

Monitoring and Detecting Atrial Fibrillation using Wearable Technology

Shamim Nemati^{1*}, Mohammad M. Ghassemi², Vaidehi Ambai³, Nino Isakadze⁴, Oleksiy Levantsevych⁴, Amit Shah^{4,5,6}, and Gari D. Clifford^{1,4,7}

Abstract—Atrial fibrillation (AFib) is diagnosed by analysis of the morphological and rhythmic properties of the electrocardiogram. It was recently shown that accurate detection of AFib is possible using beat-to-beat interval variations. This raises the question of whether AFib detection can be performed using a pulsatile waveform such as the Photoplethysmogram (PPG). The recent explosion in use of recreational and professional ambulatory wrist-based pulse monitoring devices means that an accurate pulse-based AFib screening algorithm would enable large scale screening for silent or undiagnosed AFib, a significant risk factor for multiple diseases. We propose a noise-resistant machine learning approach to detecting AFib from noisy ambulatory PPG recorded from the wrist using a modern research watch-based wearable device (the Samsung Simband). Ambulatory pulsatile and movement data were recorded from 46 subjects, 15 with AFib and 31 non symptomatic. Single channel electrocardiogram (ECG), multi-wavelength PPG and tri-axial accelerometry were recorded simultaneously at 128 Hz from the non-dominant wrist using the Simband. Recording lengths varied from 3.5 to 8.5 minutes. Pulse (beat) detection was performed on the PPG waveforms, and eleven features were extracted based on beat-to-beat variability and waveform signal quality. Using 10-fold cross validation, an accuracy of 95% on out-of-sample data was achieved, with a sensitivity of 97%, specificity of 94%, and an area under the receiver operating curve (AUROC) of 0.99. The described approach provides a noise-resistant, accurate screening tool for AFib from PPG sensors located in an ambulatory wrist watch. To our knowledge this is the first study to demonstrate an algorithm with a high enough accuracy to be used in general population studies that does not require an ambulatory Holter electrocardiographic monitor.

I. INTRODUCTION

Atrial fibrillation (AFib) is the most frequently encountered arrhythmia, with an estimated prevalence of 3% in the adult population [1] that is projected to exceed 4% by 2050 [2]. AFib is often described as a 'silent killer' because it increases the risk of ischemic stroke by 5-fold [3], yet remains asymptomatic in nearly a third of the AFib population [4], [5], [6]. The asymptomatic nature of AFib provides motivation for the development of solutions that may help patients detect the condition. Ultimately, improved detection

of silent AFib may result in significant stroke prevention using medications such as prophylactic anticoagulants, which is a well-established preventive therapy in AFib patients [7].

Classical solutions for AFib screening include hand-held ECG devices, used periodically over multiple weeks [8], [9]. As wearable technologies become increasingly affordable and utilized by the general public, AFib detection algorithms that use photoplethysmography (PPG), now serve as an attractive alternative to existing ECG-based solutions [10]. PPG is collected on most modern wearable devices and is often monitored both passively and continuously. Hence, a PPG-based AFib algorithm require less *active* participant effort, increasing ease of adoption and potential for AFib diagnosis in asymptomatic individuals. As such, they may yield better stroke prevention than current standards.

There are countless algorithms that utilize ECG for real-time AFib detection [11]. Many of these algorithms have been tested using wearable technologies and report high efficacy (96% sensitivity and 93% specificity). ECG-based solutions however require the use of multiple electrodes, making them more cumbersome for general users. More recent studies have utilized PPG measured from smart-phone cameras [12], but algorithms tailored for existing wearable devices are less well-developed. This is due to the highly variable signal quality of PPG on many wearable devices, which itself arises from (1) the variable quality of sensor technologies in the devices and (2) artifacts resulting from individual user activity. For these reasons, there is not a



Fig. 1. The Simband Smartwatch. The device contains a multi-sensor module which includes eight PPG sensors, ECG, and accelerometer, among others (bottom). The watch display allows a user to look at their data in real-time (top). More information is available at: <https://www.simband.io/documentation/>

¹Dept of Biomedical Informatics, Emory University, Atlanta, GA 30322.

²Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139.

³Philadelphia College of Osteopathic Medicine, Suwanee, GA, 30024

⁴Dept of Medicine, Emory University School of Medicine, Atlanta, GA 30322.

⁵Dept of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA 30322.

⁶Division of Cardiology, Veterans Affairs Hospital, Atlanta, GA 30322.

⁷Dept of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

*Corresponding author, Email: shamim.nemati@emory.edu

single wearable algorithm that has been developed which may be deployed to existing consumer wearable technologies, while maintaining performance levels on par with ECG-based approaches.

In this study, we report the first instance of an algorithm that utilizes an existing wearable technology (the multisensor Samsung Simband wearable wristband [13]) to detect AFib based on PPG and accelerometry signals with performance levels exceeding those of many reported ECG-based approaches.

II. METHODS

A. Data Collection

A retrospective cohort study design was utilized for this experiment. All adult patients (18-89 years old) who were hospitalized as an inpatient and were on Telemetry monitoring at Emory University Hospital (EUH), Emory University Hospital Midtown (EUHM) and Grady Memorial Hospital (GMH) were eligible to enroll into the study. The study was approved by the institutional review board (IRB) of the Emory University Hospital and Grady Memorial Hospital in Atlanta, GA. Patients were recruited at random with an over-sampling of patients in AFib, as determined by the trained telemetry and physician study coordinator. Patients were approached in their hospital rooms and were consented and enrolled in this study. The study took place from October 2015 through March 2016. Using a modern research watch-based wearable device (the Samsung Simband [13]), worn on the wrist, we recorded ambulatory pulsatile data from 46 subjects, 15 with AFib and 31 with other rhythms. Single channel electrocardiogram (ECG), multi-wavelength photoplethysmography (PPG, 8 channels) and tri-axial accelerometry (x, y, z) were recorded simultaneously at 128 Hz. Recording lengths varied from 3.5 to 8.5 minutes, depending on the time each patient had available for testing.

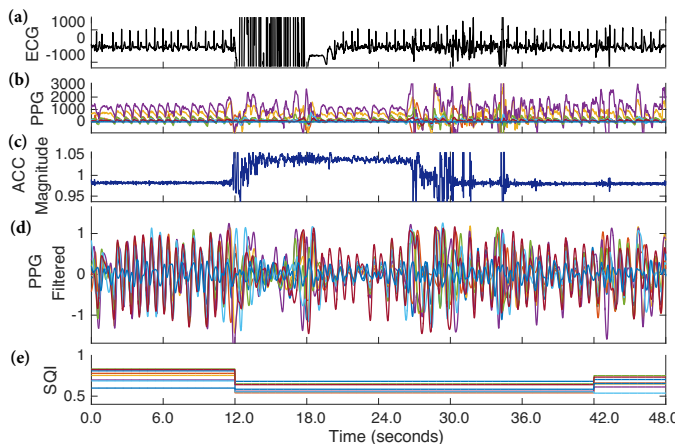


Fig. 2. An example of recorded signals and derived quantities. Continuous recording of ECG over a 48 second window, eight channels of PPG waveforms (in different colors), and magnitude of the accelerometry (ACC) signal are shown in panels (a), (b), and (c), respectively. The filtered PPG signals (as described in section II-B) are shown in panel (d). Panel (e) shows the Hjorth's purity signal quality index (SQI) of all ppg channels. Note the presence of motion-related artifacts from $\sim 12 - 42$ seconds through the record is marked by a PPG channel-specific drop in signal quality.

