34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Low Complexity Algorithm for Seizure Prediction using Adaboost

Manohar Ayinala, Keshab K. Parhi, *Fellow, IEEE*

***Abstract*— This paper presents a novel low-complexity patient-specific algorithm for seizure prediction. Adaboost algorithm is used in two stages of the algorithm: feature selection and classification. The algorithm extracts spectral power features in 9 different sub-bands from the electroencephalogram (EEG) recordings. We have proposed a new feature ranking method to rank the features. The key (top ranked) features are used to make a prediction on the seizure event. Further, to reduce the complexity of classification stage, a non-linear classifier is built based on the Adaboost algorithm using decision stumps (linear classifier) as the base classifier. The proposed algorithm achieves a sensitivity of 94.375% for a total of 71 seizure events with a low false alarm rate of 0.13 per hour and 6.5% of time spent in false alarms using an average of 5 features for the Freiburg database. The low computational complexity of the proposed algorithm makes it suitable for an implantable device.**

***Index Terms*— seizure, prediction, power spectral density, adaboost, feature selection**

I. Introduction

Epilepsy is the one of the most common serious neurological disorders in the world. Approximately, 1% of the world's population experience sporadic seizures. The quality of lives of the epileptic patients will be significantly improved with an automated seizure prediction device. Recently, there has been great progress in seizure suppression methods. Some of these approaches include deep brain stimulation therapy, etc. A closed-loop therapy system can be developed, where a seizure prediction device monitors and triggers the seizure treatment.

Recently, much of the research has been carried out on predicting and detecting seizures based on real-time analysis of electroencephalogram (EEG) data from multiple channels. Research on seizure prediction is focussed on finding features that discriminate between pre-ictal (period of time before the onset of seizure) and inter-ictal (period of time between the seizures) periods. These features include power spectral density [1], [2], auto regressive coefficients [3], wavelets [4], [5], and cross-correlation measures [6], [4].

In [3] coefficients of auto-regressive (AR) models are used as features. The prediction is based on support vector machines (SVM) classifier. A method based on multivariate signal coherence is presented in [6]. The algorithm uses space-delay correlation and covariance matrices to extract the spatiotemporal correlation structure from multichannel ECoG signals. Wavelet decomposition and cross-correlation techniques are used to predict a seizure event in [4] and [5]. Further in [5], a VLSI implementation is also proposed. Features based on spectral power in different subbands along with SVM classifier is proposed in [1] to predict a seizure onset.

Even though a lot of research has been done, these results cannot be used to realize an implantable device which can predict seizures in real-time due to their inability to achieve 1) High sensitivity and low false positive rate, and 2) Low power consumption and hardware complexity. The power consumption of a reliable seizure prediction device should be in the range of 50 μW [7]. High sensitivity and low false positive rates can be achieved using signal processing and machine learning techniques [8]. These cannot be adopted for a possible real-time implementation of seizure prediction device due to their high computational complexity. The number of features and the type of classifier used to make a prediction can have a dramatic effect on power consumption.

In this paper, we propose a seizure prediction algorithm with low computational complexity, which achieves high sensitivity and low false positive rate at the same time. The proposed seizure prediction system will be suitable for real-time implementation on an implantable device.

II. Proposed Algorithm

A. Dataset

The proposed algorithm is evaluated on the Freiburg database [9] which contains electrocorticogram (ECoG) or intracranial electroencephalogram (iEEG) recordings from 21 patients who suffer from epilepsy. The data consists of six channels sampled at 256 Hz with 16 bit analog-to-digital converters. The data records for each patient are divided into ictal and interictal records by certified epileptologists. We have chosen 16 out of the available datasets of 20 patients, who have four or more seizures. Each 20-second long window of iEEG recordings has been categorized as interictal and preictal. Fig. 1 shows the steps involved in training and testing phases.

