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

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 harddisk 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

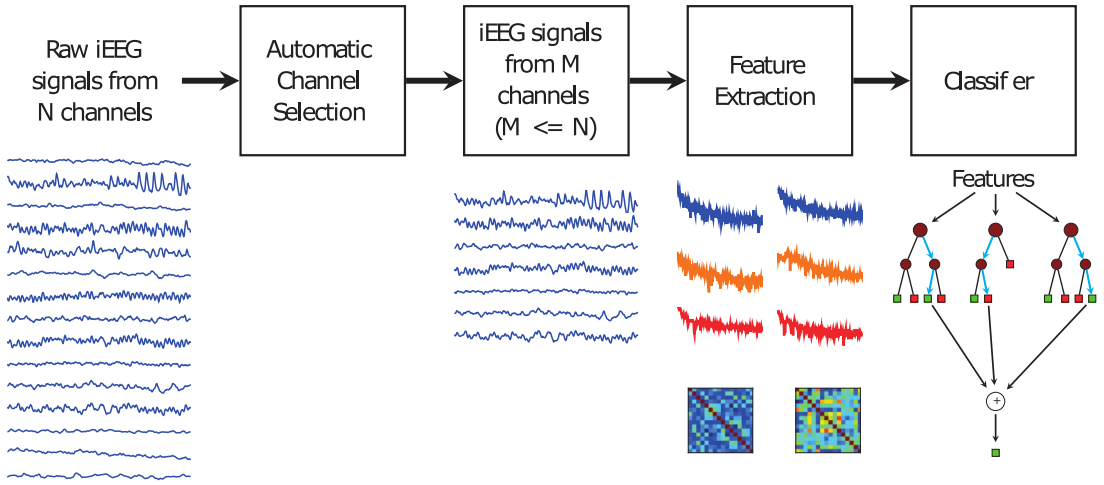

Fig. 1. Flowchart of the proposed method. Raw iEEG data from all *N* channels is fed to ACS to find *M* channels which contribute the most to a seizure. The ACS engine is executed one time only for each subject at the beginning and indexes of the *M* channels are stored on hard-disk. Feature extraction in both frequency and time domains is done on the *M* channels. Extracted features are fed to a classifier using Random Forest algorithm to discriminate interictal, ictal and early ictal epochs.

Table 2
Summary of the dataset.

Subject	No. of electrodes	Ictal data length (s)	Interictal data length (s)	Unlabeled data length (s)	Train/Test ratio
Dog–1	16	178	418	3181	0.19
Dog–2	16	172	1148	2997	0.44
Dog–3	16	480	4760	4450	1.18
Dog–4	16	257	2790	3013	1.01
Patient–1	68	70	104	2050	0.08
Patient–2	16	151	2990	3894	0.81
Patient–3	55	327	714	1281	0.81
Patient–4	72	20	190	543	0.39
Patient–5	64	135	2610	2986	0.92
Patient–6	30	225	2772	2997	1
Patient–7	36	282	3239	3601	0.98
Patient–8	16	180	1710	1922	0.98
Total		**2477**	**23445**	**32915**	**0.79**

method was tested with a hidden dataset provided by Kaggle. This dataset consists of 9.14 h of unlabeled iEEG data (Kaggle, 2014).

2.2. Automatic channels selection

The intracranial EEG data was recorded using various number of channels (16, 30, 36, 55, 64, 68, 72). Large number of channels yields higher computational complexity as it requires more data to be analyzed. This can also deteriorate the diversity of iEEG data, hence degrade the performance of seizure detection, because some channels may capture irrelevant information (Guyon & Elisseeff, 2003). One can leverage bio-medical knowledge to manually select which channels genuinely contribute to the seizure. However, it is hard, if not impossible, to disclose a set of channels that are significant for all subjects. It is required to use the expertise to analyze every subject (or group of subjects) to proclaim a list of significant channels with regards to each subject (or group of subjects) which is manifestly a time-consuming task. There have been attempts to reduce the number of channels to be analyzed (Duun-Henriksen et al., 2012; Shih, Shoeb, & Guttag, 2009) or reduce number of features extracted prior to classification (Minasyan, Chatten, Chatten, & Harner, 2010; Subasi & Ismail Gursoy, 2010). Duun-Henriksen et al. (2012) proposed an automated channel selection based on variance of EEG signal amplitude where channels with largest variance would be chosen. The detection performance using 3 channels selected by their algorithm was similar to using 3 channels selected by a clinical neurophysiologist. Shih, Shoeb, and Guttag (2009) used a greedy backward elimination algorithm to find the subset of features that results in lowest false positive rate. Seven features are extracted per each channel. The algorithm starts with all features and gradually removes the least influential ones by doing cross-validation on all subset of features. The authors were able to reduce the number of channels from 18 to 4.6 with an improvement in FPR (0.35 to 0.19/h) but degradation in sensitivity (from 99% to 97%) and detection delay (from 7.8 to 11.2 s). This approach, however, is less favorable when the number of channels per subject is too high because the number of subsets increases exponentially. Minasyan, Chatten, Chatten, and Harner (2010) performed feature selection using mutual information between individual features and output. Features with less mutual information are discarded. It is worth noting that this feature selection has to be performed not only during training but also during run-time classification. In other words, this approach induces extra processing time and makes it less suitable for portable device implementation. Subasi and Ismail Gursoy (2010) applied PCA, ICA, LDA for feature dimension reduction and used SVM classifier to distinguish between seizure and non-seizure segments. They tested with an 80 min subset of University Hospital of Bonn's dataset. They showed LDA achieved the best performance while PCA obtained the worst. However, their approaches also induce extra processing time during both training and run-time classification, similar to the work proposed by Minasyan, Chatten, Chatten, and Harner (2010).

