Seizure detection Based on Autoregressive Modeling

Mourad Adnane[1], Adel Belouchrani[1]

[1] LDCCP Lab, Electronics Department, Ecole Nationale Polytechnique, Algiers
mourad.adnane@enp.edu.dz, adel.belouchrani@enp.edu.dz

Abstract

This paper considers the use of autotoregressive (AR) modeling of Electroencephalogram (EEG) signals to discriminate between normal and epileptic EEG signals on one hand and to descriminate between seizure and seizure-free EEG signals on the other hand. Each epoch of EEG signal is modeled by an AR model of order P. Then, the obtained P AR coefficients are used in training and testing of a support vector machine (SVM) classifier. The optimal AR model order is investigated. The method is tested against a widely used EEG database and results show a classification accuracy of 100% when considering normal and epileptic EEG signals and a classification accuracy of 96.54% when considering seizure and seizure-free EEG signals. The obtained results are along with those obtained by state of the art EEG signal classifiers.

1. Introduction

The conventional methods to detect seizures involve visual inspection of electroencephalography (EEG) recrodings. The length of such recordings in the case of long term EEG (can be 24 hours long) and the number of channels (can reach 128 channel) makes this task fastidious. In addition, the development of portable EEG systems which can be used in homecare systems makes the use of automatic detection of seizures of crucial importance. Several works addressing this problem were presented and different techniques were used to realize the descrimination between differents EEG signal epochs. The main difference between these works is the chosen method for feature extraction. In [1] empirical mode decomposition is used to extract features. Entropy estimation is used in [2-4] such as discrete wavelet transform aproximate Entropy in [2] and permutation entropy in [3] and aproximate entropy used in an extreme learning machine in [4]. Local binary patterns are used in [5,6]. EEG being non stationary, time-frequency representation were used in [7,8]. Horizontal visibility graphs are used in [9] and fractional linear prediction in [10]. Several other methods are used for the extraction of descriminative features for EEG signal classification purpose and it is not possible to cite them all. Therefore a comprehensive review of EEG feature extraction methods is given in [11]. It should be noted that the problem of descrimination between normal (healthy) and epileptic EEG signals is almost addressed in term of classification accuracy. Actually, several methods reached 100% classification accuracy and it does not go under 99.5%. The reamaining issues for this classification scheme concern algorthmic complexity (and thus execution time issue) and the implementation issues. However, descriminating between different epileptic EEG signals: seizure-free and seizure EEG signals is still an ongoing issue and there is no solution that have fully addressed this problematic.

In this paper we present a method based on the use of autoregressive (AR) modeling of EEG epochs for EEG signal classification. We apply the method for both classification problems: normal against epileptic EEG signals and seizure-free against seizure EEG signals. An investigation about optimal model order is conducted. The obtained AR coefficients are used as entry features for a support vector machine (SVM) training and testing classifier. The remainder of this paper is as follows. In section 2, the materials and method are presented. The results and a discussion are presented in section 3 and finally we finish with a conclusion in section 4.

2. Materials and Methods

2.1. EEG dataset

There is a widely used and publicly available database developed within the department of epileptology at the university of Bonn [12]. This database contains 5 sets of EEG recordings noted A-E. Recordings belong to two categories: EEG signals of 5 healthy controls (A, B) and EEG signals of 5 epileptic patients (C, D, E) such as C and D include only seizure-free intervals and E includes only seizure activity. Each set consists of 100 epoch of 23.6 s length. The sampling frequency is 173.61 Hz with 12 bits depth. In this paper we use three sets such (A, D and E). Two binary classification schemes are used: 1) Descriminate between normal (A) and epileptic (E) EEG signals, 2) Descriminate between seizure-free (D) and seizure (E) EEG signals. The use of these two classification schemes is done in order to make comparisons with other reported researches that used these schemes [1-10]. Figure 1 depicts a sample from each of the three sets we are using in this paper (A, D, E).

