Fully Integrated Mixed-Signal Classifier for Cardiovascular Health Monitoring

Sumukh Prashant Bhanushali[‡], Sudarsan Sadasivuni[*], Jose Sanchez[‡], Imon Banerjee[†], and Arindam Sanyal[‡]
[*]Department of Electrical Engineering, University at Buffalo, NY, USA.
[†]Mayo Clinic, Phoenix, AZ, USA.
[‡]School of Electrical, Computer and Energy Engineering, Arizona State University, AZ, USA.
Email: spbhanus@asu.edu

2023 IEEE Biomedical Circuits and Systems Conference (BioCAS) | 979-8-3503-0026-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/BioCAS58349.2023.10388918

***Abstract*—This work presents a fully integrated on-chip classifier for on-device detection of atrial fibrillation (AFib) from electrocardiogram (ECG) signal. The ECG signals are digitized using 14-bit analog-to-digital converter followed by time-domain feature extraction. The features are provided as inputs to an in-memory computing analog, 3-layer artificial neural network (ANN) for classification into normal sinus rhythm, AFib and noisy data. On-device AI classification reduces radio-frequency (RF) transmission and extends battery life of the sensing device by performing all the analysis locally and only transmitting in case of AFib detection. Prototype test-chip is fabricated in 65nm and achieves 99.6% accuracy in classification of AFib while consuming 58.3μJ/classification.**

***Index Terms*—Machine learning, atrial fibrillation, electrocardiogram, mixed-signal classifier and in-memory computing**

I. Introduction

Atrial fibrillation (AFib) is associated with significantly high risks for major cardiovascular events, such as stroke and heart failure, and requires proper timely treatment to reduce risks of fatalities. AFib occurs sporadically and presents no symptoms for many people which makes it difficult to detect during routine clinic visits. Thus, many people with AFib are unaware of their condition. This work proposes a fully integrated classifier that can be used for continuous, at-home monitoring of patients for early AFib detection. In contrast to existing monitoring devices that transmit electrocardiogram (ECG) signals wirelessly to remote server for analysis, the proposed solution embeds machine-learning (ML) algorithms into the ECG sensor for local classification. ML can reduce wireless transmission by identifying abnormal segments in the ECG and only transmitting those segments rather than transmitting the entire raw ECG data. The proposed technique reduces latency and saves sensor energy since radio-frequency (RF) transmission typically consumes 60-90% of the entire wireless sensor power consumption [1], [2]. Wireless transmission of all ECG data is particularly energy inefficient since ECG signal is naturally sparse and its information content is much lower than its sampling rate. While there are several data compression techniques, such as compressive sensing, level-crossing sampling and adaptive resolution sampling, that can reduce RF transmission, ML can potentially achieve much higher compression ratio by exploiting sporadic nature of AFib occurrence. In addition, existing data compression techniques can be combined with ML to achieve even higher compression in RF transmission.

The proposed classifier is fabricated in 65nm CMOS process and the complete system comprising front-end analog-to-digital converter (ADC), feature extractor and neural network classifier consumes 874μW operating at 90kHz sampling rate. The classifier is demonstrated on Physionet dataset and detects AFib with mean accuracy of 99.6%. The paper is organized as follows: Section II presents the proposed architecture and the design techniques, Section III presents the measurement results on test-chips and Section IV brings up the conclusion.

II. Proposed Architecture

The system level architecture of the proposed classifier is shown in Fig. 1 alongwith timing diagram. The ECG samples are digitized by a 14-bit successive approximation register (SAR) ADC. The unit capacitor used in the SAR ADC is 2.4fF. The SAR ADC uses the bi-directional single-sided switching technique [3] to reduce power consumption. The ADC outputs are sent to a digital feature extractor which extracts 14 time-domain features from non-overlapping windows of 6000 digitized ECG samples. The extracted features are converted to analog voltages using an array of 6-bit switched-capacitor digital-to-analog converters (DACs) with 4fF unit capacitor and sent to analog 3-layer artificial neural network (ANN) for classification into normal sinus rhythm, AFib and noisy data. The DAC resolution is selected based on hyper-parameter optimization on training data.

