

Automatic detection of epileptic seizures on the intra-cranial electroencephalogram of rats using reservoir computing

Pieter Buteneers^{a,*}, David Verstraeten^a, Pieter van Mierlo^a, Tine Wyckhuys^b, Dirk Stroobandt^a, Robrecht Raedt^b, Hans Hallez^a, Benjamin Schrauwen^a

^a*Electronics and Information Systems, Ghent University,
Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium*

^b*Laboratory for Clinical and Experimental Neurophysiology, Ghent University,
De Pintelaan 185, 9000 Ghent, Belgium*

Abstract

Introduction: In this paper we propose a technique based on reservoir computing (RC) to mark epileptic seizures on the intra-cranial electroencephalogram (EEG) of rats. RC is a recurrent neural networks training technique which has been shown to possess good generalization properties with limited training.

Materials: The system is evaluated on data containing two different seizure types: absence seizures from genetic absence epilepsy rats from Strasbourg (GAERS) and tonic-clonic seizures from kainate-induced temporal-lobe epilepsy rats. The dataset consists of 452 hours from 23 GAERS and 982 hours from 15 kainate-induced temporal-lobe epilepsy rats.

Methods: During the preprocessing stage, several features are extracted from the EEG. A feature selection algorithm selects the best features, which are then presented as input to the RC-based classification algorithm. To classify the output of this algorithm a two-threshold technique is used. This technique is compared with other state-of-the-art techniques.

Results: A balanced error rate (BER) of 3.7% and 3.5% was achieved on the data from GAERS and kainate rats, respectively. This resulted in a sensitivity of 96% and 94% and a specificity of 96% and 99% respectively. The state-of-the-art technique for GAERS achieved a BER of 4%, whereas the best technique to detect tonic-clonic seizures achieved a BER of 16%.

Conclusion: Our method outperforms up-to-date techniques and only a few parameters need to be optimized on a limited training set. It is therefore suited as an automatic aid for epilepsy researchers and is able to eliminate the tedious manual review and annotation of EEG.

Keywords: reservoir computing, neural networks, EEG classification, automatic seizure detection, experimental animal models for epilepsy

*Corresponding author. Tel.: +32 9 264 33 68; fax: +32 9 264 35 94.
Email address: Pieter.Buteneers@UGent.Be (Pieter Buteneers)

1. Introduction

Epilepsy is a neurological disorder of the brain characterized by mostly recurring seizures. Approximately 1% of the world's population suffers from this illness [1]. Although a cure for this disorder has not yet been found, medication is sufficient to block the seizures in 75% of the cases [2].

1.1. Epilepsy animal models

Experimental animal models are widely used [3] to gain greater insight into the pathophysiology underlying epilepsy. Genetic absence epilepsy rats from Strasbourg (GAERS) [4] are a strain of Wistar rats that is widely accepted as a model for human absence epilepsy because it shares many clinical characteristics [5]. Kainate-induced temporal-lobe epilepsy rats on the other hand are a widely used model for human temporal lobe epilepsy [6, 7]. There is a significant clinical difference between the seizures of these two animal models. Moreover the quasi-periodic spike and wave discharges (SWDs) that occur in the electroencephalogram (EEG) during absence seizures show little resemblance to the electrographic activity during the tonic-clonic (TC) seizures in kainate-induced epilepsy rats.

When animal models are used for epilepsy research, long-term intra-cranial EEG monitoring is often done to evaluate different therapeutic methods such as new medication, deep brain stimulation [8, 9], vagus nerve stimulation and stem cell therapy. This results in large amounts of data, collected over a period of several months, from which the evolution of seizure frequency needs to be determined. To help marking the seizures, this study proposes an extension of the seizure detection method based on reservoir computing (RC) which is presented in [10].

1.2. Reservoir computing

In this work RC [11] is used as a classification algorithm. It is a training method for recurrent neural networks (RNNs) and a generalization of the Echo State Network approach introduced by H. Jaeger in [12].

Feedforward neural networks are known for their generalization properties and ability to learn complex relationships between inputs and outputs with limited training. RNNs add recurrent connections and thus time-awareness to these networks, resulting in a more biologically relevant dynamical system that can be thought to find relationships between the desired output and any past input. Traditionally, all interconnection weights between the neurons in an RNN are trained. RC on the other hand uses a randomly created neural network, called a reservoir, from which only a single linear output is trained. That way the long training time and stability issues of regular RNNs are avoided without losing the desired generalization abilities.

In the RC set-up each non-zero input sample will excite this dynamical system and push the reservoir to a new state. In practice this can be seen as a

projection of the input data to a higher dimensional space. In this high dimensional space the probability increases that the classes are linearly separable [13]. In that respect, RC strongly resembles support vector machines [14] but with a random spatio-temporal kernel as opposed to a mere spatial kernel.

To improve detection performance, the EEG-data is first preprocessed. Several features are extracted from the EEG and selected through a forward feature selection algorithm [15] (see Section 3.3.1). The output of the classification algorithm is a continuous value to which a two-threshold technique is applied. Both thresholds are optimized to correctly classify the samples as ictal or inter-ictal EEG.

Unlike the RC-method presented in [10], the proposed RC-method adds the feature selection algorithm as well as a more advanced preprocessing and thresholding technique. Moreover the performance is evaluated on a much larger dataset: 452 hours of EEG from 23 GAERS and 982 hours of EEG from 15 kainate-induced epilepsy rats as opposed to respectively 15 hours from 13 rats and 4.5 hours from 4 rats. To show the detection performance a comparison is made with seven other seizure detection methods: a human seizure detection method presented in [16], the five seizure detection methods presented for rats in [17], [18], [19], [20] and [21] and a seizure detection method based on the method presented in this study. The latter however uses a linear classifier as opposed to RC. To our knowledge, the methods in [18]-[21] are the only published methods for detecting seizures of GAERS, of which the method in [21] is currently considered as state-of-the-art. The methods presented in [20] and [21] are the only ones using features that are not applicable to kainate-induced epilepsy rats. RC has been shown to yield very good results for speech recognition [22]. It has the advantage that it is a non-linear, learning dynamical system as opposed to the mostly linear classification techniques used in literature. In addition it can be taught to ignore artifacts and to detect different types of seizures.

2. Materials

This study compares several methods for epileptic seizure detection on two different seizure types: absence seizures from GAERS and tonic-clonic seizures from kainate-induced epilepsy rats.