2.2. Autoregressive Modeling

Each epoch of EEG signal of length n is modeled by an AR model of order P as the output of recursive linear system. The input is modeled by a white noise e_n:

$$x_n = e_n + \sum_{i=1}^{P} -a_i x_{n-i} \tag{1}$$

where a_i represents the AR model coefficients and x_n represents the EEG signal epoch of length n. In the case of the used database the length of each epoch is of 23.6 s which corresponds to 2097 samples. Therefore $n = 2097$. The autoregressive modeling as presented in equation 1 represents a linear prediction of x_n using the weighted sum of previous samples and thus e_n represents the modeling error. The estimation of the AR coefficients is made by minimizing the prediction error e_n (For more details refer to [13]). By minimizing the mean squared prediction error and after some calculations we obtain the Yule-

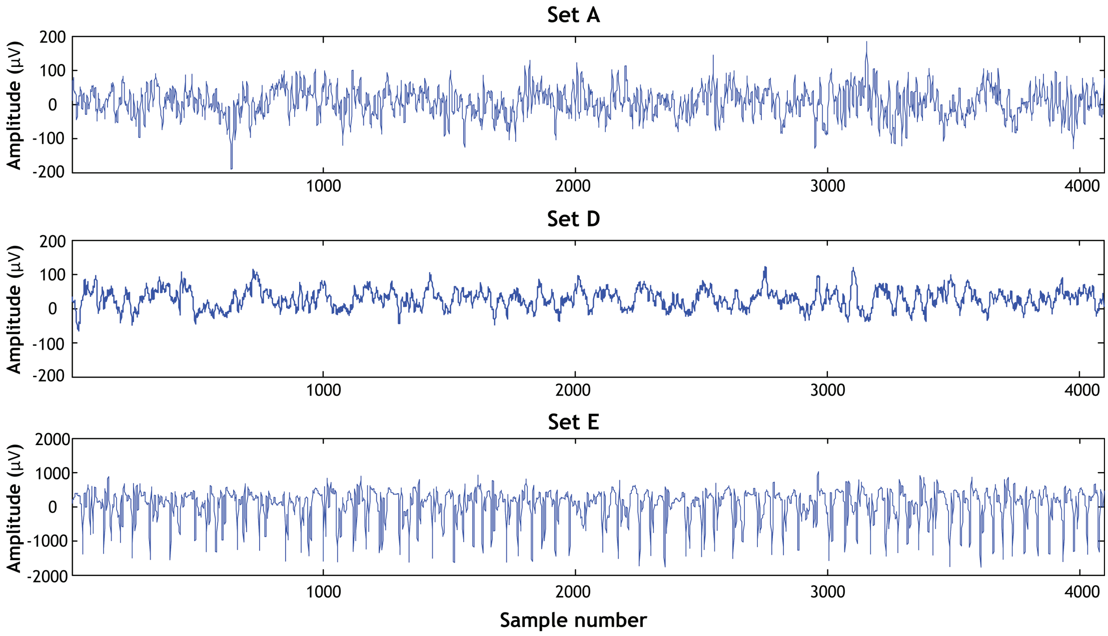

Figure 1. Sample recording of EEG signal taken from the sets A (normal), D (epileptic/seizure-free) and E (epileptic/seizure period)

Walker equations stated as:

$$\sum_{i=1}^{P} r_{xx}(i-j)a_i = -r_{xx}(j) \; for \; j = 1, ..., P \qquad (2)$$

or in a matrix form as:

$$\begin{bmatrix} r_{xx}(0) & \cdots & r_{xx}(P-1) \\ \vdots & \ddots & \vdots \\ r_{xx}(P-1) & \cdots & r_{xx}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} -r_{xx}(1) \\ \vdots \\ -r_{xx}(P) \end{bmatrix} \qquad (3)$$

Equation 3 represents a Toeplitz matrix and it can be solved using the levinson recursion algorithm (see [13]). The levinson algorithm allows the reduction of the calculation complexity from an order of P^3 to an order of P^2.

