Digital Machine Learning Circuit for Real-Time Stress Detection from Wearable ECG Sensor

Sumukh Prashant Bhanushali*, Sudarsan Sadasivuni*, Imon Banerjee† and Arindam Sanyal*
*Electrical Engineering Department, University at Buffalo, Buffalo, NY.
†Department of Bio-Medical Informatics, Emory University, Atlanta GA 30322, USA.
Email:sbhanush@buffalo.edu

***Abstract*—This paper presents a digital machine learning circuit for classifying stress condition from chest ECG signal from a wearable sensor. To minimize hardware cost, we use only 5 time-domain features that have much lower area and power consumption than frequency domain or combination of time and frequency domain features as is used conventionally. We test the time-domain features on several machine learning algorithms. Random Forest classifier shows the best classification accuracy of 0.96 with the time-domain features at an estimated power consumption of only 1.16mW at 65nm CMOS process which demonstrates feasibility of embedding a machine learning classifier in a wearable ECG sensor for real-time, continuous stress detection.**

I. Introduction

Long-term stress in humans can lead to a plethora of diseases ranging from musculoskeletal illness to digestive problems to cardiovascular disease. The severe side-effects of stress calls for continuous stress detection to increase awareness of high stress levels which may otherwise go undetected. Since stress is primarily a physiological response to a stimulus triggered by the sympathetic nervous system [1], stress can be detected from physiological signals captured using wearable sensors. Automated stress detection using wearable devices can provide real-time continuous monitoring of stress levels while being minimally intrusive. While several technologies have been developed to detect stress from wearable devices [1]–[4], these techniques rely on processing the captured physiological signals on a separate computing device and cannot provide real-time stress detection.

We propose a digital machine learning classifier (DMLC) for real-time stress detection that can fit on a wearable electrocardiograph (ECG) sensor. The DMLC needs to have small area and power consumption to fit within the constrained budget of a wearable device. To reduce power and area consumption, we propose to use only time-domain features extracted from ECG signal instead of both time and frequency domain features that are conventionally used [1], [3], [4]. As will be shown later, a key contribution of this work is the adoption of time-domain only features that results in significant area and power savings compared to frequency domain features without sacrificing classification accuracy. The rest of this paper is organized as follows: the dataset used in this work is briefly described in Section II, feature extraction and comparison of time and frequency domain features are discussed in Section III, the classifier architecture and comparison with other works on stress detection are presented in Section IV, while the conclusion is brought up in Section V.

II. Stress Dataset

We use the WESAD dataset [1] for training and validation of our stress detection model. The WESAD dataset contains multi-modal sensor data for different physiological signals, such as blood volume pulse (BVP), ECG, electrodermal activity (EDA), electromyography (EMG), respiration rate and temperature, collected from wrist and chest using different wearable sensors. The sensor data was recorded from 15 participants and annotated as baseline, stress, amusement or meditation conditions. For this work, we use ECG signal from the chest and perform binary stress versus non-stress classification. Fig. 1 shows an example ECG plot from the WESAD dataset for patient 2 for stress and no-stress conditions.

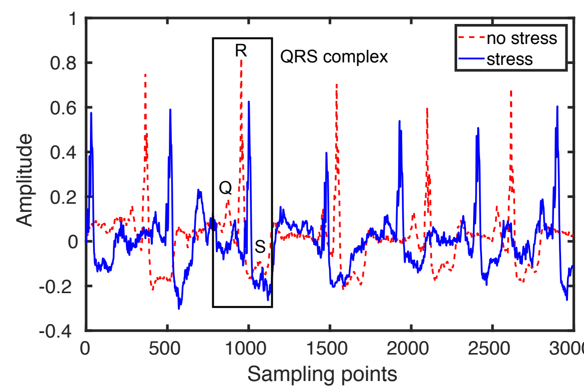

Fig. 1: Example ECG plots for stress and no-stress conditions

The sensor data for each patient is recorded for approximately 2 hours out of which the patient was under stress for roughly 10 minutes. To unskew the training data, we created overlapping segments of the sensor data such that number of segments with stress are roughly equal to number segments without stress. Out of 1020 segments, 525 segments are from ‘stress’ class, while the remaining 495 segments are from ‘no-stress’ class.

