting, closed eyes, longer duration of events, events induced by suggestion and rhythmic movement patterns [9-13] have been associated with non-epileptic events rather than epileptic seizures. In most cases, however, the diagnosis of such events still remains doubtful and agreement between physicians as to the nature of a single event may also be limited [14], resulting in misdiagnosis in around 25% of cases [2].

Despite such diagnostic uncertainty, to the best of our knowledge, only a few studies have been proposed in the literature for automated classification between epileptic and non-epileptic pathological events from EEG. Poulos et al. [15] proposed an algorithm which estimates a number of autocorrelated coefficients extracted from an appropriately selected epileptic EEG segment and examines whether these coefficients are correlated with the coefficients of the unknown EEG segments in order to classify the latest into epileptic or non-epileptic. In [16] a LVQ1 neural network was trained on an appropriately extracted set of auto-correlation coefficients (codebook) and the resulting model was used to classify the corresponding feature vectors of the unknown EEG segments. However, both the aforementioned studies do not consider different types of non-epileptic events, which constitutes a fundamental clinical problem. In our previous works [17,18], the non-epileptic class was extended to include both PNES and VVS events. In order to automatically classify epileptic and non-epileptic EEG epochs, several temporal and spectral features were extracted from different channels and combined to a large feature vector as a representative signature for each epoch.

Broadly speaking, raw EEG signals are naturally born with more than two modes (dimensions) of time and space and represented by a multi-way array (tensor). In addition, the process of feature extraction produces structured high-order multi-way arrays that are usually very high dimensional, with large amount of redundancy, while occupying only a subspace of the input space [19]. However, all the previous research works in epileptic and non-epileptic events classification treated EEG features as concatenated vectors (i.e. matrix representation with observations in the rows and features in the columns) in a very high-dimensional space neglecting the inherent structure and correlation in the original feature space [20] [21]. Although matrix representation is suitable for many datasets, it is not always a natural representation because it assumes the existence of a single target variable and lacks a means of modeling dependencies between other features [22]. Motivated by the above, in this study, we compare the commonly used matrix representation in which features are concatenated from all channels in order to capture the total spatiotemporal context with a tensor-based scheme which extracts signature features to feed the classification models. TUCKER decomposition is applied to learn the essence of original, high-dimensional domain of feature space and extract a multi-linear discriminative subspace. The proposed scheme reduced dramatically the computational complexity of the subsequent classification step, which now was performed efficiently in a lower dimensional feature space. The advantage in terms of computational cost relied on the notion that once the mapping (from the original feature space to a reduced space) was learned, its application to unknown EEG segments would only require a few matrix multiplications.

The rest of this paper is organized as follows. Firstly, Section 2 is devoted to the proposed tensor-based scheme, data description, parameterization and pre-processing of EEG signals and classification. In Section 3, the experimental results are presented and a direct comparison with a scheme using linearized feature vectors is performed along with some discussion. The final section is devoted to some concluding remarks.