2.3. Determining AR model optimal order

The AR modeling of each EEG epoch results with P AR coefficients. These coefficients are used as entry features for the training and testing classifier. The classifier used in this work is the support vector machine (SVM). The kernel used in this paper is the quadratic kernel with sequential minimal optimization method [14]. The theoretical background on the SVM method can be found in [15]. Herein, we use two different classification schemes (two distinct classifiers) such as: Normal vs. Epileptic (A and E) and Seizure-free vs. Seizure (D and E). For each scheme the AR model order that maximizes an objective function (in this paper the classification accuracy) is determined. The methodology followed for AR model order determination is as follows:

1. Vary the AR order P from 2 to 20.
2. Estimate AR coefficients for the 200 epochs.
3. Divide the data into 10 folds (the number of epochs from each class is the same) and use 9 folds for training and one fold for testing and calculate the classification accuracy.
4. Repeat the operation 10 times by changing every time the training folds and the testing fold. Then, average the classification accuracy over the 10 experiments.
5. Determine P that maximizes the classification accuracy.

This methodology is applied to the classification scheme 1 (A vs. E) and the obtained classification accuracies for the AR order varying from 2 to 20 is depicted in figure 2. Similarly, the same methodology is applied to the classification scheme 2 (D vs. E) and the obtained classification accuracies for the AR order varying from 2 to 20 is depicted in figure 3.

3. Results and Discussion

Three metrics are used to evaluate the performance of our algorithm the classification accuracy (Acc), the sensitivity (Se) and the specificity (Sp):

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \times 100 \qquad (4)$$

$$Se = \frac{TP}{TP + FN} \times 100 \qquad (5)$$

$$Sp = \frac{TN}{TN + FP} \times 100 \qquad (6)$$

where TP stands for true positives which means the number of seizure episodes correctly classified, FN stands for false negatives which means the number of seizure episodes wrongly classified (classified N), TN stands for true negatives which means the number of normal EEG episodes (or seizure-free episodes) correctly classified, and FP stands for false positives which means the number of normal EEG episodes (or seizure-free episodes) wrongly classified (classified as seizure episode). Classification accuracy is an overall metric of the algorithm. The sensitivity estimates the ability of the algorithm to detect seizures. The specificity estimates the ability of the algorithm to detect normal EEG episodes (or seizure-free episodes).
The method is applied in the scheme 1 (normal EEG vs. seizure

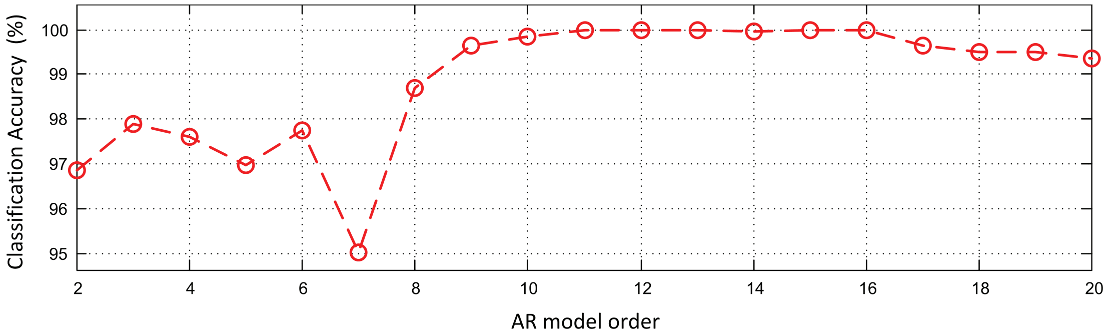

Figure 2. Classification accuracy obtained for the classification scheme 1 (A vs. E) for AR order varying from 2 to 20

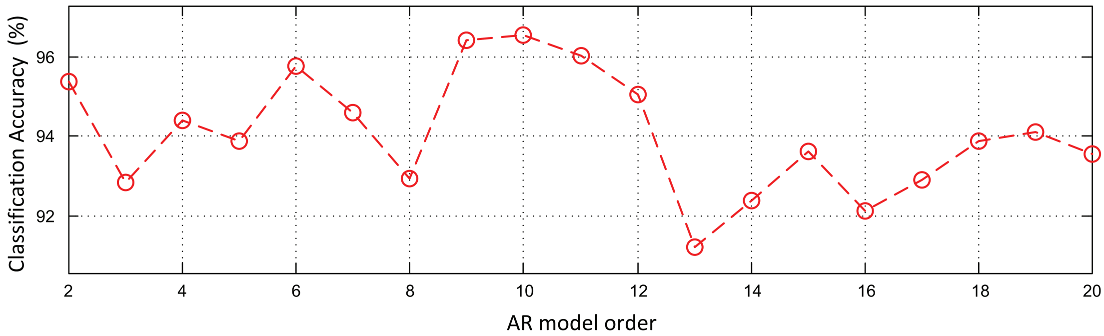

Figure 3. Classification accuracy obtained for the classification scheme 2 (D vs. E) for AR order varying from 2 to 20

