31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Age-independent Seizure Detection

Stephen Faul, Andriy Temko and William Marnane

Abstract— **This paper examines whether an appropriate algorithm, developed for use with neonatal data, could also be used, without alteration, for the detection of seizures in adults with epilepsy. The performance of a feature extraction and SVM classifier system is evaluated on databases of 17 neonatal patients and 15 adult patients. Mean ROC curve areas of 0.96 and 0.94 for neonatal and adult databases respectively show that high accuracy can be achieved independent of age. It is also shown that features contribute differently for neonatal and adult data.**

I. INTRODUCTION

Neonatal seizures and adult epileptic seizures are two areas of biomedical signal processing which have undergone major advances in recent years. With technological improvements has come the ability to address seizure detection problems with more advanced mathematical approaches.

The history of many neonatal seizure detection algorithms begins with the exploration of seizure detection in adults. There are a number of reasons for this. Adult EEG databases are freely available on the internet [1], allowing many groups to commence research in the area. Large databases of neonatal seizure data are much more difficult to gain access to. Furthermore, because the mature adult brain is better developed than the neonatal brain, seizure events in adults are accompanied by less complex waveforms than those found in neonates.

Hence, most neonatal seizure detection systems originate as adult seizure detection systems which are then altered in an attempt to provide neonatal seizure detection [2]. In this paper, it is investigated if high performance can be achieved using an algorithm designed primarily for the more complicated neonatal seizure detection problem, leading to a system which can be used for accurate seizure detection in *both* adults and neonates.

This paper will also examine the relative performance of signal processing features on adult and neonatal EEG. A feature set of 47 features is extracted from the adult and neonatal EEG databases, including time, frequency, wavelet, information theory and modeling methods. The performance of the feature set will be discussed in relation to the datasets.

As far as the authors are aware, a direct comparison of a seizure detection system on neonatal and adult datasets has not been previously carried out. This paper aims to examine some of the problems posed by age-independent seizure detection.

Supported by Science Foundation Ireland (SFI/07/SRC/I1169) and Wellcome Trust (085249/Z/08/Z)

S.Faul, A.Temko and W.Marnane are with Department of Electrical Engineering, University College Cork, Ireland stephenf@rennes.ucc.ie

II. EEG DATA

A. Neonatal Dataset

A dataset of 17 neonates is used to assess the performance of the seizure detection algorithm on neonatal EEG. The database consists of a total of 267.9 hours of scalp recorded EEG with 691 seizures with a mean duration of 4.67 minutes. This database has been previously used in the development and assessment of neonatal seizure detection algorithms [3], [4].

B. Adult Dataset

A dataset of adult EEG from patients with epilepsy is maintained and made publicly available by the University of Freiburg [1]. It contains both seizure and non-seizure intracranial recordings (only the seizure recordings used here) for 21 patients with ages ranging from 13 to 50. 3 patients (1, 18 and 19) were not used in this test as their recordings contained very short length activity (less than 10 seconds) annotated as seizure. Rather than being seizure events, activity of less than 10 seconds in length is regarded as being *interictal discharge* rather than seizure. Therefore, the algorithm discussed here is designed to detect activity of only 10 seconds in correspondence with this clinical definition. 3 other patients (5, 8 and 10) were removed from the test data because their is a large amount of artifact or measurement inconsistency throughout their recordings. Artifacts exist in some amount in all EEG recordings, caused by external sources (equipment, electrical noise, human interaction) or internal sources (muscles, eye blinks, heart). In the 3 recordings listed above however, artifacts are seen on a large scale through out the entire recording, with non-seizure, repetitive waveforms (patient 5), eye movement (patient 8) or large bursts and attenuation (patient 10). Figs. 1 and 2 show examples from these recordings. Therefore, the reduced adult database used consists of recordings from 15 patients totaling 132.7 hours. There are 62 seizures with a mean duration of 1.88 minutes and standard deviation of 1.31 minutes. Information regarding electrodes and seizure types for each patient can be seen in [1] and [5].

