www.elsevier.com/locate/cbm

Exploring temporal information in neonatal seizures using a dynamic time warping based SVM kernel

Rehan Ahmed, Andriy Temko, William P. Marnane, Geraldine Boylan, Gordon Lightbody

PII: S0010-4825(17)30021-5
DOI: http://dx.doi.org/10.1016/j.compbiomed.2017.01.017
Reference: CBM2589

To appear in: *Computers in Biology and Medicine*

Received date: 1 July 2016
Revised date: 24 January 2017
Accepted date: 25 January 2017

Cite this article as: Rehan Ahmed, Andriy Temko, William P. Marnane, Geraldine Boylan and Gordon Lightbody, Exploring temporal information i neonatal seizures using a dynamic time warping based SVM kernel, *Computer in Biology and Medicine,* http://dx.doi.org/10.1016/j.compbiomed.2017.01.017

EXPLORING TEMPORAL INFORMATION IN NEONATAL SEIZURES USING A DYNAMIC TIME WARPING BASED SVM KERNEL

Rehan Ahmed[1,2*], Andriy Temko[1,2], William P. Marnane[1,2], Geraldine Boylan[1,3] and Gordon Lightbody[1,2]

[1] Irish Centre for Fetal and Neonatal Translational Research (INFANT)

[2] Department of Electrical and Electronic Engineering, University College Cork, Ireland

[3] Department of Pediatrics and Child Health, University College Cork, Ireland

[*]**Correspondence:** Rehan Ahmed: **Tel**: 00353 21 490 3156. rehan@eleceng.ucc.ie

Abstract:

Seizure events in newborns change in frequency, morphology, and propagation. This contextual information is explored at the classifier level in the proposed patient-independent neonatal seizure detection system. The system is based on the combination of a static and a sequential SVM classifier. A Gaussian dynamic time warping based kernel is used in the sequential classifier. The system is validated on a large dataset of EEG recordings from 17 neonates. The obtained results show an increase in the detection rate at very low false detections per hour, particularly achieving a 12% improvement in the detection of short seizure events over the static RBF kernel based system.

Keywords:

Automated neonatal seizure detection, sequential classifier, fusion, Gaussian dynamic time warping

1. INTRODUCTION

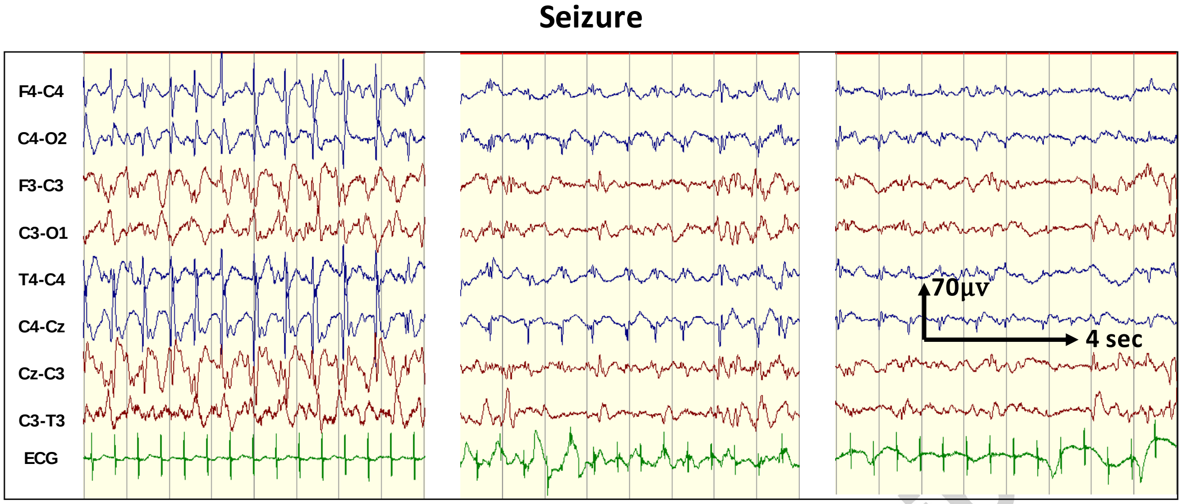
Neonatal seizures are reported to happen in 3 out of 1000 full term neonates (Lanska et al., 1995) and this figure grows to 58 - 132 per thousand in pre-term infants (Watkins et al., 1988). Failure to detect seizures and the resulting lack of treatment may result in brain damage and in severe cases, death (AL-Naddawi et al., 2011; Lynch et al., 2012). Electroencephalography (EEG) is considered to be the gold standard for the detection of neonatal seizures (Rennie et al., 2004). However, interpreting neonatal EEG requires expertise which is not available around the clock in a typical busy Neonatal Intensive Care Unit (NICU). Therefore, automatic computer-based detection of neonatal seizures could be of great help for the medical staff and has been the focus of research attention for many years.

A typical neonatal seizure detection system comprises of the following main stages: i) The signal representation stage (feature-level) is the process where relevant features are extracted from the EEG signal. ii) The classification stage (classifier level), where the extracted feature or feature vectors are assigned to the seizure or non-seizure class using a set of rules and thresholds which are either automatically derived from the data (classifier) or manually selected following the reasoning of expert neurophysiologists. iii) The post-processing stage (decision level) which involves smoothing or other transformations to offer clinicians support in decision making.

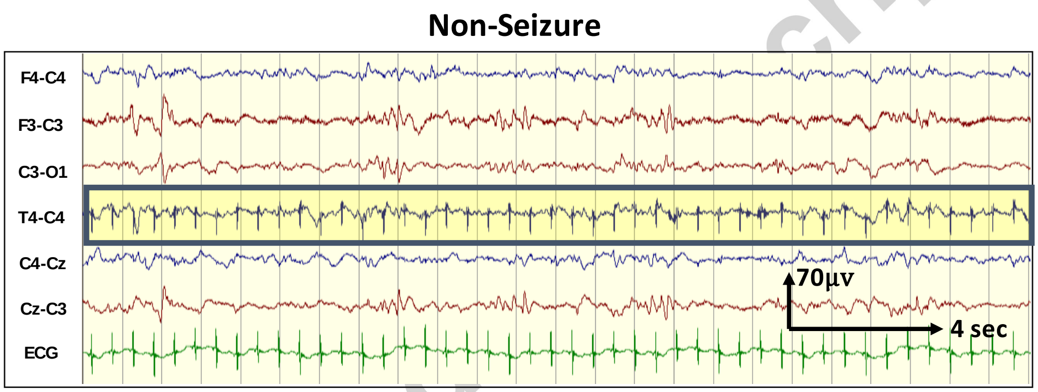
In contrast to background EEG and artifacts, seizure events in newborns change in frequency, morphology, and propagation. A comparison of seizure, normal background EEG and an Electrocardiogram (ECG) artifact is shown in Figure 1. Non-seizure background EEG lacks any structure and ECG artifact does not evolve in time (Figure 1d) whereas the attenuating amplitudes and the evolving frequency of the spikes can be clearly seen in the seizure example (Figure 1a-c). This contextual information (also referred as the temporal evolution, the signal dynamics or the sequentiality) of the neonatal seizures, has been explored in various ways in most of the automated detection systems reported to date (Aarabi et al., 2009; Ahmed et al., 2012; Bogaarts et al., 2014; Chaovalitwongse and Pardalos, 2008; Deburchgraeve et al., 2008; Nagaraj et al., 2014; Navakatikyan et al., 2006; Ocak, 2009; Shoeb and Guttag, 2010; Temko et al., 2011a; Thomas et al., 2011; Wong et al., 2007).

At the feature-level, the temporal evolution of the EEG signal within the window has been captured using template extraction and matching (Aarabi et al., 2009), wavelet transformation (Ocak, 2009) or EEG wave sequence analysis (Navakatikyan et al., 2006). Concatenation of the consecutive feature vectors in time was proposed in (Shoeb and Guttag, 2010). The Kalman filter was exploited to enhance the contrast of the extracted features to the past background activity (Bogaarts et al., 2014).

Figure 1: Comparison of a neonatal seizure, artifact and normal background EEG. (a-c) Evolution of a single seizure event from start to end. (a) Slow wave activity at start with the presence of sharp/spike components and high amplitude of the EEG. (b) As the seizure progresses the EEG becomes lower in amplitude. (c) only very low amplitude discharges are seen. (d) An example of normal background EEG. ECG artifact is shown in channel T4-C4 (highlighted). Both normal background EEG and EEG corrupted with the ECG artifact lacks any structure.



(a) Start (b) Middle (c) End



(d)

At the decision-level, the temporal structure of the neonatal seizure can be incorporated through the use of smoothing filters, applied over the classifier output, such as the moving average (Temko et al., 2011a), median filter (Nagaraj et al., 2014) or Kalman filter (Bogaarts et al., 2014).

The classifier-level techniques typically find the temporal and contextual matching between dynamic length sequences of feature vectors. Examples include support vector machines with sequential kernels (Ahmed et al., 2012; Chaovalitwongse and Pardalos, 2008) or the Hidden Markov Model (HMM) (Wong et al., 2007).

