Improved Detection and Classification of Convulsive Epileptic and Psychogenic Non-epileptic Seizures Using FLDA and Bayesian Inference

Shitanshu Kusmakar, Chandan K. Karmakar, Bernard Yan, Terence J.O'Brien, Marimuthu Palaniswami and Ramanathan Muthuganapathy

***Abstract*— A high number of patients with epileptic seizures (ES) are misdiagnosed due to prevalence of mimic conditions. The clinical characteristics of mimics are often similar to ES. The events mostly misdiagnosed are of psychogenic origin and are termed as psychogenic non-epileptic seizures (PNES). The gold standard for diagnosis of PNES is video-electroencephalography monitoring (VEM), which is a resource demanding process. Hence, need for a more object method of PNES diagnosis is created. Accelerometer sensors have been used previously for the diagnosis of ES. In this work, we present a new approach for detection and classification of PNES using wrist-worn accelerometer device. Various time, frequency and wavelet space features are extracted from the accelerometry signal. Feature compression is then performed using Fisher linear discriminant analysis (FLDA). A Bayesian classifier is then trained using kernel estimator method. The algorithm was trained and tested on data collected from 16 patients undergoing VEM. When tested, the algorithm detected all seizures with 20 false alarms and correctly classified 100% PNES and 75% ES, respectively of the detected seizures.**

I. INTRODUCTION

The diagnosis of epileptic seizures (ES) is often done on the basis of involuntary paroxysmal movements or loss of conscious behavior. Detection of ES on these basis is not enough as there are many events with overlapping characteristics. These events are called as epileptic mimics. A class of mimics that are of psychogenic origin are termed as psychogenic non-epileptic seizures (PNES). The mean latency in correct diagnosis of PNES is reported to be 5.2 years [1]. Long-term video-electroencephalography (EEG) monitoring (VEM) is the gold standard for detection of PNES but, is resource demanding. Thus, alternative methods have been attracting greater interest. In this work, we present an algorithm for automated detection of convulsive PNES and ES, using a wearable accelerometer sensor.

Previous studies have shown that various clues are present in movement data, which can differentiate ES from activities of daily living (ADL's) [2] [3] [4]. Cuppens *et. al.* [4] modeled ADL's and classified any anomalies as seizures. Poh *et. al.* [2] used combination of accelerometer and skin conductivity sensor for detection of ES. However, the use of accelerometry in classification of convulsive PNES and ES is reported rarely. Bayly *et. al.* [3] in their work showed that time frequency mapping of the limb movement data can differentiate convulsive PNES and ES. In addition, some groups have also investigated the use of surface electromyography (sEMG) [5], and EEG [6] [7] for automated classification of ES and PNES.

The classification of convulsive PNES and ES pose a challenge due to their overlapping characteristics. In this work, a novel method based on Fisher linear discriminant analysis (FLDA) and a Bayesian classifier is proposed. FLDA is a widely used discriminant criterion in face recognition problems. In our previous work [8], we had proposed classification algorithm based on k-means and SVM using dyadic scale features. In this work, we propose the use of FLDA to find the set of projection bases, such that the samples of every class could be maximally separated. A Bayesian classifier is then trained using the kernel estimator method for classifying unseen data into normal, ES, and PNES.

II. METHOD

Convulsive PNES and ES are generally associated with stereotypical limb movement. Bayly *et. al.* [3] in their work suggested that PNES events have a stable manifestation. However, the exact aetiology and characteristics of PNES remains unknown, and no direct causal relationship of PNES events have been identified. In this work, we present a framework (Fig. 1) for detection and classification of convulsive PNES using wrist-worn accelerometer sensor.

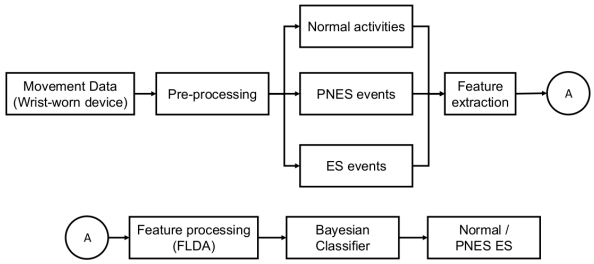


Fig. 1: Flowchart of the proposed methodology.