III. DETECTION ALGORITHM

The seizure detection algorithm proposed in this paper has been developed over a number of years in the signal processing group in University College Cork. It consists of a number of steps consisting primarily of feature extraction, classification and post-processing stages. The algorithm is described in more detail in [6] and will be only briefly described here.

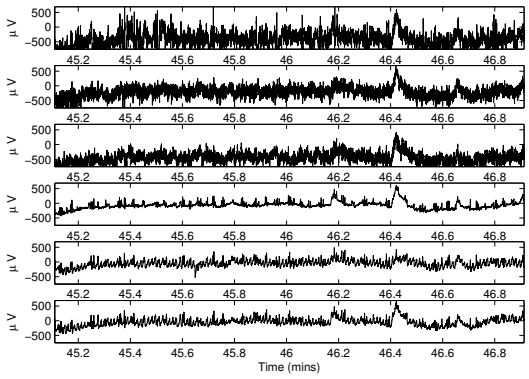

Fig. 1. Repetitive oscillations seen here in the lower 3 channels, continue throughout the EEG of patient 5.

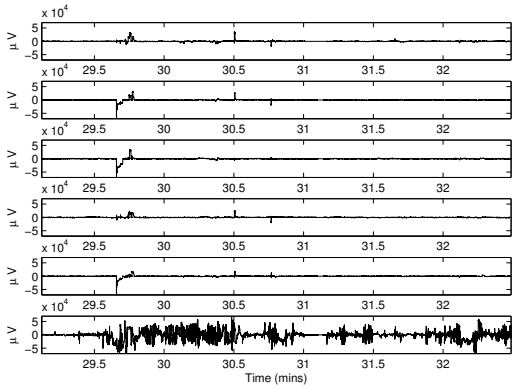

Fig. 2. Recording artifacts seen in the EEG of patient 10.

A. Feature Extraction

The EEG is first notch filtered at 50 Hz to remove line noise. This is followed by downsampling the EEG to 32 Hz. Each channel of EEG is windowed separately with a window of 8 seconds in length with an overlap of 4 seconds. 47 features are then extracted from each of these epochs. More details on these features can be found in [7] and [3]. Features are grouped into 3 classes to later determine if feature types are more dominant in any one age group. The 3 groups are *Spectral*, *Energy* and *Structural*. When a feature overlaps two groups, it is included in both of those groups. The features are listed in Table I.

B. Classification

The Support Vector Machine (SVM) is a classification method which transforms data to a higher dimensional space where a complex classification problems can be solved with linear discriminant functions. SVMs are based on using only those training patterns that are near the decision surface assuming they provide the most useful information for classification. The SVM system with Gaussian kernel used in this work has been developed for neonatal seizure detection in [6].

spectral edge frequency (80%, 90%, 95%)[1]		
dominant spectral peak frequency[1]		
singular value decomposition entropy[3]		
variance of the first and second derivatives[1,2]		
zeros crossings of the first and second derivatives[1]		
Hjorth activity, mobility and complexity[2]		
power in 2 Hz frequency bands ($0-2$Hz, $1-3$Hz,$\ldots 10-12$Hz)[1]		
normalised power in 2 Hz bands[1]		AR model fit[3]
number of minima and maxima[1]		kurtosis[3]
RMS amplitude[2]	line length[1,2]	zero crossings[1]
nonlinear energy[1,2]	spectral power[2]	entropy[3]
wavelet coefficients[1]	spectral entropy[1,3]	Fisher info[3]
skewness[3]		

TABLE I

THE FEATURES EXTRACTED FROM THE EEG DATA. 1=SPECTRAL, 2=ENERGY, 3=STRUCTURAL.

During training, each of extracted features is normalised by subtracting the mean and dividing by the standard deviation to ensure that each feature has equal prominence in the model. The normalising parameters are stored in order that the same normalisation routine can be applied to the testing data.