2. Materials and Methods

2.1 Tensor-based scheme for (non)epileptic EEG events classification

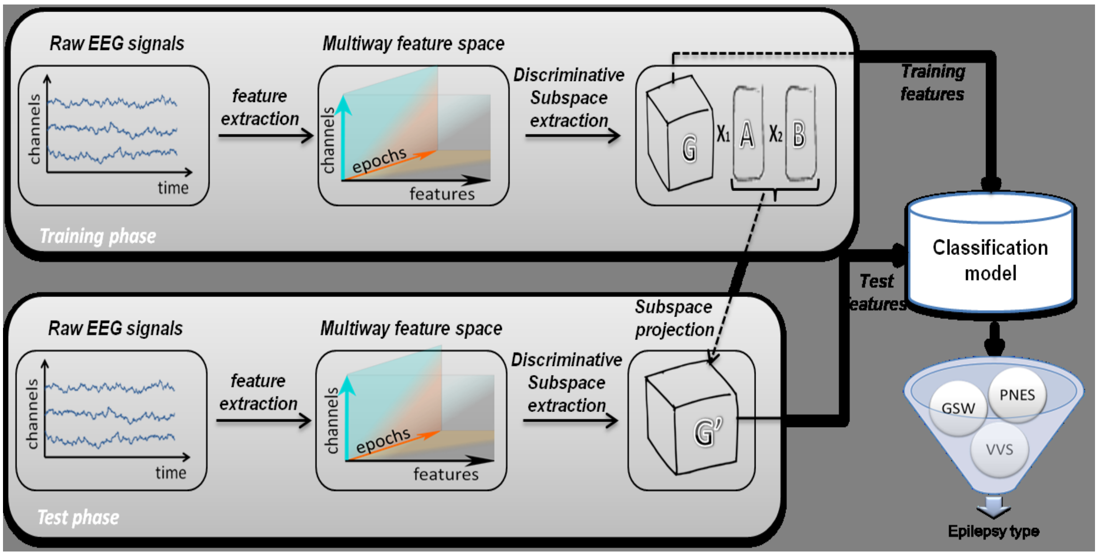

Fig. 4 The tensor-based scheme for EEG-based epileptic type classification

The block diagram of the proposed tensor-based scheme is shown in Figure 4. During the training phase a set of multichannel EEG data denoted as $S = \{s_i\}, 1 \leq t \leq T$, where T denotes the number of samples per channel and $s_i \in R^M$, where M denotes the number of channels, with known time annotations for the events of interest (i.e., PNES, VVS and GSW) was used to train binary classification models. Initially, each EEG signal was frame blocked with a Hamming window to non-overlapping frames of w samples. For each windowed frame, Q temporal and spectral EEG features (see Section 2.3 for more details) were estimated for each of the M channels, resulting in a third-order feature tensor $F_v \epsilon\ R^{M \times Q \times K}$ with K being the total number of windows. In particular, the constructed training tensor is a third-order tensor with modes the EEG channels, the features and the time epochs. Then, based on tensor decomposition, the proposed method extracts simultaneously dominant temporal, spatial and spectral information from the training data, seeking an optimal discriminative feature subspace to project the test data and drive the classification process [23]. Therefore, TUCKER decomposition [24] was applied to reserve multi-linear discriminative subspace from the training feature tensor by decomposing the training tensor to two basis factors and a low-dimensional tensor $G \epsilon R^{R_1 \times R_2 \times K}$.

In this scheme, TUCKER-2 was applied to extract the discriminative multi-linear subspace. Given the tensor $F_v \epsilon\ R^{M \times Q \times K}$, its TUCKER-2 model, expressed as a decomposition of a 3-D tensor into two basis factors and a core tensor, is defined as:

$$F_v = G \times_1 A\ \times_2 B$$

with the symbol $\times_n$ denoting the *n*-mode product of a tensor with a matrix along the mode-*n* (i.e. tensor unfolding in the direction of the *n-th* dimension) [25], the $A \in R^{M \times R_1}$, $B \in R^{Q \times R_2}$being the basis factors (projection filters) and $G \in\ R^{R_1 x R_2 \times K}$ the extracted signature features. The core tensor G consists of signature features of F_v projected onto the factor subspace spammed by A and B. Then, the low-dimensional tensor was matricized and used as signature features to train the classification model.

We used Tucker decomposition instead of canonical polyadic decomposition (CPD) [26] due to its superior flexibility. Tucker model enables all the components of each mode (dimension) to interact with each other through the mean of the core tensor, whereas, in CPD, a component in a certain mode can be linked to only a single component in another mode. Another critical issue, when applying tensor decomposition to perform data analysis, was the determination of the number of components $R_1 R_2$. Here, the values of the both parameters were set to two maximizing the classification of epilepsy type in our recordings.

During the test phase, the calculated basis factors were used as a projected filter to perform feature extraction and finally the test features were used to feed the classification model.