Incorporation of contextual information at each level has its advantages and drawbacks. The feature-level methods typically consider temporal information in the EEG on a short-time scale. The filtering methods introduced at the decision level perform smoothing over a large window and hence the short seizures may be suppressed and are frequently missed as a result (Temko et al., 2011a). The exploration of contextual information in the EEG on the classifier-level has been relatively scarce in the area of neonatal seizure detection, whereas such classifier level techniques have shown promising results in other areas of signal processing such as speech recognition (Shimodaira et al., 2002; Smith and Gales, 2002), handwriting character recognition (Bahlmann et al., 2002), and acoustic events classification (Temko et al., 2006).

The state-of-the-art neonatal seizure detection system was previously developed in our research group. It was reported that this system more frequently missed seizures of length less than one minute (Temko et al., 2011a). In our previous pilot study, (Ahmed et al., 2012), it was shown that a Gaussian Dynamic Time Warping (GDTW) kernel based system could improve seizure detection performance. However, these results were based on the EEG recording from only one patient. In this paper, we further extend the work carried out in the pilot study and report the results on a dataset of 17 neonates. This will better indicate the strengths and weaknesses of the two classifiers. Moreover, we also investigate whether incorporating a classifier that can implicitly use the contextual information and classify sequences of short-term feature vectors, could improve the detection rate of short seizures while keeping down the number of false detections. A novel system based on the fusion of the static classifier and the dynamic time warping based sequential classifier is presented here and validated on a dataset of 17 neonates.

The paper is organized as follows. The dynamic time warping technique and its usage in the Support Vector Machine (SVM) is explained first. Section 3 describes the particulars of the developed fusion based neonatal seizure detection system. Performance of the individual SVM with the Radial Basis Function (RBF) kernel and SVM with the GDTW kernel is compared with the proposed fusion based system in section 4, followed by conclusions and future work.

2. METHODS

2.1 DATASET

EEG data from 17 neonates was used in this study. These neonates were full term with the gestational age ranging from 39-42 weeks. The data was collected in the NICU of Cork University Maternity Hospital. A written consent from the parents was obtained for the use of the data for the research purposes. The multichannel EEG was recorded using a Carefusion NicOne video EEG monitor with a sampling rate of 256Hz. The 10-20 system of electrodes placement, modified for neonates, was used. Eight bipolar EEG channels were then derived (F4-

C4, C4-O2, F3-C3, C3-O1, T4-C4, C4-Cz, Cz-C3, C3-T3). The seizures were annotated based on the consensus of two expert neurophysiologists using EEG and simultaneous video recordings. These seizures were a secondary injury due to Hypoxic-Ischemic Encephalopathy (HIE), however neonates were not cooled for the HIE treatment. Table 1 shows the details of each recording of the dataset. There was a total of 261 hours of EEG data with a mean EEG recording time of 15 hours per patient. A total of 821 seizure events were present in this dataset. The recordings were not edited and no artifacts had been manually removed. This dataset is a true representation of the real-time situation in hospitals where EEG is recorded and monitored for several hours and as such it includes the unexpected adverse events that may affect the quality of EEG.

There were two types of annotations performed on this dataset; i) global annotations where the seizures were marked in time without their channel information. ii) Per-channel annotations where seizures were marked with the channel information. In this work, per channel annotations were used to train the classifiers. This allows the selection of the best representation of seizure events that can then be used as training data. Approximately 20 min (depending on the number and duration of seizure events) of per channel annotations of seizure events for every neonate were used for training.

2.2 Dynamic Time Warping

Dynamic Time Warping (DTW) is a technique used to measure the similarity between two variable length sequences (Muller, 2007). Consider two sequences $R = \{r_1, \dots, r_{N_r}\}$ and $S = \{s_1, \dots, s_{N_s}\}$ with lengths N_r and N_s, where s_i and r_i could each be a time series data point or feature vector of T dimensions in the sequences. A local distance measure, for example the Euclidean distance, between each element of the two sequences can be calculated to provide a distance matrix of size $N_r \times N_s$. Consider a warp path $W = \{w_1, \dots, w_k, \dots w_K\} \in \mathbb{W}$ of length K through the distance matrix where $w_k = (n_k, m_k)$, $n_k, m_k \in \mathbb{N}$ represents the k^{th} vertex of this path. This path is constructed under the monotonic constraints $m_{k+1} \geq m_k$ and $n_{k+1} \geq n_k$ to preserve both shape and continuity of the path (Muller, 2007). The cost D_W of a particular path $W \in \mathbb{W}$ is

$$D_W(R,S) = \frac{1}{K} \sum_{k=1}^{K} d_l(n_k, m_k), \quad (1)$$

where

$$d_l(n_k, m_k) = \left\| s_{n_k} - r_{m_k} \right\|^2, \quad (2)$$

Table 1: EEG Dataset

Patient ID	Record length (h)	Seizure events	Seizure length		
			Mean	Min	Max
1	24.1	17	1’ 40"	28"	4’ 4"
2	24.7	3	6’ 19"	1’ 4"	11’ 19"
3	22.7	170	2’ 18"	24"	10’ 55"
4	26.1	65	1’ 30"	40"	3’ 28"
5	24	51	6’ 32"	28"	31’ 11"
6	5.7	44	1’ 15"	31"	2’ 4"
7	13.2	71	1’ 58"	28"	10’ 31"
8	24.5	17	6’ 6"	40"	19’ 23"
9	24	157	5’ 27"	28"	37’ 16"
10	10.4	28	6’ 0"	24"	34’ 55"
11	6.2	14	5’ 40"	55"	7’ 47"
12	12	37	2’ 31"	31"	10’ 16"
13	12.1	25	4’ 29"	1’ 19"	12’ 48"
14	5.5	13	8’ 58"	1’ 55"	39’ 12"
15	12.2	58	2’ 17"	19"	7’ 19"
16	7.6	31	10’ 34"	2’ 23"	34’ 47"
17	6.6	20	5’ 36"	36"	23’ 23"
Total	261.7	821			

is the Euclidean norm between the feature vectors s_{n_k} and r_{m_k}. In order to find the optimal alignment path that gives the shortest distance $D_{\emptyset} = min\{D_W(S,R)\}$ in the gram matrix from point $\{s_1, r_1\}$ to $\{s_{N_s}, r_{N_r}\}$, an accumulated cost matrix D_a is calculated. Each element of this matrix is the shortest distance from the origin to that element. For example to compute the shortest accumulated distance for an element at position (i,j) in matrix D_a , the algorithm seeks the minimum value using $\min\ \{\ (d_{a(i-1,j-1)} + d_{l(i,j)})\ ,\ (d_{a(i-1,j)} + d_{l(i,j)}),\ (d_{a(i,j-1)} + d_{l(i,j)})\ \}$, where $d_{a(:,:)}$ is the accumulated distance of the previous neighboring nodes. A dynamic programming algorithm as defined in (Sakoe and Chiba, 1978) is then used to find the path that gives the shortest DTW distance in the matrix D_a. Further explanation of this method with an intuitive figure can be explored in (Ahmed, 2015).

Figure 2: An illustration of the dynamic time warping technique. Here sequences of 15 feature vectors (1 minute) of two different seizures and a non-seizure are compared. The red line indicates the warp path in the accumulated cost matrix. (a) seizure to seizure (b) seizure to non-seizure (c) A detailed view of warping process between two seizure sequences. The colored parts of each EEG signal show the matched parts of the other sequence.