For each model, 5-fold cross validation is applied on the training data in order to find the optimum parameters for each model. Once the parameters are chosen, they are used to train the final model on all the training data.

C. Post-processing

In the testing stage, the trained model is used to classify the features for each epoch of each channel. The output of the SVM is a value indicating the confidence of the decision; the higher the absolute value of the output, the higher the distance of the testing point from the separating hyper-plane and thus the higher the confidence that the model has correctly classified that epoch. The confidences are then post-processed to improve the seizure detection performance. This post-processing consists of:

- Moving average filter: A moving average filter of length 15 is applied to the classifier confidences for each channel. This technique is used to smooth the classifier output, reducing random noise while retaining a sharp step response, in effect reducing the number of false alarms.
- Threshold: The filtered outputs from the classifier are compared to a threshold value to produce a binary result.
- Channel fusion: A single decision for each epoch is then made up from the individual channels. If any one channel reports a seizure for the epoch, then that epoch is classed as seizure.
- Collar: A collar of length 8 is applied to the binary result. This process extends seizure detections by 8 epochs (36 seconds with a window of 8 seconds as mentioned in section III-A) on either side of the detection, compensating for the difficulty in detecting seizures in their very early or late stages. This improves the temporal accuracy of detections.

IV. TESTING METHODS

A. Classifier Training

The system is tested using a patient independent method. This provides the most unbiased method for testing classifiers. For each of k patients, their EEG is classified using a model trained using data from the other $k-1$ patients. This means that the test is totally blind to the characteristics of each test patient.

The training data for each patient was chosen specifically from visual inspection of each seizure. This process allows expert knowledge to be incorporated into the training process and can greatly improve the accuracy of the classification.

B. Metrics

1) General System Performance: Receiver Operating Characteristics (ROC) curves are often used to determine the performance of a classification problem. They are generally created by plotting sensitivity (the percentage of seizure epochs correctly classified) against specificity (the percentage of non-seizure epochs correctly classified). The area under this curve gives an indicator of system performance, with an area of 1 being perfect performance (100% sensitivity and specificity).

2) Feature Performance: A number of tests are carried out in order to determine if the importance of features changes with age.

- Single Feature Group (SFG) Performance: An SVM is trained as discussed in section IV-A but only using the features from a single group. The test data is then processed by the system using this single feature group SVM and the mean ROC area over all patients for each database is calculated. This is repeated for each of the feature groups. The dominance of a feature group is shown by a higher mean ROC area.
- Feature Group Removed (FGR) Performance: For each dataset, 3 classification tests are carried out, each with one feature group removed. The first test includes only those features in the *Energy* and *Structural* groups, the second test only the *Spectral* and *Structural* groups and the final test includes only the *Spectral* and *Energy* groups. The dominant feature group can then be determined as being that which causes the largest reduction in mean ROC area when removed.

V. RESULTS

A. Adult versus Neonatal Performance

Table II shows the performance of the system for the adult and neonatal databases. The average adult ROC area is 0.9409 with a standard deviation of 0.0830. This compares equivalently with the results achieved for the neonatal database, which resulted in a mean ROC of 0.9580 with a standard deviation of 0.0302. While the mean of the ROC areas for the adult database are approximately 2% lower than for the neonatal database (not statistically significant, $p = 0.2310$), the individual patient results show that higher performance is achieved in more patients for the adult database. This is also reflected in the larger standard deviation for the adult database. 5 adult patients result in a ROC area of > 0.99 whereas only 1 neonatal patient reaches the 0.99 mark. Indeed, these 5 adult patients have the top 5 highest ROC areas of all the patients in both databases. Correspondingly, 9 adult patients result in a ROC area of > 0.98 whereas only 5 neonatal patients achieve this result. At the other end of the scale, only 2 neonatal patients have a ROC of less than 0.90, whereas in the adult database there are 4 such patients, with 2 below 0.80. The reason for the poor performance in these patients (7 and 21) is as yet unknown. Both patients have long seizures and neither recording is badly inflicted with artifacts.

