A 1.5nJ/cls Unsupervised Online Learning Classifier for Seizure Detection

Adelson Chua[1], Michael I. Jordan[1], Rikky Muller[1,2]
University of California, Berkeley, Berkeley, CA, USA[1], Chan-Zuckerberg Biohub, San Francisco, CA, USA[2]

Abstract

This work presents a 1.5 nJ/classification (nJ/cls) seizure detection classifier which provides unsupervised online updates on an initial offline-trained regression model to achieve >97% average sensitivity and specificity on 27 patient datasets, including three that have >250 hours of continuous recording. The classifier was fabricated in 28nm CMOS and operates at 0.5V supply. Through hardware optimizations and low overall computational complexity and voltage scaling, the online learning classifier achieves 24x better energy per classification and occupies 10x lower area than state-of-the-art.

Keywords: seizure detection, classification, online learning

Introduction

Epilepsy is a debilitating neurological disorder affecting 1% of the world population. Recently, implantable devices that record neural activity and detect seizures have been adopted to issue warnings or to trigger neurostimulation to suppress the seizure. Such implants rely on high-accuracy classifiers to perform seizure detection at the onset of a seizure. To prolong implant battery life, energy efficiency becomes a key metric. State-of-the-art classifiers employ support vector machines (SVMs) or decision trees due to their high accuracies and relatively simple implementations. However, these classifiers can still have significant memory requirements [1,2] leading to high on-chip area and power consumptions. Moreover, over long periods of time, neural signals can drift requiring classifier retraining to maintain high accuracies [3]. Regular signal post-processing, labeling, and retraining by an expert physician can be costly and impractical. This paper presents a seizure detection classifier that incorporates continuous unsupervised online updates from an initial offline-trained model. This architecture allows the classifier to dynamically adapt to the neural signals over time to maintain high sensitivity detection without requiring any external intervention. Energy efficiency is achieved by using a computationally simple algorithm combined with architectural optimizations and supply voltage scaling.

System Overview

Fig. 1 shows the system diagram of the online learning seizure classifier. The classifier receives 16-bit digitized electroencephalography (EEG) data in 8 channels clocked at 1kS/s. Fig. 2 shows a more detailed architecture of the online learning classifier. The feature extraction unit maximizes hardware reuse by time multiplexing data from the 8 channels. Classification occurs after all features have been transferred, resulting in a classification rate of 1kHz. Two main feature classes were chosen: line length, which captures the high-amplitude and oscillatory characteristic of seizures, and spectral band power (α, 8-16 Hz; β, 16-32 Hz; γ, 32-96 Hz), which capture frequency-dependent patterns. These features are simple to calculate, leading to energy-efficient classification. Spectral band power calculations are approximated using time-domain sum of squares, eliminating the need for dedicated FFT circuits. 6th-order direct-form I IIR filters were utilized for bandpass filtering and reduced hardware by 10x compared to a FIR filter architecture.

Online Learning for Seizure Detection

Logistic regression is used as the basis for classification. The initial model is trained offline using the first 30% of the dataset to prevent overfitting. Incremental model updates on the remaining 70% are performed through stochastic gradient descent (SGD), following the equation shown in Fig. 3. Since SGD requires a label (y_t), unsupervised learning is achieved by using the classifier's own predicted seizure probability output as the label for the SGD update. To prevent erroneous updates due to misclassifications, a windowing technique is employed such that a series of high-confidence predictions are required to trigger the online model update (Fig. 3, lower right). The window size and confidence threshold are tuned offline on a patient-specific basis and are programmable on chip.

Architectural optimizations for the classifier include the multiplier reuse, which reduces compute logic requirements by ~16x, and the utilization of a 10-entry look-up table (LUT) for the sigmoid function, which impacts accuracy by <1%. Moreover, since only current feature weights (~200 bytes) are stored for retraining, the required memory is reduced by >300x compared to common SVM-based seizure classifiers [1,2].

Results

The classifier was fabricated in TSMC's 28nm HPM process. The 8kHz operating frequency enabled aggressive supply voltage scaling down to 0.5V resulting in a power consumption of 1.5μW and an energy efficiency of 1.5nJ/cls. Due to the relative computational simplicity of logistic regression coupled with architectural optimizations implemented to support online learning, digital logic and memory requirements were significantly reduced compared to state-of-the-art classifiers, leading to lower area and better energy efficiency.