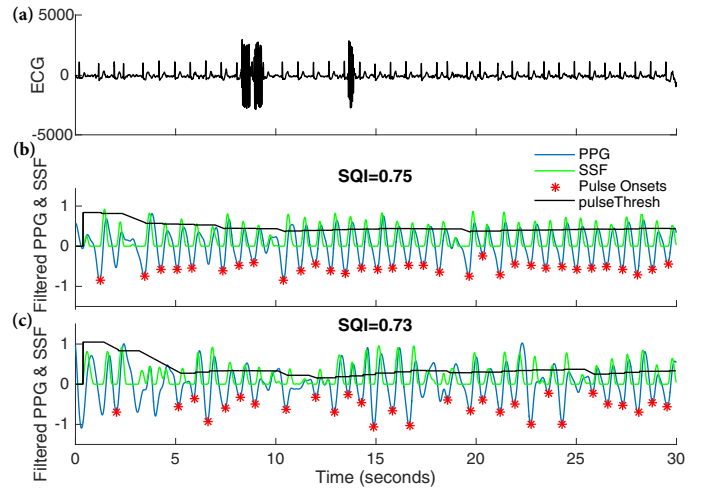


Fig. 3. Pulse onset detection. An example of a 30 seconds window of ECG and PPG recordings (two separate channels) are shown in panels (a)-(c). We utilized an adaptive technique based on the slope sum function (SSF) approach [14] to detect the onset of each PPG pulse (denoted by an asterisk *). The SQI allows us to select the PPG channels with the best quality for pulse interval estimation.

Sample waveforms are depicted in Fig. 2. Out of the total of 46 patients, 36 patients were available at the time of model construction, and therefore were included in the 10-fold cross-validation study (80% training, 20% testing, with random selection with replacement). The outcomes of the remaining 10 subjects (3 Afib and 7 non-AFib records) were kept secret by the clinical team, in order to assess the out-of-sample performance of the algorithm ("single-blinded experiment").

B. Preprocessing

Outlier rejection and amplitude normalization was performed on non-overlapping 30 seconds windows of each PPG channel. After subtracting the signal mean, we calculated the lower 1-percentile and upper 99-percentile of the PPG signal within each window segment, any value surpassing these two thresholds were set to the corresponding threshold (extreme value clipping), and the resulting segment was normalized by the segment max. Next, each PPG channel was bandpass filtered to remove frequencies outside the range of 0.2 – 10Hz, using a butterworth filter of order three. Within each window we performed PPG pulse onset detection using the slope sum function (SSF) approach [14] (see Fig. II-A). Additionally, we calculated a signal quality index (SQI) for each window using the *Hjorth's purity* quality metric (SQI_{purity}) [15], which takes a value of zero for random noise and one for a sinusoidal signal. A second SQI was calculated by taking the average value of the magnitude of the accelerometer ($ACC = \sqrt{x^2 + y^2 + z^2}$) within each 30 seconds window (SQI_{acc}).

C. Features

Since our goal was to accurately detect AFib using beat-to-beat interval variations, we calculated several metrics of

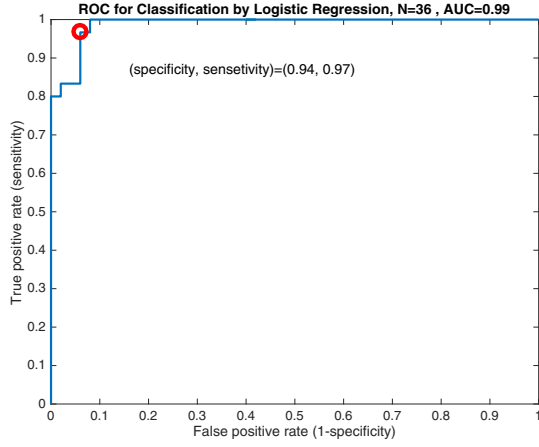


Fig. 4. Receiver operating characteristic (ROC), based on 10-fold cross-validated performance on the testing dataset.

variability/complexity of the PPG pulse onset time series. We calculated the sample entropy with the embedding dimensions $m = 1$, and 2 ($SampEn_1$ and $SampEn_2$, respectively). Next, we calculated the standard deviation of the beat-to-beat interval time series (STD), as well as a robust version of the standard deviation (STD_r), after discarding the intervals outside the 0.05-0.95 percentile range (we assumed that the extreme intervals are due to erroneous pulse onset detection). Finally, we calculated a weighted standard deviation feature, by taking the inverse of the ACC waveform within each interval as the weighing factor when calculating the standard deviation (STD_{acc}). Within each window, we used the SQI_{purity} to choose a channel with the highest signal quality, then we used the features of only that channel. Since, our goal was to make a single binary assessment of the presence of AFib per each record, we took the median of the features across all the temporal windows as our final features (7 features). Additionally, we took the min and the max of the sample entropy features (4 additional features). The complete list of all 11 features are given in the first column of Table I.