In both cases, the EEG was recorded with a custom-built amplifier (gain: 510x) connected to an NI-DAQ6259 data acquisition card (National Instruments, USA) and a PC. Afterwards this EEG was subsampled to 200Hz before being evaluated by experienced encephalographers.

All rats were treated according to guidelines approved by the European Ethics Committee (decree 86/609/EEC). The study protocols were approved by the Animal Experimental Ethical Committee of Ghent University Hospital (ECD 01/26, ECD 05/10 and ECP 04/08). The animals were kept under environmentally controlled conditions (12h normal light/dark cycles, 20-23°C and 50% relative humidity) with food and water *ad libitum*. All animals were allowed to move freely and no drugs were administered during EEG recording.

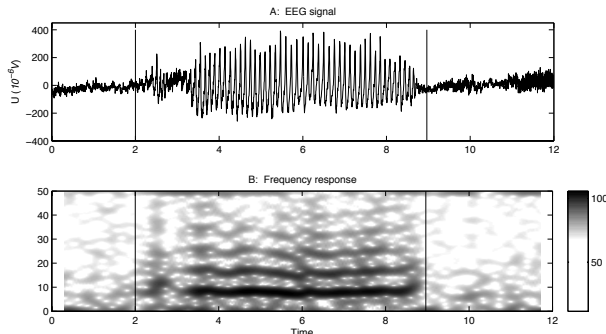


Figure 1: An example of a spike and wave discharge caused by an absence seizure in genetic absence epilepsy rats from Strasbourg. In (A) the EEG signal of one intra-cranial channel is shown. The seizure starts at time = 2 s and stops at time = 8.9 s. (B) shows the spectrogram of the EEG signal with a Hamming-window of 128 samples and an overlap of 120 samples.

2.1. Genetic absence epilepsy rats from Strasbourg

GAERS are a strain of Wistar rats that all exhibit spontaneous, generalized epileptic seizures characterized by paroxysmal unresponsiveness to environmental stimuli and cessation of ongoing activity. These absence seizures, which are displayed as synchronous spike and wave discharges (SWDs) on the EEG, occur mostly when the animal is in a state of quiet wakefulness. However, they are rare during periods of active arousal and sleep. The number of seizures and their duration increase with age, until it reaches a maximum at about 6 months. The EEG of SWDs shows a fundamental frequency in the range of 7 to 12 Hz and several harmonics (see Figure 1), an amplitude varying from 300 to 1000 μV and a duration from 0.5 to 120 s.

The animals used in these studies were all male GAERS, of 4 to 5 months old, with a weight of 300 to 350 g. After the operation to insert the electrodes all animals had one week of recovery.

Dataset A was made during a study to evaluate the effect of acute and subacute high (130 Hz) and middle high (60 Hz) frequency deep brain stimulation on the occurrence of SWDs [8]. The rats were anesthetized with ketamine/xylazine. Three epidural EEG screw electrodes, made of stainless steel, were bilaterally implanted over the frontoparietal cortex, together with one depth electrode in the anterodorsal thalamus and one reference electrode placed at lambda.

The rats from dataset B were part of a study to evaluate the effect of long-term vagus nerve stimulation with a current intensity of 1.5 mA, pulsewidth of 250 μs , a frequency of 30 Hz and a 5 min on / 30 s off cycle. In this case the rats were anesthetized with isoflurane, three epidural EEG screw electrodes were bilaterally implanted over the frontoparietal cortex, together with one reference electrode placed at lambda.

Because a little more than half of the animals were stimulated during recording, some of the EEG is contaminated with stimulation artifacts. These arti-

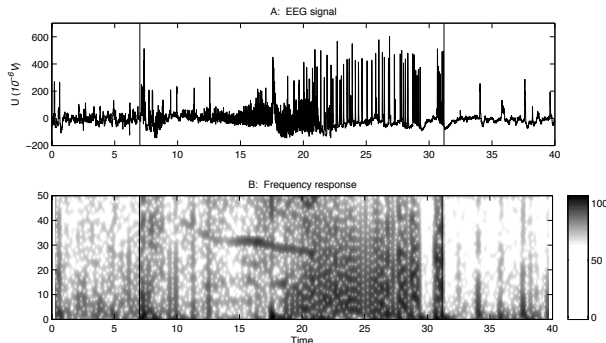


Figure 2: An example of a tonic-clonic seizure in a kainate-induced epilepsy rat. In (A) the EEG signal is shown, the seizure starts at time = 7 s and stops at time = 31.2 s. (B) shows the spectrogram of the EEG signal with a Hamming-window of 128 samples and an overlap of 120 samples.

facts were excluded from the datasets to avoid a high number of false positives. However, other artifacts such as movement and scratching artifacts were not excluded.

Waterschoot evaluated all EEG fragments visually, selected one EEG channel and marked all present SWDs with a minimum seizure length of 0.5 s. These annotations were used as the ‘gold standard’ in this study. From study A, 64.5 hours of single-channel referential EEG-data from 12 different rats was used 23% of the total time contained the 3468 seizures which lasted on average 15 seconds. Study B yielded 390 hours of single-channel referential EEG-data from 11 rats. A total number of 6183 seizures made up 4.5% of the data and lasted 10 seconds on average. Each of the seizures lasted between 0.5 and 110 s.

2.2. Kainate-induced temporal-lobe epilepsy rats

Kainic acid is a potent central nervous stimulant, isolated from the seaweed *digenea simplex*. This excitotoxic product is a glutamate analogue and systemic injection in healthy rats triggers a cascade of molecular and cellular events eventually leading to status epilepticus, followed by a period of gradual increase in seizure frequency, which eventually stabilizes. Finally, rats display the spontaneous, generalized tonic-clonic convulsions which resemble those seen in temporal-lobe epilepsy patients [23].

Spontaneous EEG seizures were recognized against background by their large amplitude (more than 3 times baseline amplitude), high-frequency EEG activity (≥ 5 Hz), with characteristic high temporal correlation and progression of spike frequency. Figure 2 shows an example of a TC-seizure.

For studies C and D healthy male Wistar rats (≥ 150 g) (Harlan, the Netherlands) were injected intraperitoneally (5 mg/kg) with kainic acid according to the protocol of Hellier et al. [6]: kainic acid was administered every hour until a stable self-sustaining status epilepticus was displayed for ≥ 3 hours. Fifty

days after the kainate injection, rats that displayed more than six clonic or tonic-clonic convulsions were selected for implantation with EEG-recording electrodes and cannulas.

