

Feasibility of atrial fibrillation detection from a novel wearable armband device



Syed Khairul Bashar, MS,* Md-Billal Hossain, BS,* Jesús Lázaro, PhD,^{†‡}
Eric Y. Ding, MS,[§] Yeonsik Noh, PhD,^{||} Chae Ho Cho, MS,* David D. McManus, MD, FHRS,[§]
Timothy P. Fitzgibbons, MD, PhD,[§] Ki H. Chon, PhD*

From the *Department of Biomedical Engineering, University of Connecticut, Storrs, Connecticut, [†]Aragon Institute for Engineering Research, University of Zaragoza, Zaragoza, Spain, [‡]Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain, [§]Division of Cardiology, University of Massachusetts Medical School, Worcester, Massachusetts, and ^{||}College of Nursing, University of Massachusetts, Amherst, Massachusetts.

BACKGROUND Atrial fibrillation (AF) is the world's most common heart rhythm disorder and even several minutes of AF episodes can contribute to risk for complications, including stroke. However, AF often goes undiagnosed owing to the fact that it can be paroxysmal, brief, and asymptomatic.

OBJECTIVE To facilitate better AF monitoring, we studied the feasibility of AF detection using a continuous electrocardiogram (ECG) signal recorded from a novel wearable armband device.

METHODS In our 2-step algorithm, we first calculate the R-R interval variability-based features to capture randomness that can indicate a segment of data possibly containing AF, and subsequently discriminate normal sinus rhythm from the possible AF episodes. Next, we use density Poincaré plot-derived image domain features along with a support vector machine to separate premature atrial/ventricular contraction episodes from any AF episodes. We trained and validated our model using the ECG data obtained from a subset of the MIMIC-III (Medical Information Mart for Intensive Care III) database containing 30 subjects.

RESULTS When we tested our model using the novel wearable armband ECG dataset containing 12 subjects, the proposed method achieved sensitivity, specificity, accuracy, and F1 score of 99.89%, 99.99%, 99.98%, and 0.9989, respectively. Moreover, when compared with several existing methods with the armband data, our proposed method outperformed the others, which shows its efficacy.

CONCLUSION Our study suggests that the novel wearable armband device and our algorithm can be used as a potential tool for continuous AF monitoring with high accuracy.

KEYWORDS Accuracy; Armband; Atrial fibrillation; Classification; Entropy; Feature selection; Premature atrial contraction; Template; Wavelet

(Cardiovascular Digital Health Journal 2021;2:179–191) © 2021 Heart Rhythm Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Atrial fibrillation (AF) is the most common arrhythmia, associated with significant morbidity. It has become a global health concern as the prevalence and associated mortality have grown significantly in the past decade.¹ AF is associated with a 5-fold increased risk of stroke, heart failure, and dementia and a 2-fold increased risk of death.² One-third of hospitalizations related to cardiac arrhythmias are due to AF-related complications and 1 in 5 of all strokes is attributed to AF.^{3,4} AF affects more than 3 million people in the United States and by 2050 it is estimated it will affect more than 7.5 million.⁵ However, the

total AF burden could be even higher, as about one-third of patients have asymptomatic or paroxysmal AF that remains undiagnosed.⁶ Continuous monitoring would increase the likelihood of early AF detection, thus allowing early treatment for primary and secondary stroke prevention.^{2,3}

Currently, research is going on for continuous AF monitoring using various modalities and smart devices, including smartwatch photoplethysmographic-based AF detection,⁷ smartphone camera-based AF detection,⁸ and AF detection using facial video monitoring.⁹ However, these approaches are not feasible for continuous monitoring, especially for paroxysmal AF. More importantly, these smart devices' recordings suffer from frequent motion artifacts, which limit the AF detection to being accurate only when the subject is not moving. However, electrocardiogram (ECG) being the gold standard for diagnosing AF, continuous monitoring of ECG is often

Address reprint requests and correspondence: Dr Ki H. Chon, Professor of Biomedical Engineering, University of Connecticut, Bronwell 217C, 260 Glenbrook Rd, Unit 3247, Storrs, CT 06269-3247. E-mail address: ki.chon@uconn.edu

