Hardware Accelerated Epileptic Seizure Detection System Using Support Vector Machine

Yahia Elhazek*, Ahmed Ibrahim*, Mostafa Amer*, Ahmed Abubakr*, Hassan Mostafa*†
*University of Science and Technology, Nanotechnology and Nanoelectronics Program, Zewail City of Science and Technology
October Gardens, 6th of October, Giza 12578, Egypt
†Electronics and Communications Department, Cairo University, Giza 12613, Egypt

***Abstract*—Epilepsy is one of the most common neurological disorders affecting millions of people and causing serious injuries such as fractures and vehicle accidents. The ability to detect and prevent the occurrence of epileptic seizures is very important to prevent such injuries. In this paper, a low complexity implantable hardware is proposed to detect patient specific seizure onsets based on support vector machine (SVM) classification. The SVM training algorithm used is the sequential minimal optimization (SMO). A number of time domain, hardware-inexpensive and discriminant features are extracted namely; Hjorth mobility, Hjorth complexity, energy, variance and coastline. These features are used to train the SVM SMO algorithm. The proposed method managed to detect 100% of the selected patients' seizures with smaller latency compared to previous work [1] while using simpler time domain features. The implantable part of the system consumes 90 μW and occupies layout area of 0.2 mm^2.**

***Keywords*—Epilepsy, seizure detection, SMO, SVM, Hardware Acceleration, Implantable chip.**

I. Introduction

Epilepsy is a neurological disorder characterized by an abnormality in the brain activity which leads to recurrent seizures or periods of instability in the brain behavior. It affects more than 50 million people worldwide making it one of the most common neurological disorders. There are different symptoms that characterize an epileptic seizure such as blank stares, uncontrollable movements, sudden jerks in the whole body, body stiffness and loss of consciousness. These symptoms appear as sudden rapid changes in the EEG signals. Long term drug treatments are usually employed with patients however more than 30% of these patients are drug resistant. When these medications fail, the alternative is to remove the area of the brain causing the seizure. However, as shown in previous work [2], epilepsy is not confined to a single area of the brain but rather it is an epileptic network where different areas of the brain interact synchronously causing these seizures. Also, the probability of becoming seizure free depends on the type of the seizure and these probabilities are between 35% and 75% which is very low.

Owing to the unpredictable nature of an epileptic seizure, it represents a major worry and a handicap to the patients. Hence, researchers and practitioners are paying great attention to developing algorithms that can detect and anticipate the occurrence of a seizure. Afterwards, a stimulus is applied to suppress the seizure using closed loop or open loop strategies. There are already commercialized devices that are implanted in the brain to treat these seizures [3]. However, these devices lack accuracy, are very expensive to afford, hard to implement and they are power hungry, which means that a patient having one of these devices in his brain must undergo a complex and expensive surgery every two years to only change its battery.

Seizure detection can be done through the analysis of EEG signals of a patient, however, these signals vary from one patient to the other. Hence, it is difficult to formulate a generic and a robust detector that can detect the onset of the seizure for all the patients. However, training a machine learning algorithm to detect the the seizure on each patient is feasible. Usually, on the onset of a seizure, the EEG signal starts to fluctuate and deviate from the behavior of the rest of the signal making this abnormal change a bio-marker characterizing the seizure. unfortunately, there are several movements that produce a similar signal, like a sudden jerk or eye blinking, thus, the machine learning algorithm should be able to differentiate between these changes and the seizure itself.