For surgery, the rats were anesthetized with isoflurane. After exposure of the skull, 10 small burr holes were drilled; six and four, respectively for study C and D, were intended for the positioning of anchor screws, the others for the electrodes and cannulas. Bipolar and quadripolar electrodes were custom-made by gluing together two or four straight-cut poly-imide coated stainless steel wires. Cannulas were only used in study D. The electrodes were led to a head stage which was fixed not only to the screws, but also to the cannulas and the skull with acrylic dental cement. Following surgery, the rats were allowed one week of recovery before continuous intracranial EEG-recording was started.

Dataset C was made during a study to evaluate the effect of long-term high frequency (130 Hz) and Poisson distributed high frequency (on average 130 Hz) deep brain stimulation on the occurrence of TC seizures [9]. For this study, a custom-made epidural electrode was screwed in the right side of the skull at the height of the frontal cortex. A similarly constructed reference electrode was placed on the right side posterior to sutura lambdoidea. Two depth electrodes were stereotactically inserted into the hippocampus (AP -5.8, DV -7.5, ML 4.7 and at ML - 4.7 to bregma [24]): a quadri-polar in the right hippocampus and a bipolar in the left hippocampus.

Wyckhuys evaluated all EEG fragments visually and marked all present seizures. Episodes of EEG during Deep Brain Stimulation were omitted from analysis as these were contaminated with stimulation artifacts. This resulted in 913 hours of four channel EEG from 11 different rats. Approximately 2.5% of this data consisted of 1541 seizures which have a duration of 9 to 240 seconds with an average of 54 seconds.

In study D the effect of introducing stem cells from foetal mice brains in the epileptogenic areas was studied on the occurrence of TC seizures. For this study, custom-made reference electrodes were placed anterior to the lambdoid structure. Bipolar recording electrodes were inserted into the left and right hippocampus (anterioposterior -4.1 mm, mediolateral 2.6 mm, dorsoventral -3.6 mm to bregma [24]). More posterior, one guide cannula was inserted in the left and one in the right hippocampus.

All EEG fragments were visually evaluated by Van Dycke and all present seizures were marked. From this study 69 hours of 4 channel EEG from 4 different rats was used. Dataset D contained 113 seizures that were located in about 2.5% of the data and lasted 23 to 360 seconds with an average of 51 seconds.

Datasets C and D consist of 4 channel intra-cranial EEG with a referential montage for each rat. All 4 were used since they had also been used during annotation and to create a generic system without the need to select the best channel. Although the EEG contained several artifacts and some channels had a low signal quality, no data was excluded from the dataset.

3. Methods

3.1. Training, validation and test sets

The complete dataset comprised 454 hours of SWD data and 982 hours of TC data. The SWD training set, 5 hours in total, consisted of 3 times 4.35 minutes of EEG per rat: the first, the midmost and the last 4.35 minutes of the complete SWD dataset that constituted least 12% of one of the two classes. This made sure that each class was well represented. The TC training set of 8h20, consisted of 10 seizures per rat together with inter-ictal EEG before and after the seizure. These parts are each randomly between 2 to 3 times the length of the seizure. The first and the last seizures are used together with 8 seizures equally distributed over the dataset. Although the data of some rats contained artifacts, the selection procedure of the training set did not consider them.

The whole dataset was used for the test set, including the parts that were selected for the training set. Although the test set included the training samples, they were not used in the evaluation. This is unconventional but was necessary to ensure that the data remained continuous. The validation set is a part of the training set that is excluded during training and is used to validate for example the current parameter setting. To obtain statistically relevant results this is cross-validated by applying several different training and validation sets for each tested set-up and averaging the results. In this study 10-fold cross-validation is used with a validation set of about 10 % of the training set.

3.2. Evaluation measures

The scoring of experienced encephalographers is the golden standard used to compare the different detection methods in this study. As error measure the balanced error rate (BER) was used. The BER is the average fraction of misclassified samples per class and is given by:

$$BER = \frac{1}{2} \left(\frac{FP}{TN + FP} + \frac{FN}{FN + TP} \right),$$

where FP, FN, TP and TN respectively stand for the number of samples that is falsely classified as positive, falsely classified as negative, truly classified as positive and truly classified as negative. This measure takes into account that most datasets, such as that used in this study, are unbalanced. To evaluate the result for each class independently, the sensitivity and specificity [25] are also measured. In fact, the BER is the average error on the sensitivity and specificity: $BER = \frac{1-Sens}{2} + \frac{1-Spec}{2}$.

3.3. Seizure detection using reservoir computing

The proposed detection method is comprises three parts: a pre-processing stage where EEG features are selected, a classification stage which is based on RC and a thresholding stage. In Figure 3 these steps are illustrated for an SWD.

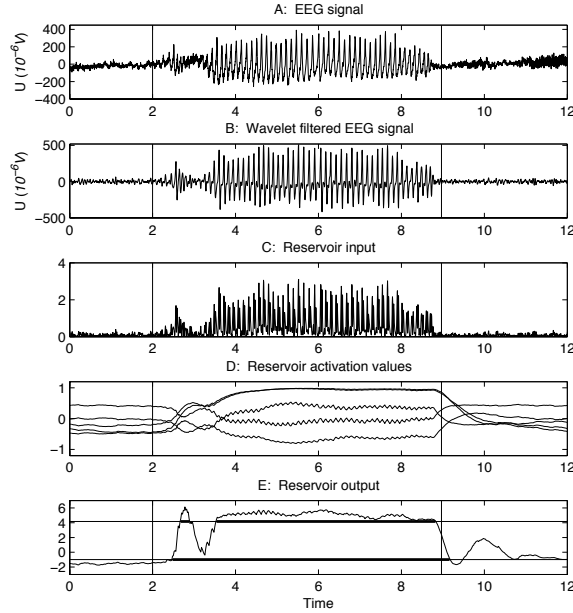


Figure 3: The RC detection method applied to EEG of GAERS. In (A) the EEG signal of a Spike and Wave Discharge (SWD) is shown. The seizure starts at time = 2 s and stops at time = 8.9 s. The preprocessing is shown in (B), the wavelet filtered signal, and (C), the reservoir input. (D) shows the activation values of 5 of the 200 reservoir neurons and (E) shows the generated output with the two thresholds. Every sample above the high threshold is a detection, illustrated with the bold part of the high threshold line (the upper horizontal line). All samples neighboring these detected samples that are above the second threshold are also considered to be part of a seizure and are illustrated by the bold part of the low threshold line.

