Epileptic seizure detection using Reservoir Computing

Pieter Buteneers, Benjamin Schrauwen, David Verstraeten, Dirk Stroobandt
ELIS research department, UGent:
Sint-Pietersnieuwstraat 41, 9000 Ghent Belgium
e-mail: pieter.buteneers@ugent.be

***Abstract*—In this paper it is shown that Reservoir Computing can be successfully applied to perform real-time detection of epileptic seizures in Electroencephalograms (EEGs). Absence and tonic-clonic seizures are detected on intracranial EEG coming from rats. This resulted in an area under the Receiver Operating Characteristics (ROC) curve of more than 0.99 on the data that was used. For absences an average detection delay of 0.3s was noted, for tonic-clonic seizures this was 1.5s. Since it was possible to process 15h of data on an average computer in 14.5 minutes all conditions are met for a fast and reliable real-time detection system.**

***Index Terms*—epilepsy, real-time seizure detection, Reservoir Computing, absences, SWDs, tonic-clonic seizures, neural networks, EEG**

I. Introduction

Epilepsy is a neurological disorder of the brain where the patient is disturbed by mostly recurring seizures. Around 1% of the world's population suffers from this illness [1]. Although a cure for this disorder has not yet been found, medication is in most cases sufficient to block the seizures.

To determine whether the applied medication is working, a doctor needs to determine the remaining epileptic activity, and thus also the seizures, on hours of recorded EEG data. Especially in the case of absence epilepsy, where the number of seizures is a very good indication, an automatic detection system is highly desired. Additionally anti-epileptic drugs are known for their side-effects. To avoid these side-effects one could apply a closed-loop system, where the medication is applied in real-time when a seizure occurs. A third application of a fast and reliable detection system is as a warning system. The environment of the patient can then be alerted when a seizure occurs so they are able to help and protect the patient in this unpleasant and dangerous episode.

In this study we show how Reservoir Computing (RC) can be successfully applied to detect epileptic seizures on EEG in real-time. RC is a training method for recurrent neural networks where only a simple linear readout function is trained for a randomly created network or reservoir.

The seizures of two different types of generalized epilepsy are detected: absence epilepsy and tonic-clonic epilepsy. Spike Wave Discharges (SWDs) are the EEG patterns that occur when a patient with absence epilepsy is having a seizure. These absences last from several seconds to a few minutes and are generally very well treatable with medication. Tonic-clonic seizures on the other hand are not as easily suppressed by medication. They last from about a minute to even more than 15 minutes and consist of two phases as can be distinguished in Fig. 3. In the tonic phase the muscles are strained followed by convulsions in the clonic phase. Although the data used for this study is intracranial rat data, it can be used as a model for human data.

In section II Reservoir Computing is explained. Section III describes the reservoir setup used to detect SWDs and evaluates the results by a comparison in detection delay and ROC curves with other methods. Section IV briefly describes the similar setup used for tonic-clonic seizure detection and gives the test results on the data set used. In the last section a conclusion is drawn with a link to future work.

II. Reservoir Computing

Reservoir Computing (RC) [2] is a very recent training method for recurrent neural networks and is based on the Echo State Network (ESN) approach which was presented by H. Jaeger in 2001 [3]. RC, which has been shown to render very good results at speech recognition [4], greatly simplifies the training of recurrent neural networks. To generate an output, a linear readout function is attached to a randomly created recurrent neural network, which is called a reservoir, and only the weights of this function are trained. This way the known stability issues of RNNs are easily avoided as well as the long training time [5].

The RC concept is shown to work with several types of neurons: spiking neurons, analog neurons... For this study Leaky Integrator Neurons [6] are used. These are basic hyperbolic tangent neurons followed by a simple first order lowpass filter.