By combining different extracted features of the brain EEG signals, a system that can analyze those features can also detect seizures in real time basis. This system is divided into two main modules: a training module and a classification module. The training module receives the raw EEG signal from the brain and trains the SVM algorithm to differentiate between the normal and pathological signals. The classification module which is an implantable chip classifies any new EEG signal into seizure(ictal) and non-seizure(non-ictal) signals. The training module is separated from the classification module because implantable devices require very low power consumption and the SVM training is very power hungry and requires a large memory to accommodate the large training matrix. Also, the training is performed on a dedicated chip not on a computer because running an SVM algorithm on a general purpose processor is inefficient, takes a lot of time, a lot of power and a large RAM given the amount of data being generated -which are used in training- from the EEG signal are large, thus, the developed ASIC chip for the training algorithm offers the advantage of hardware acceleration, instead of executing the training on a CPU or optimizing the code to work on a GPU. A hardware module is specifically designed to train the data, thus, decreasing the computational complexity, providing less latency and better throughput, increasing the processing

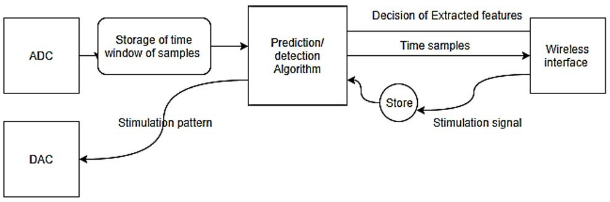


Fig. 1: Block diagram for the overall seizure detection system.

speed and accelerating the training. This paper is part of a project to develop a complete system for seizure detection using implantable components, starting from acquiring EEG signals from inter-cranial micro electrodes and analog/digital interfaces required to process neurological signals ending by brain stimulation signals that can suppress an incoming or ongoing seizure. Figure 1 shows a block diagram for the whole system and this paper is concerned with the middle block called prediction/detection algorithm.

The paper is structured in six sections. Section II describes the machine learning algorithm used in seizure detection and the features used in training the algorithm. The metrics used to assess the performance of the model are presented in Section III, the results are presented in Section IV, and the hardware description is shown in section V, Section VI concludes the paper.

II. BACKGROUND

A. Support Vector Machine

Support vector machine is a supervised machine learning technique that is based on statistical learning theory (SLT). It was first introduced by Vladimir Vapnik in 1979 [22] and usually used in binary classification problems. Given a set of training data belonging to one of two categories marked either -1 or 1, the SVM technique builds a model that can assign categories to novel data from the constructed model based on the learning algorithm. In the linear case, the margin between the two classes is defined by the distance of the hyper-plane to the nearest of the positive and negative samples. Maximizing margin can be expressed via the following optimization problem(1):

$$\min_{\vec{w},b}(0.5||\vec{w}||^2) \tag{1a}$$

subject to

$$y_i(\vec{w}_i\vec{x}_i - b) \geq 1 \tag{1b}$$

where x_i is the i-th training vector (sample), and y_i is the class of training sample. The value y_i is +1 for one class (ictal or non-ictal) and –1 for the other. This optimization problem is converted to the Lagrangian dual form which is a Quadratic Programming problem, where the objective function ψ depends on a set of Lagrange multipliers α_i.

In our case, data sets are not linearly separable. There is no hyperplane that directly splits the two classes. Fortunately, Vapnik suggested a modification which allows, but penalizes, the failure to reach the correct margin. That modification is shown in (2):

$$\begin{aligned} & min\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j K(\vec{x_i},\vec{x_j})\alpha_i\alpha_j - \sum_{i=1}^{N}\alpha_i \\ & 0 \leq \alpha_i \leq C, \forall i \\ & \sum_{j=1}^{N} y_i\alpha_i = 0 \end{aligned} \tag{2}$$

Where C is the new penalty factor and $K(x_i, x_j)$ is a kernel function that replaced the previous dot product between ω and x. Kernel functions represent dot products in higher dimensional spaces. In our work, the radial basis function (RBF) kernel is used, as it is widely used through literature for complex data like EEG signal when a better and adequate classification is needed.

B. SMO Algorithm

There are several algorithms designed to compute the Lagrange multipliers and find the maximum margin by following the optimization problem. However. Sequential Minimal Optimization(SMO) was chosen due to its simplicity and speed. SMO is an algorithm that can find the Lagrange multipliers needed to solve (2). It chooses two Lagrange multipliers and optimizes them together at a time. In order to solve for the two Lagrange multipliers, SMO first computes the constraints on these multipliers where these constraints resulted from(2).