KEY FINDINGS

- We have studied the feasibility of atrial fibrillation (AF) monitoring using a novel wearable armband device.
- We have developed a machine learning model using the intensive care unit electrocardiogram (ECG) data (MIMIC-III) and applied that model directly to the wearable armband ECG data. When tested on the armband ECG, our model achieved high performance and outperformed several other existing methods.
- The participants had satisfactory mean response about general impression of the device, the ease to use, and the comfortability of the armband device.
- This novel wearable armband device has the potential to be used for detection of asymptomatic or paroxysmal AF that remains undiagnosed.

recommended to detect paroxysmal AF episodes.^{6,10} The ECG devices have fewer motion artifact issues when compared to photoplethysmographic-based smart devices. The current options for continuous wearable ECG monitoring are Holter/event monitors and the recently developed patch devices. However, these devices have inherent disadvantages—the former require obtrusive leads, and they all use hydrogel-based electrodes with adhesives that often cause skin irritation.¹¹ To address these issues, we have recently developed a wearable armband device for continuous ECG monitoring¹² that uses dry electrodes that do not require hydrogel. This self-contained armband device can be worn on the left upper arm for continuous ECG monitoring, with no external leads, and has shown promising reliable ECG recording coverage during both night-time and daytime monitoring.¹² As a result, the purpose of this work is to analyze the feasibility of AF detection using the ECG recordings obtained from this novel wearable armband device.

AF detection using the ECG recordings from a wearable armband device has not been studied before. Continuous AF monitoring using an easy-to-wear ECG device with longer battery life would potentially help undiagnosed AF episodes and prevent AF-related complications. As a result, in this work we analyze the feasibility of AF monitoring using the recently developed novel wearable ECG device, which is placed on the upper arm. First, based on the ECG data collected from the intensive care unit (MIMIC-III [Medical Information Mart for Intensive Care III] dataset subset), an AF detection model based on R-R interval variability was developed. Moreover, to overcome false-positives owing to ectopic beats, novel density Poincaré image-based features were incorporated and fed into a 2-step machine learning approach. Finally, the developed model was tested using the wearable armband ECG data without modifying any parameters learned from the training dataset of MIMIC-III. We believe the resulting algorithm could successfully be

embedded in the armband device for real-time continuous monitoring for episodes of AF.

Methods

This study consisted of first developing an automated AF detection algorithm using archived ECG data and then testing the trained algorithm using the wearable armband ECG data. Ventricular response variability–based features including density Poincaré plot and statistical parameters are used along with machine learning classifiers to detect the AF segments. For training and validation, stored ECG recordings from intensive care units (ICU) were used, whereas for testing, the armband ECG data were used.

Description of datasets

Two different datasets were used in this study, which are described below.

MIMIC-III dataset (training data)

For the algorithm training and validation, ECG recordings obtained from 30 subjects in the MIMIC-III data were used. MIMIC-III is a large open-source medical record database publicly available in PhysioNet.¹³ MIMIC-III contains de-identified health-related data from patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012.¹⁴ However, no annotations were provided for this dataset. As a result, ECG signals were annotated by board-certified physicians specializing in AF management for our study. According to the physicians' annotations, we selected 30 subjects: 10 with AF, 10 with normal sinus rhythm (NSR), and 10 subjects with premature atrial/ventricular contractions (PACs/PVCs). Here we do not discriminate PACs from PVCs; hence they are commonly referred to as PACs here. Moreover, the 30 subjects we selected were identified to have sepsis according to the International Classification of Diseases, Ninth Revision (ICD-9) codes. The ECG recordings were sampled at 125 Hz and the measuring unit was millivolts (mV).