A. Data Collection

A prospective study was conducted at Royal Melbourne Hospital (RMH), Melbourne Australia. Patients with history of seizures admitted for VEM at RMH were recruited in the study. All patients were made to sign a consent form and the

S. Kusmakar and M. Palaniswami are with the Department of Electrical and Electronic Engineering, The University of Melbourne, Vic - 3052, Australia. skusmakar@student.unimelb.edu.au, palani@unimelb.edu.au

C. K. Karmakar is with Department of Electrical and Electronic Engineering, University of Melbourne, Vic - 3052, Australia and also with Deakin University Geelong, Vic - 3125, Australia. karmakar@unimelb.edu.au

B. Yan and T. O'Brien are with the Melbourne Brain Centre, Royal Melbourne Hospital, Dept. of Medicine, The University of Melbourne, Vic - 3052, Australia. Bernard.Yan@mh.org.au, obrientj@unimelb.edu.au

R. Muthuganapathy is with the Department of Engineering Design, Indian Institute of Technology Madras, India. mraman@iitm.ac.in

study was approved by the human research ethics committee of the RMH (HREC Project 300.259). The movement data was collected using a wrist-worn accelerometer device. The data was annotated by consulting epileptologist and epilepsy scientist while, being blinded to accelerometer traces. The annotated data served as the ground truth for supervised learning. Table I shows the overview of the collected data.

TABLE I: Table shows the patient statistics.

Demography	ES	PNES
Patients	$10 + 1^*$	$5 + 1^*$
Number of events	21	21
Age	30.00 ± 10.99	36.57 ± 14.72
Male:Female	$5 : 6$	$1 : 5$
Duration of events (seconds)	106.00 ± 117.01	210.00 ± 206.79

* The patient had co-morbid epilepsy with both ES and PNES events.

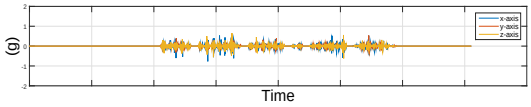
B. Pre-Processing

The raw accelerometer data is sampled at 50Hz and data is recorded for three axes with a time stamp. As movement data is recorded continuously for the complete duration of VEM. The movement data contains large packets of ADL's in comparison to seizure moves. Thus to reduce the computational workload, the raw data has to be pre-processed. The pre-processing steps have been explained in our previous work [8], and are briefly explained here.

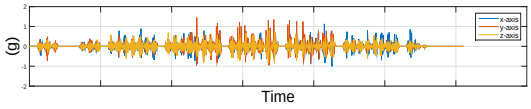
In the pre-processing steps, the data is passed through an activity and a time filter. The activity and time filtering steps selectively removes movement epoch's that are subtle or have duration $\leq 20s$. After the filtering step, all movement epoch's lying $\leq 20s$ apart are clustered as a single activity. At this stage, the movement data is divided into several activities of varied duration that includes: (1) ADL's, and (2) seizure moves (both ES and PNES). The pre-processed data corresponds to only 2% to 3% of the initial data. Fig. 2 shows the typical filtered normal, PNES and ES moves.

C. Feature Extraction

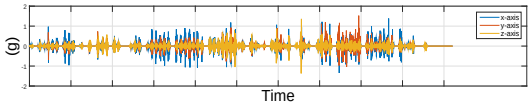
We extracted several time, wavelet and frequency domain features from the accelerometry data. Short time windows of 2.56sec duration with 50% overlap were selected in accordance with Bayly *et. al.* [3]. A set of 15 time-domain features including signal power, energy, measures of dispersion (maximum amplitude, inter-quartile range, standard deviation), measures of central tendency (mean, median, mode), skewness, kurtosis, entropy (log energy, shannon, norm, threshold), and zero-crossing rate were extracted from the four accelerometry channel (X, Y, Z, and $Resultant$). Further, it has been shown that the frequency foot-print of PNES events is rather stable over the course of an event in contrast to ES [3]. However, this frequency footprint can be scattered over a band of frequencies. Thus wavelets are employed and a six level wavelet decomposition is performed using 'Db5' as the mother wavelet. Features such as coefficient of variation of kurtosis and skewness are calculated for the first four details and approximate coefficient of the resultant accelerometry signal. In addition, sub-band power and entropy (Shannon) were also calculated for the first four details and approximate coefficient. The frequency foot-print of ADL's have a very narrow frequency band in comparison to seizure moves. Mean spectral power and energy of the norm of the movement data were also computed. In total we had a feature set comprising of 82 features.