D. Atrial Fibrillation Detection

Given the features we built an *Elastic Net* logistic model to classify each feature into one of two classes of AFib versus non-AFib. All features were normalized by removing the mean and dividing by the standard deviation of the features. These two parameters were only calculated on the training set and were used to normalize testing and validation features. The regression hyper-parameters (regularization factors) were chosen using internal 10-fold cross-validation on the training set. Predicted probabilities of outcome for each test set example was pooled across all 10 folds to calculate the final AUC.

III. RESULTS

A. 10-fold cross-validation

We utilized a 10-fold cross-validation study to assess the generalization performance of our AFib detection algorithm ($N = 36$ patients). Table I provides a summary of the fitted

Features	β	e^{β}	p -value
SQI_{purity}^{med}	-4.37 [-4.67, -1.75]	0.01 [0.01, 0.17]	<0.01
SQI_{acc}^{med}	-3.73 [-5.02, -2.74]	0.03 [0.01, 0.06]	<0.01
$SampEn_1^{med}$	-3.09 [-4.02, -2.78]	0.05 [0.02, 0.06]	<0.01
$SampEn_2^{med}$	-5.72 [-6.22, -3.32]	0.00 [0.00, 0.04]	<0.01
STD^{med}	-4.21 [-5.40, -2.81]	0.02 [0.00, 0.06]	<0.01
STD_r^{med}	-3.45 [-4.33, -2.44]	0.03 [0.01, 0.09]	<0.01
STD_{acc}^{med}	-3.57 [-4.92, -2.61]	0.03 [0.01, 0.07]	<0.01
$SampEn_1^{min}$	-5.46 [-5.97, -4.29]	0.00 [0.00, 0.01]	<0.01
$SampEn_2^{min}$	-3.10 [-4.11, -2.38]	0.05 [0.02, 0.09]	<0.01
$SampEn_1^{max}$	1.47 [0.88, 1.90]	4.37 [2.42, 6.71]	<0.01
$SampEn_2^{max}$	-1.94 [-2.13, -1.58]	0.14 [0.12, 0.21]	<0.01

TABLE I
SUMMARY OF THE LOGISTIC CLASSIFIER FEATURES AND MODEL COEFFICIENTS.

model coefficients (β), the associated odds-ratios e^{β} , and p-values. The reported medians and interquartiles in columns 2 and 3 indicate that the identified model parameters were fairly stable across the different folds. Notably, maximum of the sample entropy over all windows ($SampEn_1^{max}$) had the largest odds ratio (4.37). Similarly, $SampEn_1^{min}$ resulted in the lowest odds ratio. Nevertheless, all features were significant predictors of AFib.

As shown in Fig. 4, the pooled AUC was 0.99. Moreover, we achieved a sensitivity ($\frac{TP}{TP+FN}$) of 0.97 for a specificity ($\frac{TN}{TN+FP}$) of 0.94. At the corresponding threshold, model's accuracy ($\frac{TP+TN}{TP+FP+FN+TN}$) was 0.95. Here TP is the number of true positive, TN is the number of true negative, FP is the number of false positive and FN is the number of false negative.

B. Single-blinded experiment

We utilized an ensemble of 10 classifiers from the previous section (one per fold) to determine the presence of AFib in the 10 remaining records, by taking the majority vote over all classifiers. This method achieved an accuracy of 100% on the unseen patients, in this single-blinded experiment.

IV. DISCUSSION AND FUTURE DIRECTION

In this study, we derived and validated an algorithm to detect AFib with a high degree of accuracy using just PPG and accelerometry data from a multisensor wristband. Long term, a successful algorithm may enable superior sensitivity in diagnosing patients with AFib, and be particularly useful to the large proportion of patients who do not have symptoms and, without knowing, are at a five-fold increased risk of stroke. As opposed to twice-daily rhythm checks for AFib that were presented in the STROKESTOP study [8], our proposed algorithm, which could be integrated with multisensor wristbands, offers the added benefit of longer-term continuous monitoring, and potentially increased sensitivity of AFib detection. Because anticoagulation can decrease stroke risk by 2/3 in such patients, the potential for benefit from such screening devices is high. Given the rapid increase in at-risk persons due to the aging baby-boom generation and obesity

epidemic, this represents an increasingly large public health opportunity in prevention of stroke. Furthermore, because such wristbands are a potentially low-cost and low-burden technology, the potential economic benefits in offsetting the multi-billion dollar healthcare costs of AFib and related strokes are large as well.