We propose a novel approach for automatic channels selection (ACS) as follows. The approach is designed to be run offline in order to select channels for future online analysis. The labeled data is first transformed to obtain frequency information. Specifically, FFT

is applied onto the raw iEEG data on all N channels. FFT values are then sliced to extract data in 1-Hz bins in the range of 1–47 Hz. $\log_{10}$ is then applied to the magnitudes. The transformed data is a $N \times 47$ matrix where 47 is the number of 1-Hz bins in the range of 1–47 Hz. If the channels correlation is involved in ACS stage, it will be confusing to identify which channels are the most important based on the importance level of the correlation between each pair of channels. Therefore, the correlation among channels is disregarded in this stage. Each individual channel becomes a feature to be fed to classifiers. One or a set of classifiers determine the importance level of each feature or channel. There are several options of classifiers using different ensemble algorithms such as Gradient Boosting, AdaBoost and Random Forest. If multiple classifiers are used, the final importance level of each channel is the sum of importance values obtained from all classifiers. The measure of feature importance in this paper is implemented using scikit-learn ensemble library (Scikit-learn, 2014). The importance of a feature is estimated by how often that feature is used in split points of each individual decision tree of the ensemble classifier (Scikit-learn, 2014). It is important to note that only train dataset was involved in the ACS stage.

The output of the channel selection algorithm is a set of M channels sorted based on the level of their contribution to the detection of a seizure. In this paper, we selected the value of M through some experiments aiming at maximizing the final AUC score. Particularly, we gradually drop channels with lowest rank and check the cross-validation performance using the rest of channels. It is important to differentiate our approach to the one proposed by Shih, Shoeb, and Guttag (2009) that we are able to rank the channels prior to channel reduction. This helps us to know which channel should be dropped at each round, instead of extensively dropping one by one channel and comparing performance for all cases to decide which channel should be dropped. We also use coarse to fine approach to accelerate this selection. Typically, we start the selection by dropping 20% of channels with lowest rank until the performance decreases. Then we start next round with the number of channels present just before the aforementioned performance drops.

2.3. Feature extraction

2.3.1. Feature extraction in frequency domain

The iEEG signals from M selected channels are transformed by FFT. The transformed data then is filtered to discard high frequency noise and low frequency artifacts. Frequency range of 1–47 Hz was shown to achieve the best performance for the dataset (Hills, 2014). Eigenvalues have been used as an effective technique to discriminate ictal epochs in Hills (2014), Sardouie, Shamsollahi, Albera, and Merlet (2015) and Zhang and Parhi (2016). In order to compute eigenvalues, spectral power is primarily normalized (zero mean and standard deviation of one) along each channel before estimating cross spectral matrix (Hills, 2014). Contrary to the Hills feature extraction, we did not use cross spectral coefficients as a feature because our empirical observation shows that such feature could worsen detection accuracy. Sample recordings and corresponding power spectrum for ictal and interictal segments of Patient–1 are illustrated in Figs. 2 and 3.

The feature set in frequency domain consists of:

- Spectral power in 1 Hz bins in range of 1–47 Hz by applying $\log_{10}$ to the magnitude of FFT transformation, and
- Eigenvalues, sorted in descending order, of cross spectral matrix on all selected channels of the above spectral power.

2.3.2. Feature extraction in time domain

Raw iEEG signals are firstly re-sampled to 400 Hz. Similarly to frequency domain, filtered iEEG data is normalized to zero mean and unity standard deviation along each channel prior to computing covariance matrix and its eigenvalues. As illustrated in Fig. 4, iEEG data from 16 selected channels of Patient–1 have a very low correlation to each others in interictal states. The correlation slightly increases when seizure is at early state and becomes remarkable beyond the early state.

The feature set in time domain consists of:

- Coefficients in upper triangle of correlation matrix of iEEG signals from selected channels, and
- Eigenvalues of the correlation matrix above, sorted in descending order.

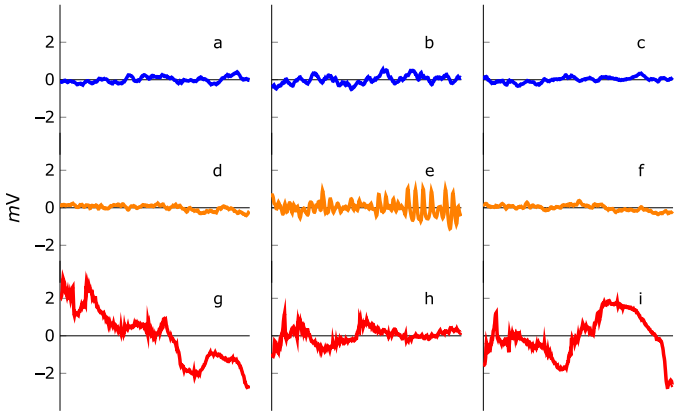

Fig. 2. Sample 1 s iEEG recordings. (a, b and c) interictal; (d, e and f) ictal at early state (within 15 s from seizure onset); (g, h and i): ictal after early state. iEEG signals presented in one column, (e.g. a, d and g) are recorded from the same channel.