A. Feature extraction

The feature extractor performs simple pre-processing steps and extracts time-domain features. While feature extraction in the literature is performed in both time and frequency domains, time-domain feature extraction is selected for this work. Time-domain features are easy to compute on-chip at low power compared to frequency-domain features which require computation of fourier transform of the ECG signal [4]. The ECG signal is normalized between [0,1] before feature extraction. Median value of the ECG signal window is calculated and subtracted from all samples in that window to remove baseline wander before feature extraction. 14 time-domain features are extracted from ECG segments with 6000 samples. The extracted features are summarized in Figure 1

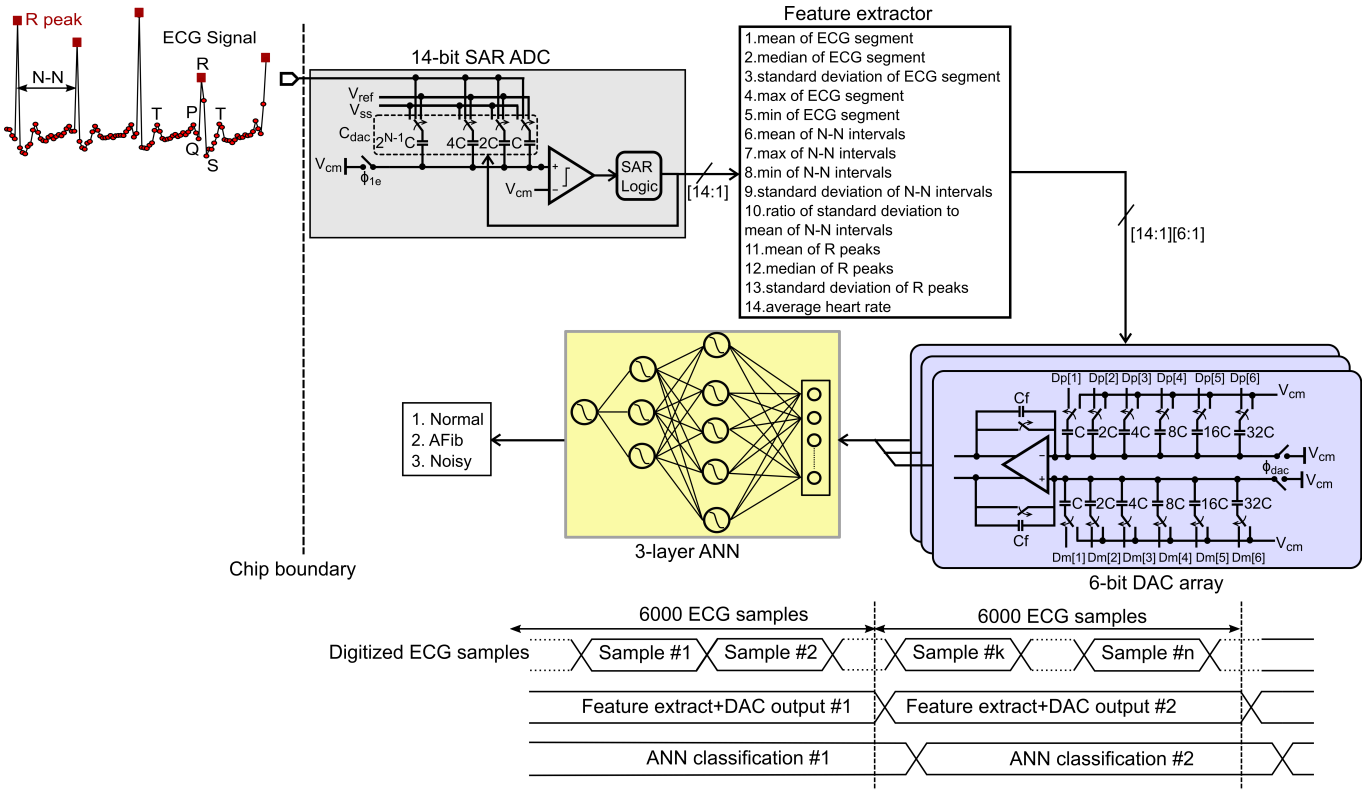


Fig. 1: Block diagram of the AFib classifier with associated timing diagram

and are measures of central tendency, dispersion, shape of distribution of a window, R peaks, N-N intervals, variance between R peaks and average heart rate. The R peaks are identified in time-domain through thresholding. The threshold for peak detection is set to 30% of the difference between maximum and minimum value for each segment.