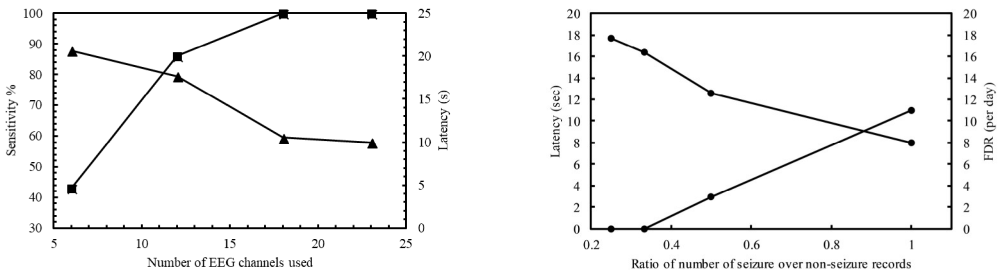
The algorithm takes one Lagrange multiplier in each iteration then searches heuristically for the other to be optimized with. Then it computes upper and lower limits on the second Lagrange multiplier until it arrives at the optimal values.

C. Feature Extraction

To be able to differentiate between the different categories of the input signal and allow the SVM algorithm to differentiate between the ictal and non-ictal states, the raw data are converted into a set of distinctive features via feature

TABLE I: List of features extracted from the Raw data

Feature	Equation
Average Energy	$E_{avg} = sum_{n=1}^{N}\frac{x^2}{N}$
Variance	$\sigma^2 = \frac{\sum(x-\mu)^2}{N}$
Hjorth Mobility	$HM = \sqrt{\frac{var(\frac{dx(t)}{dt})}{var(x(t))}}$
Hjorth Complexity	$HC = \frac{HM(\frac{dx(t)}{dt})}{HM(x(t))}$
Coastline	$x(i+1)-x(i)$



(a) (b)

Fig. 2: Effect of selection of training data and number of EEG channels

extraction block. Through these features, the model will detect the occurrence of a seizure if the features extracted are similar or comparable to the features of the ictal period in the training data set.

Table I shows the main features extracted from the raw EEG data and used to build the model where μ and σ are mean and standard deviation of x respectively, E(t) represents expected value of t, Hjorth mobility parameter represents the mean frequency of the power spectrum, Hjorth complexity represents the change in frequency, Energy is a measure of the power of the input raw EEG signal and Coastline is a measure of the variation of each data point from the previous point. These features were chosen due to their simplicity when building an RTL model with minimal power. As the feature extraction module will be a common module between the training and classification modules, and because the classification module will be implanted in the brain, this module is designed to consume the least amount of power to minimize the energy consumption.

III. Performance Metrics

Since there is no standard or unified performance assessment metrics, different studies are hard to be compared with each other due to the different assessment schemes they followed. A survey [4] was conducted on some well-known studies shows how they assessed their performance. Epoch-based metrics are viewed as application irrelevant metrics because every epoch is considered as a separate testing example regardless of the importance that its correct/incorrect classification has for a particular task. In a binary decision problem such as the seizure detection, the decision made by the classifier is represented in a structure known as a confusion matrix or contingency table. The confusion matrix has four categories: true positives (TP) are epochs correctly labelled as seizures; false positives (FP) refer to epochs incorrectly labelled as seizure; true negatives (TN) correspond to correctly labelled non-seizure epochs and finally, false negatives (FN) refer to epochs incorrectly labelled as non-seizure. On the other hand, the event-based metrics are thought to reflect the performance of a system for a specific application. Unlike the epoch-based metrics, the subsequent decisions of the same class are joined to create an event. Two scores are defined; good detection rate (GDR) or known as sensitivity in [1] is the percentage of seizure events correctly identified by the system as labelled by an expert in neonatal EEG. If a seizure was detected any time between the start and end of a labelled seizure this was considered a good detection. The other score is the number of false detections per hour (FD/h) calculated as the number of predicted seizure events in 1 h that have no overlap with actual reference seizures. To cope with the spiky nature of false detections, the metric FD/h is at times reported by joining subsequent false detections.