Neonatal		Adult	
Patient	ROC Area	Patient	ROC Area
1	0.8843	2	0.9992
2	0.9320	3	0.9944
3	0.9658	4	0.9987
4	0.9834	6	0.9475
5	0.8941	7	0.7741
6	0.9541	9	0.8748
7	0.9822	11	0.9839
8	0.9725	12	0.9826
9	0.9511	13	0.9944
10	0.9467	14	0.9981
11	0.9901	15	0.9838
12	0.9641	16	0.9562
13	0.9707	17	0.9849
14	0.9827	20	0.8979
15	0.9625	21	0.7436
16	0.9616	**Mean**	**0.9409**
17	0.9876	**Std**	**0.0830**
Mean	**0.9580**		
Std	**0.0302**		

TABLE II
ROC AREAS FOR THE NEONATAL AND ADULT DATABASES

Mean ROC curves for the adult and neonatal databases are shown in Fig. 3. The equal error rate line is also shown. This is where the sensitivity and specificity are equal. The EER for both the adult and neonatal data is 90%. The results for the system compare well to previous studies. Subasi [8] previously used a dynamic wavelet network to detect seizures in adults with epilepsy. On the 5 test patients they recorded ROC areas of 0.907 and 0.921 for two different classifier types. Slooter et al. [9] previously employed a synchronisation likelihood approach to detect seizures in adults. Their system achieved a ROC area of 0.812. Comparing these results with those reported in this paper validates the motivation to use a system designed for the more complicated problem of neonatal seizure detection to improve seizure detection performance in adults. There are a number of studies which use the same University of Freiburg data as used for the adult database in this study [5], [10]. However, all of these studies deal with prediction of seizures and there is no way to compare the results of this algorithm with their results. In terms of neonatal detection performance, Greene et al. [3] investigated a number of classifier types and architectures for neonatal seizure detection and achieved a ROC area of 0.820.

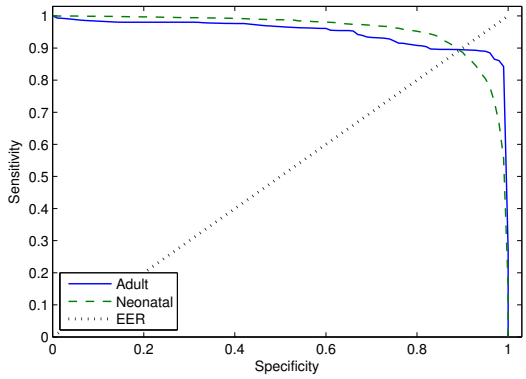

Fig. 3. Mean ROC curves and EER for the adult and neonatal databases.

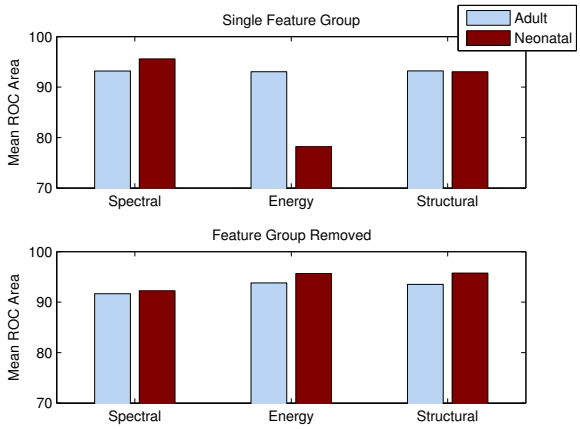

Fig. 4. Performance of each feature group for the *single feature group* (top) and the *feature group removed* (bottom).