The system was verified with two previously recorded human patient datasets: (1) a 3-patient intracranial EEG dataset from the University of Melbourne (UoM) [4] with >250 hours of recording time, and (2) a 24-patient scalp EEG dataset from CHB-MIT [5]. Fig. 4 shows the measured cumulative sensitivity vs. time of an exemplary patient (Patient 3 from UoM) with and without online learning enabled (the latter representing the performance of offline-only-trained classifier). Incorporating online learning results in an average sensitivity and specificity of 97.9% and 98.2% respectively, improving sensitivity by 6.5% on average with <1% specificity degradation over 3 patients. For the CHB-MIT dataset, the classifier achieves 97.5% and 98.2% average sensitivity and specificity over 24 subjects, comparable to state-of-the-art performance. The sensitivity for the subjects either stayed the same (6/24) or improved (18/24) by 1-3%. An improvement of >12% was observed on three subjects when compared against other state-of-the-art presenting a per-subject sensitivity breakdown [1,2]. These sensitivity improvements can be attributed to seizure pattern changes that can be tracked by the classifier's online learning scheme. Detection latency was measured to be 1.6-2.6s. Latencies <5s have demonstrated clinical efficacy in detection-triggered stimulation devices [6]. A comparison with the recent state-of-the-art seizure detection systems [7-10] is shown in Fig. 5.

Conclusion

This work is the first to support an on-chip unsupervised online learning classifier, starting from an offline-trained model, which maintains high classification sensitivity over time. Comparing the classifier with and without online model updates on long-term >250-hour datasets, sensitivity improved by 6.5% on average with <1% specificity degradation. Due to computational simplicity, architectural optimizations, and supply voltage scaling, the online learning classifier achieves 24x better classification energy efficiency at 1.5nJ/cls, while occupying 10x lower area compared with recent state-of-the-art seizure detection classifiers.

Acknowledgements This research was funded by NSF CAREER, the Wagner Foundation, and the Army Research Office under contract/grant number W911NF-16-1-0368. The authors thank the sponsors of the Berkeley Wireless Research Center. Thanks to Dr. Dean Freestone and Dr. Mark Cook for providing the patient datasets.

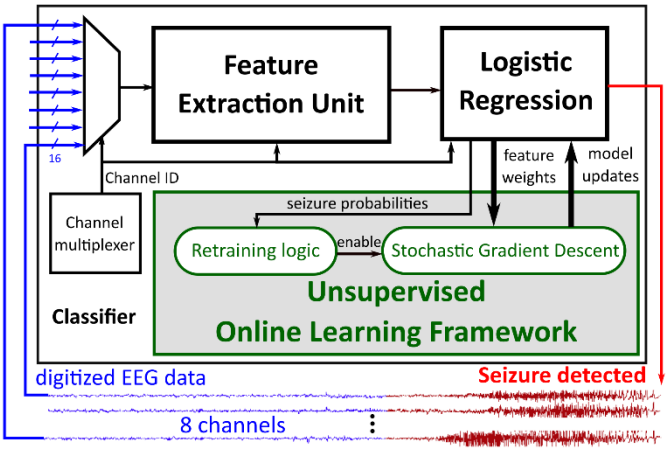


Fig. 1. Seizure detection SoC featuring the unsupervised online learning framework for incremental model updates for long-term high accuracy detection

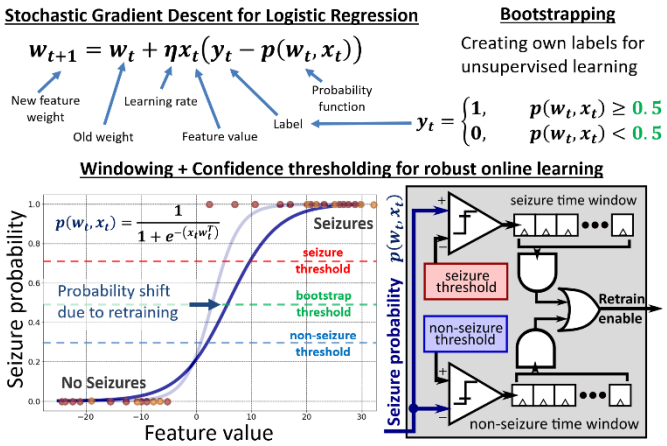


Fig. 3. Unsupervised online learning using SGD: retraining is done once a series of high confidence classifications were observed, shifting the feature values corresponding to classification threshold

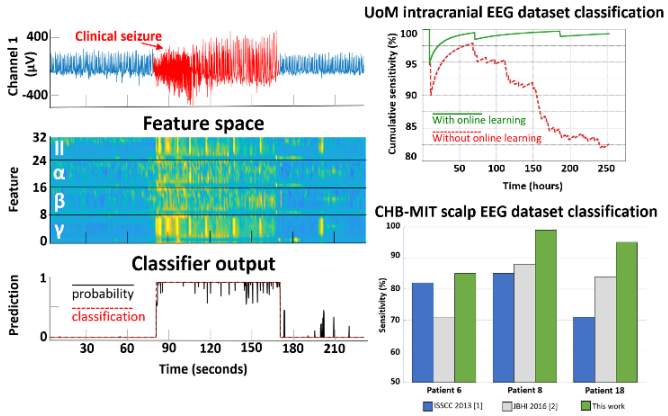


Fig. 4. Seizure classification performance on the UoM intracranial EEG (for Patient 3, 250 hrs) and CHB-MIT scalp EEG (for Patients 6, 8, 18) datasets; 70% of the datasets were used as the testing set