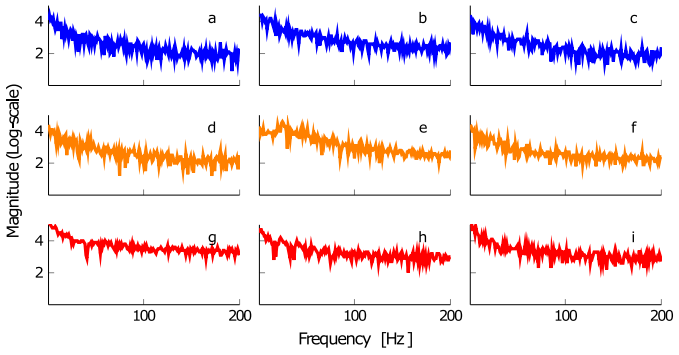

Fig. 3. Sample 1 s iEEG recordings power spectrum. (a, b and c) interictal; (d, e and f) ictal at early state; (g, h and i): ictal after early state. iEEG signals presented in one column, (e.g. a, d and g) are recorded from the same channel. Subplots in this figure are one-by-one associated with subplots in Fig. 2.

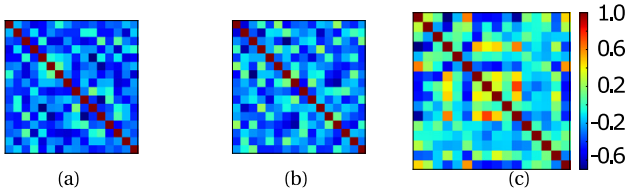

Fig. 4. Covariance matrix: (a) interictal; (b) ictal at early state; (c) ictal. Correlation between channels is very low in interictal period. The channels are more correlated after the seizure onset and highly correlated in ictal state.

2.4. Classifier

Random Forest algorithm was first proposed by Breiman (2001). The algorithm uses a large set of decision trees to acquire an average results. Random Forest has been shown with good

performance on dataset with high dimensional datasets in biology and medical fields (Cabezas, Galleguillos, & Perez-Quezada, 2016; Huynh et al., 2016; Scornet, 2016). This paper will not go in deep about its mathematical properties as they can be found in Breiman (2001); Scornet (2016) but rather on fine-tuning the parameters to achieve the highest performance with the given feature sets.

Random Forest classifier in this paper is implemented using scikit-learn library (Scikit-learn, 2014). Parameters of the classifier are reused from the approach proposed by Hills (2014) with 3000 decision trees. The classifier analyses each 1 s iEEG epoch and categorizes them into 3 classes as outputs: early ictal (ictal within 15 s from the onset), ictal, and interictal. Regarding sensitivity, specificity, F1-score and detection delay evaluation, the Random Forest classifier is adjusted from three-class classifier to binary classifier which detects whether a 1 s iEEG signal is ictal or interictal.

2.5. System evaluation

Here our method is compared with a visual inspection based focal channel selection (channels where seizures first appear) method and the variance-based method (Duun-Henriksen et al., 2012). We skip the method introduced by Shih, Shoeb, and Guttag (2009) because we have high number of channels, the greedy backward elimination method becomes impractical. For patient–4 with 72 channels, for example, we need to evaluate a factorial of 72 subsets in the worst case. Though we can stop the search when the performance of subsets starts to drop, the number of cases to be analyzed is still huge.

Metrics used to test the proposed approach are area under the receiver operating characteristic curve (*AUC*), sensitivity, specificity, F1-score and onset detection delay. To have a robust evaluation, we follow a leave-one-out cross-validation approach for each subject. If a subject has $\mathcal{N}$ seizures, $(\mathcal{N}-1)$ seizures will be used for training and the withheld seizure for validation. This round is repeated $\mathcal{N}$ times so all seizures will be used for validation exactly one time. Interictal segments are randomly split into $\mathcal{N}$ parts. $(\mathcal{N}-1)$ parts are used for training and the rest for validation. The metrics to be reported are the average of all rounds. The cross-validation will be based on the labeled iEEG dataset of 7.2 h. We will also test the system after being trained by 7.2 h of labeled iEEG data with the hidden dataset consisting of 9.14 h of unlabeled iEEG data provided by Kaggle.

The dataset used in this work is from a competition to detect whether a given 1 s iEEG segment represents a seizure and whether that segment is within the first 15 s (early) of its respective seizure. The overall *AUC* is the average of the two *AUC*s of the two detections (Kaggle, 2014), and is given by

$$AUC = \frac{1}{2}(AUC_S + AUC_E), \tag{1}$$

where,

- AUC_S is *AUC* for two classes: ictal (including early seizure) and interictal, and
- AUC_E is *AUC* for two classes: early seizure and non-early-seizure (including ictal states after 15 s from onset and interictal states).

Sensitivity, specificity and F1-score are commonly used in evaluating a seizure detection system and are given by

$$\text{Sensitivity (or Recall)} = \frac{TP}{TP+FN} \tag{2}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{5}$$

where,

- *TP* is the total number of 1 s ictal segments are correctly classified as ictal, and
- *TN* is the total number of 1 s interictal segments are correctly classified as interictal, and
- *FP* is the total number of 1 s interictal segments are wrongly classified as ictal, and
- *FN* is the total number of 1 s ictal segments are wrongly classified as interictal.

3. Results

Using same setup proposed by Duun-Henriksen et al. (2012), 3 channels are selected for each method. Fig. 5 illustrates the iEEG signal from seizure onset for a seizure of Dog–1. For this subject, selected channels using focal, variance-based, and our methods are (9, 10, 13), (3, 8, 9), and (4, 10, 12) respectively. A completed set of channels selected using the three methods is presented in Table A1. To benchmark the efficiency of the three methods, we calculated *AUC* through leave-one-out cross-validation as shown in Fig. 6 and more details in Table 3. As seen from Fig. 6, our method is better than the other two. Here we use a two-tailed signed rank test at a significance level of 0.05 to compare the three methods. Since there are multiple comparisons, the significance level should be adjusted to be $0.05/3 = 0.01667$ using Bonferroni correction. A two-tailed signed rank test on *AUC* scores of focal channel method and our method has p-value of 0.373 which indicates that the two methods has no statistical difference at the adjusted significance level of 0.01667. However, the same test between variance-based method and our method shows a significant result at p-value of 0.0076. This result confirms our channel selection method superior to the variance-based method proposed by Duun-Henriksen et al. (2012).