III. Feature Extraction and Comparison

A. *Time-domain features*

The time-domain features used in this work are derived following statistical measures of heart-rate variability (HRV)

978-1-7281-8058-8/20/$31.00 ©2020 IEEE

presented in [5]. The time-domain features are based on first order statistics (mean, standard deviation) calculated on normal-to-normal (NN) intervals which are intervals between adjacent QRS complexes (see Fig. 1) in the ECG signal. QRS complex represents depolarization of ventricles and wide QRS complexes indicate slow depolarization which may be due to dysfunction of ventricles. QRS complexes in the raw ECG signal are detected using peak detection algorithm [1]. The time-domain features used in this work are summarized in Table I and are calculated from segments of ECG signal which allows comparison of HRV during stress and non-stress conditions.

TABLE I: Summary of time-domain features

SDNN	The standard deviation of all NN intervals
RMSSD	The square root of the mean squared dierences of successive NN intervals
NN50	The number of interval differences of successive NN intervals greater than 50 ms
PNN50	The proportion derived by dividing NN50 by the total number of NN intervals
SDSD	The standard deviation of differences between adjacent NN intervals

B. Frequency-domain features

12 frequency-domain features are used in this work as shown in Table II and are based on the work in [1], [5]. Power spectral density analysis of ECG signal provides an estimate of power distribution across frequency bands and is linked with modulations of heart period. The frequency spectrum of ECG signal is computed and features are calculated on 4 frequency bands: ultra-low (ULF: 0.01-0.04Hz), low (LF: 0.04-0.15Hz), high (HF: 0.15-0.4Hz) and ultra-high (UHF: 0.4-1Hz) [1].

TABLE II: Summary of frequency-domain features

Band Energy	Energy in the 4 bands
Σ**(Band power)** all	Summation of power in the 4 bands
%**power**	Ratio of power in each band to the total power
LF Norm	Normalized LF power
HF Norm	Normalized HF power
LF/HF ratio	Ratio of LF power and HF power

C. Comparison of feature extraction techniques

We compare the time-domain and frequency-domain feature extraction in terms of hardware requirement as well as classification accuracy with different machine learning algorithms. The feature extraction circuits are synthesized in 65nm CMOS process, and uses a 700Hz clock which is the frequency used for sampling the ECG signals in the WESAD dataset. All the feature extraction circuits operate from 1V power supply. The time-domain features are calculated on 5 minutes segments of ECG signals. The frequency-domain features are extracted from 2^{18}-point FFT of ECG signal. Fig. 2 shows the synthesized circuit for SDNN feature-extraction circuit which consumes the highest power among the time-domain feature-extraction circuits.

The synthesized feature-extraction circuits are compared in Table III. The time-domain feature extraction circuit consumes

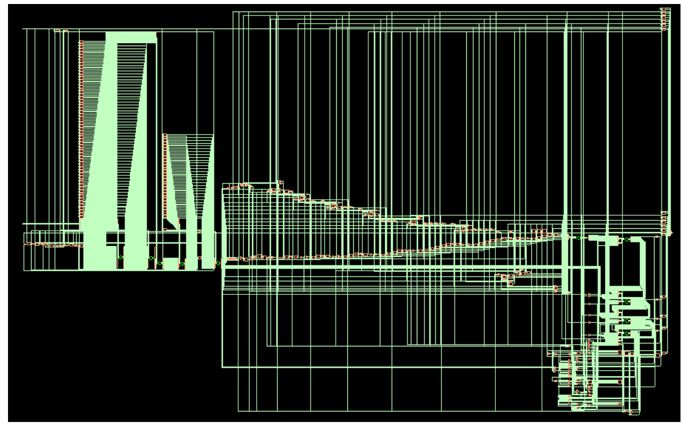

Fig. 2: Synthesized circuit for SDNN calculation

62μW power and 0.098mm^2 area which are $5\times$ and $910\times$ less than that of frequency-domain feature extraction circuit respectively. The area and power consumption for frequency-domain feature extraction is dominated by the FFT computation circuit [6]. The combined feature-extraction circuit using both time and frequency domain features consume 393μW power and 82.02mm^2 area.