2.2. Data Description and Pre-processing

The previously described classification methodology was evaluated on multi-parametric recordings performed under the ARMOR project [27]. All EEG data were recorded at the Department of Clinical Neurophysiology and Epilepsies in St. Thomas' Hospital in London. The data consisted of 105 generalized seizures (epileptic group) and 21 (19 PNES and 2 VVS) seizure-like events (non-epileptic groups) from 11 different patients. All participants had at least one of their typical epileptic or non epileptic events captured during the recording procedure. The epileptic group, consisted of patients with known diagnosis of idiopathic generalized epilepsy, manifested clinically with absence seizures and they had at least one clinical episode captured during the recording associated with generalized spike wave discharges on the EEG. The non-epileptic group included patients who had sustained a vasovagal syncope (2 patients) or a psychogenic non-epileptic attack (5 patients) during their monitoring. Patients with focal seizures were excluded from this analysis.

The recordings were performed using conventional AgCl EEG electrodes positioned according to the extended international 10–20 system. After the completion of data acquisition, each recorded dataset was visually inspected for noise and motion artifacts and a subset of the main EEG channels was selected for analysis including the following channels: Fp2, F8, F4, T4, C4, A2, P4, T6, O2, Fp1, F7, F3, A1, C3, T3, P3, T5, O1, Fz, Cz, Pz. Notch filtering was applied to attenuate interference at 60 Hz and its harmonics from power lines. Also, baseline correction and re-sampling at 250 Hz was applied in order to obtain a common resolution level for all data coming from different patients and acquisition systems. The recordings were manually annotated by expert Neurologists of the Kings College London. Only epochs during paroxysmal events were considered for training and for testing.

2.3 Parameterization of EEG signals

The parameterization of the brain signals was based on the temporal and spectral information in the EEG channels. Initially, the incoming EEG signals $s_i \in R^M$ were frame blocked to epochs with a sliding Hamming window of length $w = 2\ sec$ and without time-overlap between successive epochs. A large number of hybrid features were investigated including statistical features such as minimum/maximum value, mean, variance, standard deviation, percentiles (25%, 50%-median, 75%), interquartile range, mean absolute deviation, range, skewness, and kurtosis [28-30]. In addition, several studies have supported that the number of 'zero crossings' in the EEG is thought to change during seizure activity [28,31]. Here, the zero crossing rate iis calculated as the sum of all positive zero crossings for each epoch of the zero-meaned EEG. Spectral features including 6-th order autoregressive-filter (AR), power spectral density, frequency with maximum and minimum amplitude, the power of continuous wavelet transform using symlet 5 mother wavelet of scale 25 and 32, the power of discrete wavelet transform with mother wavelet function Daubechies 16 and decomposition level equal to 8 were also used. For each EEG epoch, fifty-five $Q = 55$ features in total were analysed for each one of the $M = 21$ channels. All feature vectors were first normalized. The extracted features were derived from linear and nonlinear signal analysis and all have been employed in EEG applications in the past.

2.4 Classification

Aiming to evaluate the ability of the extracted signature features to discriminate between epileptic and non-epileptic events widely used classifiers were used, namely the random forest (RF) [32], the k-nearest neighbors (KNN) algorithm and its weighted (wKNN) version [33], and support vector machines (SVM) using the sequential minimal optimization algorithm [34], linear discriminant analysis (LDA) [35] and BayesNet [36] were investigated. For the KNN and wKNN classifiers, the Euclidean distance was selected as the distance metric. After testing the parameter space, $k = 9$ was chosen empirically. Moreover, the Gaussian radial basis function (RBF) for the SVM kernel was used. Polynomial-based kernels were also considered, but their performance was considerably lower than the RBF kernel. The values of the soft margin parameter $C = 20$ and the scaling factor $\gamma = 0.1$ were found to offer optimal classification performance after a grid search at all combinations of $C = \{1.0, 5.0, 10.0, 20.0, 30.0\}$ and $\gamma = \{0.001, 0.01, 0.5, 1.0, 2.0\}$.

3. Experimental Results and Discussion

The tensor-based classification scheme presented in Section 2.1 was applied to the EEG dataset described in 2.2 in order to be compared with our previous classification scheme [17] that uses matrix representation with linearized feature vector of dimensionality $M \times Q = 21 \times 55$ features for each frame of the training and tests sets. In this matrix representation features are concatenated from all channels in order to capture the total spatiotemporal context. For both schemes, evaluation was performed in a leave-one-out cross-validation setting. Specifically, each time one subject was left-out for testing, while the rest of the subjects were used for training. For the left-out subject, all epochs between seizure onset and offset were used as testing samples. **Error! Reference source not found.** shows the number of epochs that were extracted from each subject during the seizure(s).