3.3.1. Preprocessing

For preprocessing the relevant features from the EEG [26, 27] are selected during training with a forward feature selection algorithm. The EEG features used in this study were: a filter bank of Butterworth filters ranging from 1 to 30 Hz with a bandwidth of 2 Hz, a set of Daubechies 4 wavelet filters (level 2 to 6), the first derivative, the energy of the signal and the energy in the theta (4 to 8 Hz), alpha (8 to 12 Hz), beta (12 to 30 Hz) and gamma (>30 Hz) bands. Each of these features is designed in such a way that they result in a signal sampled at 200 Hz after extraction.

For each of the features independently, this signal is then subdivided in non-overlapping intervals, with length L , from which the foreground input signal FG is calculated as the mean absolute value of these intervals. Since there is a high variability in signal amplitudes between different rats, a background signal is

Algorithm 1 Forward Feature Selection

```
rankedFeatures = sortByIncrError(features);
featureSet = rankedFeatures(1);
for i = 2 to length(rankedFeatures) do
    set = {featureSet; rankedFeatures(i)};
    if error(set) < error(featureSet) then
        featureSet = set;
    end if
end for
```

estimated for each of the EEG-features to serve as a reference level for rescaling. It is estimated from the foreground signal, for each hour of EEG, as follows: $BG = \text{median}(FG)$. This is a quantile-based estimation technique as proposed in [28], which is based on the assumption that epileptic seizures occur less than half of the time.

As input for the classification algorithm $I = \{\frac{FG_{n,f}}{BG_{n,f}}\}, \forall n \in N, f \in FS$ is used. Here FS is the selected feature set and N is the list of EEG-channels. The features are selected using a forward feature selection procedure [15]. The pseudo-code of this algorithm is shown in Algorithm 1. In a first stage, a list of features is constructed, ranking them from small to large BER. In the second stage, a set of selected features FS is constructed. In the first step, the best performing feature from the ranking is added to FS . In the following steps the next feature from the ranking is added to this set if that results in a smaller BER. Otherwise, the feature set remains unchanged.

For each dataset, the following procedure was followed to achieve to the optimal pre-processing set-up. During this phase, the RC parameters described in [10] were used (see Section 3.3.2). First the feature set was selected using the described feature selection algorithm with an interval length $L = 0.5$ s. In a next step the interval length L was optimized for L in the range $[0.005, 0.01, 0.02, 0.05, \dots, 2]$ s. The feature set that was selected for the SWD datasets consisted only of the level 3 Daubechies 4 wavelet filter with $L = 0.02$ s. For the tonic-clonic dataset, only the energy in the beta band was selected with $L = 0.2$ s.

3.3.2. Reservoir computing

The operation of the reservoir is shown in Figure 4 and can be described as follows. We use $\mathbf{x}[k]$ to represent the current activation values of the neurons in the reservoir at time k , $\mathbf{u}[k]$ as the input vector, $\mathbf{y}[k]$ for the desired output and $\hat{\mathbf{y}}[k]$ for the output generated by the RC system. The inputs of the neurons in the reservoir are connected with the bias as well as with the input and the output of all the neurons in the reservoir. The weights of these connections are represented respectively by the weight matrices \mathbf{W}_{bias} , \mathbf{W}_{inp} and \mathbf{W}_{res} . If n is the number of neurons these matrices respectively have the following shapes: n -by-1, n -by- N and n -by- n with N equal to the number of inputs. The elements of the bias weight matrix \mathbf{W}_{bias} and the internal weight matrix \mathbf{W}_{res}

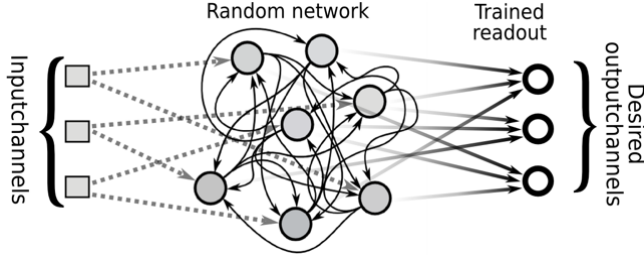


Figure 4: A schematic representation of reservoir computing

are initially uniformly distributed between -1 and 1. All elements in the input weight matrix \mathbf{W}_{inp} are randomly set to $\frac{-1}{N}$ or $\frac{+1}{N}$ with the number of inputs N being equal to the number of EEG channels multiplied by the number of selected features. The sparseness of these matrices as well as the initialization process are however, not critical for the performance. If basic sigmoid neurons, a weighted sum followed by the hyperbolic tangent function, are used the state update equation is given by:

$$\mathbf{x}[k+1] = \tanh(\mathbf{W}_{res}\mathbf{x}[k] + \mathbf{W}_{inp}\mathbf{u}[k] + \mathbf{W}_{bias}).$$

In this work, we used leaky integrator neurons, i.e. basic sigmoid neurons followed by a first-order low-pass filter. The state equation now becomes:

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

In this equation γ represents the leak rate which sets the cutoff frequency of the low-pass filter in the neurons. This extra parameter of leaky integrator neurons is used to tune the reservoir memory and timescales [29].

To generate the output, the following equation is used:

$$\hat{\mathbf{y}}[k] = \mathbf{W}_{out} \begin{bmatrix} \mathbf{x}[k] \\ 1 \end{bmatrix},$$

in which '1' represents the connection from the bias to the output. The weights of the output weight matrix \mathbf{W}_{out} are trained using ridge-regression [30], which minimizes the following loss function:

$$f_{loss} = \|\hat{\mathbf{y}} - \mathbf{y}\|^2 + \lambda \|\mathbf{W}\|^2,$$

with $\hat{\mathbf{y}}$ and \mathbf{y} respectively representing all the states of the computed and desired output, and λ being the regularization parameter. This parameter adds an extra cost to the size of the weights, to avoid overfitting and ensure proper generalization. Minimizing the loss function results in the following formula to calculate the optimal output weights:

$$\mathbf{W}_{out} = (\mathbf{x}^T \mathbf{x} + \lambda \mathbf{1})^{-1} \cdot \mathbf{x}^T \mathbf{y}.$$

The optimal λ is found through cross-validation on the training set and selecting the one with the smallest validation error.

