Expert Systems With Applications 86 (2017) 199–207

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Supervised learning in automatic channel selection for epileptic seizure detection

Nhan Duy Truong[a], Levin Kuhlmann[b,c], Mohammad Reza Bonyadi[d,e], Jiawei Yang[f], Andrew Faulks[g], Omid Kavehei[a,*]

[a] School of Engineering, Royal Melbourne Institute of Technology (RMIT), Melbourne, VIC 3000, Australia
[b] Centre for Human Psychopharmacology, Swinburne University, Hawthorn, VIC 3122, Australia
[c] Neuroengineering Laboratory, Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3010, Australia
[d] Centre for Advanced Imaging (CAI), University of Queensland, QLD 4072, Australia
[e] Optimization and Logistics Group, University of Adelaide, SA 5005, Australia
[f] Wenzhou Medical University, 268 Xueyuan West Rd, Wenzhou, China
[g] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Clayton South, VIC 3169, Australia

ARTICLE INFO

Article history:
Received 1 February 2017
Revised 21 April 2017
Accepted 20 May 2017
Available online 29 May 2017

Keywords:
Seizure detection
iEEG
Random Forest
Automatic channel selection

ABSTRACT

Detecting seizure using brain neuroactivations recorded by intracranial electroencephalogram (iEEG) has been widely used for monitoring, diagnosing, and closed-loop therapy of epileptic patients, however, computational efficiency gains are needed if state-of-the-art methods are to be implemented in implanted devices. We present a novel method for automatic seizure detection based on iEEG data that outperforms current state-of-the-art seizure detection methods in terms of computational efficiency while maintaining the accuracy. The proposed algorithm incorporates an automatic channel selection (ACS) engine as a pre-processing stage to the seizure detection procedure. The ACS engine consists of supervised classifiers which aim to find iEEG channels which contribute the most to a seizure. Seizure detection stage involves feature extraction and classification. Feature extraction is performed in both frequency and time domains where spectral power and correlation between channel pairs are calculated. Random Forest is used in classification of interictal, ictal and early ictal periods of iEEG signals. Seizure detection in this paper is retrospective and patient-specific. iEEG data is accessed via Kaggle, provided by International Epilepsy Electro-physiology Portal. The dataset includes a training set of 6.5 h of interictal data and 41 min in ictal data and a test set of 9.14 h. Compared to the state-of-the-art on the same dataset, we achieve 2 times faster in run-time seizure detection. The proposed model is able to detect a seizure onset at 89.40% sensitivity and 89.24% specificity with a mean detection delay of 2.63 s for the test set. The area under the ROC curve (*AUC*) is 96.94%, that is comparable to the current state-of-the-art with *AUC* of 96.29%.

1. Introduction

Epileptic seizure affects nearly 1% of global population but only two thirds can be treated by medicine and approximately 7 – 8% can be cured by surgery (Litt & Echauz, 2002). Therefore, seizure onset detection and subsequent seizure suppression becomes important for the patients that cannot be cured by neither drug nor surgery. Early detection can allow early electrical stimulation to suppress the seizure (Echauz et al., 2007). In this paper, we focus on how to effectively and reliably detect seizure onset based on iEEG patterns. Note that cause and treatment of epilepsy is beyond the scope of this paper.

EEG has been commonly used in brain-computer interface thanks to the convenient real-time readings and high temporal resolution of EEG signals (Zeng & Song, 2015; Zhang, Yang, & Guan, 2013). In recent years, EEG has provided a promising possibility to detect and even predict an epileptic seizure (Fatichah, Iliyasu, Abuhasel, Suciati, & Al-Qodah, 2014; Kuhlmann et al., 2009; Osorio & Frei, 2009; Parvez & Paul, 2015; Saab & Gotman, 2005; Tieng, Kharatishvili, Chen, & Reutens, 2016). For seizure detection, Fatichah, Iliyasu, Abuhasel, Suciati, and Al-Qodah (2014) used

* Corresponding author.
E-mail addresses: nhanduy.truong@rmit.edu.au (N.D. Truong), lkuhlmann@swin.edu.au (L. Kuhlmann), reza@cai.uq.edu.au (M.R. Bonyadi), yangjw@wibe.ac.cn (J. Yang), andrew.faulks@csiro.au (A. Faulks), omid.kavehei@rmit.edu.au (O. Kavehei).

http://dx.doi.org/10.1016/j.eswa.2017.05.055
0957-4174/

Table 1
Summary of existing EEG-based seizure detection methods .

Reference	EEG type	No. of patients	No. of seizures	Data duration		Patient-specific	Split data for training	Testing sensitivity	FDR[a]	Mean detection delay
				ictal	interictal					
Saab and Gotman (2005)	scalp	44	195	1012 h[b]		No	64%	76%	0.34/h	9.8s
Kuhlmann et al. (2009)	scalp	21	88	525 h[b]		No	70%	81%	0.60/h	16.9s
Wang et al. (2016)	scalp	10	44	72 min	121 h	Yes	80%	91.44%	99.34%	n/a
Zabihi et al. (2016)	scalp	24	161	2.55 h	169 h	Yes	25%	88.27%	93.21%	n/a
Fatichah, Iliyasu, Abuhasel, Suciati, and Al-Qodah (2014)	intracranial[c]	n/a	n/a	39.3 min	2.62 h	n/a	90%	94.55%	98.41%	n/a
Hills (2014)	intracranial	12	48	41 min	6.5 h	Yes	50%	91.33%	94.02%	3.17s
Parvez and Paul (2015)	intracranial	21	87	58 h	490 h	n/a	80%	100%	97%	n/a