Table 1 Number of seizures and number of seizure epochs per subject

Subject	*Class*	*# Epochs*	*# Seizures*
1	GSW	59	52
2	GSW	29	19
3	GSW	16	14
4	GSW	19	20
5	PNES	1	1
6	PNES	1	1
7	PNES	1	1
8	PNES	13	13
9	PNES	3	3
10	VVS	45	1
11	VVS	18	1

Throughout the paper, the classification accuracy defined as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

where true positives are denoted as TP, true negatives as TN, false positives as FP and false negatives as FN, served as the primary performance metric. Here we consider the epileptic class as the positive and the non-epileptic class (PNES or VVS) as the negative.

The accuracy of the proposed scheme for classifying between epileptic (GSW) and non-epileptic (VVS, PNES) events, for the investigated classifiers, is shown in Fig. 5. The overall highest accuracy was 97.7% achieved by the KNN classifier, with the second highest being 96.1% using SVM as a classification model. On the other hand, the overall highest accuracy achieved by the scheme with the matrix representation of the linearized feature vectors was 86% for the BayesNet classifier. As can be seen, the performance of the tensor-based scheme was considerably higher compared to the scheme where the original features (without tensor decomposition) were utilized to drive the classifiers. In particular, the system performance was increased by approximately 13%. It seems that the high dimensionality of the training samples in the scheme without tensor decomposition is not appropriate for datasets with limited number of instances such as our dataset.

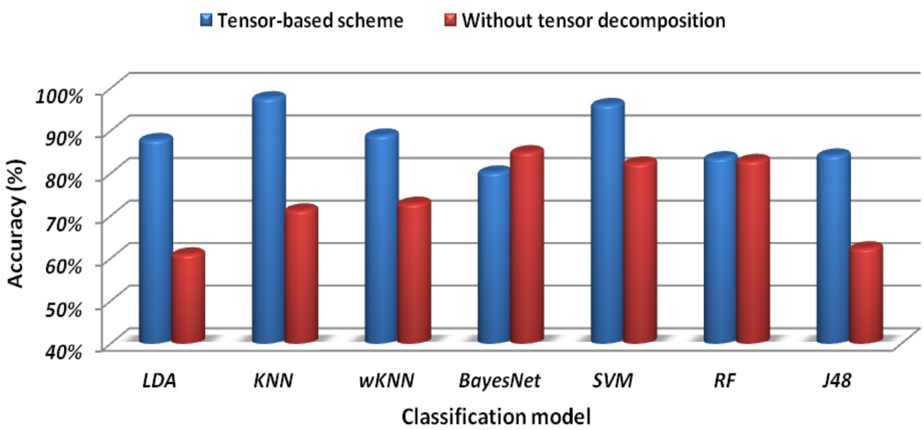

Fig. 5 Classification accuracy for the different classifiers with and without tensor decomposition.

In a further step, to make a fair comparison of the two schemes, we tried to optimize the scheme without tensor decomposition by performing dimensionality reduction following two different strategies: principal component analysis (PCA) and feature selection by feature ranking.

PCA is a transformation that finds the optimal linear combinations of the features, in the sense that they represent the data with the highest variance in a feature subspace, without taking the intra-class and inter-class variances into consideration separately. The reduced dimension of the feature vectors is determined by observing the eigenvalues of the covariance matrix of the feature vectors sorted in descending order. The largest eigenvalues that constitute a high percentage of the total variance (e.g. 99%) of the principal components and account for much of the variability of the data are selected. The eigenvectors corresponding to the selected eigenvalues are used to form the transformation matrix, resulting in feature vectors with reduced dimensionality. PCA was performed on the feature matrix and the performance in terms of accuracy for different number of retained eigenvectors so as different amounts of variation are kept, was evaluated.