In this study RC, is used as a classifier. An interval of length L (see Section 3.3.1) is considered part of a seizure if more than half of the samples in that interval are part of a seizure. The requested output $\mathbf{y}[k]$ is then set to $-\frac{N_{pos}+N_{neg}}{N_{neg}}$ if there is no seizure at the interval mapped to time k and to $+\frac{N_{pos}+N_{neg}}{N_{pos}}$ if there is a seizure. N_{pos} and N_{neg} are respectively the number of positive and negative examples in the training set. This compensates for the unbalance in the dataset when optimizing the mean square error [31].

The processing power of reservoirs is the greatest when they operate at the edge of stability [32]. Therefore, the spectral radius, the largest norm of the eigenvalues of \mathbf{W}_{res} , the weights of \mathbf{W}_{inp} and \mathbf{W}_{bias} and the leak-rate are scaled to achieve optimal reservoir dynamics [33]. The optimal scaling factors are determined using Monte-Carlo simulations on the training set to get statistically relevant results. For this study, the average BER over 10 reservoirs was used to evaluate a certain parameter set. Each parameter was optimized independently without changing the other parameters. The following parameters, used in [10], were set as initial values: 200 neurons, a spectral radius of 1.3, an input to reservoir scaling of 0.16, a bias scaling of 2.8 and a leak rate of $\gamma = 0.06$ for the SWDs and $\gamma = 0.003$ for the tonic clonic seizures. After optimization only the leak rate differed considerably from the value reported in [10]. For both datasets $\gamma = 0.05$ was optimal. This is not unexpected because the system in [10] lacked the pre- and post-processing methods presented here. More specifically, the different length of the averaging interval L ensured that the leak rate is universal for the two different types of seizures.

3.3.3. Two-threshold

The output of the reservoir $\hat{\mathbf{y}}[k]$ is a continuous output that approximates the desired output $\mathbf{y}[k]$ as shown in Figure 3. A two-threshold setup was used to transform this output to a discrete classification label. Each sample higher than the first threshold is considered as part of a seizure, together with all direct neighboring samples higher than the second threshold. All other samples are regarded as regular EEG.

The optimal thresholds were determined during training for each of the datasets. After applying the thresholds, the output was upsampled to 200 Hz, which is the original sample rate of the input.

3.4. Other detection methods

In this section, the other seven seizure detection methods are explained in more detail. Method 1 by Osorio and Frei [16] is in fact a detection method proposed for human seizures, but it relies on features that are shared with epileptic seizures from rats and has been previously applied on rat data in [18] and [21]. Method 2 by White [17] has been specifically designed for detecting epileptic seizures in rats. To our knowledge, this is the most widely used technique because it can easily be implemented and performs reasonably well. Methods 3

by Van Hese [18] and 4 by Westerhuis [19] have been intentionally designed to detect SWDs in GAERS but the features on which they rely have been used by other seizure detection methods for humans as well as for animal models. Methods 5 by Fanselow [20] and 6 by Van Hese [21] have also been designed to detect SWDs in GAERS but rely on features that are specific for GAERS and will therefore not be compared for detecting the epileptic seizures of the kainate-induced epileptic seizures. Method 7 is a linear detection method, which is also introduced in this study, based on the presented RC detection method.

To allow a fair comparison between the different methods, interval sizes and/or other free parameters were optimized on the same training sets.

3.4.1. Method 1 - Osorio-Frei

In [16] Osorio and Frei presented a detection method for human epileptic seizures. In a first step, the EEG (sampled at 240 Hz) is decomposed using a FIR-filter based approximation of the Daubechies 4 level-3 wavelet, which shows similarity to a 5 - 45 Hz bandpass filter. From this signal the foreground signal FG is determined through median filtering over a moving window given by:

$$FG_k = \text{median}\{y_k^2, y_{k-1}^2, \dots, y_{k-L_1+1}^2\},$$

where L_1 is the time window. This foreground signal is then compared to a background signal BG which is generated by passing the FG_k sequence through a decimation filter, followed by another median filter, and then an exponential forgetting filter. BG_k is given by:

$$\begin{cases} (1 - \lambda)\text{median}\{FG_k, \dots, FG_{k-(L_2-1)s}\} + \lambda BG_{k-1} & \text{if } k = ps, \\ BG_{k-1} & \text{if } p(s-1) \leq k < ps, \end{cases}$$

where $p = 0, 1, 2, \dots, s = L_1/4$, L_2 is the overlap and $\lambda = 0.999807$. In the next step, a dimensionless ratio is computed as follows:

$$R_k = \max_{1 \leq n \leq N} \left\{ \frac{FG_k^{(n)}}{BG_k^{(n)}} \right\},$$

where N denotes the number of channels. If this ratio R_k is above a threshold δ_{OF} , sample k is considered part of a seizure.

To optimize the algorithm for rats, a sample rate of 200Hz was used as opposed to 240Hz, and the following parameters were optimized during training: L_1 , L_2 and δ_{OF} .

3.4.2. Method 2 - White

In [17] White et al. presented 4 epileptic seizure detection methods for detecting seizures in continuous rat EEG. In this study the method with the best results in [17] was used for a comparison. The method consists of two basic building blocks: an autocorrelation metric, R_{Wh} , to exploit the similarity between spikes in a seizure, as well as a spike frequency metric, f_{sp} , to improve specificity.

$$R_{Wh} = \sum_{i=0}^{99} (HV_i - LV_i),$$

with high value and low value i :

$$\begin{aligned} HV_i &= \min\{\max\{S_i\}, \max\{S_{i+1}, S_{i+2}\}\} \\ LV_i &= \max\{\min\{S_i\}, \min\{S_{i+1}, S_{i+2}\}\}, \end{aligned}$$

here S_i is a non-overlapping interval of the EEG-signal with length L .

The spike frequency is determined slightly differently in this study than in [17] because this resulted in better results on both datasets. A spike detection was obtained using a linear regression method with a width of 16 ms and thresholding the slope with a threshold equal to the 90th percentile of a robust exponential fit (using the standardized median [28]) to the histogram of positive slope values.

A detection was made if the number of spikes exceeded $\delta_{Wh,sp}$ and the autocorrelation metric exceeded $\delta_{Wh,R}$ in a non-overlapping interval of length L . For the TC dataset, the autocorrelation metric and the number of spikes were summed over the different channels. During training the following parameters were optimized: L , $\delta_{Wh,R}$ and $\delta_{Wh,sp}$.

3.4.3. Method 3 - energy

In [18] two different methods for detecting epileptic seizures in SWDs were proposed: TSM0 γ and an energy method. The first and best performing method was later optimized to Method 6 in [21]. The method based on the energy of the EEG-signal will be used in this study because it is seizure type non-specific since the energy of the EEG-signal is increased during both seizure types.

