Algorithm Characterization and Implementation for Large Volume, High Resolution Multichannel Electroencephalography Data in Seizure Detection

***Abstract*—Ubiquitous bio-sensing for personalized health monitoring is slowly becoming a reality with the increasing availability of small, diverse, robust, high fidelity sensors. This oncoming flood of data begs the question of how we will extract useful information from it. In this paper we explore the use of a variety of representations and machine learning algorithms applied to the task of seizure detection in large volue of high resolution, multi-channel EEG data. We explore classification accuracy, computational complexity and memory requirements with a view toward understanding which approaches are most suitable for such tasks as the number of people involved and the amount of data they produce grows to be quite large. In particular, we show that layered learning approaches such as Deep Belief Networks excel along these dimensions. We also present the implementation of these algorithms on different hardware approaches including Virtex-7 FPGA, GPUs and 65 nm-CMOS ASIC.**

I. Introduction

Evidence based, personalized health care depends crucially on large volumes of data about both individuals and populations. It is easy to imagine a near future in which it is common to wear a number of bio-sensors that continuously monitor various aspects of our physiological state, including heart rate, blood pressure, eye movement, brain activity, and many others. There are two aspects of this enterprise - gathering the data and doing something useful with it. Our starting point is the data, and we ask how it is possible to accurately and efficiently extract information from it for purposes of identifying health states. This leads to the related issues of how to represent large volumes of medical time series so that the information they carry about health state is exposed, and what algorithms are best to extract that information. In this paper we focus on these issues in the context of seizure detection.

Classifier	Memory Req.	Computation Req.	Memory Req.	Comput. Req.
SF	0	$19W + 16\alpha_K W + 10$	-	-
KNN	$TR(CM+1)$	$3T(CM+N) + (N+1) + SF$	10,000x	1,096.5x
CNN	$\alpha_{CNN}TR(CM+1)$	$3\alpha_{CNN}T(CM+N) + (N+1) + SF$	2,500x	274.8x
SVM	$\alpha_{SVM}TR(CM+2)$	$2CM + \alpha_{SVM}T + 5 + SF$	502.5x	1.086x
LR	$R(CM+2)$	$2CM + 5 + SF$	1x	1x

TABLE I

Comparison of memory and computational complexity requirements for simple feature extraction (SF), KNN, CNN, SVM and LR classifiers. The last two columns show the relative memory and computation requirements, respectively, for each classifier relative to logistic reqression, which did the best for both requirements.

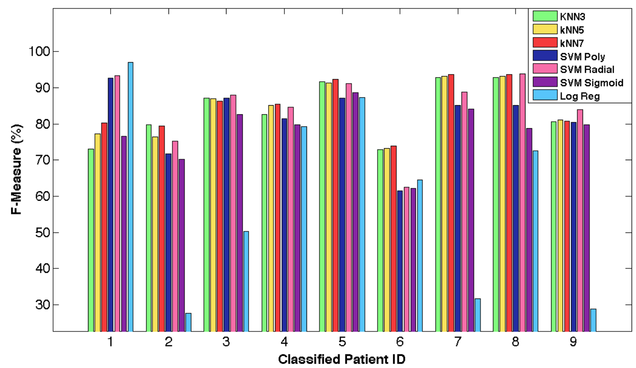

Fig. 1. Comparison of different classifiers when other patients data are used for training

A. Deep Belief Networks and Classifiers used in this work

Three classifiers are used in this work to compare the detection accuracy and computational and memory requirements. These classifiers are: K-nearest neighbor (KNN) with 3, 5, and 7 neighbors, Support Vector Machines (SVM) with sigmoid, radial basis function, and polynomial kernels, and logistic regression. We also look at two state-of-the-art methods for featurizing time series data, Symbolic Aggregate approXimation (SAX) and deep belief networks (DBNs). SAX lies at the heart of many of the most powerful algorithms for time series clustering, classification, motif discovery, and anomaly detection. It is particularly well-suited for a low-power implementation in hardware, but has yet to be extended to multi-variate time series, which this project will do. While SAX has a long track record with physiological data, DBNs are just beginning to show promise. They naturally deal with multi-variate time series well, thus making it possible to integrate multiple signals to overcome noise, but they are less amenable to low-power implementations.

B. Computational and memory complexity requirements

Besides the ability for the classifiers to accurately predict seizures, it is also necessary for the classifiers to minimize complexity since they will be running on a low-power, embedded sensor device in ambulatory setting. Since the device can be trained offline, the complexity comes in the form of memory required to store the classifier model and computation required to classify an incoming test vector. Table I summarizes the memory and computational complexity for each of the classifiers. The memory and computation required for all the simple features is denoted as SF.