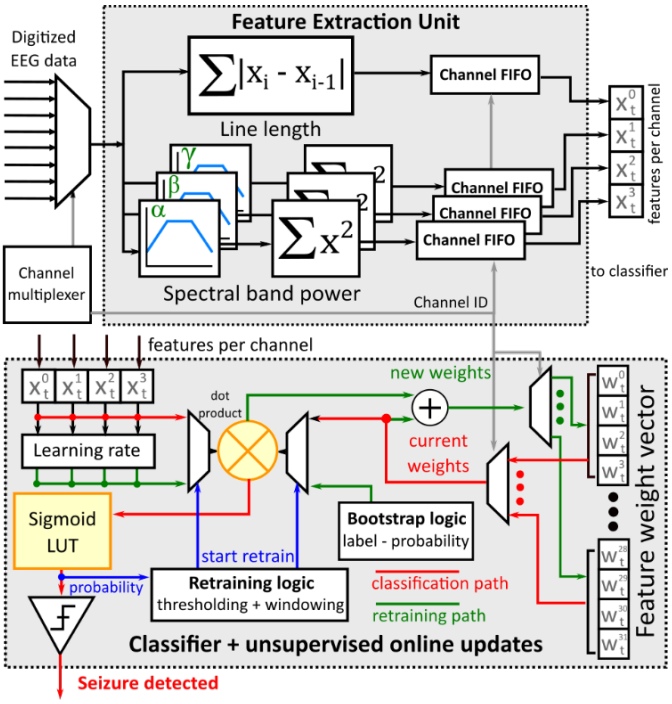
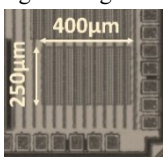


Fig. 2. Digital back end architecture (feature extraction unit and logistic regression classifier + online learning). Multipliers are reused for classification and retraining, and 10-entry LUT is used to approximate the logistic regression sigmoid function.



	JETCAS 2018[7]	VLSI 2018[8]	ISSCC 2020[9]	ISSCC 2020[10]	This Work	
Dataset	ieeg.org	-	EPILEPSIAE	CHB-MIT	UoM	CHB-MIT
Channels	32	16	8	8	8	
Classifier	Decision trees	Non-linear SVM	BrainForest	Coarse/Fine LS-SVM	Logistic regression + SGD for updates	
Training method	Supervised Offline	Supervised **Online**	Supervised Offline	Supervised Offline	Initial Offline + Unsupervised Online	
Sensitivity (%)	83.7	96.1	96.7	97.8	97.9	97.5
Specificity (%)	88.1	- [a]	- [b]	99.7	98.2	98.2
Latency (s)	1.79	0.71	-	<0.3	2.6	1.6
Technology (nm)	65	40	65	180	28	
Supply Voltage (V)	0.8	0.58	1.2	1.5	0.5	
Energy Efficiency (nJ/cls)	41.2	170,000	36	14,200	**1.5**	
Power (μW)	-	1,900	9.6	1.16	1.5	
Classifier Area (mm²)	1	4.5	1 [c]	3.51	**0.1**	

a: Reported 0.34 false alarms per hour
b: Reported 0.80 false alarms per hour
c: Estimated from chip photo

Fig. 5. Chip micrograph and comparison table vs. recent state-of-the-art seizure detection classifiers

References

[1] J. Yoo, *JSSC,* 2013, pp. 214-228 [2] C. Zhang et al., *JBHI,* 2016, pp. 996-1007 [3] A. Chua, et al., *EMBC,* 2020, pp.528-531 [4] L. Kuhlmann et al., *Brain*, vol. 141, pp. 2619-2630 [5] A. Goldberger, et al., *PhysioNet*, 2000, [Online] pp. e215-e220. [6] J. Echauz, et al., *Epilepsia*, 2001, pp 35-36 [7] M. Shoaran, et al., *JETCAS,* 2018, pp.693-707 [8] S. Huang, et al., *VLSI,* 2018, pp.259-260 [9] G. O'Leary, et al., *ISSCC,* 2020, pp. 402-404 [10] Y. Wang, et al., *ISSCC,* 2020, pp. 406-408