The proportion of retained variance for different number of retained eigenvectors was computed after sorting the eigenvectors in decreasing order of eigenvalues by

$$Variance = \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{j=1}^{m} \lambda_j}$$

where λ_i is the eigenvalue for the i-th principal component, r the number of retained eigenvectors, and and m the number of total number of components. Figure 6 shows the retained variance as a function of the number of retained eigenvectors.

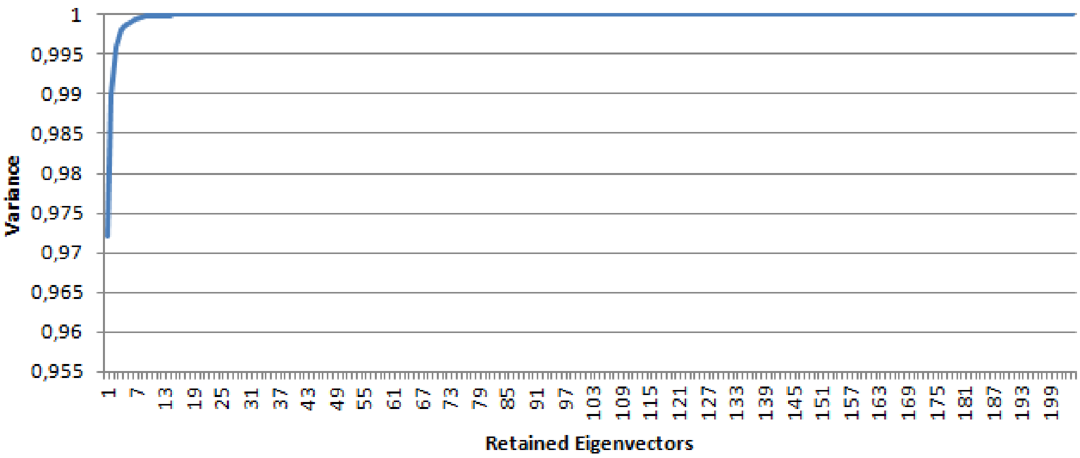

Fig. 6 Classification Retained Variance for different number of PCA retained eigenvectors.

As can be seen, 198 eigenvectors are required in order to achieve 100% variance while in order to achieve 99,9% variance, only 5 retained eigenvectors are required. The classification performance for different number of retained eigenvectors with respect to the BayesNet classifier is shown in Table 2. For the evaluation of the scheme without tensor decomposition we selected the BayesNet classifier, since it was the one that reached the highest accuracy. The maximum accuracy, which is 85,85%, is achieved when 7 components are retained (99,9% retained variance). Although PCA does not improve the accuracy of the scheme without tensor decomposition, it provides an accuracy which is almost equal to the initial one obtained with a feature vector of significantly lower dimensionality.

Retained Components	**Retained Variance**	***Accuracy***
1	97,2%	60,98%
2	99,0%	78,54%
3	99,6%	75,61%
4	99,8%	70.24%
5	99,9%	78,05%
6	99,9%	83.90%
7	99,9%	**85,85%**
198	100%	82,44%

Table 2 Classification performance for different amounts of retained variance with respect to the scheme without tensor decomposition

Note that these results are produced by applying PCA without any standardization of the data before performing the analysis. In such a case, since the different features are not measured on the same scale and PCA is performed on the non-standardized features, each principal component is dominated by a single or a few features, the one(s) with the highest variance resulting somehow to an ordering of the features by their variance. In this case very few components explain all the variance in the data. On the other hand, when z-score is used before PCA, the other components contribute as well to the explanation of the data variance, since standardizing implies assigning equal importance to all variables. As a result, when standardizing the data many more principal components are required to achieve the same variance (in order to achieve 99% variance, 115 retained eigenvectors are required). However, when standardizing the data, it seems that the additional components introduce noise resulting to significantly reduced classification accuracy.

As an alternative strategy for dimensionality reduction, we examined the discriminative power of the extracted features for the classification of epileptic and non-epileptic EEG events by feature ranking. The t-test was used for estimating the importance of each feature in binary classification. In this study, ranking is performed by following a leave-one-out strategy on the available subjects. Specifically, for each leave-one-out experiment, feature ranking is performed using the t-test in each training subset. This means that for each leave-one-out experiment the retained features may be different. The performance of the method, in terms of accuracy for different number of N-best features (N=10, 20, 30, ..., 1150) using the BayesNet classifier that had shown the best performance for the scheme without tensor decomposition are shown in Figure 7.