Wearable armband data (testing data)

The armband ECG data were collected using a recently developed wearable armband device.¹² This device is designed to be worn on the upper left arm. It consists of 3 pairs of carbon-black dry electrodes that record 3 different ECG channels and 1 electromyogram channel simultaneously. [Figure 1A](#) and [1B](#) shows the prototype of the armband device and the 3 pairs of electrodes, whereas [Figure 1C](#) and [1D](#) shows the dry electrode pairs and their connections. [Figure 1E](#) shows a subject wearing the armband device (without the cover) on the upper left arm.

This preliminary study data consisted of 12 participants (over 21 years of age) with AF recruited from the University of Massachusetts Memorial Health Care system. All participants were approached for informed consent prior to a scheduled elective cardioversion or ablation procedure to treat their AF. Trained research staff instructed participants on the

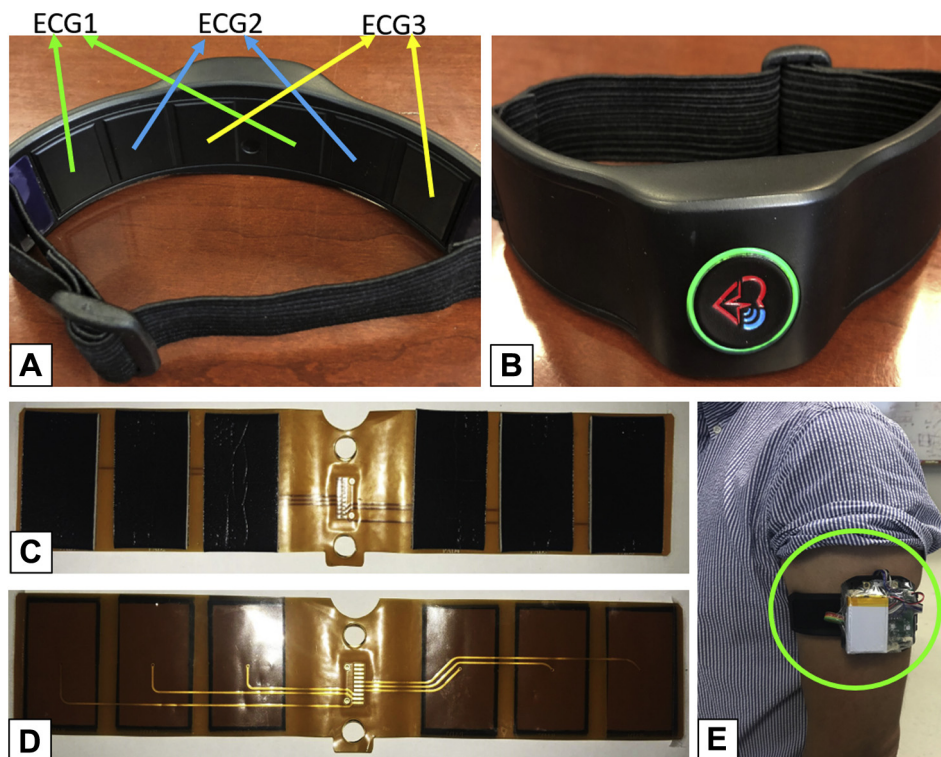


Figure 1 Wearable armband device prototype. **A:** Backside of the device with 3 pairs of electrodes. **B:** Front view of the armband device. **C,D:** Three pairs of carbon-black dry electrodes and their connections, which are incorporated into the armband in panel A. **E:** A subject wearing the device on the upper left arm (without the cover).

appropriate use of the wearable armband device and discussed study procedures. After obtaining consent, a research staff member placed the ECG armband on the left upper arm of the participants while they were simultaneously monitored on ECG telemetry prior to their cardiac procedure.

Research staff returned to the participant's bedside after their procedure and prior to hospital discharge, placed the ECG armband and an FDA-cleared mobile cardiac telemetry patch (Cardiac Insight Inc, Bellevue, WA) on the participant, and asked each participant to wear both devices for 14 days. At the conclusion of the 14-day follow-up period, participants were asked to return both the armband and the ECG patch. The ECG recordings from the patch were used as the reference. According to the reference telemetry patch data, among the participants, 2 had persistent AF whereas 10 participants had non-AF rhythms (including PAC and NSR). This protocol was approved by the institutional review board of the University of Massachusetts Memorial Hospital (H00013799). The research reported in this paper adhered to the Human Research: Helsinki Declaration as revised in 2013.