We now compare the efficacy of the proposed method with the current state-of-the-art method proposed by Hills (2014) on the same dataset. The algorithm was implemented in Python 2.7 in Ubuntu 14.04 LTS. Random forest classifier was implemented using scikit-learn library (Scikit-learn, 2014). FFT was performed with numpy library. All simulations were performed on a workstation with CPU Xeon (4 cores enabled) and 16 GB of RAM. Using our approach, average number of channels can be reduced from 35.1 to 10.3. Consequently, training time and test time are improved by 39.2% and 49.7% respectively (see Table 4). With 49.7% reduction in test time, our seizure detection system is 2 times faster at run-time than Hills approach. Therefore, the automatic channel selection is promising for real-time seizure detection application.

Table A2 describes the comparison between the state-of-the-art and proposed method on *AUC*, sensitivity, specificity, F1-score and onset detection delay for the modified leave-one-out cross-validation applied to the training set. Regarding sensitivity and specificity evaluation, the Random Forest classifier is adjusted from three-class classifier to binary classifier which detects whether a 1 s iEEG signal is ictal or interictal. The classifier's outputs range from 0 to 1 indicating how likely the input signal is ictal. The threshold of the classifier's output used to separate whether a 1 s iEEG segment is ictal or interictal was determined per subject. The value of threshold was selected to achieve the balance between sensitivity and specificity (i.e., the higher threshold value yields the higher specificity but the lower sensitivity and vice versa). Moreover, the threshold selection must result in similar specificity scores across the methods to have a meaningful sensitivity and mean detection delay comparisons. As seen from Table A2, our

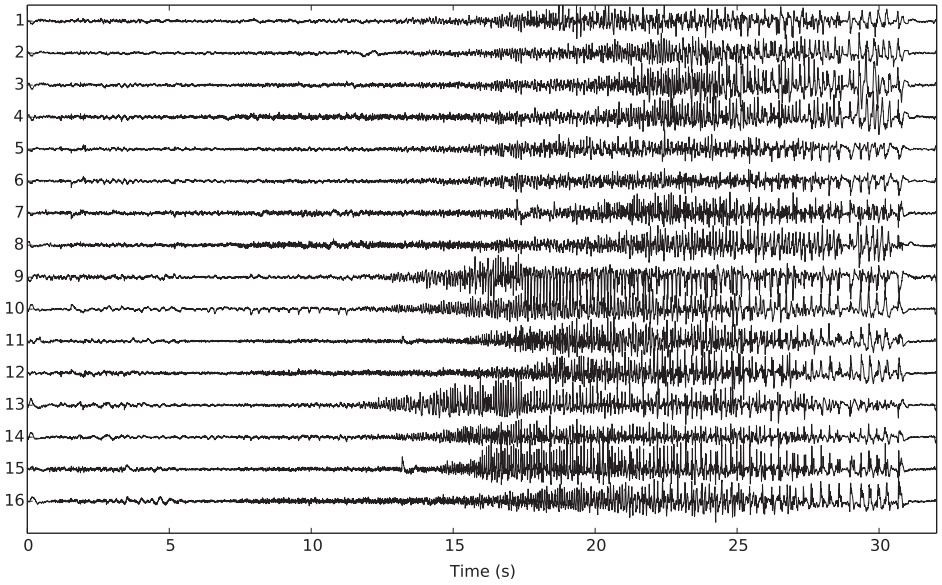

Fig. 5. iEEG recording of a seizure since its onset from Dog–1. Selected channels using focal, variance-based, and our methods are (9, 10, 13), (3, 8, 9). and (4, 10, 12) respectively.

Table 3
Comparison among three channel selection methods.

Subject	Focal channel			Variance-based			Proposed method		
	AUC_E (%)	AUC_S (%)	AUC (%)	AUC_E (%)	AUC_S (%)	AUC (%)	AUC_E (%)	AUC_S (%)	AUC (%)
Dog-1	95.02	98.46	96.74	95.03	98.47	96.75	94.79	98.44	96.62
Dog-2	92.32	95.20	93.76	91.31	96.74	94.03	94.35	96.30	95.32
Dog-3	95.72	98.92	97.32	92.05	98.02	95.03	95.90	98.88	97.39
Dog-4	98.76	91.11	94.94	98.49	89.15	93.82	98.79	94.62	96.71
Patient-1	94.29	96.46	95.38	71.72	89.77	80.75	82.03	93.49	87.76
Patient-2	99.22	98.75	98.98	99.20	99.04	99.12	99.17	98.72	98.95
Patient-3	88.62	92.75	90.68	87.89	92.25	90.07	87.61	94.68	91.15
Patient-4	97.05	97.05	97.05	94.82	94.82	94.82	100	100	100
Patient-5	76.69	88.31	82.50	65.86	77.16	71.51	69.73	83.83	76.78
Patient-6	97.33	99.51	98.42	74.48	87.75	81.11	97.56	99.61	98.58
Patient-7	79.22	79.41	79.32	71.13	76.75	73.94	89.21	94.24	91.72
Patient-8	81.29	97.99	89.64	27.37	84.76	56.07	80.06	97.71	88.88
Average	91.29	94.49	92.89	80.78	90.39	85.59	90.77	95.88	93.32