The EEG signal is first filtered using a band pass Butterworth filter from 5 to 30Hz. Then the energy for non-overlapping intervals with length L is calculated as follows:

$$E(k) = \frac{1}{L} \sum_{i=0}^{L-1} (x(kL - i))^2,$$

where x denotes the EEG-signal. If $E(k)$ exceeds a threshold δ_E a seizure is detected.

For this method, the interval length L and threshold δ_E were optimized during training. For the TC dataset, the $E(k)$'s of each of the channels were summed.

3.4.4. Method 4 - Westerhuis

In [19] an SWD detection method was proposed, in which the steepness of the EEG signal was taken into account. First the EEG signal is filtered using a band pass Butterworth filter from 5 to 30Hz. Then the steepness is calculated for each sample as the first-order difference $d(i)$. In the following step, the

maximum steepness is determined for each non-overlapping interval of length L as:

$$D(k) = \max_{i=0, \dots, L-1} \{d(kL - i)\}.$$

If this $D(k)$ exceeds a threshold $\delta_{W_{est}}$ a detection is made.

Again L and $\delta_{W_{est}}$ were optimized during training and the $D(k)$'s of each channel of the TC dataset were summed together.

3.4.5. Method 5 - Fanselow

The SWD detection method proposed in [20] uses the maximum absolute amplitude of the EEG-signal as a main feature. In a first step the EEG signal is filtered using a band pass Butterworth filter from 5 to 30Hz. Then the amplitude of the signal is evaluated in overlapping windows of length L_1 as:

$$A(k) = \max_{i=0, \dots, L_1-1} \{x(kL_2 - i)\},$$

where L_2 is the step size which results in an overlap of length $L_1 - L_2$. Once again if $A(k)$ is above a threshold δ_F a detection is made.

For this detection method, the parameters L_1 , L_2 and δ_F were optimized during training.

3.4.6. Method 6 - Van Hese

In [21] an SWD detection method for GAERS was presented that exploits the fact that SWDs are quasi-periodic signals. They have a fundamental frequency from 7 to 12 Hz and several harmonics (see Figure 1). In a first step, the Short Term Fourier Transform (STFT) is applied to non-overlapping and zero-padded intervals, which results in a spectrogram. In a second step, the background spectrum is determined for each of the frequency components as proposed in [28]. In a following step, called the harmonic analysis, spectral peaks, higher than the background spectrum, are used to determine the fundamental frequency and the harmonics at approximately 2, 3, or more times that frequency. If these signals contain an energy above a certain threshold, the interval is considered part of an SWD. For a more detailed explanation we refer to [21], from which the implementation was also used in this study. During training only the threshold was optimized.

3.4.7. Method 7 - linear seizure detection

In this study we also propose a linear seizure detection method. For this technique, only the reservoir is omitted from the method presented in Section 3.3 without changing the pre-processing and thresholding method. This means that the features, the interval length L (see Section 3.3.1) and the thresholds (see Section 3.3.3) are selected during training using the linear classifier. Note that this method can be applied to evaluate the advantage of using RC as a classifier.

Table 1: The average Sensitivity, Specificity, BER and standard deviation on the complete SWD and TC datasets.

SWD method	Sensitivity (std)	Specificity (std)	BER (std)
RC	0.965 (0.005)	0.933 (0.009)	0.050 (0.004)
RC (best)	0.964 (0.020)	0.962 (0.023)	0.037(0.018)
1 Osorio-Frei	0.845 (0.111)	0.893 (0.056)	0.130 (0.054)
2 White	0.963 (0.033)	0.887 (0.078)	0.075 (0.037)
3 energy	0.754 (0.332)	0.905 (0.105)	0.171 (0.154)
4 Westerhuis	0.756 (0.326)	0.884 (0.133)	0.180 (0.147)
5 Faselow	0.759 (0.337)	0.873 (0.136)	0.184 (0.150)
6 Van Hese	0.961 (0.046)	0.959 (0.071)	0.040 (0.045)
7 linear	0.973 (0.023)	0.887 (0.069)	0.070 (0.037)
TC method	Sensitivity (std)	Specificity (std)	BER (std)
RC	0.921 (0.006)	0.981 (0.012)	0.049 (0.008)
RC (best)	0.939 (0.041)	0.991 (0.009)	0.035(0.024)
1 Osorio-Frei	0.791 (0.121)	0.888 (0.048)	0.160 (0.075)
2 White	0.806 (0.135)	0.759 (0.124)	0.218 (0.095)
3 energy	0.687 (0.283)	0.667 (0.245)	0.323 (0.106)
4 Westerhuis	0.653 (0.293)	0.757 (0.181)	0.295 (0.105)
7 linear	0.813 (0.148)	0.940 (0.051)	0.123 (0.078)

4. Results

Even though most methods are able to detect epileptic seizures of both types, namely SWDs and TC seizures, they do not necessarily have the same optimal parameters in both cases. Therefore the two datasets were considered independently and the different methods were evaluated using the balanced error rate.

4.1. Overall performance

In order to evaluate the overall detection performance each method was trained on the complete training set and tested on the complete test set. Table 1 presents the results for the different methods. For ‘RC (best)’ and the other methods the average and std are computed over the different rats in the dataset. Each rat thus has the same influence on the results, independent of the amount of data that was recorded for this rat. Out of 10 trained reservoirs, ‘RC (best)’ is the reservoir that performed best on the training set. For the results of RC, the average and the standard deviation (std) are computed over 10 reservoirs with a different random weight initialization.

4.2. Generalisation properties

To evaluate how each detection method performs on data from unseen rats, only the data from the training sets was used. In Table 2 the results are pre-

Table 2: Performance on new rats. This table contains the 10-fold cross-validated average BER and standard deviation, trained on data from respectively 21 and 14 rats of the SWD and TC training sets and validated on the data from the other rats in that set.