Proposed method

Preprocessing

First the ECG signals from the MIMIC-III ICU database were divided into 2-minute nonoverlapping segments. Next, these ECG recordings were checked for noise artifacts using an automated noise detection algorithm.¹⁵ For the armband

ECG signals, since there were 3 simultaneous recordings from the 3 pairs of electrodes, at the beginning of the preprocessing a single channel ECG signal was reconstructed from those 3 channels. This reconstruction was based on a recently developed algorithm that calculates several signal quality indices, and based on those, a single-channel ECG was obtained using the 3 armband channels. Details of this reconstruction algorithm are described in work by Lázaro and colleagues.¹² Figure 2A through 2C shows the 3 simultaneous armband ECG signals after band-pass filtering (3–25 Hz) and Figure 2D shows the reconstructed single-channel ECG signal using the algorithm of Lazaro and colleagues¹² for a sample 20-second armband ECG data segment. From this point onwards, the term “armband ECG” refers to this reconstructed single-channel ECG signal. Similar to the MIMIC ECG recording, this reconstructed ECG was then divided into 2-minute nonoverlapping segments.

For both the MIMIC and the armband ECG signals, once the clean segments were detected and reconstructed by combining only the good ECG channel signals, the R peaks of the ECG signals were detected using a highly accurate peak detection algorithm based on the variable frequency complex demodulation algorithm.¹⁶ Finally, from the detected R peaks, the heart rate (HR) was calculated by taking the reciprocal of the R-R interval series.

Feature extraction

AF is a supraventricular tachyarrhythmia that is characterized by chaotic contraction of the atria. AF leads to irregularly

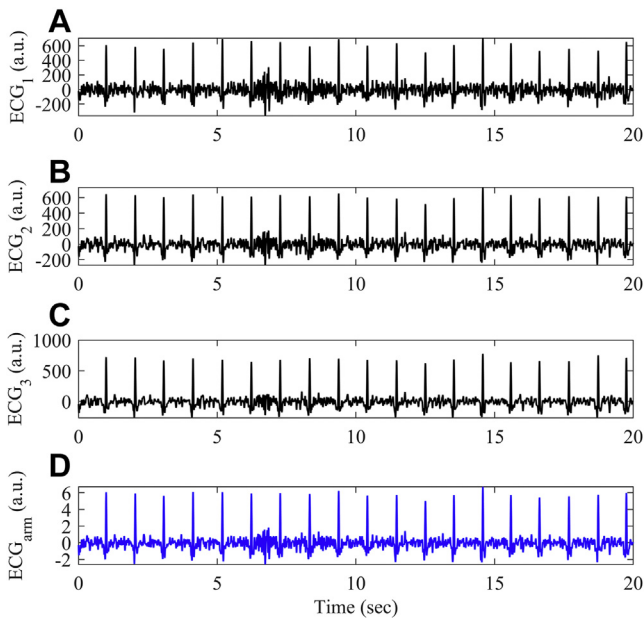


Figure 2 Sample 20-second data from the wearable armband. **A–C:** Three simultaneous electrocardiogram (ECG) signals obtained from the 3 channels of the armband. **D:** Reconstructed single-channel ECG.

irregular R-R intervals with increased beat-to-beat variability and complexity.¹⁷ During an AF episode, the HR varies randomly, compared to the normal heart rhythm (NSR), during which the HR has regular variation.

Figure 3 shows examples of 20-second NSR and AF ECG segments from the MIMIC-III ICU data and the corresponding HR. From the figure, it can be seen that the NSR HR (shown in Figure 3B) varies in a small range (70–73 beats per minute); however, the HR during the AF episode (shown in Figure 3D) varies randomly in a large range (100–160 beats per minute). Similarly, Figure 4A and 4B shows a sample 20-second NSR ECG segment and the corresponding HR obtained from the armband ECG, whereas Figure 4C and 4D shows a sample AF ECG segment and the corresponding HR from the armband.