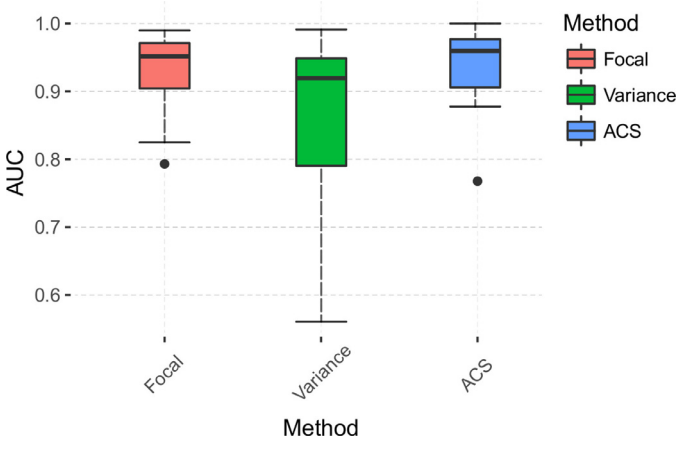

Fig. 6. Channel selection method comparison on overall AUC scores. The dots are for outliers.

proposed method achieved better score on all metrics though not significant. However, the proposed method yields a considerable improvement in mean onset detection delay. Onset detection delay indicates the time in seconds after that the classifier can detect a seizure onset. Delay is 1 s if the first 1 s ictal iEEG segment at seizure onset can be correctly detected. Since iEEG signals are divided into 1 s epochs, the minimum onset detection delay could be achieved is 1 s. Our work has a mean detection delay of 3.31 s which is comparable with that of Hills method. Fig. 7 demonstrates the advantages of the proposed method in terms of processing time, number of channels to be analyzed and detection delay.

We also test our method with the unlabeled dataset from the Kaggle competition. Labels for this dataset is not publicly available. The Kaggle competition organizers provided us the labels so we are able to evaluate the performance metrics per subject. Table A3 describes the seizure detection performance on the unlabeled dataset. It is non-trivial to note that all the thresholds were kept the same as they were during the cross-validation. A two-tailed signed rank test on the *AUC* between Hills and our methods result in a p-value of 0.6599 which means the difference is not significant at p-value < 0.05. F1-scores of the two methods are comparable. Our method has sensitivity of 89.40%, specificity of 89.24% and mean detection delay of 2.63 s. Since our proposed method has better specificity but worse sensitivity, comparison on detection delay may not be meaningful here. However, both methods achieve a good mean detection delay at less than 3 seconds. Finally, the overall *AUC* score across all subject of our method is 96.94%, slightly higher than that of Hills method at 96.29%.

Table 4
Comparison between state-of-the-art and proposed method on computational efficiency.

Subject	No. of electrodes	*M*	Training data (min)	Test data (min)	Hills (2014)		Proposed method			Training time improvement	Test time improvement
					Training[b] (s)	Test[c] (s)	ACS[a] (s)	Training[b] (s)	Test[c] (s)		
Dog–1	16	9	9.9	53	24.9	18	6.3	22.8	13.5	8.43%	25.00%
Dog–2	16	10	22	50	66	15.3	15.4	56.1	11.8	15.00%	22.88%
Dog–3	16	8	87.3	74.2	365.5	22.6	78.1	245	14.7	32.97%	34.96%
Dog–4	16	13	50.8	50.2	159.8	16.7	37.7	135.7	14.9	15.08%	10.78%
Patient–1	68	16	2.9	34.2	19	54.9	4.9	14.5	16.3	23.68%	70.31%
Patient–2	16	11	52.4	64.9	138.4	41.8	39.7	97	25.5	29.91%	39.00%
Patient–3	55	8	17.4	21.4	120.2	49.5	32.4	53	13.7	55.91%	72.32%
Patient–4	72	4	3.5	9.1	23.5	32.1	7.9	12.1	8.7	48.51%	72.9%
Patient–5	64	16	45.8	49.8	652.1	136.3	119.7	197	35.5	69.79%	73.95%
Patient–6	30	8	50	50	202.7	57.9	57.5	73.4	20.8	63.79%	64.08%
Patient–7	36	13	58.7	60	523.8	87.2	119.8	223.6	30	57.31%	65.6%
Patient–8	16	8	31.5	32	82.5	21.9	23.1	51.2	12.2	37.94%	44.29%
Average	**35.1**	**10.3**								38.19%	**49.67%**

[a] Automatic channel selection (ACS) time.
[b] Training time includes time for feature extraction and classifier training.
[c] Test time includes time for feature extraction and classification.

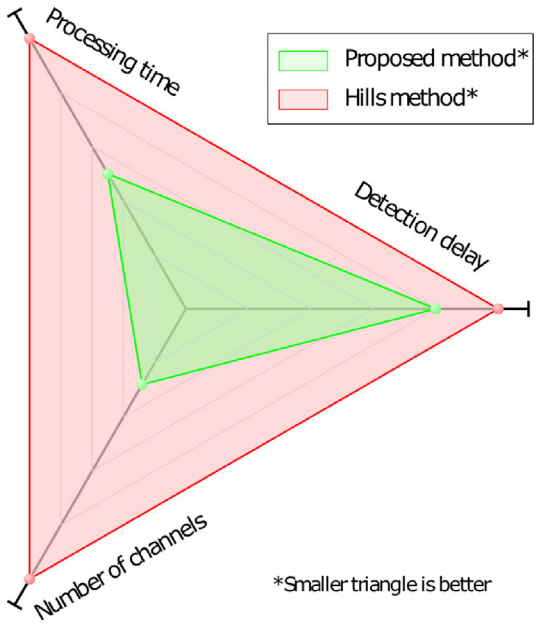

Fig. 7. Comparison between Hills method and proposed method in terms of detection delay, number of processed channels and processing time.