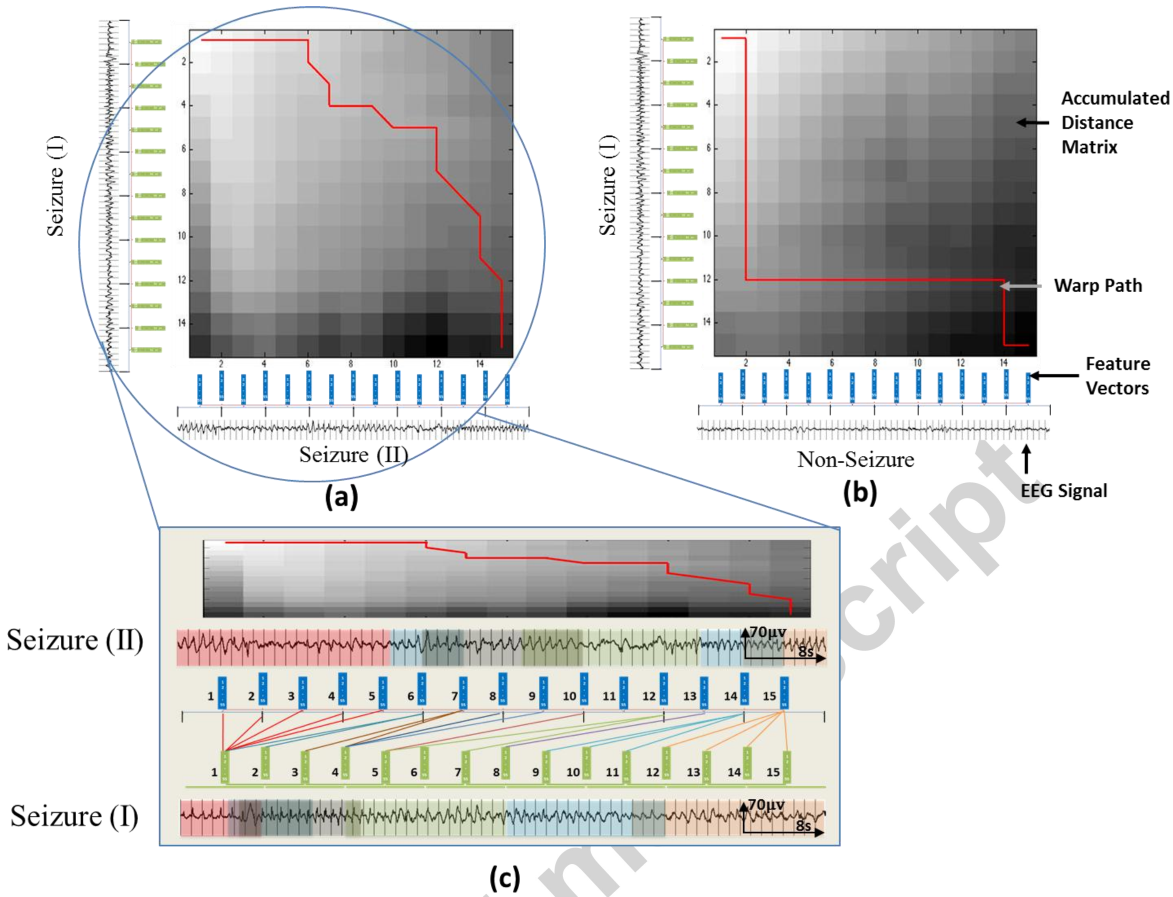


Figure 2 shows an illustration of the DTW method where a seizure sequence is compared to both another seizure and then a non-seizure sequence. It can be seen, that when a seizure sequence is compared to a non-seizure sequence, the warp path becomes longer and moves away from the diagonal. The warp path is near the diagonal and hence shorter when a seizure is compared to another seizure sequence. Figure 2c shows a detailed view of the sequence matching process using DTW. It can be seen, that the seizure sequence (II) became attenuated after the first 8 seconds whereas the seizure sequence (I) kept its repetitiveness and because of this mismatch the warp path was horizontal (far from the diagonal) for the first 6 epochs.

2.3 Gaussian DTW kernel in SVM

SVM has shown state of the art performance in many pattern recognition areas (Depeursinge et al., 2010; Fauve et al., 2007; Joachims, 1998). SVM is a binary classifier and it uses a kernel function to map the data to a higher dimensional space where the separation of the data is easier. Kernels could be thought of as a measure of similarity between input data points. Many kernel functions have been defined in the past, such as RBF, polynomial and sigmoid kernel functions, each targeting different types of data. Most of these kernel functions are restricted to the comparison of only one data-point at a time. However, as many machine learning problems need to explore the temporal and sequential information, a different breed of SVM kernel was developed to measure the similarity between variable length sequences of data points (Bahlmann et al., 2002; Campbell et al., 2006; Lodhi et al., 2002; Shimodaira et al., 2002). The DTW based kernel is explored here for the seizure detection problem.

The classical SVM classifier uses a hyperplane to separate the input data. Consider a two-class problem, with a pre-labelled training set $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ where $y_i \in \{-1, +1\}$ are the labels and $\mathbf{x}_i \in \Re^T$ are the feature vectors. In SVM classification, a test vector $\mathbf{x}$ is assigned a class by evaluating,

$$f(\mathbf{x}) = sign\left(\sum_{i \in I_{SV}} \alpha_i y_i \mathbb{K}(\mathbf{x}, \tilde{\mathbf{x}}_i) + b\right) \tag{3}$$
$$= sign\left(\theta_{svm}(\mathbf{x})\right),$$

where α_i are Lagrange multipliers, b is the bias, I_{SV} is the set of support vectors retained after training the SVM and $\tilde{\mathbf{x}}_i$ is the i^{th} support vector in this set. θ_{svm} is the RBF-SVM distance. Further details on SVM training can be found in (Burges, 1998). $\mathbb{K}$ is the kernel function of the SVM that maps the input data onto a higher dimensional feature space. A commonly used kernel function is the Gaussian RBF kernel defined as,

$$\mathbb{K}(\mathbf{x}, \tilde{\mathbf{x}}_i) = \exp\left(-\frac{\|\mathbf{x}-\tilde{\mathbf{x}}_i\|^2}{2\sigma^2}\right). \tag{4}$$

This kernel can be adapted to now represent the similarity of two sequences by replacing the Euclidean distance in Eq. 4 by the DTW distance $D_\emptyset$, (Bahlmann et al., 2002). This yields the following Gaussian Dynamic Time Warping kernel based classifier (GDTW-SVM) for a test sequence S,

$$f(S) = sign\left(\sum_{i \in I_{SV}} \alpha_i y_i \exp\left(-\frac{D_\emptyset(S, \tilde{R}_i)}{2\sigma^2}\right) + b\right) \tag{5}$$
$$= sign\left(\theta_{GDTW}(\mathrm{S})\right).$$

Here $\tilde{R}_i$ is the i^{th} retained support sequence and θ_{GDTW} is the GDTW-SVM distance. In this manner, the SVM with a GDTW kernel will be able to classify variable length sequences according to their DTW distances and the support vector concept is now replaced with support sequences. It should be noted that the GDTW kernel is not a positive semi definite (PSD) kernel. However, it has shown excellent performance in contextual classification problems across a wide range of pattern recognition areas (Chaovalitwongse and Pardalos, 2008; Temko et al., 2006). We will further discuss this property of GDTW kernel in detail in section 4 and will show that the use of this kernel for seizure detection produces stable results.

3 NEONATAL SEIZURE DETECTION SYSTEM

3.1 PRE-PROCESSING AND FEATURE EXTRACTION

An overview of the complete neonatal seizure detection system is shown in Figure 3. The raw EEG is first down-sampled from 256Hz to 32Hz with an anti-aliasing filter set at 12.8 Hz. Filtered EEG in each channel is then segmented into 8s epochs with a 50% overlap using a sliding window. Fifty-five different features are then extracted from each epoch. These features, as outlined in Table 2, are derived from the frequency, time and information theory domains. The usability of these features has been validated in a number of previous studies on neonatal seizure detection (Gotman et al., 1997; Greene et al., 2008; Temko et al., 2011a; Thomas et al., 2010), adult seizure detection (Faul et al., 2009), grading background EEG (Ahmed et al., 2016) and neurological outcome prediction (Doyle et al., 2010). This set of simple features provides a detailed picture of the signal statistics required for the neonatal seizure detection problem. The extracted feature vectors are then fed to the classification stage.

3.2CLASSIFICATION

Figure 3: An overview of the complete proposed neonatal seizure detection system.

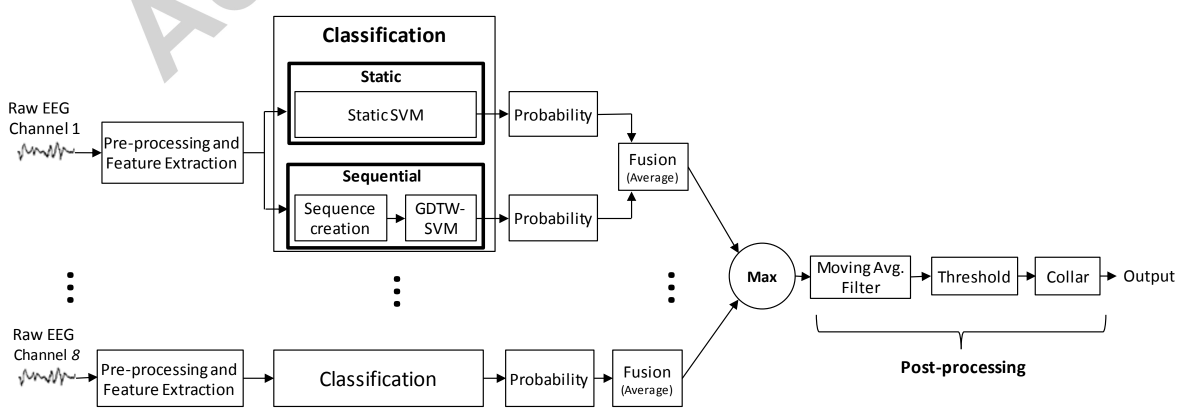


Table 2:Features extracted from the 8 seconds EEG epoch

Domains	Feature List
Frequency	• Total power (0-12Hz) • Peak frequency of spectrum • Spectral edge frequency (80%, 90%, 95%) • Power in 2Hz width sub-bands (0-2Hz, 1-3Hz, ...10-12Hz) • Normalized power in sub-bands • Wavelet energy
Time	• Non-linear line length • Number of maxima and minima • Root mean squared amplitude • Hjorth parameters • Zero crossings (raw epoch, Δ and ΔΔ*) • Autoregressive modeling error (model order 1-9) • Skewness • Kurtosis • Nonlinear energy • Variance (Δ and ΔΔ)
Information theory	• Shannon entropy • Singular value decomposition entropy • Fisher information • Spectral entropy

* Δ= 1st derivative and ΔΔ = 2nd derivative of the raw epoch

The classification stage uses two separate classifiers, a static-SVM (RBF-SVM) classifier as used in (Temko et al., 2011a) and a GDTW-SVM based sequential classifier. Each channel is classified separately. The static classifier (RBF-SVM) classifies a single feature vector at a time. A Gaussian RBF kernel is used inside the static-SVM classifier. More description on this classification method can be found in (Temko et al., 2011a).