B. Feature Extraction

Preprocessing step is done to remove the artifacts such as line noise, electrical noise, and movement artifacts in iEEG data. When features are extracted, the spectral power in the bands of 47-53 Hz and 97-103Hz are excluded to remove the power line noise. Further, bipolar and time-differential methods have been used to reduce the effect of other types of artifacts in iEEG recordings [1]. The space-differential measurement provides common-mode rejection to reduce line noise and movement artifacts that are common to all the electrodes. Time-differential method is used to normalize the spectral power in high and low frequency bands.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (email: ayina004@umn.edu; parhi@umn.edu).

Feature extraction consists of computing the spectral power in different frequency bands in a 20-second long window of iEEG with 50% overlap. This provides a prediction of a seizure every 10 seconds. Spectral bands are selected based on standard iEEG frequency bands but the wide gamma band is split into four bands: delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), gamma bands (30-47Hz, 53-75Hz, 75-97Hz, 103-128Hz). The power in each band is normalized by the total power and is included as the last feature. Features are extracted from 4 space-differential signals, which makes a total of 36 features for a 20-second window.

C. Feature Selection

The complexity of the seizure prediction algorithm is proportional to the number of features required to make a prediction. We propose to find a subset of most discriminating features out of 36 features through the process of feature selection. Feature selection has been known in the machine learning community for many decades. It finds a subset of features which contain essentially most of the relevant information for making decisions. This subset of features may result in an increase in the performance accuracy of the model as irrelevant features are not taken into consideration [10]. Adaboost [11] based feature selection algorithms have been proposed in the literature [12]-[14]. Most of these algorithms did not provide a method to sort the features, but instead used the implicitly selected features in each iteration. We propose a simple criterion based on which features will be ranked.

Algorithm 1 Adaboost

Given : $(x_1, y_1), ..., (x_N, y_N)$
where $x_i \in X, y_i \in Y = \{-1, +1\}$
Initialize: T, h_t
model = Adaboost(X,y,h_t, T)
model contains $\alpha_t, 1 \leq t \leq T$
decision = $sign(\sum_{t=1}^{T} \alpha_t h_t(x))$

Adaboost algorithm takes input data (training set: $(x_i, y_i), 1 \leq i \leq N$), where each x_i belongs to some domain or instance space X, and each label y_i is in the label set $Y = \{-1, +1\}$. Adaboost calls a given weak or base learning algorithm repeatedly over T iterations. Algorithm 1 shows the basic flow of the Adaboost algorithm. The reader may refer to [11] for the complete algorithm. The base classifier for feature ranking process [15] is defined as follows:

$$h(x) = \{ \begin{array}{c} -1 \text{ if } x_k < v \\ 1 \text{ if } x_k \geq v \end{array} \quad (1)$$

where k is a parameter indicating the input variable used to create the split and v is the splitting value. That is, k indicates the feature and v denotes the threshold to differentiate between the two classes. This base classifier is called a "decision stump" as it consists of a classification tree with tree depth of one (a single split decision and two terminal nodes). Parameters k and v are selected to minimize the cost function using a greedy optimization strategy.

1) Ranking Algorithm: Adaboost with decision stumps as the base classifier inherently performs feature selection. In each iteration, the algorithm selects the most discriminating feature (one with the lowest weighted error) for the corresponding weights. Further, adaboost algorithm generates weight (α) for that particular classifier which signifies the performance of that individual base classifier. In this case, as the base classifier is a decision stump, α signifies the discriminating power of that particular feature used in that iteration. We propose a ranking algorithm based on this observation. Features are sorted by assigning a weight (wt) for each feature. The weight of each feature is computed using α values the algorithm has generated. The outline of the ranking algorithm in presented in Algorithm 2. If a particular feature is not selected at all, then the corresponding weight will be zero.