B. ANN model training and circuit design

A 3-layer ANN is used as the classifier. The ANN has 20 neurons in the first hidden layer, and 6 neurons in the second hidden layer. The ANN is designed using switched-capacitor circuits and amplifiers as shown in Fig. 2. Multiply-and-accumulate (MAC) operations are performed using switched-capacitor circuits and custom nonlinear activation functions are realized using single-stage five-transistor operational transconductance amplifiers (OTAs). The ANN weights are encoded as capacitor values in the MAC circuits thus performing in-memory computation (IMC) for matrix multiplications. Switched-capacitor IMC is selected in this work since it has higher linearity compared to current-domain MAC in widely used SRAM IMC circuits. The trade-off with switched-capacitor IMC is that the ANN weights are hard-coded and cannot be re-tuned after fabrication. The ANN weights are quantized to 4-bit in the hidden layers, and 6-bit in the output layer. The weight quantization is done during the training iterations to preserve accuracy during testing. 4fF unit capacitor, with mismatch standard deviation of 0.4%, is selected as LSB weight in the MAC circuits to ensure that classification accuracy remains greater than 99% even in presence of random mismatch.

The timing diagram for the 3 layers are shown in Fig. 2. During ϕ_{11}, the extracted features are sampled on capacitors in the first-layer, and the top plates of the capacitors are shorted during ϕ_{12}. Thus, at the end of ϕ_{12}, the voltages at the positive and negative inputs of the differential amplifier in a neuron in the first layer are given by $[Vdd - \sum_i X_p[i]C[i]/\sum_i C[i]]$ and $[V_{dd} - \sum_i X_m[i]C[i]/\sum_i C[i]]$ respectively which are MAC outputs of the matrix multiplication of input vectors X_p and X_m and weight vector C applied to that neuron. The MAC outputs are applied to custom activation function realized through dynamic OTA as shown in Fig. 2. The OTA is clocked with ϕ_{b1} phase and turned off after the completion of ϕ_{11} and ϕ_{12} phases to conserve power. Output of the first hidden layer is sampled on the capacitors in neurons in the second hidden layer during ϕ_{21} and output of the second hidden layer is evaluated during ϕ_{22}. Finally, output of the second hidden layer is sampled on the capacitors in the output layer during $\phi 31$ and the classifier output is evaluated during ϕ_{32}. As in the first hidden layer, amplifiers in the second hidden layer and output layer are turned on only during ϕ_{b2} and ϕ_{b3} respectively.

The hidden layers use custom tanh activation and the output layer uses custom softmax activation. Custom tanh activation is realized through a fully-differential OTA while custom softmax activation is realized through a differential OTA with single-ended output as shown in Fig. 2. Current-