4. Discussions

We presented a seizure detection method based on a novel approach for automatic iEEG channel selection that provides comparable performance to the state-of-the-art method for the dataset considered. Although this leads to an extra overhead computing time in the beginning, the impact overall processing time is negligible because the channel selection need only be computed offline for each subject, before any future online seizure detection would be performed. One may argue that different channels may provide better performance over time; hence, channel selection is necessary over time. For the current dataset which was linked to a Kaggle seizure detection competition and downloaded from Kaggle (2014), the precise information about the times of the inter-seizure and seizure windows is not available. Therefore, investigating channel selection over time is not possible with this dataset. In our approach, data collected in real-time implementation would be transferred from a seizure control implant for off-line processing to update approximately every 1–6 months the optimal seizure detection channel for use during real-time implementation. The advantages of the automatic channel selection, on the other hand, are remarkable. Firstly, redundant and unrelated iEEG signals are eliminated which helps to improve efficacy of seizure detection system. Secondly, since the amount of data to processed is reduced, the processing time is also reduced. Gain in computational complexity becomes visible and significant for subjects with large number of channels. For instance, by reducing number of channels to be analyzed from 64 to 16 for Patient–5 (see Table 4, classification time can be improved by 74%.

Our channel selection method showed significantly better performance compared to the variance-based method proposed by Duun-Henriksen et al. (2012). Although the focal channel method has comparable performance with our method, it is more demanding to select focal channels for subjects with large number of channels. We can use our channel selection method to automatically select channels whereas selecting focal channels requires visual inspection by a neurophysiologist. In the scenario where we continuously collect data and re-select the best channels every 1–6 months, an offline automated approach may be more cost/time effective with regards to person hours and neurophysiologist time. In other words, there are more offline computation hours but more importantly less time spent by clinical staff labeling data so they can pay more attention to the other needs of their patients. Moreover, our method has an important advantage over approaches proposed by Minasyan, Chatten, Chatten, and Harner (2010) and Subasi and Ismail Gursoy (2010) since our automatic channel selection only runs during training phase, not in run-time classification.

Spectral power, correlation matrix and its eigenvalues on iEEG channels in both frequency and time domains have been shown as important features in seizure detection using iEEG recordings. The proposed subject-specific approach has a mean seizure onset detection delay of 2.63 s that is critical, for example, for an electrical stimulator to suppress the seizure on time.

5. Conclusion

Detection of seizure, especially at its early state, is crucial for patients who cannot be treated by drugs or surgery. Precise seizure detection allows electrical stimulation to timely interrupt the alteration of consciousness and subsequent convulsions. Although high performing seizure detectors are available, translating state-of-the-art seizure detection methods into battery-saving hardware implementations in implantable seizure control devices requires greater

gains in computational efficiency. This paper proposed automatic channels selection engine as a mechanism to adequately determine most informative iEEG recordings prior to feature extraction. The engine gave rise to significant computational efficiency improvements on subjects having large number of recording channels. The overall results of the proposed method were comparable with that of the state-of-the-art while it save 49.4% of the processing time and reduced the average number of channels requiring analysis by 71%, both critical factors for real-world applications.

Acknowledgment

The authors appreciate Dr Benjamin H. Brinkmann support from Mayo Systems Electrophysiology Lab for providing information on some unlabeled datasets. N. Truong and O. Kavehei acknowledge financial support from The Commonwealth Scientific and Industrial Research Organisation (CSIRO) via agreement PN 50041400. J. Yang acknowledges National Natural Science Foundation of China for their financial support under Grant 61501332.

Appendix

Table A1–A3

Table A1
Channels selected using the three channel selection methods.

	Focal channel	Variance-based	Our method
Dog–1	(9, 10, 13)	(3, 8, 9)	(4, 10, 12)
Dog–2	(1, 5, 13)	(1, 2, 5)	(1, 9, 12)
Dog–3	(8, 13, 14)	(1, 4, 9)	(7, 13, 14)
Dog–4	(1, 5, 13)	(2, 7, 9)	(7, 8, 15)
Patient–1	(11, 13, 14)	(1, 27, 44)	(19, 27, 30)
Patient–2	(1, 2, 3)	(1, 2, 3)	(1, 2, 3)
Patient–3	(5, 9, 14)	(5, 9, 11)	(5, 6, 26)
Patient–4	(7, 10, 15)	(26, 31, 47)	(37, 45, 66)
Patient–5	(5, 9, 12)	(10, 12, 49)	(9, 18, 25)
Patient–6	(2, 8, 15)	(2, 9, 18)	(15, 23, 24)
Patient–7	(8, 11, 15)	(7, 8, 28)	(26, 28, 36)
Patient–8	(3, 10, 11)	(2, 3, 4)	(3, 10, 11)

Table A2
Comparison between state-of-the-art and proposed method on *AUC*, sensitivity (SEN), specificity (SPE) and F1-score for binary classification of seizure and non-seizure states for the modified leave-one-out cross-validation of the training set. Threshold values were chosen to achieve best SEN-SPE balance and similar specificity scores between the two methods. This helps the sensitivity and mean detection delay comparisons meaningful.