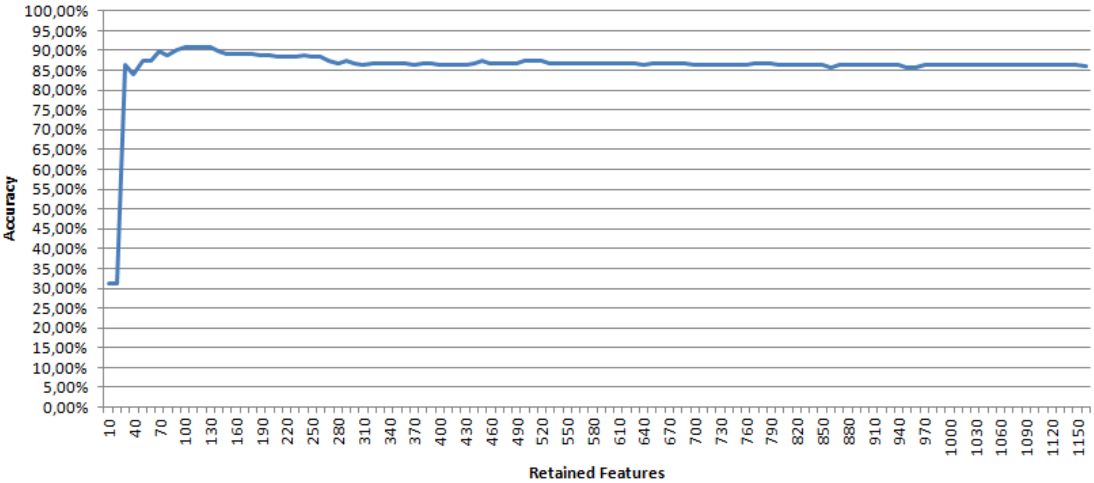

Fig. 7 Classification Accuracy for the scheme without tensor decomposition for different subsets of N-best features (N=10, 20, .., 1150).

As can be seen in the above figure the highest classification accuracy is achieved when a small subset of discriminative features is used. Specifically, the scheme without tensor decomposition achieves its highest accuracy (90,73%) for a subset of 100 best features. Such results indicate the superiority of the tensor-based scheme even after the optimization of the scheme without tensor decomposition through feature selection.

Finally, in order to examine the ability of the tensor-based scheme using the best performed classifier (KNN) to discriminate each type of pathological events from the others, we performed pair-wise classification of all possible pairs of the events (GSW-PNES, GSW-VVS, and PNES-VVS). The achieved accuracies are presented in Table 3.

Table 3 System performance (with and without tensor decomposition) in terms of accuracy for all the pairwise classification problems

Pairwise classification	Tensor-based scheme	Without tensor decomposition
GSW-PNES	0,991	0,901
GSW-VVS	0,983	0,903
PNES-VVS	0,886	0,760

As can be seen, the discrimination of PNES and VVS pathological events was the most difficult problem (87% accuracy). The above might be attributed to the nature of the different types of the pathological events. In general, generalized spike waves were very specific ictal neurophysiological patterns, presenting much more consistent features (compared to the other types) and consequently making their detection an easier task. On the other hand, PNES has no specific EEG patterns but was frequently accompanied by muscular artifacts presenting great variability across subjects. Similar variability appears even between consecutive epochs of VVS examples, since there were several changes that happen successively in time during such an episode (beta / alpha → theta → delta → lower voltage rhythms → isoelectric suppression). It seems that the variability in the feature values of the PNES and VVS epochs was high (with respect to the available training data) impeding the learning of a discrimination model.

Although direct comparison with other studies was not possible due to the different characteristics of each dataset (e.g. different seizure types, lack of PNES or VVS examples in most studies or use of single channel data), the achieved classification accuracy was higher than the one reported in the literature [15, 16]. Regarding our previous work on the same dataset [17,18], again the proposed tensor-based scheme achieves higher accuracy.