For the sequential classifier (GDTW-SVM), sequences of 15 feature vectors are created. This corresponds to an EEG signal of 64s in duration. A shift of one epoch is used to make the next sequence. Therefore, the GDTW classifier will give a probability output with each new epoch. Therefore, the output of both classifiers remains synchronized.

The distances θ_{GDTW} and θ_{SVM} provided from each classifier (Eq. 3 and Eq. 5) are then converted into posterior probabilities using Platt's method (Platt, 1999) of applying a sigmoid function defined as,

$$P(seizure|\mathbf{x}) = \frac{1}{1+\exp(\lambda_{SVM}\theta_{SVM}+\delta_{SVM})} \tag{6}$$

$$P(seizure|S) = \frac{1}{1+\exp(\lambda_{GDTW}\theta_{GDTW}+\delta_{GDTW})} \,, \tag{7}$$

where θ_{SVM} and θ_{GDTW} are the SVM distances to the separating hyperplane from each classifier. The parameters $\{\lambda_{SVM}, \delta_{SVM}\}$ and $\{\lambda_{GDTW}, \delta_{GDTW}\}$ of each sigmoid function are estimated over the training dataset (Platt, 1999).The probabilistic outputs from both classifiers are then fused together using a simple averaging function. The same process is repeated for all 8 channels.

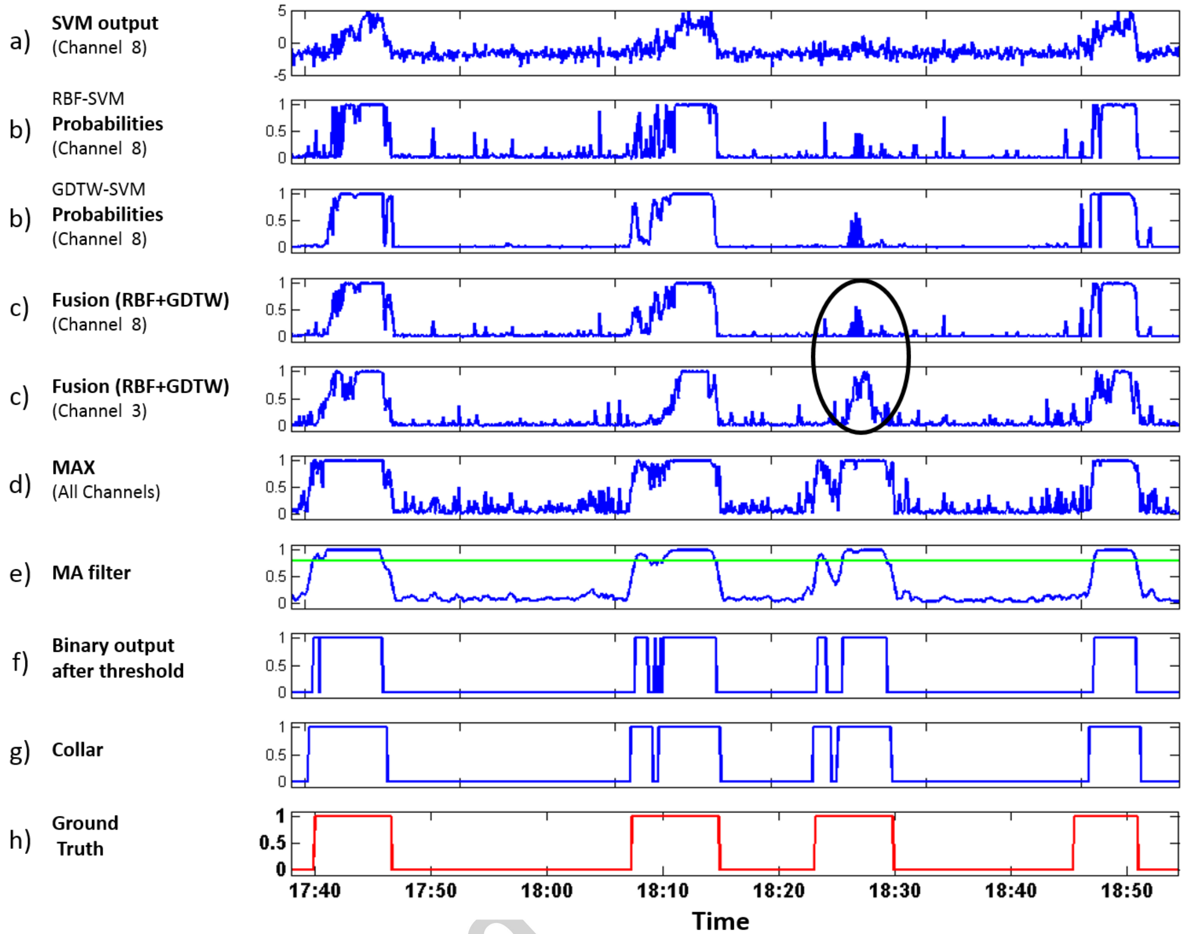
3.3 Post-processing

Figure 4 shows the effects of various post-processing steps on the output from the classifier stage. The fused probabilistic outputs from each channel are first combined by applying a MAX operator (Figure 4d). This process is similar to the way seizures are clinically annotated; if a seizure event is found in one channel then the whole epoch is marked as a seizure. A Moving Average Filter (MAF) is applied to this fused probabilistic stream (Figure 4e). The output from this stage is then compared to a threshold and a binary decision is made i.e. 0: non-seizure and 1: seizure (Figure 4f). When a seizure is detected, the decision of the whole seizure event is extended on either side using a collar operation (including a set number of epochs on either side of the decision) (Figure 4g). This operation provides some compensation for the smoothing operation of the MAF and for the possible difficulties in detecting the start and end of a seizure.

3.4 Performance Assessment

In order to assess the performance of the proposed system, a Leave One Patient Out (LOPO) cross validation was used. The classification model was created using the training data of 16 patients and the resulting model was tested on the remaining unseen patient's recording. This process was repeated until every patient had been used as the test patient. Mean results across all test patients were then reported. It is well known that LOPO provides the most unbiased assessment of the system's performance to match real-life conditions (Vapnik and Kotz, 1982). For the purpose of reporting and comparing the performance of different configurations of the system, two types of metric were used: epoch based sensitivity, specificity & precision and event based Good Detection Rate (GDR) & False Detections per hour (FD/h).

Figure 4: Effects of different post processing steps. (a) The raw output of the SVM classifier. (b) The raw outputs of Static-SVM and GDTW-SVM converted to probability using the sigmoid function. (c) Fusion of the probabilities of both classifiers in channel 8 and channel 3. It can be seen that the encircled seizure event is not completely present in channel 8. (d) Therefore, a MAX operation is performed on all the channels. (e) The smoothed probabilities after a 9-tap moving average filter is applied. (f) The binary output resulting from applying a threshold of 0.8 to the filtered probabilities of seizure. (g) The collar operation is performed on the detected seizure events which increases the duration of all seizures on either side by 7 epochs (h) The neurophysiologist annotations, where 1 indicates seizure.



Sensitivity is the percentage of correctly identified seizure epochs whereas specificity is the percentage of correctly detected non-seizure epochs. The precision is defined as the percentage of identified seizure epochs that are correct. The area under the Receiver Operating Characteristic (ROC) curve is reported by computing the sensitivity and specificity at different thresholds applied on the final probability. The threshold is gradually moved from 0 to 1. Moreover, for an automated system to be of any clinical significance, the specificity needs to be very high. Therefore, we also report the ROC90 as adopted by our earlier study (Temko et al., 2013). ROC90 is the area under the ROC curve where the specificity is greater than 90%. Furthermore, to test if the improvement achieved by the proposed fusion based method was significant, the ROC90 areas of the RBF-SVM and the fusion based system from each patient were compared according to the method outlined in (Robin et al., 2011). This method uses the bootstrapping technique to estimate the standard error of the paired partial ROC curves. In this study, the p-values are reported after 1000 stratified bootstrap cycles.

A disadvantage of the ROC curve is its insensitivity to the amount of data in each class (Davis and Goadrich, 2006). Such is usually the case in seizure detection problems, where non-seizure epochs are significantly greater in number as compared to seizure epochs. Therefore, the Precision-Recall/sensitivity (PR) curve is also reported which better represents the capability of a classifier to detect seizure events, taking into account the amount of non-seizure data present in a particular recording.

The events based metrics allow for the assessment of the system's performance from the perspective of individual seizure events. In this work, we report the GDR (which is the percentage of detected seizure events) obtained at the cost of a given number of false detections per hour.

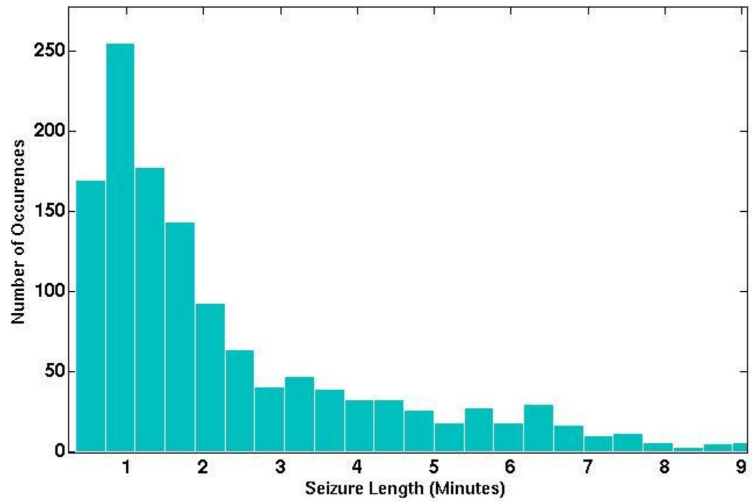
3.5 Model Selection and Training

There are two SVM parameters; kernel parameter σ and generalization parameter C that need to be tuned during the training.