The operation of the reservoir can be described as follows: If we use $\mathbf{x}[k]$ to represent the current activation for each of the neurons in the reservoir we can calculate the next reservoir state $\mathbf{x}[k+1]$ using the following equation:

$$\mathbf{x}[k+1] = (1-\gamma)\cdot\mathbf{x}[k] + \gamma\cdot\tanh(W_{res}^{res}\mathbf{x}[k] + W_{inp}^{res}\mathbf{u}[k] + W_{bias}^{res})$$

In this equation γ represents the leak rate and is used to set the cutoff frequency of the lowpass filter in the neurons, $\mathbf{u}[k]$ is the input vector and W_a^b represents the randomly generated weight vector from layer a to b. To generate the output $\hat{\mathbf{y}}[k]$ the following equation is used:

$$\hat{\mathbf{y}}[k] = W_{res}^{out}\mathbf{x}[k] + W_{bias}^{out}$$

Here $\hat{\mathbf{y}}[k]$ is used instead of $\mathbf{y}[k]$ because the latter is often used as the desired output. In these previous equations only

W^{out}_{res} and W^{out}_{bias} are trained for which in this study a ridge regression based algorithm was used.

To render good results the reservoir itself is just slightly tweaked [5]. For W^{res}_{res} and W^{res}_{inp} the optimal connection fraction is determined. This factor states the percentage of weights that are not zero. The processing power of reservoirs is greatest when it operates at the edge of stability [7]. Therefore the spectral radius (the largest absolute eigenvalue) of W^{res}_{res} is scaled together with W^{res}_{inp}. The elements in W^{res}_{bias} are all equal to 1 and are also scaled. These connection factors and scaling parameters together with γ are optimized by a Monte-Carlo simulation to get statistically relevant results.

Before presenting the data to the reservoir, it is preprocessed to gain better results. Therefore the detection system has for each of the seizure types a specific though simple preprocessing step, which will be discussed in the relevant sections.

III. SWD DETECTION

A. The reservoir setup

As preprocessing, the best channel is manually selected from the EEG and resampled from 200Hz to 100Hz. For each of these datasets the mean absolute value is determined. This value is used as a scaling factor which results in a mean absolute value of 1 for each of the datasets.

As shown in Fig. 1 an absence seizure has a main frequency around 8 Hz and several harmonics. The best results were obtained by using the absolute values of the following signals as input for the reservoir: the first derivative and the frequency bands around 8, 16 and 24 Hz.

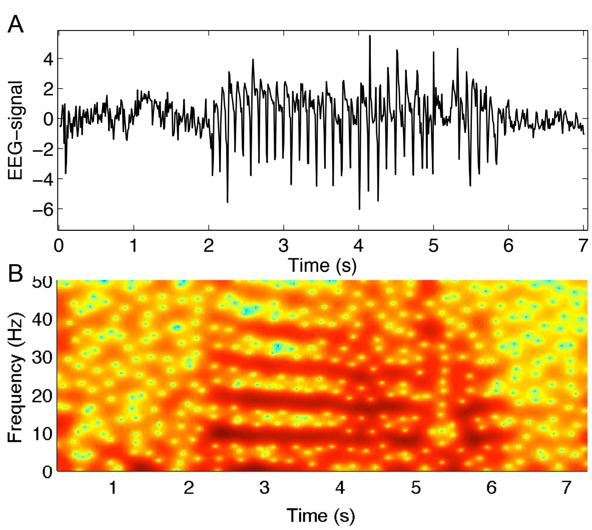

Fig. 1. Absence seizure (A) with its frequency content (B). An absence seizure has a main frequency around 8Hz and several harmonics. These harmonics together with the first derivative are used as input to the reservoir.

The reservoir that is used consists of 200 neurons which are randomly connected to the input layer and each other with a connection fraction of 90%. For the spectral radius of the reservoir 1.3 was used and the connections from the input to the reservoir are scaled up to 1.6. These values might seem unstable (see section II), but due to the bias (of 2.8) and the non-linearity in the neurons this results in stable reservoirs that operate in the highly non-linear area of the hyperbolic tangent functions. The optimal value for the leak rate was found to be $\gamma = 0.06$.