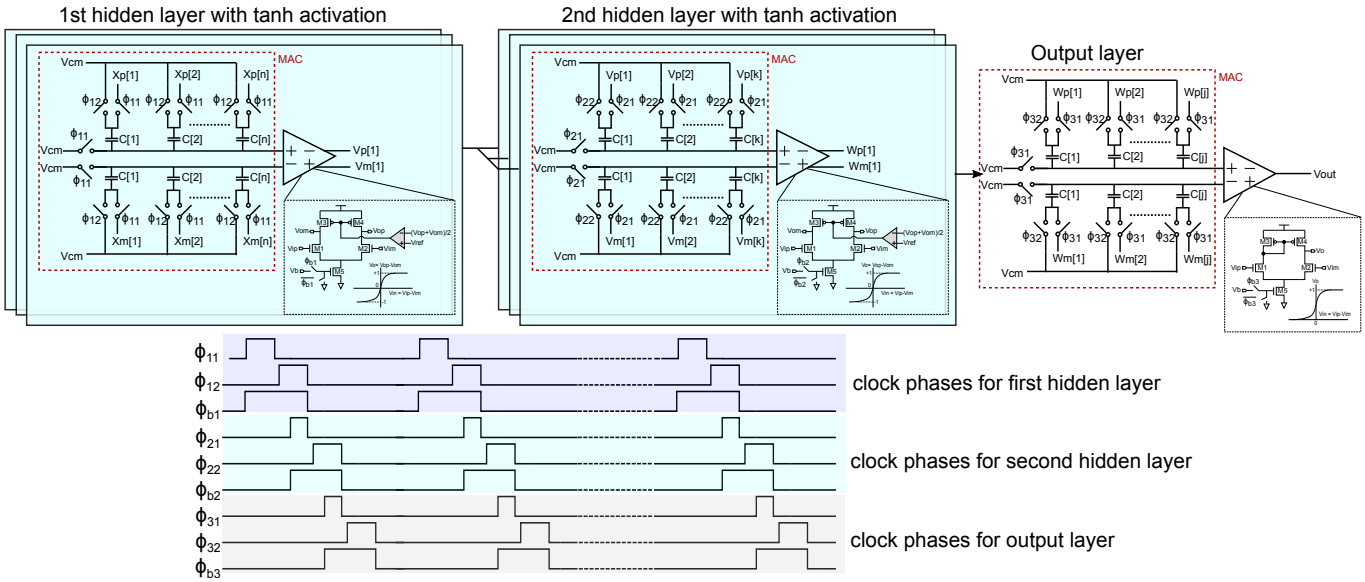
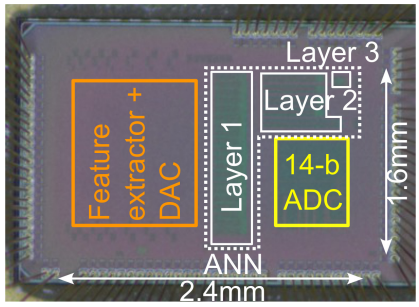


Fig. 2: Circuit schematic of the 3-layer ANN with timing diagram

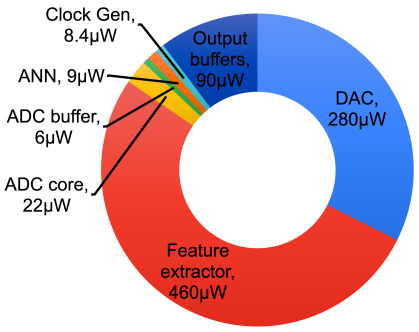
domain common-mode feedback (CMFB) loop is used to ensure constant output common mode voltage for the fully differential OTA. Current-domain CMFB also reduces area of the fully differential OTA compared to resistive CMFB which requires large feedback resistors, and reduces area of neurons by more than $4\times$ compared to that in [5]. The custom analog activation functions resemble their ideal, mathematical counterparts, but are not exactly the same. To ensure good matching between software ANN model and IC measurements, we use a hardware-software co-design methodology in which amplifier transfer curves, and their derivatives, are used to train the ANN model iteratively [6]. Stochastic gradient descent is used to optimize the ANN model by minimizing the loss function at each epoch. Offset in the amplifiers is calibrated in the foreground by adjusting the classifier decision thresholds on the training set.

III. Measurement Results

Fig. 3a) shows the micro-photograph of test-chip fabricated in 65nm. The feature extractor is digitally synthesized while the rest of the blocks are custom designed. The core area of the classifier is 3.84mm^2. The ECG samples are applied to the test-chip at 90kHz using National Instrument Data Acquisition Box to accelerate the measurements. The analog classifier output is captured with oscilloscope and sent to computer for analysis. The ADC and feature extractor operates at 90kHz while the DAC and the ANN operates at 15Hz. The power breakdown of the test-chip is shown in Fig. 3b). The feature extractor consumes the most power at 460μW followed by the DAC at 280μW and together these two blocks consume 84.5% of the total chip power. The 14-bit ADC consumes 22μW and the ANN consumes 9μW while the remaining power is consumed by buffers and clock generator. With a total power consumption of 874μW, the classifier consumes 58.3μJ/inference.