As this random variation of R-R intervals is the most prominent and easy-to-obtain characteristic of AF, we have used 2 commonly used features—the root mean square of successive differences (RMSSD) and sample entropy (SampEn)—to discriminate between the NSR and AF ECG segments.

RMSSD is a commonly used statistical parameter to measure the beat-to-beat variability of HR. The value of RMSSD is expected to be higher in ECG segments containing AF episodes than in ECG segments containing only NSR, as AF exhibits higher variability than does the regular rhythm.^{17,18} Since heart rate can vary a lot across subjects, the RMSSD value of each 2-minute segment is normalized by the mean of the HR of that segment to counter this variability among different subjects and segments.

SampEn measures the irregularity and randomness of a time series.¹⁹ Since AF subjects have high variability and randomness in their R-R intervals, the value of SampEn is

expected to be higher for the AF subjects. In this study, to estimate SampEn, template length = 1 and tolerance = 60 ms were selected per recommendation in Lake and Moorman.²⁰

However, PACs/PVCs are common benign causes of rhythm irregularity and their frequent occurrences can mimic the irregular beat pattern typical of AF, thus resulting in false-positive AF detection.²¹ To discriminate between AF and PACs, we have used the previously proposed density Poincaré plot approach. PACs occur when an ectopic foci originating in an atrium leads to premature activation of the atria prior to sinoatrial node activation, whereas a PVC occurs when a similar process occurs in a ventricle.²²

Figure 5A shows a sample 20-second PAC ECG segment from the MIMIC data and the corresponding HR is shown in Figure 5B. Figure 5C and 5D show similar examples from the armband data. From Figure 5B and 5D, it is evident that the PAC HR has more variation compared to the previously shown NSR HR (shown in Figures 3B and 4B), especially where the PAC beats occur. As a result, this kind of ECG segment will be incorrectly detected as AF based on RMSSD and SampEn.

To discriminate the PACs from possible AF segments, we used density Poincaré plots derived from the differences of HR.

Feature extraction from density Poincaré plot

To detect the repetitive ectopic beat patterns in the PAC rhythms, Poincaré trajectory-based methods are commonly used in literature.^{22,23} Poincaré plots are generated from the differences of the heart rates. Figure 6A shows a Poincaré plot obtained from a sample 2-minute PAC heart rate, whereas Figure 6D shows a Poincaré plot for an AF segment. The PAC Poincaré plot has a triangular shape (ie, a kite shape) for the repetitive beat pattern, whereas the AF Poincaré plot has a random pattern because the HR varies randomly.

However, these kinds of binary Poincaré plots can misdetect the PAC patterns if the triangular shape is not obvious owing to heart rate trajectories not being perfectly overlapped with each other or owing to noisy beats. Moreover, the binary Poincaré plots do not provide any information on how many times the trajectory lines are overlapped. We are interested in repeated patterns of trajectory lines, as they provide consistency and repeatability of PAC patterns. As a result, recently proposed novel “density” Poincaré plot is used here to better capture the overlapped trajectory information.²⁴ This plot contains how many times the trajectory lines are overlapped, which indicates the density of the overlap; hence, the plot is defined as the density Poincaré plot. Details of the density Poincaré plot can be found in Bashar and colleagues.²⁴ Figure 6B and 6E shows the density Poincaré plots from both the PAC and AF binary Poincaré plots shown in Figure 6A and 6D, respectively. From the figure, it can be seen that the density plots have varying colors that correspond to different densities. The density distribution