As training the output weights were optimized to generate an output of 1 when the considered sample is part of a seizure and -1 in the opposite case. To classify the output a simple threshold is used. If an output sample is higher than the threshold it is considered part of a seizure, in the other case not part of a seizure. To gain better results, this binary output is be post-processed: a sample is considered part of a seizure if the output is above the threshold for more than 300 consecutive samples (3s).

B. Evaluation

A comparison was made with three other detection methods for SWDs. The first one was presented by Van Hese et al. in [8] and considers the energy in an EEG signal to distinguish between normal EEG and a seizure. The second method was presented by Westerhuis et al. in [9] and uses the first derivative of the EEG to detect seizures. The last one was presented by Fanselow et al. in [10] and uses the amplitude of the EEG signal. All three methods use intervals in which the EEG signal is tested and state that a seizure has to last at least a certain amount of intervals. To make a fair comparison the length of these intervals was optimized on the same data as was used to train the reservoir.

In order to generate ROC curves the specificity and sensitivity is measured for each threshold value between the maximum and minimum output value. In ROC curves sensitivity is plotted versus specificity and thus the higher the Area Under Curve (AUC) the better the performance.

Every result is based on the output of the complete dataset, in total 15 hours and 17 minutes of EEG-data coming from 13 different Genetic Altered Epilepsy Rats from Strasbourg, GAERS rats [11]. The data contains seizures that last from about 8 seconds to a little more than 2 minutes with a frequency of about 1.5 seizures every minute. The reservoir and the intervals are trained on only 10% of this data which was enough to achieve good results. For testing all data was used.

In Fig. 2 the ROC curves are shown for the three methods together with the presented Reservoir Computing approach. It is clear that previous work is outperformed by the method using Reservoir Computing which results in a maximum AUC of 0.992.

The detection delay is defined as the time needed for the reservoir output to become higher than the threshold value after the start of a seizure. To make this consistent the threshold value for which specificity and sensitivity are equal (the Equal Error Rate or EER point) is used. Since the methods used for a comparison are not created with the intention to detect in real-time, they don't perform well on detection speed. All methods including the presented method resulted in a detection delay of more than 3 seconds. This is mainly because

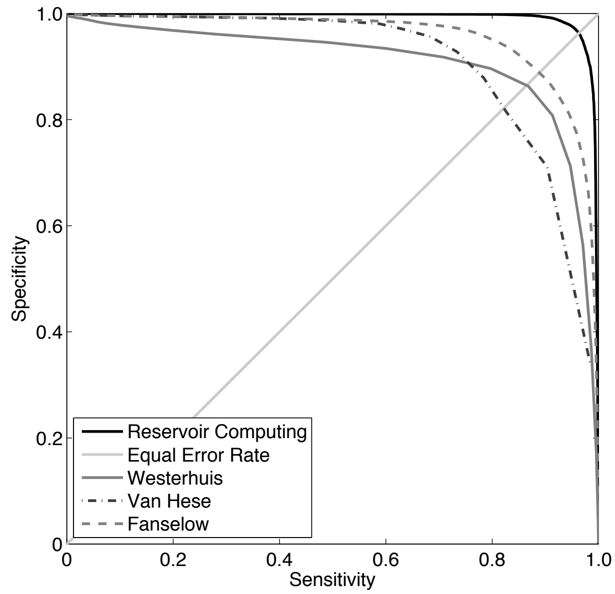

Fig. 2. ROC curves showing the better performance achieved using Reservoir Computing with a AUC of 0.992. The method by Van Hese et al. results in a AUC of 0.906, the method by Westerhuis et al. in a AUC of 0.910 and the method by Fanselow et al. in a AUC of 0.956.