(a) Typical normal event



(b) Typical ES event



(c) Typical PNES event

Fig. 2: Accelerometer traces of typical (a) Normal, (b) ES, and (c) PNES activity obtained after pre-processing.

D. Fisher Linear Discriminant Analysis

Fisher linear discriminant analysis or linear discriminant analysis is a linear transformation technique, used to compress the data and improve the class separability. FLDA is mostly regarded as a feature pre-processing technique before classification. FLDA can be used to compresses the data to $c - 1$ dimensions, where c is the number of classes. FLDA maximizes the discriminant that gives a large separation between the projected class means, while giving a small variance within each class, so as to minimize the class overlap. The projection of the data onto a lower dimension with good class separability prevents overfitting and reduces computational efficiency. The fisher criterion function is given by equation (1). The criterion function $J(\cdot)$ is maximized with respect to weight vector to find the specific choice of direction for the projection of the data.

$$J(\cdot) = \frac{w^T S_B w}{w^T S_w w} \tag{1}$$

where, S_B is the between class covariance matrix and S_w is the total within class covariance matrix. If, x_n is the d-dimensional feature vector, then the d-dimensional sample mean is given by equation 2 and S_B and S_w are given by equation 3 and 4.

$$m_i = \frac{1}{n_i} \sum_{n \in c_i} x_n \tag{2}$$

$$S_B = \sum_{i \in c_i} (m_{i+1} - m_i)(m_{i+1} - m_i)^T \tag{3}$$

$$S_w = \sum_{n \in c_i} (x_n - m_i)(x_n - m_i)^T \tag{4}$$

E. Kernel Density Estimators

A Bayesian classifier is build using a non-parametric kernel estimator method. A classifier based on Bayesian approach uses equation (5) to estimate the posterior probability.

$$p(\frac{\omega_i}{x}) = \frac{p(\frac{x}{\omega_i})p(\omega_i)}{p(x)} \tag{5}$$

The estimation of posterior requires calculating the probabilities $p(\frac{x}{\omega_j})$ and $p(\omega_j)$. The kernel estimator method finds these probabilities by putting a window (with defined width) using a kernel function. A window function is placed on every datum and the number of observations falling inside the window are determined. The probability density estimate or the parzen window estimate is defined as shown in equation (6).

$$p(x) = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{V}k(\frac{x - x_n}{h}) \tag{6}$$

where, N is the total number of data points, $V = h^D$ is the volume of the Parzen window and $k(\cdot)$ is the kernel function. However, the kernel estimator method might suffer from artificial discontinuities at the boundary of the window function. An alternative is to choose a smoother density model using a smoother kernel function like Gaussian. The Gaussian kernel density function is shown in equation (7).

$$p(x) = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{(2\pi h^2)^{\frac{D}{2}}}\exp\frac{-\parallel x - x_n \parallel^2}{2h^2} \tag{7}$$

where, h represents the standard deviation of the Gaussian components.

III. Results and Discussion

In total 700 ADL's, 21 PNES and 21 ES moves were recorded from 16 convulsive epileptic patients. The data was randomly split into a training set ($\approx 60\%$) and a test set ($\approx 40\%$) with approximately equal class proportion. The seizure data is representative of a class imbalanced data therefore, oversampling method SMOTE [9] was employed on the training set to reduce the class imbalance. FLDA was then employed as a discriminant criterion to perform compression of the input feature space thus, preventing overfitting of the data. The test data was then projected onto the bases obtained from the training set (Fig. 3). The approach presented here, has been developed keeping in mind the challenges observed in diagnosis of mimic conditions such as PNES.

The reduced feature set is then used to train a Bayesian classifier using a non-parametric kernel density estimation method. Non-parametric approaches does not make any assumptions about the probability distribution of the data and assumes parameters based on the training data. However, non-parametric approaches can be computationally slower than parametric approaches, owing to the fact that all training data is stored in non-parametric form. Thus, the use of FLDA with non-parametric approaches is largely justified.