EEG) and the classification accuracy as shown by the Fig.2 reached 100% for AR model order 11, 12 and 15. It should be noticed that these results are average results obtained in the case of 10-fold cross validation technique which shows the power of AR modeling in descriminating between normal and epileptic EEG episodes. On the other hand, the method is applied to the scheme 2 (seizure-free EEG vs. seizure EEG) and the results as shown by the Fig.3 reached its best classification accuracy for the AR model order 10 (Acc=96.54%, Se=96.99%, Sp=96.1%). This is a good result in comparison to the results obtained in literature. The table 1 shows a comparison between results obtained by our method applied to the database of the university of Bonn [12] and results of 10 researches using the same database. As we can see in Table 1, the obtained results for the classification scheme 1 (normal vs. epileptic) reached 100% for the proposed method which is the same accuracy obtained by researchers in [1,2,8,9] and better than two researcher results that did not reach 100% classification accuracy [5,7]. As we mentioned earlier the results obtained by our method is the average over 10 different folds which means different training sets and testing sets for each classification which means the performance is a strong one since it obtained 100% in the different 10 cases. In addition, these classification accuracy values are obtained for several AR model orders (11, 12, 15) and are nearly 100 % for neighboring orders (99.98% for order 13 and 99.97% for order 14 and 99.99% for order 16). In other words the use of AR modeling of EEG signal with order going from 11 to 16 presents great ability of generalization for the descrimination between normal and epileptic EEG.

On the other hand, results shown in Table 1 for the classification scheme 2 (seizure-free vs. seizure) reached 96.54 % for the proposed method which is better than results of 7 researchers presented in table 1 [2-5,7, 9-10] and slightly inferior to three researcher results presented in table 1 [1,6,8]. However, the methods presented in these 3 papers are costly in term of calculation (Artificial neural network [1], Local binary pattern [6], Time-frequency representation [8]). This is not the case of our method where the estimation of AR coefficients are made recursively which make our method more suitable for implementation. Overall, these results show that the issue of descriminating between seizure-free periods and seizure periods automatically is not yet addressed. Autoregressive modeling did not get that power of descrimination such as in the case of classification scheme 1. The AR modeling can be combined to other methods to increase the classification accuracy. However, the computational complexity will increase. Future investigation should be turned to other modeling methods applied to EEG and the comparison between their respective descrimination ability and their respective performances.

4. Conclusion

In this paper we present a method for the classification of electroencephalogram (EEG) according to two different schemes where the first one concerns normal against epileptic EEG signals and the second one concerns seizure-free against seizure EEG signals. The method is based on AR modeling of the EEG epochs. The optimal AR model order is determined for each classification scheme. Results show that the classifier based on AR modeling for the case of classification scheme 1 is very strong and can be generalized for other EEG datasets. The classifier obtained good results in the second classification scheme. However, it is still far from solving the issue of detecting seizures in an epileptic EEG recording. Future investigation should focus on the comparison with other modeling methods in term of performance and descrimination ability.

Table 1. Comparison of classification performance with different researchers' methods for EEG classification schemes