TABLE III: Comparison of feature extraction techniques

Feature	Area (mm^2)	Power (μW)
Time	0.098	62
Frequency	81.92	331
Combined	82.02	393

To compare classification accuracy with the 3 feature extraction techniques (time, frequency and time+frequency combined), we use 5 machine learning algorithms and calculate classification accuracy for each algorithm for the different feature extraction techniques. The machine learning algorithms used are linear discriminant analysis (LDA), random forest (RF), support vector machine (SVM), and artificial neural network (ANN) with 2 and 3 hidden layers. We used Matlab for implementing the machine learning algorithms. We used 100 base estimators for the RF classifier. We used gaussian kernel for the SVM classifier. For the ANNs, we used tanh activation function for hidden layers and softmax activation function for output layer. The 2 hidden-layer ANN has 4 neurons in the first hidden layer and 3 neurons in the second hidden layer, while the 3 hidden-layer ANN has 5 neurons in the first hidden layer, 5 neurons in the second hidden layer and 3 neurons in the third hidden layer. We used 80% of the data for training the machine learning models, and 20% of the data for testing. To calculate classification accuracy, the dataset was randomly split into training and test sets and this process was repeated 20 times.

Fig. 3 graphically summarizes comparison between different machine learning algorithms and feature extraction techniques using box-plot. The time-domain, frequency-domain, and combined feature extraction techniques are denoted by the labels ‘T’, ‘F’ and ‘TF’. The median accuracy is shown with horizontal line inside the box-plots, while the mean accuracy is represented by a circle inside the box-plots. Whiskers in the box-plot denote maximum and minimum observations, while '+' denotes outliers in the observations. As seen in Fig. 3, classification accuracy of time-domain features is comparable to accuracy obtained using frequency-domain or combined features. The RF classifier has the highest mean and median accuracy among all the machine learning algorithms considered in this work. Mean accuracy of RF classifier using time-domain, frequency-domain and time-frequency combined features are 0.94, 0.96 and 0.98 respectively. While mean classification accuracy using time-domain features is 0.02 less than that with frequency-domain features, time-domain feature extractor circuit consumes 910× lower area and 5× lower power than frequency-domain feature extractor circuit. Hence, we have used time-domain features for developing our stress classifier that can be integrated with wearable ECG sensor.

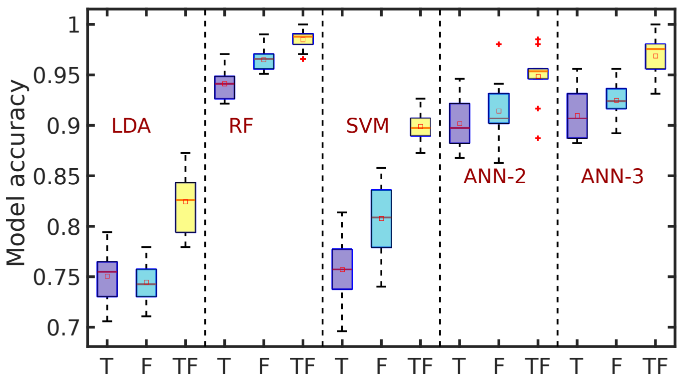

Fig. 3: Classification accuracy of different machine learning models for different feature extractors

D. Feature importance

Table IV shows the feature importance values for the time and frequency domain features computed using neighborhood component analysis (NCA) which calculates the average leave-one-out classification error. The time-domain features SDNN and NN50 are the most important features which also verifies the premise that only time-domain features are adequate for stress detection. To verify the feature importance scores, we repeated classification task with the RF classifier and 3 time domain features: SDNN, NN50 and PNN50. The mean classification accuracy is 0.92.

Fig. 4 shows the worst case confusion matrix with the 3 time-domain features. The worst-case accuracy and sensitivity for stress classification are both 0.87.