Two classification models were learned from the training data. A Parzen window method and a gaussian kernel function were used for estimating the probability density model. The value of $h = 0.265$ was chosen from quantitative analysis of the training data. The learned models were then tested on the test set comprising of 16 seizures (8 PNES, and 8 ES events) and 280 ADL's. Table II shows the confusion matrix when Parzen window is chosen for density estimation.

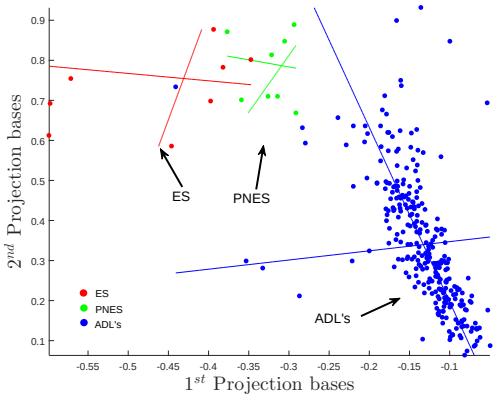


Fig. 3: A FLDA scatter plot of samples labeled as ADL's, PNES and ES corresponding to the test set.

TABLE II: Confusion matrix for Parzen window as estimator.

Class	ADL's	PNES moves	ES moves
ADL's	251	27	2
PNES moves	0	8	0
ES moves	0	3	5

We observed that, all seizure events are correctly detected by the algorithm. However, three ES events are misclassified as PNES. A total of 29 false alarms were also obtained (Table II). Majority of the false alarms were raised due to misclassification of ADL's as PNES move. Whereas, all the PNES events are correctly identified using the Parzen window estimation method. Indicating, a close packed structure of PNES events. Thus, all PNES events can be said to have characteristics similar to each other. The results also suggest that ADL's and ES events form distinct classes with least overlap, and ES events can be characterized as anomalies in the data (Fig. 3, and Table II). These findings are in agreement with Cuppens *et. al.* [4], and Poh *et. al.* [2] who distinguished ES events as anomalies.

TABLE III: Confusion matrix for Gaussian as a kernel function.

Class	Normal moves	PNES moves	ES moves
Normal moves	260	18	2
PNES moves	0	8	0
ES moves	0	2	6

The performance of the proposed algorithm further improves with the use of a Gaussian kernel (Table III) (Table IV). The Gaussian kernel with same parameter h, outperforms Parzen window method (Fig. 4). A Gaussian kernel accounts for the discontinuities created at the edges of the window, thus obtaining a smoother density model so

that the density is normalized correctly. Therefore, reducing the number of false alarms to 20. Table IV shows the performance measures of the algorithm with Gaussian as the kernel function.

TABLE IV: Table showing the performance measures of the algorithm with Gaussian as a kernel function.

Measure	ADL's	PNES moves	ES moves
Sensitivity	0.92	1.00	0.75
Specificity	1.00	0.93	0.99
Precision	1.00	0.28	0.75
F1-score	0.96	0.44	0.75
Overall Accuracy	0.92		

The proposed algorithm detected all 16 (8 ES, 8 PNES) events in the test set. In addition, the algorithm correctly classified all 8 PNES events (sens: 100%) and 6 of 8 ES events (sens: 75%) (Table IV). Further, our results also validate the hypothesis that convulsive ES events can be differentiated from mimics based on the differing evolutionary patterns of the events [3].

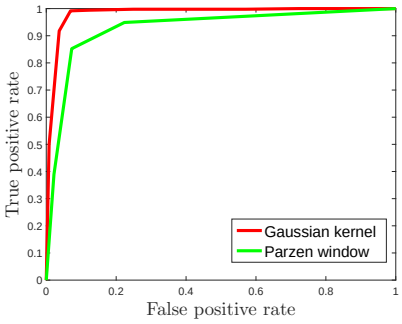


Fig. 4: The Red and green curve represents the ROC curves of system with Parzen window and Gaussian kernel function as estimation method. The Gaussian kernel outperforms Parzen window method.

The results shows the potential of the proposed algorithm in detection and classification of convulsive PNES and ES using wrist-worn accelerometer sensor. Table V shows the comparison of the proposed algorithm with existing PNES classification algorithms.

TABLE V: Comparison of the proposed algorithm with other PNES classification algorithms.