IV. Results

We employed the method of [1] in obtaining the performance of the system to be compared with his work. Evaluation method is based on event based metrics which are sensitivity (GDR) , latency, false detection rate. Performance can be affected with various parameters, such as the ratio between number of seizure records used in training to that of non-seizure records as shown in figure 2.The results of the training algorithm were compared to [1], as shown in table II, for the same selected patients and yielded an average of sensitivity (GDR) of 100%, a latency of 3.31 seconds and an FDR of 1.4 per day in the event-based metrics using a kernel RBF and [1] achieved an average of sensitivity (GDR) of 100%, a latency of 4.35 seconds and an FDR of 1 per day in the event-based metrics, while in the epoch-based metrics a sensitivity of 62.14%, a specificity of 99.93% and an accuracy of 99.5% were achieved in this work.

V. Hardware Description

A. Feature Extraction Hardware

The hardware implementation of features used and how they are processed starting by collecting raw EEG data up to final data normalization after which the data is ready to be classified as shown in figure 3. Raw data normalization was used in order to reduce the number of bits needed in the fixed point notation used, so as to reduce power consumption and hardware area. Most of the hardware expensive blocks, such as multipliers, are implemented in serial way that exploits the long waiting periods between each two samples acquired from EEG (1/256 sec). That implementation yields 75 μW power for the different features in table I .

TABLE II: Comparison between this work and [1] results.

Patient #	Event based sensitivity [GDR]			Latency [seconds]			FDR/Day		
	Linear SVM	RBF SVM	[1]	Linear SVM	RBF SVM	Shoeb	Linear SVM	RBF SVM	[1]
1	6/6	6/6	6/6	5.43	2.29	3.67	14.00	2.00	3.00
3	7/7	7/7	7/7	9.43	3.43	2.43	5.00	5.00	1.00
7	3/3	3/3	3/3	14.00	4.00	5.33	11.00	0.00	0.00
9	4/4	4/4	4/4	9.00	4.00	8.25	0.00	0.00	0.00
10	7/7	7/7	7/7	8.57	2.86	2.083	14.00	0.00	1.00
Mean				9.29	3.31	4.35	8.80	1.40	1.00

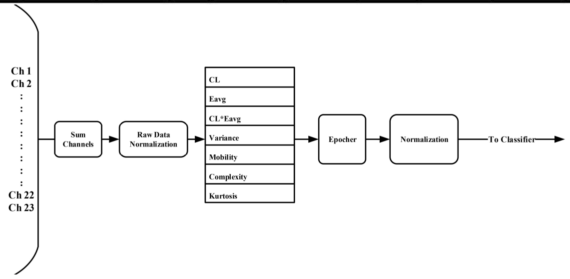


Fig. 3: Block diagram for feature extraction hardware flow.

B. The Trainer and Classifier

Figure 4 shows a block diagram for the hardware implementation of SMO-SVM as a complete trainer. If the configuration used as a standalone trainer, i.e. it needs a memory to store the training data, its power consumption reaches 14 mW, which is why the training can not be done inside the brain. On the other hand, if it's used as a real time classifier only using predefined parameters it consumes only 15 μW. Figure 5 shows the layout of the classifier part along with the features extraction hardware using UMC 0.18 μm technology.