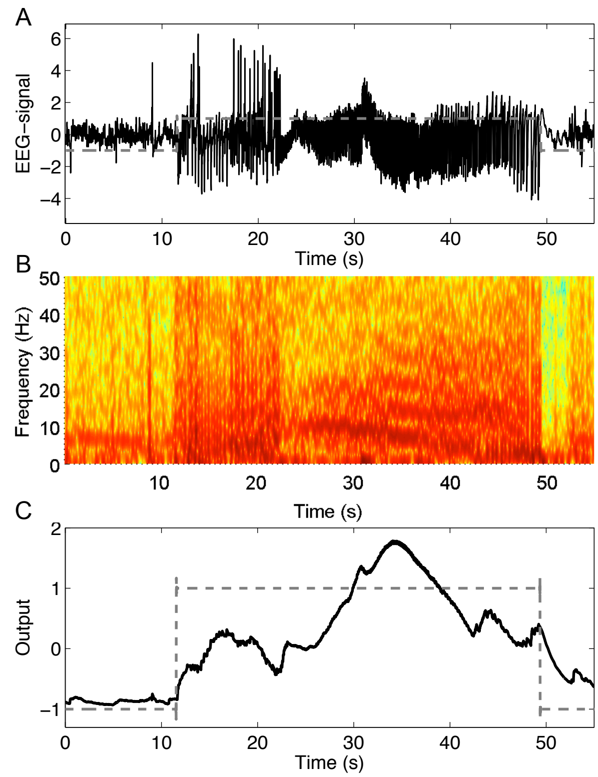

Fig. 3. Tonic-clonic seizure (A) with its frequency content (B) which doesn't show a main frequency for the seizure. The slow decay of the generated output (C) illustrates the slowness of the reservoir.

they all use a similar post-processing method. If however this post-processing step is abandoned, rather interesting results are achieved for RC and the method by Fanselow et al. The presented method then results in an AUC of 0.987, which is just slightly lower than 0.992, with an average detection delay of 0.3 seconds and 0.2 standard deviation. The method by Fanselow without post-processing results in an AUC of 0.936, instead of 0.956, an average detection delay of only 0.24 seconds and with a standard deviation of 0.3 seconds.

Just to show the processing power of a single reservoir the surface without pre- and post-processing was determined. Because EEG can differ much in amplitude, the data was however only rescaled as explained before. This simplification resulted in a AUC of 0.977. Even this outperforms the results achieved by the other methods.

IV. Detection of tonic-clonic seizures

A. The reservoir setup

Although tonic-clonic seizures create a very different pattern in the EEG, as shown in Fig. 3.A, these seizures can be detected with almost the same setup as for SWDs. As shown in Fig. 3.B, the frequency content doesn't contain the same harmonic structure as SWDs. Still it is clear that the higher frequency bands contain the most information.

For this type of seizures four channels were used instead of one. Each of them was again resampled, from 500Hz to 100Hz, and rescaled by dividing by its mean absolute value. To separate the high frequency content the simplest first order high-pass filter, the first derivative, is applied to all four of them.

To optimize the parameters we started from the same reservoir of 200 neurons as used in the previous section. The only setting that differed much from the previous setup was the leak rate. A clearly lower optimum was achieved for $\gamma = 0.003$, which results in a lower cutoff frequency for the lowpass filter. A post-processing method wasn't used in this setup.

Fig. 3.C shows the output generated on the seizure plotted in Fig. 3.A. The figure illustrates that a low threshold (about -0.5) is needed to get good results and that most false-positive results have their origin in the slow decay of the output after the seizure. This was not the case with SWDs and is caused by the lower cutoff frequency of the lowpass filter in the neurons. A lower cutoff frequency results in a slower change in neuron activations and thus a slower reservoir.

B. Evaluation

In total 4 hours and 23 minutes of data of 4 different rats was used. The shortest seizure, shown in Fig. 3.A, lasted about 40 seconds, the longest a little less than 12 minutes. This is remarkably longer than SWDs. Since the data studied contained only one tonic-clonic seizure every 40 minutes only the seizures together with some non-seizure data before and after the seizure were cut out of the data. This resulted in 20% seizure data and 80% non-seizure data. Only 20% of the total data set was needed for training to get good results, the other data was used to test the performance.