Analysis of heart rate dynamics has been previously shown to discriminate among AFib, normal sinus rhythm and sinus rhythm with frequent ectopy [16]. Machine learning approaches has been proposed to optimally weight different ECG features for accurate detection of AFib [17]. Moreover, Oster and colleagues developed an Open source Java-based ECG analysis software and Android app for AFib screening [18]. Wearable technologies such as the one presented here have the potential to induce a paradigm shift in traditional home-monitoring strategies that normally require physician adjudication and burdensome devices that also result in significant delays between arrhythmia and diagnosis [19]. In particular, application of signal quality indices and sensor fusion techniques [20], [15] has the potential to circumvent the Achilles' heel of the wearables technology, that is motion and muscle artifacts. To this end, Simband's multisensor technology provides a unique testbed for application of signal processing and machine learning techniques that utilize accelerometry and sensor fusion to achieve a robust performance under realistic device utilization scenarios.

Our future works include improving PPG pulse onset detection and inclusion of morphological features and other metrics of time series irregularity [21]. Moreover, although here we only considered the problem of AFib detection using multiple PPG channels and accelerometry, Simband's recent addition of an ECG sensor provides a unique opportunity to compare the ECG-based AFib detection techniques against our proposed method. We hypothesize that ECG and PPG provide complementary information, since various types of recording artifacts tend to affect each of these sensors differently. Finally, although our cross-validation and single-blinded experiments have proved very promising, we continue to collect more patient data and aim to test the reproducibility of these results on a larger cohort (including several hundred patients) and across multiple Simband devices.

ACKNOWLEDGMENT

S.N. is grateful for an NIH early career development award in biomedical big data science (1K01ES025445-01A1). A.S. is grateful to the American Heart Association (SDG-20593449) and National Institutes of Health (UL1-TR-000454, KL2-TR-00045, K23-HL-127251). M.G. Is grateful to the Salerno foundation for their support of his PhD studies. The Simband was provided on loan at no charge by The Samsung Electronics. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Samsung or the funding agencies.