Subject	Hills (2014)							Proposed method							
	AUC_E (%)	AUC_S (%)	*AUC* (%)	SEN (%)	SPE (%)	F1 (%)	Delay (s)	Thres.	AUC_E (%)	AUC_S (%)	*AUC* (%)	SEN (%)	SPE (%)	F1 (%)	Delay (s)	Thres.
Dog–1	98.53	99.69	99.11	96.79	99.29	97.49	1.60	0.30	97.79	99.43	98.77	96.79	98.81	96.96	1.60	0.29
Dog–2	97.47	99.66	98.57	87.68	99.30	91.00	2.00	0.28	99.58	99.58	98.58	90.86	98.87	91.41	1.00	0.26
Dog–3	97.76	99.58	98.67	95.21	98.17	89.20	2.08	0.28	96.71	99.25	98.08	92.08	98.34	88.16	3.08	0.35
Dog–4	99.86	97.15	98.51	57.06	99.28	69.49	1.00	0.26	99.70	97.34	98.56	69.54	98.03	73.11	1.00	0.20
Patient–1	90.94	98.13	94.54	91.03	96.15	92.27	4.50	0.41	95.06	98.32	97.27	92.31	97.12	93.64	4.00	0.59
Patient–2	99.37	99.50	99.43	94.72	99.00	88.14	1.33	0.26	99.34	99.40	99.42	94.72	99.03	88.49	1.33	0.26
Patient–3	88.50	95.34	91.92	78.00	92.44	74.87	6.00	0.37	88.25	93.61	90.25	76.21	92.16	73.42	7.57	0.45
Patient–4	100	100	100	100	100	100	1.00	0.21	99.63	99.63	100	95.00	97.89	88.31	1.00	0.50
Patient–5	84.20	88.82	86.51	58.52	99.08	56.88	18.00	0.23	87.72	91.19	90.45	62.96	99.20	62.16	5.00	0.22
Patient–6	98.73	99.82	99.27	97.35	98.48	90.11	2.25	0.22	98.66	99.79	99.28	96.89	99.49	95.39	2.25	0.36
Patient–7	87.33	90.10	88.72	66.63	99.35	65.49	7.33	0.14	92.31	94.61	93.20	66.63	99.69	66.64	7.33	0.27
Patient–8	81.46	97.82	89.64	92.78	97.13	84.36	4.50	0.22	78.78	98.03	88.52	94.44	97.54	86.74	4.50	0.26
Average	93.68	97.14	95.41	84.65	98.14	83.27	4.30		94.46	97.52	96.03	85.70	98.01	83.70	3.31	

Table A3
Comparison between state-of-the-art and proposed method on *AUC*, sensitivity (SEN), specificity (SPE) and F1-score for binary classification of seizure and non-seizure states for the unlabeled dataset. The threshold values were kept the same as being used during cross-validation. The total *AUC* is the *AUC* estimated across all the subjects.

Subject	Hills (2014)								Proposed method							
	AUC_E (%)	AUC_S (%)	*AUC* (%)	SEN (%)	SPE (%)	F1 (%)	Delay (s)	Thres.	AUC_E (%)	AUC_S (%)	*AUC* (%)	SEN (%)	SPE (%)	F1 (%)	Delay (s)	Thres.
Dog–1	98.80	99.63	99.22	93.08	99.47	91.64	2.75	0.30	97.70	99.12	98.41	92.45	99.57	92.16	2.75	0.29
Dog–2	97.65	96.86	97.26	100	32.27	13.63	1.00	0.28	92.98	95.45	94.22	97.37	40.67	14.89	1.50	0.26
Dog–3	94.51	98.47	96.49	93.50	97.43	85.19	1.70	0.28	92.62	97.69	95.16	89.75	97.56	83.68	1.70	0.35
Dog–4	99.79	99.61	99.70	100	67.07	20.82	1.00	0.26	99.76	98.92	99.34	100	65.75	20.18	1.00	0.20
Patient–1	98.15	99.32	98.74	98.70	98.46	91.22	1.14	0.41	97.07	99.80	98.44	92.02	99.52	93.17	1.57	0.59
Patient–2	98.93	99.73	99.33	96.51	99.07	91.32	1.50	0.26	98.98	99.72	99.35	96.07	99.05	90.91	1.50	0.26
Patient–3	73.37	95.01	84.19	90.62	88.29	61.21	2.50	0.37	85.39	96.14	90.77	85.16	95.58	75.69	3.00	0.45
Patient–4	67.50	67.50	67.50	48.00	80.73	28.40	1.00	0.21	61.84	61.84	61.84	52	84.38	33.99	1.00	0.50
Patient–5	88.10	95.75	91.93	86.90	98.33	80.89	6.50	0.23	80.60	93.71	87.16	88.10	90.92	51.75	6.00	0.22
Patient–6	98.10	99.86	98.98	93.18	99.39	92.76	1.00	0.22	98.50	99.90	99.20	88.18	99.93	93.27	3.00	0.36
Patient–7	99.83	99.99	99.91	100	98.89	95.24	1.00	0.14	99.81	99.98	99.90	98.89	99.72	98.21	1.00	0.27
Patient–8	86.48	98.15	92.32	92.78	97.93	87.21	7.50	0.22	86.59	98.08	92.34	92.78	98.22	88.36	7.50	0.26
Average	91.77	95.82	93.80	91.11	88.11	69.96	2.38		90.99	95.03	93.01	89.40	89.24	69.69	2.63	
Total		96.29								96.94						

References

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5–32. doi:10.1023/A:1010933404324.

Cabezas, J., Galleguillos, M., & Perez-Quezada, J. (2016). Predicting vascular plant richness in a heterogeneous wetland using spectral and textural features and a random forest algorithm. *IEEE Geoscience and Remote Sensing Letters, 13*(5), 646–650. doi:10.1109/LGRS.2016.2532743.

Duun-Henriksen, J., Kjaer, T. W., Madsen, R. E., Remvig, L. S., Thomsen, C. E., & Sorensen, H. B. D. (2012). Channel selection for automatic seizure detection. *Clinical Neurophysiology, 123*(1), 84–92. doi:10.1016/j.clinph.2011.06.001.