4. Conclusions

In this paper, a tensor-based scheme was proposed to discriminate different pathological events, The scheme incorporates spatial-spectral-temporal features extracted from EEG brain activity, and TUCKER decomposition to extract a multi-linear discriminative subspace. The proposed scheme was compared against the commonly used matrix representation in which features are concatenated from all channels in order to capture the total spatiotemporal context. Experimental results demonstrated that using multi-linear models, we were able to extract signature features of EEG recordings for discriminating various type of epileptic events with higher accuracy. Even after the optimization of the scheme without tensor decomposition through feature selection, the proposed tensor-based scheme still presented higher classification accuracy.

Acknowledgement

The authors would like to thank Dr. Iosif Mporas and Giorgos Drakopoulos for their valuable comments and discussion about this work. This study was partially funded by the EC under the FP7/2007-2013 with grant ARMOR, Agreement Number 287720 and European Research Council Grant Diocles (ERC-STG-259112)..

6. References

[1] W. Hoefnagels, G. Padberg, J. Overweg, E. Van der Velde, and R. Roos, "Transient loss of consciousness: the value of the history for distinguishing seizure from syncope," Journal of neurology, vol. 238, pp. 39-43, 1991.

[2] A. Moya, R. Sutton, F. Ammirati, J.-J. Blanc, M. Brignole, J. B. Dahm, et al., "Guidelines for the diagnosis and management of syncope (version 2009) The Task Force for the Diagnosis and Management of Syncope of the European Society of Cardiology (ESC)," European heart journal, vol. 30, pp. 2631-2671, 2009.

[3] F. Brigo, R. Nardone, and L. G. Bongiovanni, "Value of tongue biting in the differential diagnosis between epileptic seizures and syncope," Seizure, vol. 21, pp. 568-572, 2012.

[4] D. Smith, B. Defalla, and D. Chadwick, "The misdiagnosis of epilepsy and the management of refractory epilepsy in a specialist clinic," Qjm, vol. 92, pp. 15-23, 1999.

[5] M. Chapman, P. Iddon, K. Atkinson, C. Brodie, D. Mitchell, G. Parvin, et al., "The misdiagnosis of epilepsy in people with intellectual disabilities: A systematic review," Seizure, vol. 20, pp. 101-106, 2011.

[6] A. Krumholz, "Nonepileptic seizures: diagnosis and management," Neurology, vol. 53, pp. S76-83, 1998.

[7] T. Lewis, "A lecture on vasovagal syncope and the carotid sinus mechanism," British medical journal, vol. 1, p. 873, 1932.

[8] R. S. Fisher, W. v. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, et al., "Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," Epilepsia, vol. 46, pp. 470-472, 2005.

[9] S. G. Jones, T. J. O'Brien, S. J. Adams, R. Mocellin, C. J. Kilpatrick, R. Yerra, et al., "Clinical characteristics and outcome in patients with psychogenic nonepileptic seizures," Psychosomatic medicine, vol. 72, pp. 487-497, 2010.

[10] C. Opherk and L. J. Hirsch, "Ictal heart rate differentiates epileptic from non-epileptic seizures," Neurology, vol. 58, pp. 636-638, 2002.

[11] J. Bayly, J. Carino, S. Petrovski, M. Smit, D. A. Fernando, A. Vinton, et al., "Time- frequency mapping of the rhythmic limb movements distinguishes convulsive epileptic from psychogenic nonepileptic seizures," Epilepsia, vol. 54, pp. 1402-1408, 2013.

[12] S. Benbadis, K. Johnson, K. Anthony, G. Caines, G. Hess, C. Jackson, et al., "Induction of psychogenic nonepileptic seizures without placebo," Neurology, vol. 55, pp. 1904-1905, 2000.

[13] P. Alboni, F. Dipaola, P. Coppola, and R. Furlan, "Clinical Presentation and Diagnosis of Vasovagal Syncope," in Vasovagal Syncope, ed: Springer, 2015, pp. 123-135.