VI. Conclusion

A Support Vector Machine (SVM) was used to detect the seizure onset and a training algorithm namely the "Sequential

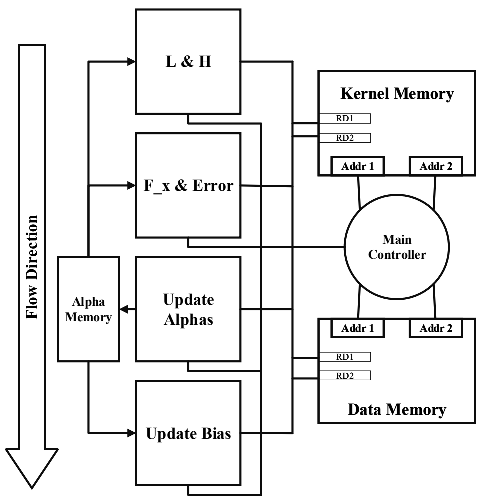


Fig. 4: Block diagram for SMO-SVM hardware modules.

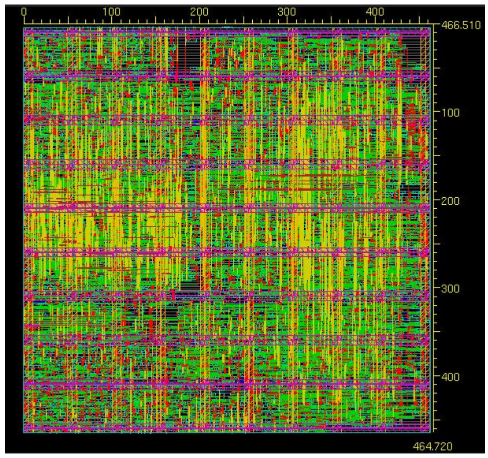


Fig. 5: Layout of the implantable classifier and feature extraction hardware parts.

Minimal Optimization" (SMO) was used in the model development and training the SVM algorithm to detect the seizure. A number of discriminant features were used and the most distinct features were selected in training the model to detect and classify the seizure states from non-seizure states. The trainer, features and classifier HLL description was converted to HDL language, the produced hardware was tested against the simulation results and their classification decisions was the same with maximum difference in classification function of 0.06. Total classification power and features extraction cost was 90 μW.

Acknowledgment

The authors would like to thank the supporting and funding agencies.This work was partially funded by ONE Lab at Zewail City of Science and Technology and at Cairo University, NTRA, ITIDA, ASRT, and NSERC.

References

[1] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 975-982), 2010.

[2] L. Lemieux, J. Daunizeau, and M. Walker, "Concepts of connectivity and human epileptic activity," Frontiers in systems neuroscience, vol. 5, p. 12, 2011.

[3] Medtronic. Deep Brain Stimulation systems. Available: https://www.medtronic.com/us-en/healthcare-professionals/products/neurological/deep-brain-stimulation-systems.html.

[4] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "Performance assessment for EEG-based neonatal seizure detectors," Clinical Neurophysiology, vol. 122, no. 3, pp. 474-482, 2011.

[5] A. Elnabawy, H. Abdelmohsen, M. Moustafa, M. Elbediwy, A. Helmy, and H. Mostafa, "A Low Power CORDIC-Based Hardware Implementation of Izhikevich Neuron Model", IEEE International NEW Circuits and Systems Conference (NEWCAS 2018), Montreal, Quebbec, Canada, pp. 130-133, 2018.

[6] K. Helal, A. Y. Aboelmakarem, A. – M. B. R. Fouad, T. S. Kamel, K. A. M. Nageeb, M. M. K. Mohamed, M. M. Abdelrhman, Y. Ismail, and H. Mostafa, "Low-Power High-Accuracy Seizure Detection Algorithms for Neural Implantable Platforms", IEEE International Conference on Microelectronics (ICM 2017), Beirut, Lebanon, pp. 231-234, 2017.

[7] A. H. Hassan, H. Mostafa, Y. Ismail, and A. M. Soliman, "A Low-Power High-Efficiency Inductive Link Power Supply for Neural Recording and Stimulation SoC", American Scientific Publisher (ASP) Journal of Low Power Electronics (JLOPE), vol. 14, issue 1, pp. 129-139(11), 2018.