For each iteration of LOPO, the data of 16 patients were used to perform internal cross validation and train the classifier. In order to find the best parameters for the GDTW-SVM a 3-fold cross validation was applied on the training data. In a 3-fold cross validation, the training data is divided into three partitions. These partitions are patient independent, meaning that each fold had the data from one of three separate patient groups. An SVM model is trained using the data from two folds and then tested on the remaining one unseen fold. This process is repeated 3 times, first for a range of C parameters and then for σ kernel parameters. The accuracy and variance of the results obtained on the 3 test folds then determines the most suitable SVM parameters. A similar 5-fold cross validation was used for the RBF-SVM. It was empirically observed that the kernel parameter for the GDTW-SVM did not significantly change, therefore a lower number of folds were used for it.

Figure 5 shows the histogram of the number of seizures and their respective lengths in our dataset. It can be seen that the modal seizure length is around 1 minute. For GDTW-SVM, the sequence length was therefore fixed at 15 epochs for both the seizure and non-seizure classes motivated by the fact that the modal seizure length is 60 seconds. In cases where a seizure was less than 15 epochs, then to complete the sequence, some non-seizure epochs were added on both sides of the seizure epochs.

Figure 5:Histogram showing the length of seizures.



4 Results and Discussion

In this section the performance of the classifiers obtained on the continuous recordings of 17 patients (as described in Table 1) is presented. The use of post processing methods on the probabilistic output of the classifiers does not allow for the observation of the real behavior of a classifier. Therefore, the raw performance of each classifier will be described first and then the performance of the proposed fusion based system with the post processing stage added (MA filter and collar) will be presented. Moreover, it must be noted that the performance reported below is the mean performance of the systems over all patients.

The ROC curves for the individual classifiers and for the fusion based classifier is shown in Figure 6. It can be seen that the GDTW-SVM outperforms the RBF-SVM in terms of the epoch based metric. The GDTW-SVM based system achieves an ROC90 area of 71.9% as compared to 69.8% attained by the static RBF-SVM classifier. Moreover, the fusion of the raw probabilities yields the highest mean ROC90 area of 75.2%. The ROC90 area achieved by the individual patients are presented in Table 3 along with their respective p values from significance test (as calculated using he bootstrapping method proposed in (Robin et al., 2011)). It can be seen, that the fusion based system improves the ROC 90 in 16/17 patients and this improvement is significant (p-value<0.05).

Figure 7 shows the PR curves of the classifiers. The GDTW-based system achieved a higher PR area of 79.6% as compared to 77.9% for the RBF-SVM based system. Moreover, the fusion of the both classifier’s probabilities further increased the PR area by 3%.

The performance of the classifiers using event based metrics is presented in Figure 8. Although, individually the GDTW-SVM based classifier performed better in epoch based metrics it did not however achieve a higher event detection rate. For example, given a threshold of 0.5 FD/h, the GDTW based classifier detected 69% of the total seizure events as compared to 73% detected by the RBF-SVM. However, the fusion of both classifiers significantly increased the detection rate to 82% at 0.5 FD/h. This improvement using the fusion of two classifiers indicates that although the detection rate of the GDTW-SVM classifier is low, it however detected events which were complementary to the RBF-SVM.

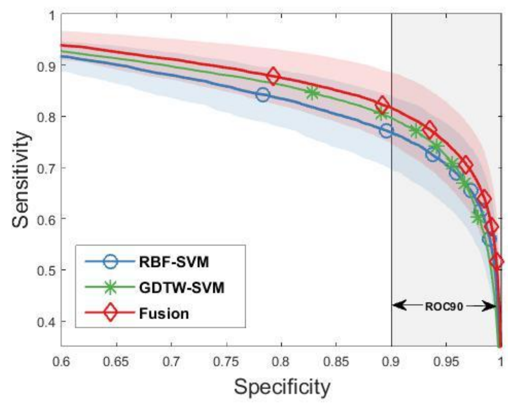


Figure 6:ROC curves with highlighted ROC90 area without post-processing. AUC: RBF-SVM=69.8%, GDTW-SVM=71.9%, Fusion=75.2%. The 95% confidence interval of ROC curves of RBF-SVM and Fusion based system are presented using the shaded area along the line.

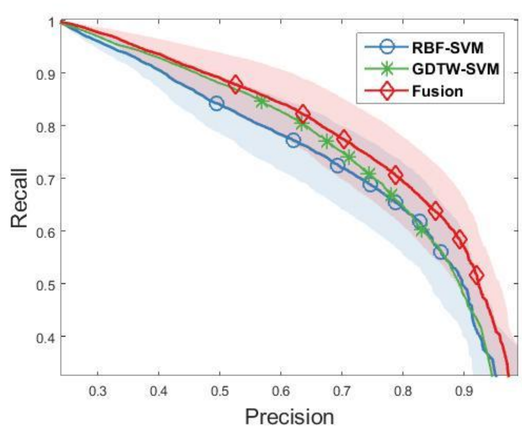


Figure 7: Precision-Recall curves without post-processing. AUC: RBF-SVM=77.9%, GDTW-SVM=79.6%, Fusion=82.7%. The 95% Confidence interval of PR curves of RBF-SVM and Fusion based system are presented using the shaded area along the line.

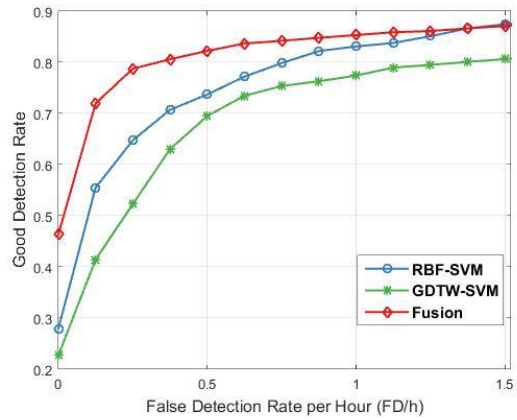


Figure 8: Good detection rate at the expense of false detections per hour of systems without post-processing.

Table 3: Per patient results of the baseline RBF-SVM and Fusion based classifier without post-processing. The area under the curve of ROC90 along with their respective p values is presented.

Pat ID	AUC ROC90 x10		*P values
	RBF-SVM	Fusion	
1	40.9	**55.1**	0.000
2	75.3	**82.1**	0.000
3	72.9	**73.0**	0.007
4	66.2	**70.8**	0.000
5	51.8	**56.9**	0.000
6	59.3	**66.0**	0.000
7	69.7	**74.4**	0.000
8	68.1	**76.5**	0.000
9	78.6	**81.7**	0.000
10	**58.3**	56.1	0.000
11	83.8	**87.1**	0.000
12	69.8	**73.5**	0.000
13	76.8	**83.1**	0.000
14	78.9	**86.1**	0.000
15	77.4	**79.5**	0.002
16	86.3	**90.1**	0.000
17	72.5	**85.9**	0.000
Mean	69.8	**75.2**	

* Significance test was performed to compare the ROC90 from the baseline RBF-SVM and the fusion based system

In order to get further insight into this complementary behavior, the detected events provided by each system were examined. However, to do so, a common threshold for all patients needs first to be fixed on their respective probability outputs. The probability traces resulting from the two different classifiers could not be directly compared because each classifier assigns the probability to a single epoch in a different way. For example, one classifier may assign the probability to the seizure epochs in the range of 0.4-0.6 and non-seizure in the range of 0.1-0.3 whereas another classifier may assign the probability 0.7-0.8 and 0.4-0.6 to the respective classes. Now if a common threshold of 0.5 is used on the probability traces of these classifiers to compare the number of seizures detected by each classifier, then the first classifier will lose out as some of the seizure epochs would have a probability of less than 0.5. Another method to assign a common threshold is by limiting the number of false detections and then comparing the number of correctly detected seizure events at this threshold. In this work, a threshold of 0.5 FD/h is used and then the GDR of each patient in analyzed. The second column of Table 4 shows the mean percentage of total seizure events detected by each classifier. The remaining columns show the unique seizure events that were detected by a classifier and were missed by the other. For example, the GDTW-SVM detected 69.4% of total seizure events, of which 15.5% were unique events that were not detected by the RBF-SVM system at the same threshold. Whereas the RBF-SVM detected 73.7% of total seizure events, of which 25% were unique event detections that were missed by the GDTW-SVM. With the fusion of both systems output, the total detection rate substantially increases to 82.2%. Moreover, out of all of these detected events, 24.9% of the events were new detections as compared to the GDTW-SVM and 17.4% were new as compared to the RBF-SVM. The agreement between GDTW-SVM and RBF-SVM on false detected events was only 12%. This indicates that GDTW-SVM does provide some complementary behavior and the fusion of the two, results in the detection of events that were missed by the individual classifiers.