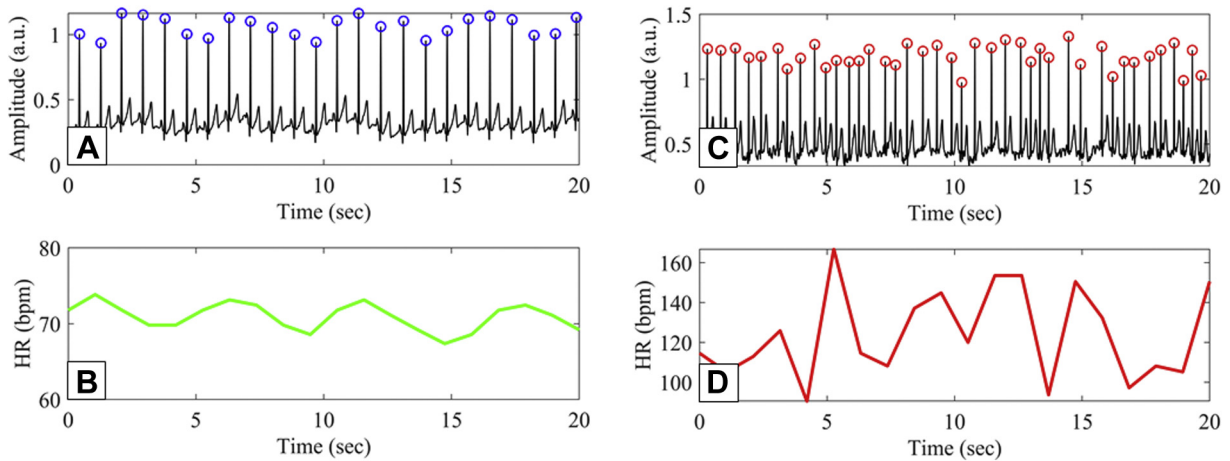


Figure 3 Examples from the MIMIC intensive care unit data. **A**: Sample 20-second normal sinus rhythm electrocardiogram (ECG) recording, and **B**: The corresponding heart rate (HR). **C**: Sample 20-second atrial fibrillation ECG segment, and **D**: the corresponding HR. The QRS peaks are denoted by blue and red circles, respectively.

is better discernible in the 3D density plot shown in [Figure 6C](#) and [6D](#).

From the density Poincaré images, 2 different image domain-based feature extraction methods are used: image template correlation and discrete wavelet transform.²⁵

Template correlation-based features

Image template correlation is a well-known and commonly used method in pattern recognition and image processing.²⁶ The underlying idea is to first calculate the 2D correlation coefficient between a template image and the input image. Next, the correlation value is used as a feature to discriminate PACs from AF. Prior to calculating the correlation, both the template and input images were normalized in $[0, 1]$ range. From the MIMIC-III subset, 12 density Poincaré images are selected as templates by visual inspection and these

template images are used for feature extraction; hence, these are removed from the dataset. The density Poincaré templates include 6 from AF and 6 from PACs. [Figure 7](#) shows some of the representative templates: [Figure 7A–7C](#) for AF and [Figure 7D–7F](#) for PACs.

Discrete wavelet transform-based features

Discrete wavelet transform (DWT) is another popular and well-established technique for extracting features from signals and images, especially with various applications in biomedical signal and image processing.^{27,28} In DWT, the input signal or image is analyzed at different scales and thus time-frequency spectral information can be obtained. In this study, 2D DWT was performed on the input density Poincaré images. At each level of 2D DWT, 4 different sub-bands are obtained, which are