Algorithm 2 Feature ranking

Given : $(x_1, y_1), ..., (x_m, y_m)$
where $x_i \in X, y_i \in Y = \{-1, +1\}$
Initialize: T, h_t to be decision stump
model = Adaboost(X,y)
for $f = 1 \rightarrow$d **do**
  id = index of the iteration feature(f) is selected
  $wt(f) = \alpha_t(id)^2$
end for
rank = index(sort(wt))

D. Classification

In this step, the computed (selected) features are classified into two classes, preictal (+1) and interictal (-1) using a machine learning algorithm. Even though SVMs have demonstrated impressive performance in seizure prediction [1], [3], [6], the computational complexity of the final decision function depends on the type of kernel used during the training process. Table I presents the complexity analysis of SVMs with three popular kernels [15] (linear, 2_{nd} order polynomial and RBF), where d denotes the number of features and N_{sv} denotes the number of support vectors generated during the training process. We can observe that among three kernels, RBF kernel requires the highest number of computations while linear kernel requires the lowest. The high computational complexity of the RBF kernel makes it unsuitable for implementing in an implantable device. The best choice would be linear-SVM for reducing the complexity of the seizure prediction algorithm.

Even though, the complexity of the linear-SVM is low, the performance may be degraded if the features used are not linearly separable. We propose to build a non-linear decision function using a combination of linear decision functions (in general linear decision functions are less computationally complex). The decision stumps can act as linear classifiers and can be boosted using the Adaboost algorithm. The main motivation of using Adaboost is its low complexity hardware implementation and the final decision boundary can be non-linear as well. Further, model selection is not required for

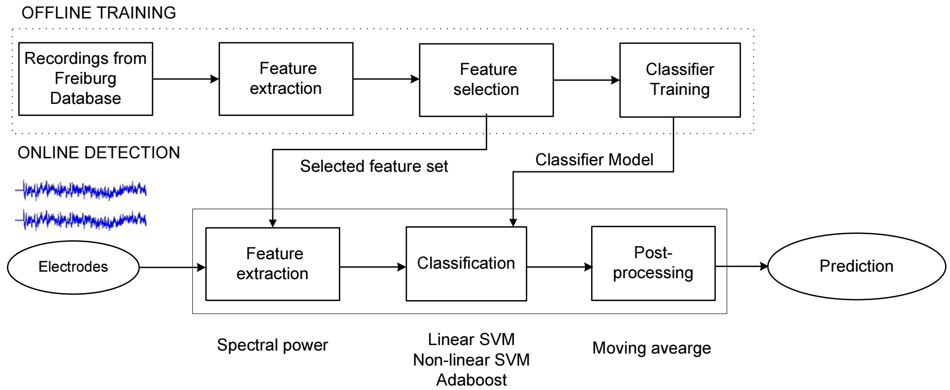

Fig. 1. Flow chart of the proposed seizure prediction algorithm

TABLE I

COMPLEXITY ANALYSIS OF SVM AND ADABOOST CLASSIFIERS

Classifier	# ADD	# MUL	# WORDS
SVM Linear	d	d	d
SVM Polynomial (p=2)	d^2	$d(d+1)$	d^2
SVM RBF	$2N_{sv}d$	$N_{sv}d$	$N_{sv}(d+1)$
Adaboost (decision stumps)	T	0	$2T$
*SVM-RBF requires extra N_{sv} exponent operations Adaboost requires extra T comparison operations			

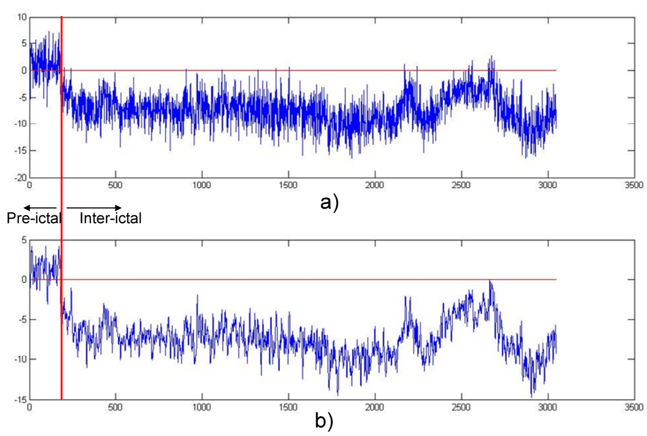

Fig. 2. Post processing the classifier output with 5-tap moving average filter. a) Classifier output b) After post-processing.

the Adaboost algorithm which is an advantage compared to the SVM.