Post-processing was then applied using the MAF and collar on the output of the proposed fusion based system. The results were compared to the baseline static RBF-SVM based system with post-processing, as proposed in (Temko et al., 2011a). The baseline system was compared to several existing alternatives and superior performance was reported. This comparison was possible because the baseline system used the same dataset. Unfortunately, comparison with other systems is usually not possible or may not provide significant insight because of the different types of datasets used and metrics reported. Another study of our group has thoroughly discussed the challenges of comparing different seizure detection system (Temko et al., 2011b). The length of the MAF and collar for both fusion and RBF-SVM systems was chosen to maximize the ROC90 area (which turns out to provide also the maximum for the PR area). A

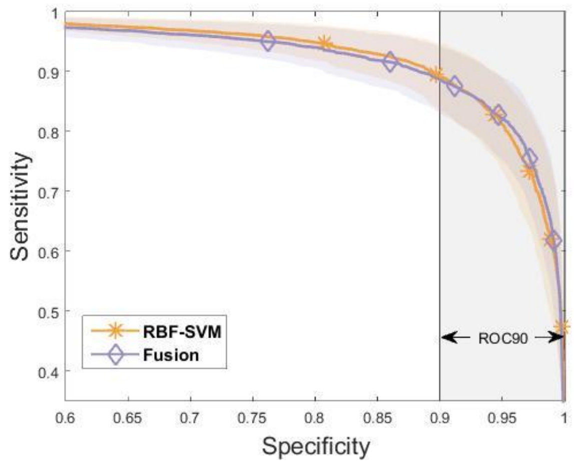
MAF of 9 epochs and a collar of 7 epochs was selected for the fusion based system. For the RBF-SVM, a MAF of 15 epochs and a collar of 7 epochs was selected.

Table 4: Percentage of detected events unique to the other classifier with the threshold set to 0.5FD/h.

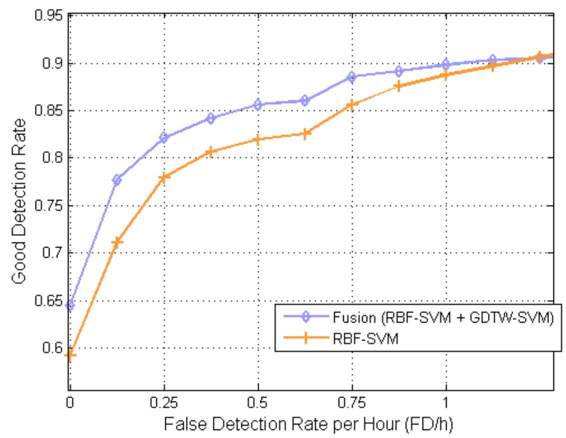
	Total detected events (%)	Unique decisions to (%)	
		GDTW-SVM	RBF-SVM
GDTW-SVM	69.4	0	15.5
RBF-SVM	73.7	25	0
Fusion	82.2	24.9	17.4

With the addition of a post-processing stage, both systems achieved ROC90 area of 82.6% (Figure 9a). The PR area for the fusion based system and the RBF-SVM with post-processing was 88.8% and 88.7% respectively. Although the presented fusion approach with post-processing achieved a similar performance to the static RBF-SVM on epoch based metrics it did however improve the event based metric and managed to achieve a higher detection rate at the cost of significantly lower false detections per hour. It can be seen in Figure 9b that the fusion based system detects 65% of the seizures at no false detections which is a 6% improvement over the previous system. The proposed system achieves 82.6% GDR at 0.25 FD/h (1 false alarm in 4 hours) which means it correctly detects 39 more seizure events as compared to the RBF-SVM based system. The proposed system detected 90% of the seizure events at 1 FD/h.

Figure 9 (a): ROC curves with confidence intervals and highlighted ROC90 area **with post-processing stage added**. ROC90 AUC RBF-SVM =82.6%, Fusion=82.6% (b) Good detection rate of the systems at the expense of false detections per hour **with the post-processing stage**.



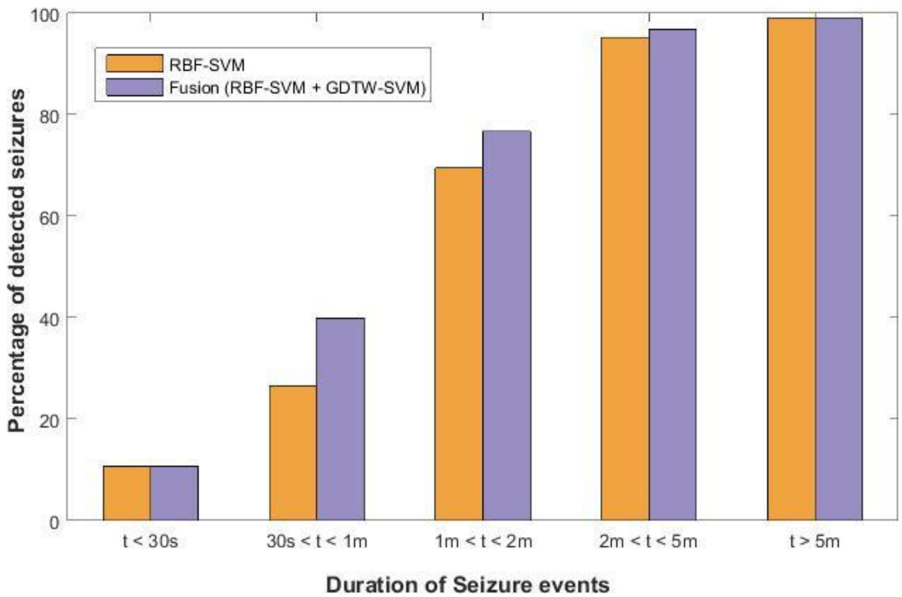
(a)



(b)

As the main motivation of the proposed technique was to improve the detection of short seizure events, it is therefore important to see how the system performed on different lengths of seizures. Figure 10 shows a comparison of the percentage of detected seizures for both the RBF-SVM baseline system and the proposed fusion system (both with the post processing stage included) at a stringent threshold of 0.25 FD/h. It can be seen, that the proposed fusion based system improved the detection rate in all categories except for the very short duration seizures of less than 30 seconds. Most importantly, a 12% improvement in the detection of short seizures of length 30sec - 1 min was obtained which were previously being missed by using only the static RBF-SVM based system. The proposed system also detected 7% more seizures of durations between 1-2 minutes. It must be noted that, this score is reported at a very stringent and low false detection per hour of 0.25 FD/h (1 false detection in 4 hours). The detection rate increases if we allow more false detections per hour for example as shown in Figure 9b that the overall GDR increases to 86% at 0.5 FD/h. At this threshold GDR for seizures 30sec-1min also increases to 50%.

Figure 10: Percentage of seizures detected according to their lengths by the proposed fusion based system and the RBF-SVM both **with post-processing**. These results represent the GDR at 0.25 FD/h. Total number of seizures in each group are 28, 128, 223, 253, 189 respectively.



Some representative examples of correctly detected short seizures are shown in Figure 11. Figure 11(a, b) shows two events where temporal evolution in both amplitude and frequency can be clearly seen. This contextual information helped the system to detect the events at a lower false detection rate. On the other hand, Figure 11c shows an event where the seizure did not evolve temporally and was still detected by the fusion-based system. Lastly, Figure 11d presents a correctly detected short event where the amplitude of the EEG remained very low. This event had a low probability of seizure from the RBF-SVM and was further suppressed by the MAF stage; however, the GDTW based system yielded a high probability because of the number of epochs being classified simultaneously. Thus, the fusion of the two systems helped to raise the probability to a level where it was easily detected.

The main performance difference using the GDTW based classifier was due to its ability to not only take into account the contextual information in the EEG but also the temporal evolution which is the main characteristic of a neonatal seizure. The work presented in this article is just related to one type of dynamic kernel. We had in fact also tried a Gaussian Mixture Model (GMM) supervector based kernel in the SVM. The GMM-Supervector also explores the contextual information given in the sequence. It however, does not explore the sequentiality/temporal characteristics of the feature vectors present in the input sequence but only considers their statistics. The GDTW-classifier outperformed the GMM-supervector based classifier (AUC ROC90 75.2% vs 69.5%). The GMM-supervector based system has been employed in a previous study on grading background EEG and is explained in detail in (Ahmed et al., 2016).

Despite a promising improvement in the detection rate using the proposed fusion, the system did not significantly increase the performance when the post-processing stage was included. This behavior does not come by surprise. Indeed, the post-processing proposed in (Temko et al., 2011) is developed for static-SVM. MAF is a very coarse method to explore the contextual information at decision level and it does not accentuate the benefits of the proposed fusion. Therefore, a different post-processing technique is needed for further improvement. Moreover, the strength of GDTW-SVM lies in classifying variable length sequences which has not yet been fully explored in this work. Therefore, it is suggested here that future work should focus on the pre-processing of the EEG signal using methods such as diariazation (Temko et al., 2014; Tranter and Reynolds, 2006). In our work, we intend to use this technique to adaptively segment the EEG recording into a number of smaller homogenous states which will then be classified by a dynamic classifier. The computation time required for the GDTW-SVM to classify one sequence of 15 epochs (64 seconds) was 1.64 seconds (Machine specs: Processor=Intel core i3 @3.10 GHz, RAM=8GB, Operating system=Windows 7).