An average AUC of 0.993 was achieved with a detection delay of on average 1.5 seconds. In 85% of the cases the seizures were detected within 3 seconds. This is a factor 5 slower than with SWD detection but still quite fast. To be used in a closed-loop environment the detection system needs to be be able to detect in less than a second and even be able to predict the seizures in order to suppress them with fast-working medication.

V. Conclusion and future work

We have shown that it is possible to detect absence and tonic-clonic seizures on rat data with Reservoir Computing in a very reliable way. For both seizure types it resulted in a surface under the ROC curve of more than 0.99. Therefore it can be used as a reliable aid to evaluate EEGs where the amount of seizures needs to be determined.

The proposed method is not only reliable, but it is also fast: 85% of the tonic-clonic seizures are detected within 3 seconds. Absences were detected with even less delay. Here 85% of the seizures were detected within 0.5 seconds.

Both conditions, reliability and low detection delay, enable the use of Reservoir Computing in a real-time warning or even a real-time treatment system for absence epilepsy in rats. However to achieve the same for tonic-clonic seizures more research needs to be done to further reduce the detection delay. Although good results were achieved for rat data, it is also paramount to test this system on human data before considering a real-time warning or treatment system.

In order to thoroughly evaluate the presented method it is necessary to compare the given results with more methods than the ones used here, certainly for tonic-clonic seizures where no comparison was made. Also a larger data set is needed for a better comparison.

References

[1] JF. Annegers. *The treatment of epilepsy: principle and practice*, chapter "The epidemiology in epilepsy", pages 165–172. Baltimore: Williams and Wilkins, 1996.

[2] B. Schrauwen, D. Verstraeten, and J. Van Campenhout. An overview of reservoir computing: theory, applications and implementations. In *Proceedings of the European Symposium on Artifical Neural Networks (ESANN)*, 2007.

[3] H. Jaeger. The "echo state" approach to analysing and training recurrent neural networks. Technical Report GMD Report 148, German National Research Center for Information Technology, 2001.

[4] D. Verstraeten, B. Schrauwen, and D. Stroobandt. Isolated word recognition using a Liquid State Machine. In *Proceedings of the 13th European Symposium on Artificial Neural Networks (ESANN)*, pages 435–440, Evere, May 2005. D-side Publications.

[5] H. Jaeger. Tutorial on training recurrent neural networks, covering BPTT, RTRL, EKF and the "echo state network" approach. Technical Report GMD Report 159, German National Research Center for Information Technology, 2002.

[6] H. Jaeger, M. Lukosevicius, and D. Popovici. Optimization and applications of echo state networks with leaky integrator neurons. *Neural Networks*, 20:335–352, 2007.

[7] R. A. Legenstein and W. Maass. Edge of chaos and prediction of computational performance for neural microcircuit models. *Neural Networks*, pages 323–333, 2007.

[8] P. Van Hese, J.-P. Martens, P. Boon, S. Dedeurwaerdere, I. Lemahieu, and R. Van de Walle. Detection of spike and wave discharges in the cortical EEG of genetic absence epilepsy rats from Strasbourg. *Physical and Medical Biology*, 48:1685–1700, 2003.

[9] F. Westerhuis, W. Van Schaijk, and G. Van Luijtelaar. Automatic detection of spike-wave discharges in the cortical EEG of rats. In *Measuring Behavior '96, International Workshop on Methods and Techniques in Behavioral Research (Utrecht, The Netherlands 16-18 Oct. 1996)*, 1996.

[10] E. Fanselow, P. Ashlan, and A. Nicolelis. Reduction of pentylenetetrazole-induced seizure activity in awake rats by seizure-triggered trigeminal nerve stimulation. *Journal of Neuroscience*, 20:8160–8168, 2000.

[11] C. Marescaux, M. Vergnes, and A. Depaulis. Genetic absence epilepsy in rats from Strasbourg - a review. *Journal of Neural Transmission*, 35:37–69, 1992.