Researchers	Year	Methods	Classes	Acc (%)
Djemili et al. [1]	2016	Empirical mode decomposition and artificial neural network	A-E D-E	100 97.7
Kumar et al. [2]	2014	DWT-based Aproximate entropy and artificial neural network	A-E D-E	100 95
Kaya et al. [5]	2014	1D-local binary pattern	A-E D-E	99.5 95.5
Samiee et al. [7]	2015	Rational Discrete Short-Time Fourier Transform	A-E D-E	99.8 95
Gao et al. [8]	2017	Visibility Graph from Adaptive Optimal Kernel Time-Frequency Representation	A-E D-E	100 98
Zhu et al. [9]	2014	Fast weighted horizontal visibility algorithm	A-E D-E	100 93
Nicolaou et al. [3]	2012	Permutation Entropy and Support Vector Machines	D-E	82.88
Yuan et al. [4]	2011	Extreme learning machine and nonlinear features	D-E	96.5
Sunil Kumar et al. [6]	2015	Local binary patterns	D-E	98.33
Joshi et al. [10]	2014	Fractional linear prediction	D-E	95.33
Proposed method	2017	Autoregressive modeling and Support vector machine	A-E D-E	100 96.54

5. References

[1] R. Djemili, H. Bourouba, M. C. Amara Korba, "Application of empirical mode decomposition and artificial neural network for the classification of normal and epileptic EEG signals", *Biocybern Biomed Eng*, vol. 36, no.1, pp: 285-291, 2016.

[2] Y. Kumar, M. L. Dewal, R. S. Anand, "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network", *Signal Image Video P*, vol. 8, no.7, pp: 1323-1334, 2014.

[3] N. Nicolaou, J. Georgiou, "Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines", *Expert Syst Appl*, vol. 39, no.1, pp: 202-209, 2012.

[4] Q. Yuan, W. Zhou, S. Li, D. Cai, "Epileptic EEG classification based on extreme learning machine and nonlinear features", *Epilepsy Res*, vol. 96, no.1-2, pp: 29-38, 2011.

[5] Y. Kaya, M. Uyar, R. Tekin, S. Yildirim, "1D-local binary pattern based feature extraction for classification of epileptic EEG signals", *Appl Math Comput*, vol. 243, pp: 209-219, 2014.

[6] T. Sunil Kumar, V. Kanhangad, R. B. Pachori, "Classification of seizure and seizure-free EEG signals using local binary patterns", *Biomed Signal Proces*, vol. 15, pp: 33-40, 2015.

[7] K. Samiee, P. Kovacs, M. Gabboudj, "Epileptic Seizure Classification of EEG Time-Series Using Rational Discrete Short-Time Fourier Transform", *IEEE T BIO-MED ENG*, vol. 62, no.2, pp: 541-562, 2015.

[8] Z-K. Gao, Q. Cai, Y-X. Yang, N. Dong, S-S. Zhang, "Visibility Graph from Adaptive Optimal Kernel Time-Frequency Representation for Classification of Epileptiform EEG", *Int J Neural Syst*, vol. 27, no.4, pp: 1750005 (1-12), 2017.

[9] G. Zhu, Y. Li, Peng (Paul) Wen, "Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm", *Comput Meth Prog Bio*, vol. 115, no.2, pp: 64-75, 2014.

[10] V. Joshi, R. B. Pachori, A. Vijesh, "Classification of ictal and seizure-free EEG signals using fractional linear prediction", *Biomed Signal Proces*, vol. 9, pp: 1-5, 2014.

[11] A. T. Tzallas, M. G. Tsipouras, D. G. Tsalikakis, E. C. Karvounis, L. Astrakas, S. Konitsiotis, M. Tzaphlidou, "Automated Epileptic Seizure Detection Methods: A Review Study, Epilepsy - Histological, Electroencephalographic and Psychological Aspects, Dr. Dejan Stevanovic (Ed.)", *InTech,* Rijeka, Croatia, 2012.

[12] R. G. Andrzejak, K. Lehnertz, F.Mormann, C. Rieke, P. David, C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state", *Phys Rev E*, vol. 64, no.6, Article ID 061907, 8 pages, 2001.

[13] S. V. Vaseghi, "Linear prediction models, (2nd ed.)., Advanced Digital Signal Processing and Noise Reduction, ch.8", *Wiley,* New York, USA, 2002.

[14] M. Adnane, A. Belouchrani, "Heartbeats classification using QRS and T waves autoregressive features and RR interval features", *Expert Syst*, DOI: 10.1111/exsy.12219, 2017.

[15] C. Cortes, V. Vapnick, "Support-vector networks", *Mach Learn*, vol. 20, no.3, pp: 273-297, 1995.