The GDTW kernel is widely criticized for not being a valid SVM kernel (Cuturi et al., 2007; Gudmundsson et al., 2008; Lei and Sun, 2007). This may hinder the SVM algorithm from converging to an optimal and unique solution. This means that the SVM may produce a different model even if the C and σ parameters remain the same; hence, the reported performance may not be repeatable. An experiment was conducted to see if the GDTW-SVM produces a unique solution and therefore if the reported results are repeatable. Five different training sets were used and 10 SVM models were generated from each training set. This resulted in 50 SVM models. The SVM-light implementation (Joachims, 2008) used in this study, works by dividing the training data into smaller working sets and then finding the alpha values (Joachims, 1999). The training data was therefore, shuffled in each iteration before training a new SVM model so that the SVM algorithm always starts with a different working set**.** The C and σ parameters were kept unchanged. These SVM models were then tested on a dataset of

850 sequences (190 seizure and 660 non-seizure) from a patient not used in the training datasets. The results with their mean and 95% confidence intervals are presented in Table 5. For all the training sets the number of support vectors varied only by ± 2. Similarly, the difference in the bias for 10 iterations of each training set was either very small ±0.001 or there was no change at all. In terms of ROC area of the test data, there was no change over all the iterations of any training set. Moreover, the overall change in ROC area across all the training sets was also very low (standard deviation = 0.27). This indicates that the presented results of our experiments are stable and repeatable.

Table 5: Results of GDTW-classifier on 10 shuffled iterations of each of the 5 training sets.

Training set	Number of support vectors	Bias	ROC area on test data
	Mean (95% CI)		
1	1364 (1363 - 1365)	0.966 (0.966 - 0.967)	97.6 (97.6 - 97.6)
2	1048 (1047 - 1048)	1.027 (1.027 - 1.027)	97.7 (97.7 - 97.7)
3	1041 (1040 - 1042)	1.092 (1.092 - 1.093)	97.1 (97.1 - 97.1)
4	1104 (1104 - 1104)	1.091 (1.091 - 1.091)	97.4 (97.4 - 97.4)
5	1324 (1323 - 1324)	1.038 (1.038 - 1.038)	97.9 (97.9 - 97.9)

4 CONCLUSION

This study presented an approach to exploit the temporal and contextual information in the neonatal seizure and background EEG. A DTW based kernel function is used which enabled the SVM to process the sequences of short term feature vectors extracted from 8 seconds EEG epochs. A patient-independent neonatal seizure detection system based on the combination of a static (RBF-SVM) and a sequential (GDTW-SVM) classifier was thus proposed. Results indicate that the fusion of the output of both classifiers leads to significant improvement in both epoch and event based metrics. The proposed fusion based system was also tested with a post-processing stage. The results show a promising improvement in the detection rate of the seizures at a significantly low setting for false detections per hour. Particularly, the system increased the detection rate of short seizure events of less than 1 minute by 12%. However, it is expected that this improvement can be accentuated with better post-processing methods tailored to the proposed fusion based method.

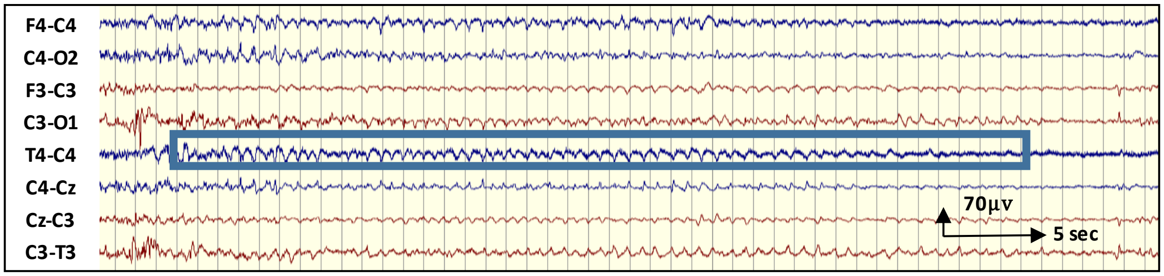
Conflict of interest

None Declared

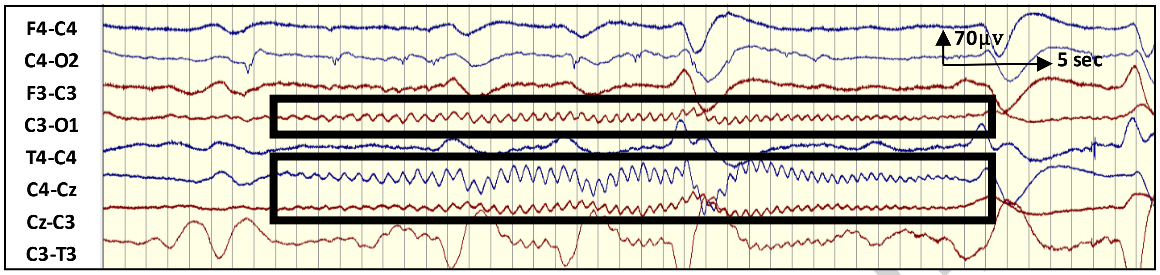
ACKNOWLEDGMENT

This research was supported by a Science Foundation Ireland Principal Investigator Award (SFI 10/IN.1/B3036) and a Science Foundation Ireland Centres Programme Award (12/RC/2272).

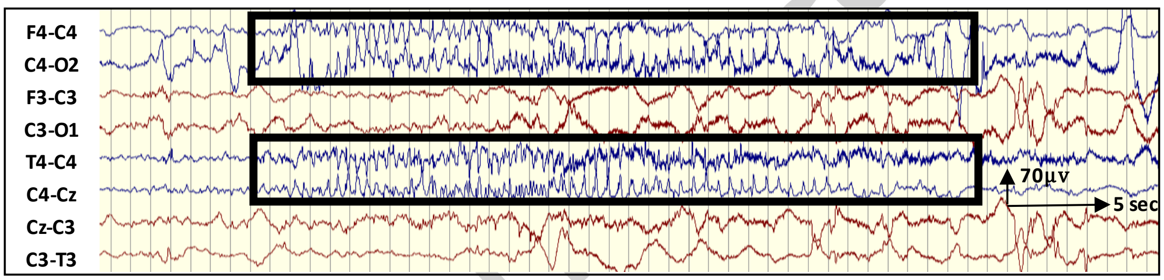
Figure 11: Examples of new short seizures correctly classified by the fusion based system at 0.25 FD/h. An example of short seizure with (a) temporal evolution (duration=41 sec) (b) oscillatory and temporal evolution (duration 37 sec) (c) high frequency (duration 40 sec) (d) very low amplitude (duration=37 sec)


F4-C4
C4-O2
F3-C3
C3-O1
T4-C4
C4-Cz
Cz-C3
C3-T3
70µv
5 sec

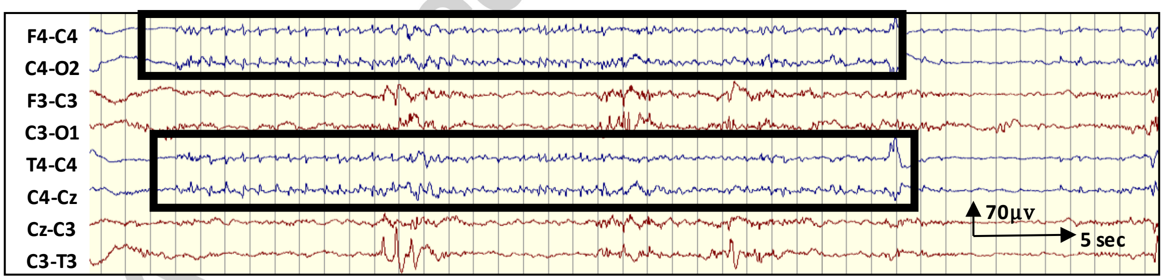
(a)


F4-C4
C4-O2
F3-C3
C3-O1
T4-C4
C4-Cz
Cz-C3
C3-T3
70µv
5 sec

(b)


F4-C4
C4-O2
F3-C3
C3-O1
T4-C4
C4-Cz
Cz-C3
C3-T3
70µv
5 sec

(c)


F4-C4
C4-O2
F3-C3
C3-O1
T4-C4
C4-Cz
Cz-C3
C3-T3
70µv
5 sec

(d)

REFERENCES

Aarabi, A., Kazemi, K., Grebe, R., Moghaddam, H.A., Wallois, F., 2009. Detection of EEG transients in neonates and older children using a system based on dynamic time-warping template matching and spatial dipole clustering. Neuroimage 48, 50–62.

Ahmed, R., 2015. Dynamic classifiers for neonatal brain monitoring. University College Cork, Cork, Ireland.

Ahmed, R., Temko, A., Marnane, W., Boylan, G., Lighbody, G., 2012. Dynamic time warping based neonatal seizure detection system, in: Proceedings of IEEE Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 4919–4922.

Ahmed, R., Temko, A., Marnane, W., Lightbody, G., Boylan, G., 2016. Grading hypoxic–ischemic encephalopathy severity in neonatal EEG using GMM supervectors and the support vector machine. Clinical Neurophysiology 127, 297–309.