[a] False detection rate (FDR) or specificity.
[b] Duration of ictal and interictal were not provided separately.
[c] Intracranial EEG for seizure class and both intracranial and extracranial for non-seizure class.

a combination of principle component analysis (PCA) and neural network with fuzzy membership function that can achieve accuracy rate up to 97.64%. Tieng, Kharatishvili, Chen, and Reutens (2016) combined wavelet de-noising with adapted Continuous Wavelet Transform in their algorithm and were able to achieve sensitivity of 96.72% and specificity of 94.69% with EEG data from mice. Another remarkable method is to transform EEG signals into images so as to leverage image processing techniques (Parvez & Paul, 2015). This approach was able to obtain 98.91% sensitivity and 94.35% specificity. Zabihi et al. (2016) reconstructed EEG phase spaces using time-delay embedding method and PoinCare section. The phase spaces were then reduced by PCA before being fed to linear discriminant analysis (LDA) and Naive Bayesian classifiers. This approach achieved 88.27% sensitivity and 93.21% specificity in seizure detection.

Shoeb (2009) deployed 8 filters spanning the frequency range of 0.5–24 Hz for each 2-s EEG epoch of all channels, then concatenated 3 epochs to form a feature set to be fed to a SVM classifier. This approach was tested with the CHB-MIT EEG dataset and was able to detect 96% of 163 test seizures with a mean detection delay of 4.6 seconds. Using the same CHB-MIT dataset, EEG signal was transformed into an image representation using 2-D projection of the patient electrodes and the magnitude of 3 different frequency bands spanning the range of 0–49 Hz of each 1 s block of EEG signal (Thodoroff, Pineau, & Lim, 2016). The recurrent convolutional neural network took 30 consecutive blocks as inputs to perform feature extraction and classification. The patient-specific detectors in this method have comparable performance compared to the proposed method by Shoeb (2009).

Prominent feature extraction techniques consider characteristics in both frequency and time domain. As an efficient tool for time-frequency-energy analysis, wavelet-based filters were used to extract a ratio of seizure content of the short foreground in comparison with the background (Osorio & Frei, 2009; Saab & Gotman, 2005). Saab and Gotman (2005) applied Bayes' formula on extracted features to estimate the probability of seizure in EEG signals. This method achieved an impressively short onset detection delay of 9.8 s with 76% sensitivity and 0.34/h false positive rate. Kuhlmann et al. (2009) extended Saab and Gotman's method by combining extra features to find a superior detector. Their method was able to achieve a sensitivity of 81%, a false positive rate of 0.60/h, and a median detection delay of 16.9 s on a dataset of 525 h of scalp EEG data.

The current state-of-the-art seizure detection method proposed by Hills (2014) for the dataset considered here is implemented and extended in this paper. The dataset is derived from a Kaggle seizure detection competition in which Hills (2014) scored *AUC* of 96.29% and announced as the winner. Description of the dataset is provided in Section 2.1. In this paper, we significantly enhanced computational efficiency of Hill's method by employing an automatic channel selection algorithm. This enabled us to process data as accurately with reduced number of channels. Table 1 summarizes the existing EEG-based seizure detection methods in recent years. We have made the research's source code publicly available on GitHub via https://goo.gl/Bc89mJ.

The remainder of this paper is organized as follows. In Section 2, after describing the dataset, we propose automatic channel selection engine that helps to reduce the number of channels to be processed. This section also presents spatio-temporal feature extraction and Random Forest classifier used for seizure detection. Section 3 evaluates the performance of the proposed model with comparison against the state-of-the-art method on the same dataset. Section 4 concludes the achievement of the paper.

2. Proposed method

The intracranial EEG data was recorded on multiple subjects with varying number of channels and sampling rates. We propose an automatic channel selection engine to filter out channels which are less relevant to seizure. The engine accepts raw iEEG data, their corresponding labels, and the number of channels to be selected, *M*, and determines indexes of channels that are most relevant for seizure detection. Indexes of these *M* channels are stored on hard-disk so the engine only needs to be executed one time at the beginning for each subject. Feature extraction was performed in both frequency and time domain on the selected channels. Information extracted in frequency and time domains was concatenated and fed to a Random Forest classifier. Fig. 1 presents flowchart of the proposed method.

2.1. Dataset

Dataset being analyzed in this paper is obtained from Kaggle (2014). Intracranial EEG signals were recorded from 4 dogs and 8 patients with epileptic seizures. Recordings were sampled at 400 Hz from 16 electrodes for dogs, and sampled at 500 Hz or 5 kHz from varying number of electrodes (ranging from 16 to 72) for humans. The data was pre-organized into 1 s iEEG epochs annotated as ictal for seizure states or interictal for seizure-free states. Interictal data was captured not less than one hour before or after a seizure onset and randomly chosen from the recorded data. Each ictal segment also came with the time in seconds between the seizure onset and first data point of the segment. The training dataset is consisted of 41 min of ictal data and 6.5 h of interictal data. Summary of the training dataset is presented in Table 2. Note that early ictal state in this paper is the ictal state occurring within the first 15 s from the seizure onset. The proposed