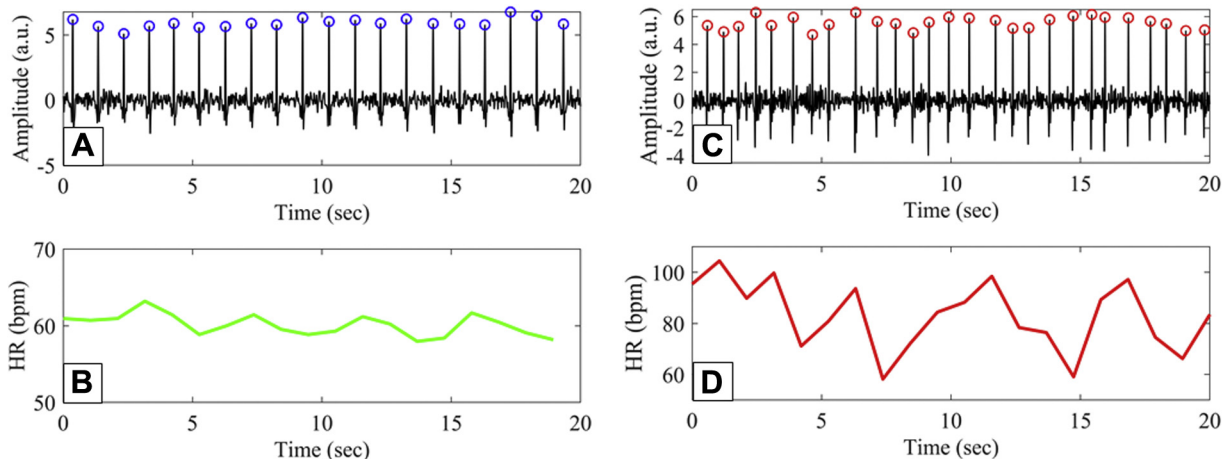


Figure 4 Examples from the wearable armband electrocardiogram (ECG) data. **A**: Sample 20-second normal sinus rhythm ECG recording, and **B**: the corresponding heart rate (HR). **C**: Sample 20-second atrial fibrillation ECG segment, and **D**: the corresponding HR. The QRS peaks are denoted by blue and red circles, respectively.

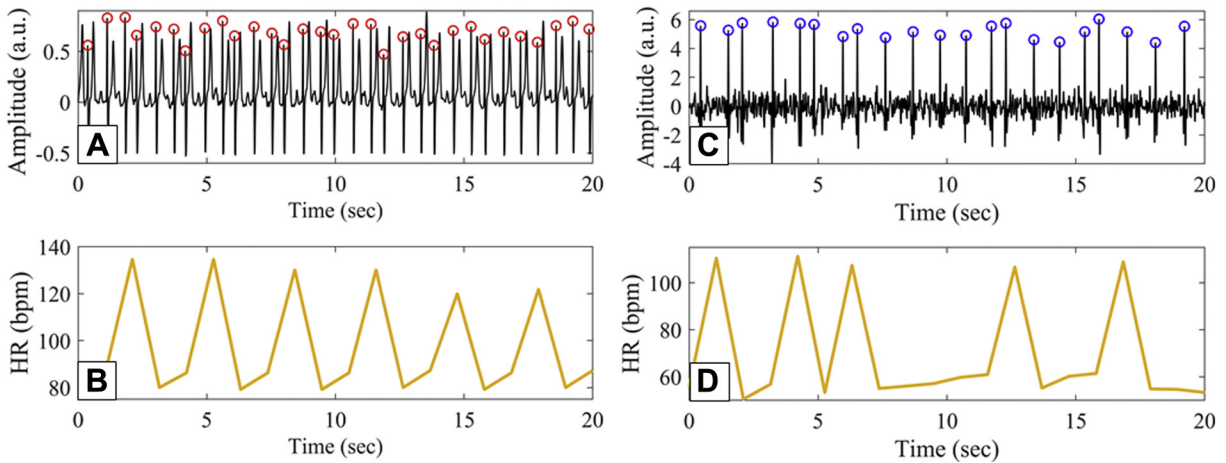


Figure 5 A,B: Example of a 20-second premature atrial contraction (PAC) electrocardiogram (ECG) segment (A) and the corresponding heart rate (HR) obtained from the MIMIC data (B). C,D: Example of a 20-second PAC ECG segment (C) and the corresponding HR obtained from the wearable armband ECG data (D).

called approximation, horizontal, vertical, and diagonal coefficients. In this study, Daubechies 4 wavelets with 2 levels of decomposition are applied to each density Poincaré image.

Figure 8 shows the 4 DWT coefficient matrices (approximation, horizontal, vertical, and diagonal) for

each of the 2 levels obtained from a sample AF density Poincaré image. Figure 9 shows the similar matrices for a sample PAC image.