AL-Naddawi, M.N., Hameed, N.N., Kadum, M.J., Al-Dabbas, N.W., 2011. Clinical types and possible etiologies of neonatal seizures: A hospital based study. J. Fac. Med. Baghdad 53, 1–5.

Bahlmann, C., Haasdonk, B., Burkhardt, H., 2002. Online handwriting recognition with support vector machines - a kernel approach., in: Proceedings of 8th International Workshop on Frontiers in Handwriting Recognition, 2002. pp. 49–54.

Bogaarts, J., Gommer, E., Hilkman, D., van Kranen-Mastenbroek, V.H.J.M., Reulen, J.P.H., 2014. EEG feature pre-processing for neonatal epileptic seizure detection. Annals of Biomedical Engineering 42, 2360–2368.

Burges, C.J., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121–167.

Campbell, W.M., Sturim, D.E., Reynolds, D., Solomonoff, A., 2006. SVM based speaker verification using a GMM supervector kernel and NAP variability compensation, in: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2006. pp. 97–100.

Chaovalitwongse, W., Pardalos, P., 2008. On the time series support vector machine using dynamic time warping kernel for brain activity classification. Cybernetics and Systems Analysis 44, 125–138.

Cuturi, M., Vert, J.-P., Birkenes, O., Matsui, T., 2007. A kernel for time series based on global alignments, in: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2007. IEEE, pp. 413–416.

Davis, J., Goadrich, M., 2006. The relationship between Precision-Recall and ROC curves, in: Proceedings of the 23rd International Conference on Machine Learning. ACM, pp. 233–240.

Deburchgraeve, W., Cherian, P., De Vos, M., Swarte, R., Blok, J., Visser, G.H., Govaert, P., Van Huffel, S., 2008. Automated neonatal seizure detection mimicking a human observer reading EEG. Clinical Neurophysiology 119, 2447–2454.

Depeursinge, A., Iavindrasana, J., Hidki, A., Cohen, G., Geissbuhler, A., Platon, A., Poletti, P.-A., Müller, H., 2010. Comparative performance analysis of state-of-the-art classification algorithms applied to lung tissue categorization. Journal of digital imaging 23, 18–30.

Doyle, O., Temko, A., Murray, D., Lightbody, G., Marnane, W., Boylan, G., 2010. Predicting the neurodevelopmental outcome in newborns with hypoxic-ischaemic injury, in:

Proceedings of IEEE Annual International Conference of Engineering in Medicine and Biology Society, (EMBC) 2010. IEEE, pp. 1370–1373.
Faul, S., Temko, A., Marnane, W., 2009. Age-independent seizure detection, in: Proceedings of IEEE Annual International Conference of Engineering in Medicine and Biology Society, (EMBC) 2009. pp. 6612–6615.
Fauve, B.G., Matrouf, D., Scheffer, N., Bonastre, J.-F., Mason, J.S., 2007. State-of-the-art performance in text-independent speaker verification through open-source software. IEEE Transactions on Audio, Speech, and Language Processing 15, 1960–1968.
Gotman, J., Flanagan, D., Zhang, J., Rosenblatt, B., 1997. Automatic seizure detection in the newborn: methods and initial evaluation. Electroencephalography and Clinical Neurophysiology 103, 356–362.
Greene, B.R., Faul, S., Marnane, W.P., Lightbody, G., Korotchikova, I., Boylan, G.B., 2008. A comparison of quantitative EEG features for neonatal seizure detection. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 119, 1248–1261.
Gudmundsson, S., Runarsson, T.P., Sigurdsson, S., 2008. Support vector machines and dynamic time warping for time series, in: Proceedings of IEEE International Joint Conference on Neural Networks, (IJCNN) 2008. IEEE, pp. 2772–2776.
Joachims, T., 2008. SVM Light implementation [WWW Document]. SVM Light. URL https://www.cs.cornell.edu/people/tj/svm_light/index.html, Accessed on: 2012-01-12 (accessed 12.1.12).
Joachims, T., 1999. Making large-Scale SVM Learning Practical, in: Schölkopf, B., Burges, C., Smola, A. (Eds.), Advances in Kernel Methods - Support Vector Learning. MIT Press, Cambridge, MA, pp. 169–184.
Joachims, T., 1998. Text categorization with support vector machines: Learning with many relevant features, in: European Conference on Machine Learning. Springer, pp. 137–142.
Lanska, M.J., Lanska, D.J., Baumann, R.J., Kryscio, R.J., 1995. A population-based study of neonatal seizures in Fayette County, Kentucky. Neurology 45, 724–732.
Lei, H., Sun, B., 2007. A study on the dynamic time warping in kernel machines, in: Signal-Image Technologies and Internet-Based System, 2007. SITIS'07. Third International IEEE Conference on. IEEE, pp. 839–845.
Lodhi, H., Saunders, C., Shawe-Taylor, J., Cristianini, N., Watkins, C., 2002. Text classification using string kernels. The Journal of Machine Learning Research 2, 419–444.
Lynch, N.E., Stevenson, N.J., Livingstone, V., Murphy, B.P., Rennie, J.M., Boylan, G.B., 2012. The temporal evolution of electrographic seizure burden in neonatal hypoxic ischemic encephalopathy. Epilepsia 53, 549–557.
Muller, M., 2007. Dynamic Time Warping, in: Information Retrieval for Music and Motion. Springer Berlin Heidelberg, pp. 69–84.
Nagaraj, S.B., Stevenson, N.J., Marnane, W.P., Boylan, G.B., Lightbody, G., 2014. Neonatal seizure detection using atomic decomposition with a novel dictionary. IEEE Transactions on Biomedical Engineering 61, 2724–2732.
Navakatikyan, M.A., Colditz, P.B., Burke, C.J., Inder, T.E., Richmond, J., Williams, C.E., 2006. Seizure detection algorithm for neonates based on wave-sequence analysis. Clinical Neurophysiology 117, 1190–1203.
Ocak, H., 2009. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Systems with Applications 36, 2027–2036.

Platt, J.C., 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, in: Advances in Large Margin Classifiers, Neural Information Processing Series. MIT Press, pp. 61–74.

Rennie, J.M., Chorley, G., Boylan, G.B., Pressler, R., Nguyen, Y., Hooper, R., 2004. Non-expert use of the cerebral function monitor for neonatal seizure detection. Archives of Disease in Childhood - Fetal and Neonatal Edition 89, F37–F40.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics 12, 1.

Sakoe, H., Chiba, S., 1978. Dynamic programming algorithm optimization for spoken word recognition. IEEE Transactions on Acoustics, Speech and Signal Processing 26, 43–49.

Shimodaira, H., Noma, K., Nakai, M., Sagayama, S., 2002. Dynamic time-alignment kernel in support vector machine, in: Advances in Neural Information Processing Systems 14. MIT Press, pp. 921–928.

Shoeb, A.H., Guttag, J.V., 2010. Application of machine learning to epileptic seizure detection, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10). pp. 975–982.

Smith, N., Gales, M.J., 2002. Using SVMs and discriminative models for speech recognition, in: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2002. IEEE, pp. 77–80.

Temko, A., Boylan, G., Marnane, W., Lightbody, G., 2013. Robust neonatal EEG seizure detection through adaptive background modeling. International Journal of Neural Systems 23, 1350018.

Temko, A., Marnane, W., Boylan, G., Lightbody, G., 2014. EEG “diarization” for the description of neonatal brain injuries, in: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2014. IEEE, pp. 5844–5848.

Temko, A., Monte, E., Nadeu, C., 2006. Comparison of sequence discriminant support vector machines for acoustic event classification, in: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2006. IEEE, pp. 721–724.

Temko, A., Thomas, E., Marnane, W., Lightbody, G., Boylan, G., 2011a. EEG-based neonatal seizure detection with support vector machines. Clinical Neurophysiology 122, 464–473.

Temko, A., Thomas, E., Marnane, W., Lightbody, G., Boylan, G., 2011b. Performance assessment for EEG-based neonatal seizure detectors. Clinical Neurophysiology 122, 474–482.

Thomas, E.M., Temko, A., Lightbody, G., Marnane, W.P., Boylan, G.B., 2011. Advances in automated neonatal seizure detection, in: New Advances in Intelligent Signal Processing, Studies in Computational Intelligence. Springer Berlin Heidelberg, pp. 93–113.

Thomas, E.M., Temko, A., Lightbody, G., Marnane, W.P., Boylan, G.B., 2010. Gaussian mixture models for classification of neonatal seizures using EEG. Physiol. Meas. 31, 1047.

Tranter, S.E., Reynolds, D.A., 2006. An overview of automatic speaker diarization systems. IEEE Transactions on Audio, Speech, and Language Processing 14, 1557–1565. doi:10.1109/TASL.2006.878256

Vapnik, V.N., Kotz, S., 1982. Estimation of dependences based on empirical data. Springer-verlag New York.

Watkins, A., Szymonowicz, W., Jin, X., Yu, V.V.Y., 1988. Significance of seizures in very low-birthweight infants. Developmental Medicine & Child Neurology 30, 162–169.
Wong, S., Gardner, A.B., Krieger, A.M., Litt, B., 2007. A stochastic framework for evaluating seizure prediction algorithms using hidden Markov models. Journal of Neurophysiology 97, 2525–2532.