E. Post-processing

We have observed some isolated false positives at the end of classification step as shown in Fig. 2. Post-processing is applied to eliminate these isolated false positives and false negatives. We applied 5-tap moving average filter to smoothen out these isolated events. The final prediction will be made using the filter output.

III. RESULTS AND DISCUSSION

Fig. 1 shows the block diagram for seizure prediction system along with training process. During the training process, features are ranked using the proposed ranking algorithm. Different classification models are built using SVM-Linear and Adaboost classifier with decision stumps for different sizes of feature sets (ranging from 1 to 36). We use double-cross validation to ensure in-sample optimization and out-of-sample testing.

A. Performance Analysis

Table II lists sensitivity, false positive % and number of false positives per hour for 16 patient data sets containing 4 or more seizures. Using all the 36 features and the Adaboost classifier, we achieved an average sensitivity of 94.375 and average FP% of 6.48 as measured by on-duration with 30-min on period for each prediction [1]. The results reported here compare favorably to previously published results. The algorithms from the prior literature requires SVM-RBF which is very computationally complex while the proposed algorithm achieves similar results with low complexity Adaboost classifier.

We can observe that using an average of 5 features, the proposed algorithm is able to achieve high sensitivity and low false positive rate. The detailed feature selection results are presented in Table III. The # Features column represent the number of features it required to achieve the given sensitivity and false positive rate. Further, Table II also shows the results with SVM-Linear using best 5 and 10 features. It can seen that the performance using SVM classifier with linear kernel degrades the performance. The chosen features may not be linearly separable which leads to the lower performance with linear kernel.

B. Complexity Analysis

Table II also analyzes the seizure prediction algorithms from the recent literature in terms of number, type of features and the classifier used. The number of features used is more than 20 in the prior art using SVM-RBF classifier. The hardware complexity of the SVM-RBF classifier is proportional to $N_{sv} * d$, where N_{sv} is the number of support

TABLE II

COMPARISON OF SEIZURE PREDICTION ALGORITHMS

	# Pat	# Sz	Prediction Horizon	Sens (%)	FP/hr	Feature	# Features	Classifier
[6]	19	83	30	90.8	0.094	Correlation	20	SVM-RBF
[3]	9	18	15	100	0.17	AR coefficients	36	SVM-RBF
[1]	18	80	30	97.5	0.27	Spectral Power	36	SVM-RBF
Proposed	16	71	30	94.375	0.13	Spectral Power	36	Adaboost
Proposed	16	71	30	91.25	0.27	Spectral Power	36	SVM-Linear
Proposed	16	71	30	**94.375**	**0.14**	Spectral Power	**4.8125**	**Adaboost**
Proposed	16	71	30	67.1825	0.15	Spectral Power	5	SVM-Linear
Proposed	16	71	30	85.625	0.19	Spectral Power	10	SVM-Linear

vectors and d is the dimensionality of the feature vector. The number of support vectors (N_{sv}) depends on the size of the training data set. We observed N_{sv} varying anywhere between 1000 to 3000 during the training process. The computational complexity of SVM-Linear and Adaboost with decision stumps are similar except that the former requires multiplication operation and the later requires comparison operation. The power consumption and the hardware area of these two classifiers are shown in Fig. 3. We can observe that Adaboost is a better option between the two when operating at same conditions.

TABLE III

PERFOMANCE OF THE PROPOSED SEIZURE PREDICTION ALGORITHM

Patient #	Sen%	FP/hr	FP %	# Features
1	100	0	0	3
3	100	0	0	4
4	100	0	0	2
5	60	0.7917	36.3657	8
6	100	0.0833	4.1667	4
7	100	0	0	2
9	100	0.25	12.5	4
10	100	0.1667	8.33	4
11	75	0	0	4
12	100	0	0	5
14	75	0.0833	4.1667	8
15	100	0.1667	8.333	8
16	100	0.1667	8.333	8
17	100	0.125	6.25	5
18	100	0.1667	8.33	3
21	100	0.1667	8.333	5
Mean	94.375	0.135	6.487	4.8125

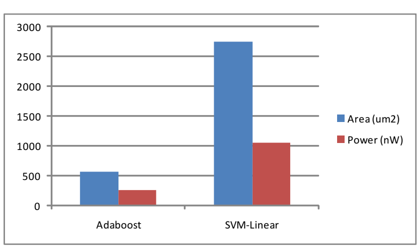

Fig. 3. Comparison of power and area of SVM-Linear and Adaboost circuits. Circuits are synthesized using 65nm technology operating at 1V Vdd and 1MHz clock frequency.