B. Feature Performance

Fig. 4 shows the performance of each feature group in the SFG and FGR tests. There is a distinct difference in the feature group contributions for adult and neonatal seizure detection. In the SFG tests it can be seen that the mean ROC for the energy group is considerably lower than for the others for the neonatal data, whereas each group contributes almost equally in adults. It is also seen that the spectral features are the largest contributor in the neonatal tests. This fact is confirmed in the FGR tests. When the spectral features are removed the lowest mean ROC results in both the neonatal and adult data. On comparing the results on removing the energy group or the structural group, it can be seen that there is no significant performance difference in each case. With the neonatal data this is significant, given that the energy group alone has significantly lower performance than the structural group, but when the energy and spectral groups are used together (Structural FGR) they have equivalent performance to the spectral and structural groups together (Energy FGR).

Together this information shows three main characteristics of the feature groups. Firstly, the spectral features are the most dominant in both adult and neonatal data. This is to be expected given the repetitive nature of seizures and the variety of spectral measures used. Secondly, the energy group performs significantly worse than the others with neonatal data, but there is less mutual information between it and the spectral group than between the spectral and structural groups. Therefore, when used with spectral features, the energy group adds significant information to the system. Finally, to answer one of the questions first posed at the beginning of the paper, there is a significant difference in the performance of features between adult and neonatal data.

VI. CONCLUSION

This paper has addressed the problem of age-independent seizure detection. This is the first study of this type known to the authors. The results shown in Table II show that it is possible to accurately detect seizures in both neonatal and adult data with one algorithm. The performance figures for both neonatal and adult data compare favourably to previous studies. It has also been shown that features extracted from the EEG contribute differently to neonatal and adult seizure detection. A medical device using this algorithm could potentially provide accurate seizure detection to patients of all ages.

REFERENCES

[1] Albert-Ludwigs-Universitat Freiburg. [Online]. Available: https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database

[2] J. Gotman, D. Flanagan, J. Zhang, and B. Rosenblatt, "Automatic seizure detection in the newborn: Methods and initial evaluation," *Electroenceph. Clin. Neurophysiol.*, vol. 103, pp. 356–362, 1997.

[3] B. R. Greene, W. P. Marnane, G. Lightbody, R. B. Reilly, and G. B. Boylan, "Classifier models and architectures for EEG-based neonatal seizure detection." *Physiol Meas*, vol. 29, no. 10, pp. 1157–1178, Oct 2008. [Online]. Available: http://dx.doi.org/10.1088/0967-3334/29/10/002

[4] E. M. Thomas, A. Temko, G. Lightbody, and G. B. Boylan, "A gaussian mixture model based statistical classification system for neonatal seizure detection," in *Submitted to IEEE Workshop on Machine Learning for Signal Processing (MLSP)*, 2009.

[5] B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, J. Timmer, and A. Schulze-Bonhage, "Do false predictions of seizures depend on the state of vigilance? A report from two seizure-prediction methods and proposed remedies." *Epilepsia*, vol. 47, no. 12, pp. 2058–2070, Dec 2006. [Online]. Available: http://dx.doi.org/10.1111/j.1528-1167.2006.00848.x

[6] A. Temko, E. Thomas, G. Boylan, W. Marnane, and G. Lightbody, "An SVM-based system and its performance for detection of seizures in neonates," submitted to 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2009.

[7] S. D. Faul, "Automated neonatal seizure detection," Ph.D. dissertation, University College Cork, 2007.

[8] A. Subasi, "Epileptic seizure detection using dynamic wavelet network," *Expert Systems with Applications*, vol. 29, pp. 343–355, 2005.

[9] A. J. C. Slooter, E. M. Vriens, F. S. S. Leijten, J. J. Spijkstra, A. R. J. Girbes, A. C. van Huffelen, and C. J. Stam, "Seizure detection in adult ICU patients based on changes in EEG synchronization likelihood." *Neurocrit Care*, vol. 5, no. 3, pp. 186–192, 2006. [Online]. Available: http://dx.doi.org/10.1385/NCC:5:3:186

[10] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. Schulze-Bonhage, and J. Timmer, "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic," *Physica D*, vol. 194, pp. 357–368, 2004.