REFERENCES

- [1] L. Friberg and L. Bergfeldt, "Atrial fibrillation prevalence revisited," *Journal of internal medicine*, vol. 274, no. 5, pp. 461–468, 2013.
- [2] Y. Miyasaka, M. E. Barnes, B. J. Gersh, S. S. Cha, K. R. Bailey, W. P. Abhayaratna, J. B. Seward, and T. S. Tsang, "Secular trends in incidence of atrial fibrillation in olmsted county, minnesota, 1980 to 2000, and implications on the projections for future prevalence," *Circulation*, vol. 114, no. 2, pp. 119–125, 2006.
- [3] P. A. Wolf, R. D. Abbott, and W. B. Kannel, "Atrial fibrillation as an independent risk factor for stroke: the framingham study," *Stroke*, vol. 22, no. 8, pp. 983–988, 1991.
- [4] J. S. Healey, S. J. Connolly, M. R. Gold, C. W. Israel, I. C. Van Gelder, A. Capucci, C. Lau, E. Fain, S. Yang, C. Bailleul *et al.*, "Subclinical atrial fibrillation and the risk of stroke," *New England Journal of Medicine*, vol. 366, no. 2, pp. 120–129, 2012.
- [5] S. H. Hohnloser, D. Pajitnev, J. Pogue, J. S. Healey, M. A. Pfeffer, S. Yusuf, and S. J. Connolly, "Incidence of stroke in paroxysmal versus sustained atrial fibrillation in patients taking oral anticoagulation or combined antiplatelet therapy: an active w substudy," *Journal of the American College of Cardiology*, vol. 50, no. 22, pp. 2156–2161, 2007.
- [6] A. J. Camm, G. Corbucci, and L. Padeletti, "Usefulness of continuous electrocardiographic monitoring for atrial fibrillation," *The American journal of cardiology*, vol. 110, no. 2, pp. 270–276, 2012.
- [7] R. G. Hart, O. Benavente, R. McBride, and L. A. Pearce, "Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis," *Annals of internal medicine*, vol. 131, no. 7, pp. 492–501, 1999.
- [8] E. Svennberg, J. Engdahl, F. Al-Khalili, L. Friberg, V. Frykman, and M. Rosenqvist, "Mass screening for untreated atrial fibrillation the strokestop study," *Circulation*, vol. 131, no. 25, pp. 2176–2184, 2015.
- [9] M. Aronsson, E. Svennberg, M. Rosenqvist, J. Engdahl, F. Al-Khalili, L. Friberg, V. Frykman-Kull, and L.-Å. Levin, "Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ecg recording," *Europace*, p. euv083, 2015.
- [10] A. Carpenter and A. Frontera, "Smart-watches: a potential challenger to the implantable loop recorder?" *Europace*, p. euv427, 2016.
- [11] S. Dash, K. Chon, S. Lu, and E. Raeder, "Automatic real time detection of atrial fibrillation," *Annals of biomedical engineering*, vol. 37, no. 9, pp. 1701–1709, 2009.
- [12] D. D. McMANUS, J. W. CHONG, A. SONI, J. S. SACZYNSKI, N. ESA, C. NAPOLITANO, C. E. DARLING, E. BOYER, R. K. ROSEN, K. C. FLOYD *et al.*, "Pulse-smart: Pulse-based arrhythmia discrimination using a novel smartphone application," *Journal of cardiovascular electrophysiology*, vol. 27, no. 1, pp. 51–57, 2016.
- [13] "Samsung Simband's official website." [Online]. Available: <https://www.simband.io/>
- [14] A. V. Deshmane, "False arrhythmia alarm suppression using ecg, abp, and photoplethysmogram," Ph.D. dissertation, Citeseer, 2009.
- [15] S. Nemati, A. Malhotra, and G. D. Clifford, "Data fusion for improved respiration rate estimation," *EURASIP journal on advances in signal processing*, vol. 2010, p. 10, 2010.
- [16] M. Carrara, L. Carozzi, T. J. Moss, M. de Pasquale, S. Cerutti, M. Ferrario, D. E. Lake, and J. R. Moorman, "Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy," *Physiological Measurement*, vol. 36, no. 9, p. 1873, 2015.
- [17] R. Colloca, A. E. Johnson, L. Mainardi, and G. D. Clifford, "A support vector machine approach for reliable detection of atrial fibrillation events," in *Computing in Cardiology Conference (CinC), 2013*. IEEE, 2013, pp. 1047–1050.
- [18] J. Oster, J. Behar, R. Colloca, Q. Li, Q. Li, and G. D. Clifford, "Open source java-based ecg analysis software and android app for atrial fibrillation screening," in *Computing in Cardiology Conference (CinC), 2013*. IEEE, 2013, pp. 731–734.
- [19] M. P. Turakhia, A. J. Ullal, D. D. Hoang, C. T. Than, J. D. Miller, K. J. Friday, M. V. Perez, J. V. Freeman, P. J. Wang, and P. A. Heidenreich, "Feasibility of extended ambulatory electrocardiogram monitoring to identify silent atrial fibrillation in high-risk patients: The screening study for undiagnosed atrial fibrillation (study-af)," *Clinical cardiology*, vol. 38, no. 5, pp. 285–292, 2015.
- [20] Q. Li, R. G. Mark, and G. D. Clifford, "Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a kalman filter," *Physiological measurement*, vol. 29, no. 1, p. 15, 2007.
- [21] C. Liu, K. Li, L. Zhao, F. Liu, D. Zheng, C. Liu, and S. Liu, "Analysis of heart rate variability using fuzzy measure entropy," *Computers in biology and Medicine*, vol. 43, no. 2, pp. 100–108, 2013.