IV. Classifier Design

RF classifier is selected as the machine learning model for stress classification based on the accuracy comparison in Fig. 3. RF classifier is an ensemble classifier which decides class of predictors by aggregating results of different decision trees or estimators. Fig. 5 shows classification accuracy versus number of estimators. Classification accuracy is calculated on test set by randomly splitting the entire dataset into training and test sets, and repeating this procedure 20 times. It can be seen from Fig. 5 that the median/mean classification accuracy does not change significantly as the number of estimators is varied from 5 to 60. To reduce power and area of the classifier, 20 estimators are used for stress detection using RF since it has the highest mean accuracy and smallest spread.

TABLE IV: Feature importance values

Feature		Importance
Time	**SDNN**	4.38
	RMSSD	0.03
	NN50	3.86
	PNN50	1.01
	SDSD	0.02
Frequency	**Energy:ULF**	3.75
	Energy:LF	0.00
	Energy:HF	0.00
	Energy:UHF	0.01
	Total power	1.30
	% Power:ULF	3.41
	% Power:LF	0.29
	% Power:HF	1.67
	% Power:UHF	2.76
	LF Norm	0.30
	HF Norm	2.1
	LF/HF	0.00

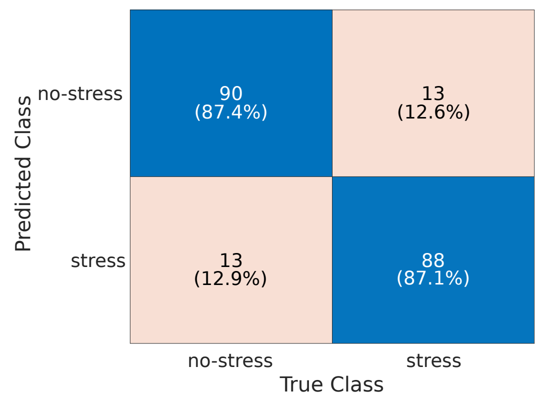

Fig. 4: Worst case confusion matrix with 3 time-domain features

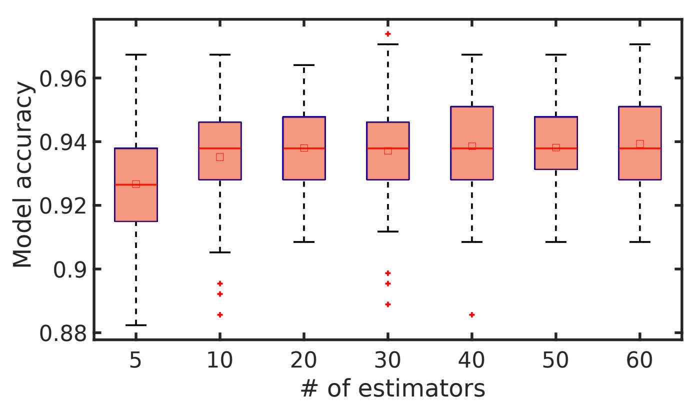

Fig. 5: Classification accuracy of RF classifier versus number of estimators

Fig. 6 shows classification accuracy versus segment size with all 5 time-domain features. Classification accuracy is calculated on test set by randomly splitting the entire dataset into training and test sets, and repeating this procedure 20 times as for Fig. 5. The mean classification accuracy is the highest at 0.96 for segment size of 3 minutes. Reducing the segment size also proportionally reduces power and area consumption for time-domain feature extractors. Fig. 7 shows the worst-case confusion matrix for the RF classifier with the 5 time-domain features. The worst-case stress classification accuracy is 0.92 while the worst-case sensitivity and specificity are 0.94 and 0.90 respectively. Sensitivity is a measure of the classifier's ability to correctly detect stress in patients, while specificity refers to the ability to correctly reject healthy patients without stress. The high sensitivity and specificity values for our classifier shows that it is effective in detecting stress.