IV. CONCLUSIONS

A low complexity patient specific algorithm is proposed that extracts spectral power based features from EEG recordings. A new feature ranking algorithm is proposed to rank the features. Non-linear classifier is built using Adaboost with decision stumps as base classifier, which makes it computationally less expensive compared to non-linear SVMs. The algorithm achieves high sensitivity and low false positive rate comparable to the previously published results but at a much lower computational complexity. Future work will focus on reducing the complexity in the feature extraction step.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Yun-Sang Park and Prof. Tay Netoff for their early contributions to this work on feature computation [1].

REFERENCES

[1] Y. Park, L. Luo, K.K. Parhi and T. Netoff, ”Seizure prediction with spectral power of EEG using cost-sensitive support vector machines,” *Epilepsia*, 52(10), pp. 1761-1770, Oct. 2011.

[2] Y. Park, T. Netoff, K. Parhi, ”Seizure prediction with spectral power of time/space-differential EEG signals using cost-sensitive support vector machine,” *IEEE ICASSP*, pp.5450-5453, March 2010.

[3] L. Chisci, A. Mavino, et. al, ”Real-Time epileptic seizure prediction using AR models and support vector machines,” *IEEE Trans. on Biomedical Engineering,* vol.57, no.5, pp.1124-1132, May 2010.

[4] C. C. B. Suarez, E. T. Fonoff, et. al, ”Wavelet transform and cross-correlation as tools for seizure prediction,” *IEEE EMBC*, pp.4020-4023, Sept. 2010.

[5] S. Hung, C. Chao, S. Wang, et. al, ”VLSI implementation for epileptic seizure prediction system based on wavelet and chaos theory,” *IEEE Region Conference TENCON,* pp.364-368, Nov. 2010.

[6] J. R. Williamson, D. W. Bliss, D.W. Browne, ”Epileptic seizure prediction using the spatiotemporal correlation structure of intracranial EEG,” *IEEE ICASSP,* pp.665-668, May 2011.

[7] A. Shoeb, et al., ”A micro support vector machine based seizure detection architecture for embedded medical devices,” *IEEE EMBS*,pp. 4202-4205. Sept 2009.

[8] P. W. Mirowski, et al., ”Comparing SVM and convolutional networks for epileptic seizure prediction from intracranial EEG,” *IEEE Workshop on MLSP,* pp.244-249, Oct. 2008.

[9] Freiburg EEG database. Available from: https://epilepsy.uni-freiburg.de/freiburg-seizureprediction-project/eeg-database.

[10] G. H. John, R. Kohavi, and K. Pfleger, ”Irrelevant features and the subset selection problem,” *International Conference on Machine Learning*, pp. 121-129, 1994.

[11] Y. Freund, R. Schapire, ”Experiments with a new boosting algorithm,” *International Conference on Machine Learning*, pp. 148-156, 1996.

[12] L. Furst, S. Fidler, A. Leonardis, ”Selecting features for object detection using an Adaboost-compatible evaluation function,” *Patter recognition letters*, pp. 1603-1612, 2008.

[13] P. Silapachote, D. R. Karuppiah, A. R. Hanson, ”Feature selection using adaboost for face expression recognition,” *Intl. Conference on Visualization, Image, and Image Processing*, Spain, Sept. 2004.

[14] D. D. Le, ”Feature selection by adaBoost for SVM-based face detection,” *Forum on Information Technology*, pp. 183-186, 2004.

[15] V. Cherkassky, F. Mulier, *Learning from data, Concepts, Theory and Methods*, Wiley.