[14] W. Hoefnagels, G. Padberg, J. Overweg, and R. Roos, "Syncope or seizure? A matter of opinion," Clinical neurology and neurosurgery, vol. 94, pp. 153-156, 1992.

[15] M. Poulos, F. Georgiacodis, V. Chrissikopoulos, and A. Evangelou, "Diagnostic test for the discrimination between interictal epileptic and non-epileptic pathological EEG events using auto-cross-correlation methods," American journal of electroneurodiagnostic technology, vol. 43, pp. 228-240, 2003.

[16] S. Papavlasopoulos, M. Poulos, and A. Evangelou, "Feature Extraction from Interictal Epileptic and Non-Epileptic Pathological EEG Events for diagnostic Purposes using LVQ1 Neural Network," in Proceedings of seventh International Conference on Mathematics Methods in Scattering Theory and Biomedical Technology, 2005.

[17] E. Pippa, E. Zacharaki, I. Mporas, V. Megalooikonomou, V. Tsirka, M. Richardson, et al., "Classification of epileptic and non-epileptic EEG events," in Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on, 2014, pp. 87-90.

[18] E. Pippa, E. Zacharaki, I. Mporas, V. Megalooikonomou, V. Tsirka, M. Richardson, et al., "Improving Classification of Epileptic and Non-Epileptic EEG Events using Feature Selection," Neurocomputing, 2015.

[19] G. Shakhnarovich and B. Moghaddam, "Face recognition in subspaces," in Handbook of Face Recognition, ed: Springer, 2011, pp. 19-49.

[20] D. Tao, X. Li, W. Hu, S. Maybank, and X. Wu, "Supervised tensor learning," in Data Mining, Fifth IEEE International Conference on, 2005, p. 8 pp.

[21] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H.-J. Zhang, "Multilinear discriminant analysis for face recognition," Image Processing, IEEE Transactions on, vol. 16, pp. 212-220, 2007.

[22] M. Barnathan, V. Megalooikonomou, C. Faloutsos, S. Faro, and F. B. Mohamed, "TWave: high-order analysis of functional MRI," Neuroimage, vol. 58, pp. 537-548, 2011.

[23] A. H. Phan and A. Cichocki, "Tensor decompositions for feature extraction and classification of high dimensional datasets," Nonlinear theory and its applications, IEICE, vol. 1, pp. 37-68, 2010.

[24] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," Psychometrika, vol. 31, pp. 279-311, 1966.

[25] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation: John Wiley & Sons, 2009.

[26] R. Bro, "PARAFAC. Tutorial and applications," Chemometrics and intelligent laboratory systems, vol. 38, pp. 149-171, 1997.

[27] A. p. W.-s. h. www.armor-project.eu/.

[28] M. J. van Putten, T. Kind, F. Visser, and V. Lagerburg, "Detecting temporal lobe seizures from scalp EEG recordings: a comparison of various features," Clinical neurophysiology, vol. 116, pp. 2480-2489, 2005.

[29] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings," Information Technology in Biomedicine, IEEE Transactions on, vol. 14, pp. 186-197, 2010.

[30] M. Valderrama, S. Nikolopoulos, C. Adam, V. Navarro, and M. Le Van Quyen, "Patient-specific seizure prediction using a multi-feature and multi-modal EEG-ECG classification," in XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010, 2010, pp. 77-80.

[31] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, "Anticipating epileptic seizures in real time by a non- linear analysis of similarity between EEG recordings," Neuroreport, vol. 10, pp. 2149-2155, 1999.

[32] L. Breiman, "Random forests," Machine learning, vol. 45, pp. 5-32, 2001.

[33] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," Machine learning, vol. 6, pp. 37-66, 1991.

[34] J. Platt, "Fast training of support vector machines using sequential minimal optimization," Advances in kernel methods—support vector learning, vol. 3, 1999.

[35] B. Scholkopft and K.-R. Mullert, "Fisher discriminant analysis with kernels," in Proceedings of the 1999 IEEE Signal Processing Society Workshop Neural Networks for Signal Processing IX, Madison, WI, USA, 1999, pp. 23-25.

[36] G. F. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data," Machine learning, vol. 9, pp. 309-347, 1992.