SWD method	val. BER (std)	train BER (std)
RC	0.034 (0.006)	0.026 (0.006)
RC (best)	0.032(0.008)	0.018 (0.003)
1 Osorio-Frei	0.109 (0.058)	0.101 (0.006)
2 White	0.071 (0.049)	0.055 (0.003)
3 energy	0.146 (0.101)	0.141 (0.011)
4 Westerhuis	0.164 (0.106)	0.155 (0.012)
5 Fanselow	0.162 (0.137)	0.153 (0.015)
6 Van Hese	0.045 (0.038)	0.035 (0.003)
7 linear	0.093 (0.011)	0.052 (0.004)
TC method	val. BER (std)	train BER (std)
RC	0.085 (0.029)	0.024 (0.006)
RC (best)	0.058(0.044)	0.017 (0.007)
1 Osorio-Frei	0.150 (0.059)	0.147 (0.005)
2 White	0.182 (0.059)	0.176 (0.005)
3 energy	0.302 (0.091)	0.289 (0.007)
4 Westerhuis	0.257 (0.058)	0.256 (0.004)
7 linear	0.090 (0.010)	0.086 (0.001)

sented for each of the methods on the SWD and tonic-clonic dataset. Each method is trained on the train data of respectively 21 and 14 rats from the training set. The BER is evaluated on a validation set which contains the rats that were not included in the training set, respectively 2 and 1 rats. To avoid statistical anomalies the results were 10-fold cross-validated, so that 10 different training and validation sets were created. We also made sure that each animal was at most once part of the validation set. The std is computed over the different folds for RC (best), i.e. the reservoir that performed best on the training set, as well as for the other methods. In the case of RC, 10 reservoirs were used for each training and validation set to average out the randomness in creating the reservoir weight matrices. The standard deviation (std) for the results of RC were computed as the average std of each fold.

5. Discussion

From Table 1 we can deduce that RC performs very well on both datasets, with an average BER of 0.037 and 0.035 on the SWD and TC dataset respectively for the reservoir that performed best on the training set. This is a significant improvement over the linear method without RC and outperforms all other techniques. Even though the reservoir with the lowest train error is

not necessarily the one that performs best on the test set, we noticed that in practice this assumption holds for the technique presented.

Method 6, the SWD detection method that is presented by Van Hese in [21] and is currently considered state-of-the-art, results in an average BER of 0.040 with a standard deviation of 0.045. The presented technique results in a BER of 0.037 with standard deviation of only 0.018. The significantly higher standard deviation of method 6 is caused by artifacts and glitches in the EEG and some irregular and less periodic SWDs in dataset B (see Section 2.1). The presented method however, shows to be more robust. We can therefore conclude that the presented technique significantly outperforms the method in [21] and is even capable of performing very well on both the SWD and TC datasets.

Table 2 reveals that most methods can generalize very well to data from rats that were not represented in the training set. The table also shows that RC performs best on both datasets. The small training error of RC indicates that the presented method in this study is very good at learning the training set. The significantly larger deterioration in performance on the tonic-clonic validation set as well as the high standard deviation for ‘RC (best)’ indicate that there is a high dependence on the examples in the training set for the presented technique.

Even though method 2 by White et al. [17] is widely used by researchers to detect kainate-induced tonic-clonic seizures it is outperformed by the simpler and faster method 1 developed by Osorio and Frei. This is probably due to the noisy nature of the tonic-clonic EEG and not due to the alterations we made to the algorithm which actually resulted in a slight improvement of the results.

Methods 3 to 5 from Section 3.4 achieve the worst results on both performance tests. The high standard deviation is mainly because these simple methods are highly dependent on the amplitude of the signal. Due to electronic shortcomings in the EEG recording system, these amplitudes can vary considerably. The linear method that was also presented in this study is in fact a more sophisticated version of these methods because it solves the amplitude dependence and adds more features.

One could argue that method 1 should be fitted with a filter adapted for the specific types of seizures used in this research. This was proposed by Haas et al. in [34] where patient-specific filters significantly improved detection results of the original Osorio-Frei method presented in [16]. We did not implement the extension proposed in [34] because the linear method discussed in this paper can also be considered as an extension to method 1. The linear method uses even multiple features instead of one. As a result of this adaptation, it performed significantly better than method 1 but the results are not as good as the performance of RC. However, the Osorio-Frei method is designed for online detection of epileptic seizures, which is not possible in the current RC set-up. Therefore, it is the subject of ongoing research.

The feature set that was selected for the SWD datasets consisted of only the level 3 Daubechies 4 wavelet filter. For the tonic-clonic dataset, only the energy in the beta band was selected as feature set. For the linear method, the algorithm selected more than one feature, which resulted in a slight improvement over selecting only one feature. This might indicate that RC does not pro-

cess multiple signals well or that more than one feature results in overfitting. However, further research on this matter is still needed.

If the presented method were to be used in a clinical study, the following simple guidelines need to be followed. The training set should contain examples of the different types of seizures that are to be expected in the dataset. In practice we noticed that the following number of examples was sufficient for a dataset of approximately 10 rats: 5 examples of ictal and inter-ictal EEG from each rat, spread in time over the complete dataset, with a length of about 4 minutes. To find the best performing reservoir we noticed that at least 1 in 5 reservoirs had a performance on the training and test set comparable to the performance of ‘RC (best)’, i.e. the reservoir that performed best on the training set. The reservoir parameters even showed to be independent of the type of seizures. Only the output weights and the following four parameters needed separate optimization: the averaging interval L , the feature set (see Section 3.3.1) and the optimal thresholds (see Section 3.3.3).

To train the system and process the 1400 hours of EEG it takes a 2.6 GHz Core 2 Quad machine (with 8 GB RAM) 13 hours of running slightly optimized Matlab code. A well trained experienced encephalographer is at least 3 times slower, being able to process 4 hours of EEG from 8 rats simultaneously in one hour.

The presented technique is thus very accurate, easy to train and able to process the data within a reasonable time limit. In our opinion this makes the system qualified to serve as a clinical tool to mark EEG data.

6. Conclusion

We presented a novel method for marking epileptic seizures in the EEG of rats and compared it with existing techniques. The performance was evaluated and compared to several detection methods from literature using a large database. It consisted of 454h of data, including absence seizures, from 23 GAERS rats and 852h of data, containing tonic-clonic seizures, from 15 kainate-induced epilepsy rats. The presented method outperforms all other tested seizure detection methods on the absence seizure data as well as on the tonic-clonic seizure data. A balanced error rate of respectively 3.7% and 3.5% was achieved. This resulted in a sensitivity of 96% and 94%, respectively and a specificity of 96% and 99%. The presented method thus achieves excellent performance on both datasets.

Even though this method contains more parameters than the other methods considered here, only a few of them need to be retrained to perform seizure detection on a different type of seizures. To recreate the results presented, only about 5 seizure examples together with some inter-ictal EEG are needed to optimize the parameters and training.