Classification Algorithm		Sensitivity	Specificity	Accuracy
Pippa *et. al.* [7]	Bayes Net (EEG)	94.00%	98.00%	95.00%
Poulos *et. al.* [6]	Cross-correlation (EEG)	83.00%	90.00%	86.00%
Beniczky *et. al.* [5]	Zero crossings (sEMG)	95%	96.00%	95%
Kusmakar *et. al.* [10]	GMM (ACM)	87.50%	91.00%	91.00%
Proposed Algorithm	FLDA (ACM)	100%	93.00%	92.00%

EEG: electroencephalography, sEMG: surface electromyography, ACM: accelerometer.

A Bayesian approach based on non-parametric kernel density estimators is advantageous as no computation is involved in the training phase. Non-parametric approaches only require storing of the training samples. However, this also becomes a limitation as the computational cost of finding the probability density increases linearly with data. These, limitations can however be overcome using compression strategies like FLDA, which also enhance classification performance of the algorithm.

IV. Conclusion

In this work an approach based on FLDA and Bayesian inference using a non-parametric kernel density estimation method is employed for automated detection and classification of convulsive PNES, ES from accelerometry signal. When tested, the algorithm correctly detected all seizure events (8 ES, and 8 PNES) and 260 (92.85%) of 280 ADL's with 20 false alarms. In addition, the algorithm correctly classified 8 (100%), and 6 (75%) of the detected seizure events as PNES and ES, respectively. The proposed system can provide diagnostic monitoring of epileptic patients in a non-invasive and timely manner.

References

[1] S. G. Jones, T. J. O'Brien, S. J. Adams, R. Mocellin, C. J. Kilpatrick, R. Yerra, J. H. Lloyd, and D. Velakoulis, "Clinical characteristics and outcome in patients with psychogenic nonepileptic seizures," *Psychosomatic medicine*, vol. 72, no. 5, pp. 487–497, 2010.

[2] M.-Z. Poh, T. Loddenkemper, C. Reinsberger, N. C. Swenson, S. Goyal, M. C. Sabtala, J. R. Madsen, and R. W. Picard, "Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor," *Epilepsia*, vol. 53, no. 5, pp. e93–e97, 2012.

[3] J. Bayly, J. Carino, S. Petrovski, M. Smit, D. A. Fernando, A. Vinton, B. Yan, J. R. Gubbi, M. S. Palaniswami, and T. J. O'Brien, "Time-frequency mapping of the rhythmic limb movements distinguishes convulsive epileptic from psychogenic nonepileptic seizures," *Epilepsia*, vol. 54, no. 8, pp. 1402–1408, 2013.

[4] K. Cuppens, P. Karsmakers, A. Van de Vel, B. Bonroy, M. Milosevic, S. Luca, T. Croonenborghs, B. Ceulemans, L. Lagae, S. Huffel *et al.*, "Accelerometry-based home monitoring for detection of nocturnal hypermotor seizures based on novelty detection," 2014.

[5] S. Beniczky, I. Conradsen, M. Moldovan, P. Jennum, M. Fabricius, K. Benedek, N. Andersen, H. Hjalgrim, and P. Wolf, "Automated differentiation between epileptic and nonepileptic convulsive seizures," *Annals of neurology*, vol. 77, no. 2, pp. 348–351, 2015.

[6] M. Poulos, F. Georgiacodis, V. Chrissikopoulos, and A. Evangelou, "Diagnostic test for the discrimination between interictal epileptic and non-epileptic pathological eeg events using auto-cross-correlation methods," *American journal of electroneurodiagnostic technology*, vol. 43, no. 4, pp. 228–240, 2003.

[7] E. Pippa, E. I. Zacharaki, I. Mporas, V. Tsirka, M. P. Richardson, M. Koutroumanidis, and V. Megalooikonomou, "Improving classification of epileptic and non-epileptic eeg events by feature selection," *Neurocomputing*, vol. 171, pp. 576–585, 2016.

[8] J. Gubbi, S. Kusmakar, A. Rao, B. Yan, T. O'Brien, and M. Palaniswami, "Automatic detection and classification of convulsive psychogenic non-epileptic seizures using a wearable device," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, no. 99, pp. 1–1, 2015.

[9] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[10] S. Kusmakar, R. Muthuganapathy, B. Yan, T. J. O'Brien, and M. Palaniswami, "Gaussian mixture model for the identification of psychogenic non-epileptic seizures using a wearable accelerometer sensor," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 1006–1009.