(a)



(b)

Fig. 3: a) Die micro-photograph and b) power breakdown

The test-chip is evaluated on retrospective patient ECG data from Physionet. The dataset has ECG of 8528 patients with 5050 patients with normal sinus rhythm, 738 patients with AFib and 2740 noisy ECG signals. The dataset is randomly partitioned into 70% split for training and 30% split for testing. The classifier decision thresholds are set by maximizing the accuracy of training set on each chip. Fig. 4 shows the

distribution of scores for each of the 3 class with decision thresholds and measured confusion matrix on the test set for 3 chips. The 3 chips have a mean accuracy of 99.6% with standard deviation of 0.35%.

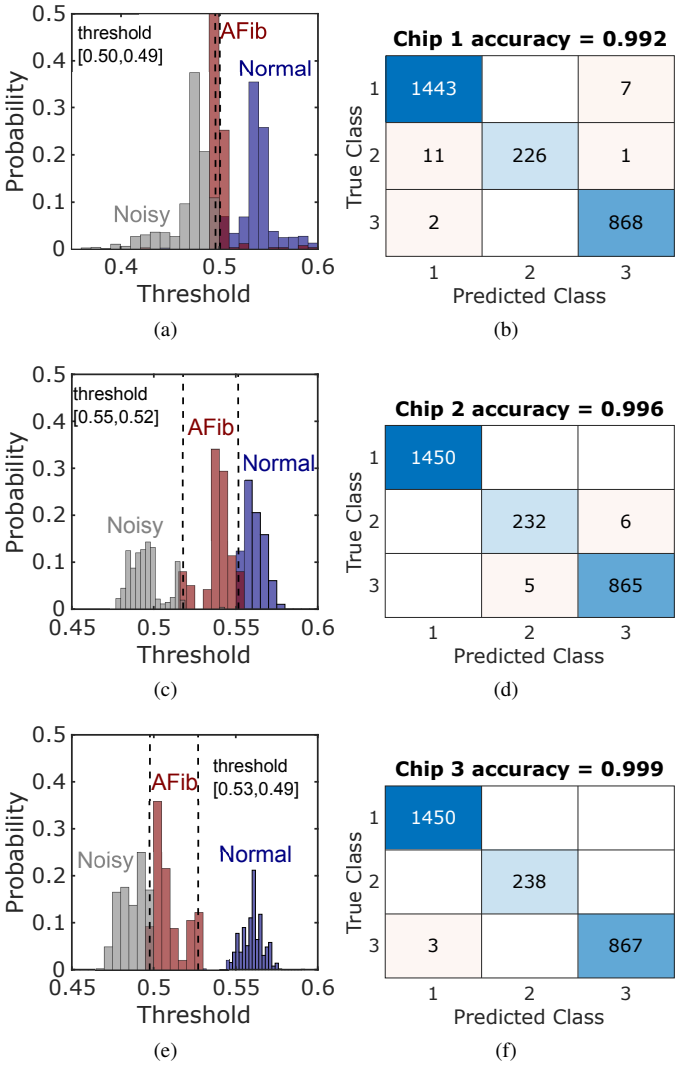


Fig. 4: Measured decision thresholds and confusion matrix for chips 1-3 with normal sinus rhythm as class 1, AFib as class 2 and noise as class 3

Table I compares this work with state-of-the-art ECG processors for cardiac arrythmia detection task. The proposed classifier achieves state-of-the-art accuracy, and lower energy/inference for the on-chip ANN classifier compared to state-of-the-art digital ANN classifiers. The spiking neural network (SNN) classifier has lower energy/inference at the cost of lower accuracy.