Because the discussed method is a learning method, we have shown that it is capable of dealing with artifacts and glitches in the EEG, and with differences between seizures and different seizure types. It is also reasonable to assume

that it is applicable to detect the seizures of other animal models, with similar seizures, if a relevant training set is chosen. However, at this time we do not have access to such data to verify this hypothesis.

Acknowledgements

The authors thank Liesbeth Waterschoot and Annelies Van Dycke for providing marked EEG-data.

This work was partially funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) and the BOF-GOA Project Home-MATE funded by the Ghent university Special Research Fund.

References

- [1] J.F. Annegers. *The treatment of epilepsy: principle and practice*, chapter "The epidemiology in epilepsy", pages 165–172. Baltimore: Williams and Wilkins, 1996.
- [2] WA. Hauser. *Epilepsy, a comprehensive textbook*, volume 1, chapter Prevalence and incidence, pages 47–58. Lippincot-Raven, Philadelphia, 1998.
- [3] S. Dedeurwaerdere. *Neuromodulation in experimental animal models of epilepsy*. PhD thesis, Ghent University, 2005.
- [4] C. Marescaux, Vergnes M., and Depaulis A. Genetic Absence Epilepsy in Rats from Strasbourg - A Review. *Journal of Neural Transmission*, 35:37–69, 1992.
- [5] L. Danober, C. Deransart, A. Depaulis, M. Vergnes, and Marescaux C. Pathophysiological mechanisms of genetic absence epilepsy in the rat. *Progress in Neurobiology*, 55:27–57, 1998.
- [6] J.L. Hellier, P.R. Patrylo, P.S. Buckmaster, and F.E. Dudek. Recurrent spontaneous motor seizures after repeated low-dose systemic treatment with kainate: assessment of a rat model of temporal lobe epilepsy. *Epilepsy Research*, 31:73–84, 1998.
- [7] Y. Ben-Ari and R. Cossart. Kainate, a double agent that generates seizures: two decades of progress. *Trends in Neuroscience*, 23:515–528, 2000.
- [8] L. Waterschoot, S. Dedeurwaerdere, and T. Wyckhuys. Stimulation of the anterodorsal thalamus in genetic absence epilepsy rats from Strasbourg (GAERS). In *Epilepsia*, editor, *7th European Congress on Epileptology*, volume 47, page 70, Helsinki, Finland, 2006. Blackwell Publishing.
- [9] T. Wyckhuys, P. Boon, R. Raedt, B. Van Niewenhuyse, K. Vonck, and W. Wadman. Suppression of hippocampal epileptic seizures in the kainate rat by poisson distributed stimulation. *Epilepsia*, 51:2297–2304, 2010.

- [10] P. Buteneers, B. Schrauwen, D. Verstraeten, and D. Stroobandt. Real-time Epileptic Seizure Detection on Intra-cranial Rat Data using Reservoir Computing. In *Advances in Neuro-Information Processing*, volume 5506/2009, pages 56–63, 2009.
- [11] David Verstraeten, Benjamin Schrauwen, Michiel D’Haene, and Dirk Stroobandt. A unifying comparison of reservoir computing methods. *Neural Networks*, 20:391–403, 2007.
- [12] 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.
- [13] T.M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE transactions on electronic computers*, 14(3):326–334, 1965.
- [14] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.
- [15] Y. Saeys, I. Inza, and P. Larrañaga. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, 2007.
- [16] I. Osorio, M. Frei, and S. Wilkinson. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. *Epilepsia*, 39:615–627, 1998.
- [17] A. White, P. Williams, D. Ferraro, S. Clark, S. Kadam, and F. Dudek et al. Efficient unsupervised algorithms for the detection of seizures in continuous EEG recordings from rats after brain injury. *Journal of Neuroscience Methods*, 152:255–266, 2006.
- [18] 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. *Physics in Medicine and Biology*, 48(12):1685–1700, 2003.
- [19] F. Westerhuis, W. Van Schaijk, and G. Van Lujtelaar. 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, 1996. Utrecht University.
- [20] 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.
- [21] P. Van Hese, J-P. Martens, L. Waterschoot, P. Boon, and I. Lemahieu. Automatic Detection of Spike and Wave Discharges in the EEG of Genetic Absence Epilepsy Rats from Strasbourg. *IEEE Transactions on Biomedical Engineering*, 56(3):706–717, 2009.
- [22] D. Verstraeten, B. Schrauwen, D. Stroobandt, and J. Van Campenhout. Isolated word recognition with the liquid state machine: a case study. *Information Processing Letters*, 95(6):521–528, 2005.

- [23] S. Baraban, editor. *Animal Models of Epilepsy: Methods and Innovations*, volume 40 of *Neuromethods*. Humana Press, New York, 2009.
- [24] G. Paxinos and C. Watson. *The rat brain in stereotaxic coordinates. Fourth Edition*. Academic Press, San Diego, 1998.
- [25] C. Metz. Basic principles of ROC analysis. *Seminars in nuclear medicine*, 8:283–298, 1978.
- [26] N. Päävinen, S. Lammi, A. Pitkänen, J. Nissinen, M. Penttonen, and T. Grönfors. Epileptic seizure detection: A nonlinear viewpoint. *Computer Methods and Programs in Biomedicine*, 79:151–159, 2005.
- [27] R.P. Costa, P. Oliveira, G. Rodrigues, B. Leitão, and A. Dourado. Epileptic seizure classification using neural networks with 14 features. In I. Lovrek, R. Howlett, and L. Jain, editors, *Knowledge-Based Intelligent Information and Engineering Systems. 12th International Conference, KES 2008*, pages 281–288, Zagreb, Croatia, 2008. University of Zagreb, Springer.
- [28] V. Stahl, A. Fisher, and R. Bippus. Quantile based noise estimation for spectral subtraction and Wiener filtering. In *International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Istanbul, Turkey, 2000. IEEE.
- [29] H. Jaeger, M. Lukosevicius, and D. Popovici. Optimization and applications of echo state networks with leaky integrator neurons. *Neural Networks*, 20:335–352, 2007.
- [30] A.N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet math Dokl*, 4:1035–1038, 1963.
- [31] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification - Second Edition*, chapter Linear Discriminant Functions, pages 242–243. John Wiley and Sons, Inc., New York, 2001.
- [32] R. A. Legenstein and W. Maass. Edge of chaos and prediction of computational performance for neural microcircuit models. *Neural Networks*, pages 323–333, 2007.
- [33] 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.
- [34] S. Haas, M. Frei, and Oso. Strategies for adapting automated seizure detection algorithms. *Medical Engineering & Physics*, 29:895–909, 2007.