Echauz, J., Georgoulas, G., Smart, O., Gardner, A., Firpi, H., Litt, B., & Vachtsevanos, G. J. (2007). Monitoring, signal analysis, and control of epileptic seizures: A paradigm in brain research. In *Mediterranean conference on control & automation* (pp. 1–6). doi:10.1109/MED.2007.4433785.

Fatichah, C., Iliyasu, A. M., Abuhasel, K. A., Suciati, N., & Al-Qodah, M. A. (2014). Principal component analysis-Based neural network with fuzzy membership function for epileptic seizure detection. In *International conference on natural computation* (pp. 186–191). doi:10.1109/ICNC.2014.6975832.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*, 1157–1182.

Hills, M. (2014). Seizure Detection Using FFT, Temporal and Spectral Correlation Coefficients, Eigenvalues and Random Forest. *Technical Report*. Github.

Huynh, T., Gao, Y., Kang, J., Wang, L., Zhang, P., Lian, J., & Shen, D. (2016). Estimating CT image from MRI data using structured random forest and auto-Context model. *IEEE Transactions on Medical Imaging, 35*(1), 174–183. doi:10.1109/TMI.2015.2461533.

Kaggle (2014). UPenn and Mayo Clinic's Seizure detection challenge. https://www.kaggle.com/c/seizure-detection.

Kuhlmann, L., Burkitt, A. N., Cook, M. J., Fuller, K., Grayden, D. B., Seiderer, L., & Mareels, I. M. Y. (2009). Seizure detection using seizure probability estimation: Comparison of features used to detect seizures. *Annals of Biomedical Engineering, 37*(10), 2129–2145.

Litt, B., & Echauz, J. (2002). Prediction of epileptic seizures. *The Lancet Neurology, 1*(1), 22–30. doi:10.1016/S1474-4422(02)00003-0.

Minasyan, G. R., Chatten, J. B., Chatten, M. J., & Harner, R. N. (2010). Patient-specific early seizure detection from scalp EEG. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society, 27*(3), 163.

Osorio, I., & Frei, M. (2009). Real-time detection, quantification, warning, and control of epileptic seizures: The foundations for a scientific epileptology. *Epilepsy & Behavior, 16*(3), 391–396.

Parvez, M. Z., & Paul, M. (2015). Epileptic seizure detection by exploiting temporal correlation of electroencephalogram signals. *IET Signal Processing, 9*(6), 467–475. doi:10.1049/iet-spr.2013.0288.

Saab, M. E., & Gotman, J. (2005). A system to detect the onset of epileptic seizures in scalp EEG. *Clinical Neurophysiology, 116*(2), 427–442. doi:10.1016/j.clinph.2004.08.004.

Sardouie, S., Shamsollahi, M., Albera, L., & Merlet, I. (2015). Denoising of ictal EEG data using semi-Blind source separation methods based on time-Frequency priors. *IEEE Journal of Biomedical and Health Informatics, 19*(3), 839–847. doi:10.1109/JBHI.2014.2336797.

Scikit-learn (2014). Ensemble methods. http://scikit-learn.org/stable/modules/ensemble.html.

Scornet, E. (2016). Random forests and kernel methods. *IEEE Transactions on Information Theory, 62*(3), 1485–1500. doi:10.1109/TIT.2016.2514489.

Shih, E. I., Shoeb, A. H., & Guttag, J. V. (2009). Sensor selection for energy-efficient ambulatory medical monitoring. In *Proceedings of the 7th international conference on mobile systems, applications, and services*. In *MobiSys '09* (pp. 347–358). New York, NY, USA: ACM. doi:10.1145/1555816.1555851.

Shoeb, A. H. (2009). *Application of machine learning to epileptic seizure onset detection and treatment*. Massachusetts Institute of Technology Ph.D. thesis.

Subasi, A., & Ismail Gursoy, M. (2010). EEG Signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications, 37*(12), 8659–8666. doi:10.1016/j.eswa.2010.06.065.

Thodoroff, P., Pineau, J., & Lim, A. (2016). Learning robust features using deep learning for automatic seizure detection.

Tieng, Q. M., Kharatishvili, I., Chen, M., & Reutens, D. C. (2016). Mouse EEG spike detection based on the adapted continuous wavelet transform. *Journal of Neural Engineering, 13*(2), 26018.

Wang, G., Sun, Z., Tao, R., Li, K., Bao, G., & Yan, X. (2016). Epileptic seizure detection based on partial directed coherence analysis. *IEEE Journal of Biomedical and Health Informatics, 20*(3), 873–879. doi:10.1109/JBHI.2015.2424074.

Zabihi, M., Kiranyaz, S., Rad, A., Katsaggelos, A., Gabbouj, M., & Ince, T. (2016). Analysis of high-Dimensional phase space via poincare section for patient-Specific seizure detection. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24*(3), 386–398. doi:10.1109/TNSRE.2015.2505238.

Zeng, H., & Song, A. (2015). Optimizing single-Trial EEG classification by stationary matrix logistic regression in brain-Computer interface. *IEEE Transactions on Neural Networks and Learning Systems, pp*(99), 1–13.

Zhang, H., Yang, H., & Guan, C. (2013). Bayesian learning for spatial filtering in an EEG-Based brain-Computer interface. *IEEE Transactions on Neural Networks and Learning Systems, 24*(7), 1049–1060.

Zhang, Z., & Parhi, K. K. (2016). Low-complexity seizure prediction from ieeg/seeg using spectral power and ratios of spectral power. *IEEE Transactions on Biomedical Circuits and Systems, 10*(3), 693–706. doi:10.1109/TBCAS.2015.2477264.