IV. Conclusion

This work has presented a fully integrated classifier for AFib detection that achieves 99.6% accuracy. Power consumption of the classifier is limited by the feature extractor and the DAC. Since the feature extractor in this work derives time-domain features from R peaks, power of the feature can be improved by leveraging sparsity of ECG signal to downsample inputs to the feature extractor. The DAC power can be further reduced by adopting dynamic bias for the amplifiers used in the DAC.

TABLE I: Comparison with state-of-the-art ECG processors for arrythmia detection task

	This work	TCAS–II 2021 [7]	TBioCAS 2019 [8]	TBioCAS 2019 [9]	TBioCAS 2022 [10]
Process (nm)	65	180	180	–	28
Area (mm^2)	3.84	0.75	0.92	–	0.54
Classifier	ANN	ANN	ANN	ANN	SNN
Feature	on-chip	off-chip			
Computation	mixed-signal	digital			
Accuracy	99.6%	98%	99.3%	99.4%	93.7%
Class #	3	5			
Power (μW)	9 (ANN) 874 (system)	1.3[1]	13.3[1]	13400[1]	–
Energy/ inference(μJ)	0.6 (ANN) 58.3 (system)	1.8[1]	3.2[1]	–	0.3[1]

[1]only for classifier

Acknowledgment

This material is based on research sponsored by Air Force Research Laboratory under agreement number FA8650-18-2-5402. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government.

References

[1] Y. Luo, K.-H. Teng, Y. Li, W. Mao, Y. Lian, and C.-H. Heng, "A 74-μW 11-Mb/s wireless vital signs monitoring SoC for three-lead ECG, respiration rate, and body temperature," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 5, pp. 907–917, 2019.

[2] X. Zhang, Z. Zhang, Y. Li, C. Liu, Y. X. Guo, and Y. Lian, "A 2.89-μW clockless wireless dry-electrode ECG SoC for wearable sensors," in *IEEE Asian Solid-State Circuits Conference (A-SSCC)*, 2015, pp. 1–4.

[3] L. Chen, A. Sanyal, J. Ma, and N. Sun, "A 24-μW 11-bit 1-MS/s SAR ADC with a bidirectional single-side switching technique," in *IEEE European Solid State Circuits Conference (ESSCIRC)*, 2014, pp. 219–222.

[4] S. P. Bhanushali, S. Sadasivuni, I. Banerjee, and A. Sanyal, "Digital machine learning circuit for real-time stress detection from wearable ECG sensor," in *IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2020, pp. 978–981.

[5] S. Sadasivuni, S. P. Bhanushali, S. S. Singamsetti, I. Banerjee, and A. Sanyal, "Multi-Task Learning Mixed-Signal Classifier for In-situ Detection of Atrial Fibrillation and Sepsis," in *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2021, pp. 1–4.

[6] S. T. Chandrasekaran, A. Jayaraj, V. E. G. Karnam, I. Banerjee, and A. Sanyal, "Fully Integrated Analog Machine Learning Classifier Using Custom Activation Function for Low Resolution Image Classification," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 3, pp. 1023–1033, 2021.

[7] Q. Cai, X. Xu, Y. Zhao, L. Ying, Y. Li, and Y. Lian, "A 1.3 μW event-driven ANN core for cardiac arrhythmia classification in wearable sensors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 9, pp. 3123–3127, 2021.

[8] Y. Zhao, Z. Shang, and Y. Lian, "A 13.34 μW event-driven patient-specific ANN cardiac arrhythmia classifier for wearable ECG sensors," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 2, pp. 186–197, 2019.

[9] N. Wang, J. Zhou, G. Dai, J. Huang, and Y. Xie, "Energy-efficient intelligent ECG monitoring for wearable devices," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 5, pp. 1112–1121, 2019.

[10] H. Chu, Y. Yan, L. Gan, H. Jia, L. Qian, Y. Huan, L. Zheng, and Z. Zou, "A Neuromorphic Processing System with Spike-Driven SNN Processor for Wearable ECG Classification," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 16, no. 4, pp. 511–523, 2022.