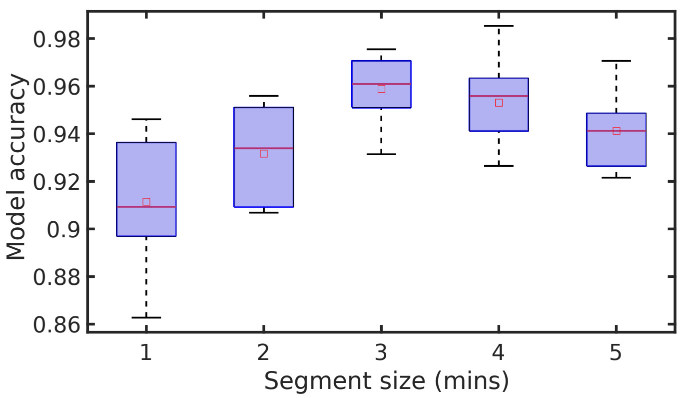

Fig. 6: Classification accuracy of RF classifier with time-domain features versus segment size

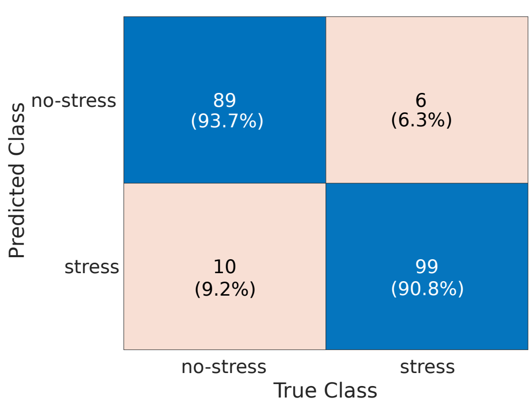

Fig. 7: Worst case confusion matrix for RF classifier

Table V summarizes the performance of our stress-detection classifier. Power of the RF classifier is estimated from [7]. Table VI compares this work with other stress-detection works that use different modalities to detect stress. Our classifier has the highest binary classification accuracy with time-domain features extracted from ECG sensor signal.

V. Conclusion

This work has presented a digital machine learning classifier that can be integrated with a wearable ECG sensor for real-time, continuous stress detection using only 5 time-domain features as opposed to frequency domain features which are computationally expensive. The time-domain feature extractor consumes 5× lower power and 910× smaller area than frequency-domain feature extractor when synthesized in 65nm CMOS process. The proposed classifier system can be used for stress detection which can help in early screening of several preventable diseases.

TABLE V: Stress-detector performance summary

Process (nm)	65
# of classes	2
Mean accuracy	0.96
Power (mW)	Feature extraction: 0.06
	Classifier: 1.1

TABLE VI: Comparison with other works

	Signal	Classifier	Feature type	Acc(%)
[2]	skin-conductance (SC) motion, cell-phone	k-NN	T	87
[1]	ECG	LDA	T+F	85
[8]	Speech	SVM	T+F	72
[9]	EEG	SVM	T+F	89
[4]	ECG, EMG, SC respiration	least square classifier	T+F	80
[3]	multiple modalities	SVM	T+F	78
This work	ECG	RF	T	96

T: time-domain; F: frequency-domain

References

[1] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, 2018, pp. 400–408.

[2] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in *IEEE Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 671–676.

[3] J. Aigrain, M. Spodenkiewicz, S. Dubuiss, M. Detyniecki, D. Cohen, and M. Chetouani, "Multimodal stress detection from multiple assessments," *IEEE Transactions on Affective Computing*, vol. 9, no. 4, pp. 491–506, 2016.

[4] J. Wijsman, B. Grundlehner, H. Liu, H. Hermens, and J. Penders, "Towards mental stress detection using wearable physiological sensors," in *IEEE Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 1798–1801.

[5] A. J. Camm, M. Malik, J. T. Bigger, G. Breithardt, S. Cerutti, R. Cohen, P. Coumel, E. Fallen, H. Kennedy, R. Kleiger *et al.*, "Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," 1996.

[6] M. Mahdavi, O. Edfors, V. Öwall, and L. Liu, "A low latency and area efficient FFT processor for massive MIMO systems," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, pp. 1–4.

[7] H. Nakahara, A. Jinguji, T. Fujii, and S. Sato, "An acceleration of a random forest classification using Altera SDK for OpenCL," in *IEEE International Conference on Field-Programmable Technology (FPT)*, 2016, pp. 289–292.

[8] M. Soury and L. Devillers, "Stress detection from audio on multiple window analysis size in a public speaking task," in *Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 529–533.

[9] V. Vanitha and P. Krishnan, "Real time stress detection system based on EEG signals," 2016.