From each of the DWT coefficient matrices, the spectral energy and Shannon entropy were calculated. Shannon entropy is a statistical measurement of disorder and

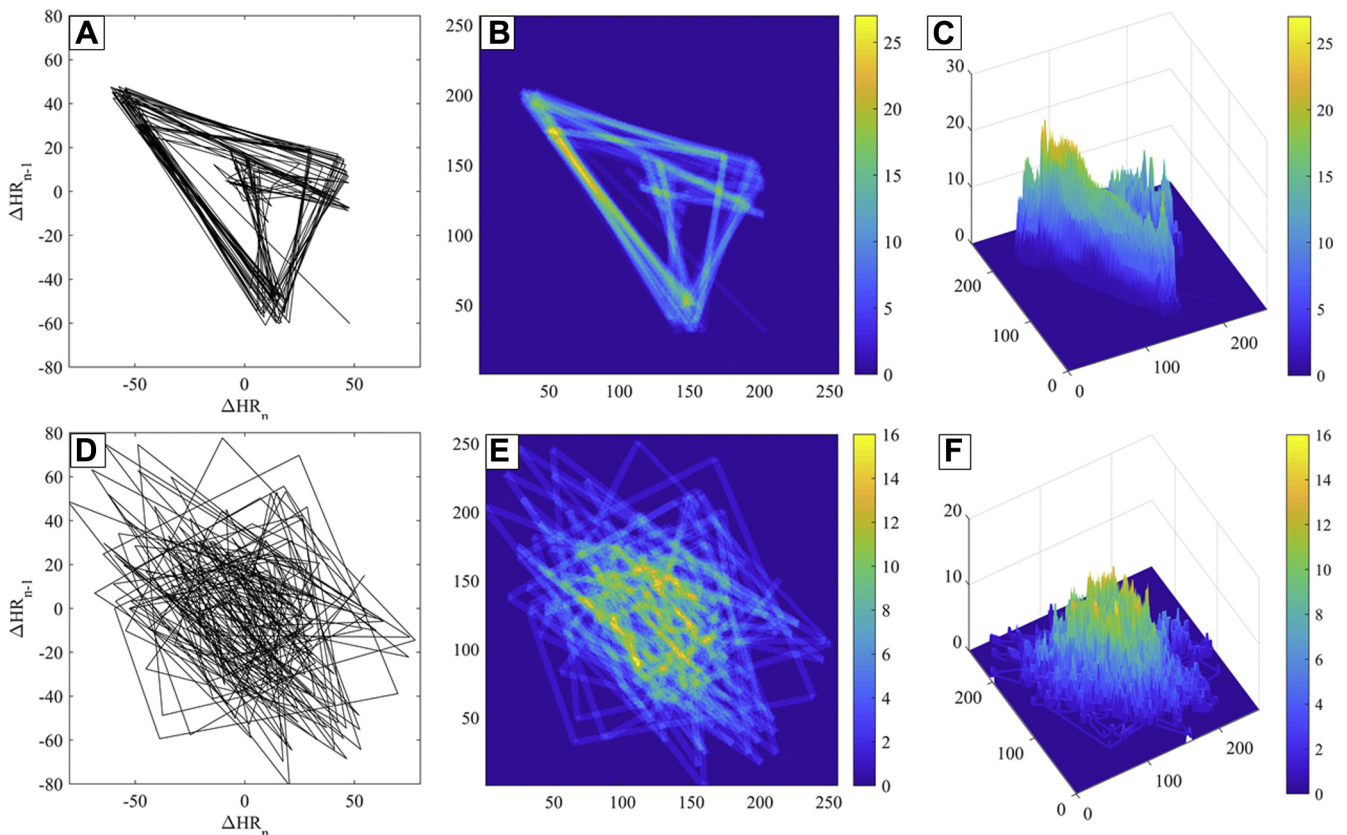


Figure 6 A: Binary Poincaré plot, B: density Poincaré plot, and C: 3D density plot obtained from a sample 2-minute premature atrial contraction segment. D-F: Similar plots for a sample 2-minute atrial fibrillation segment.

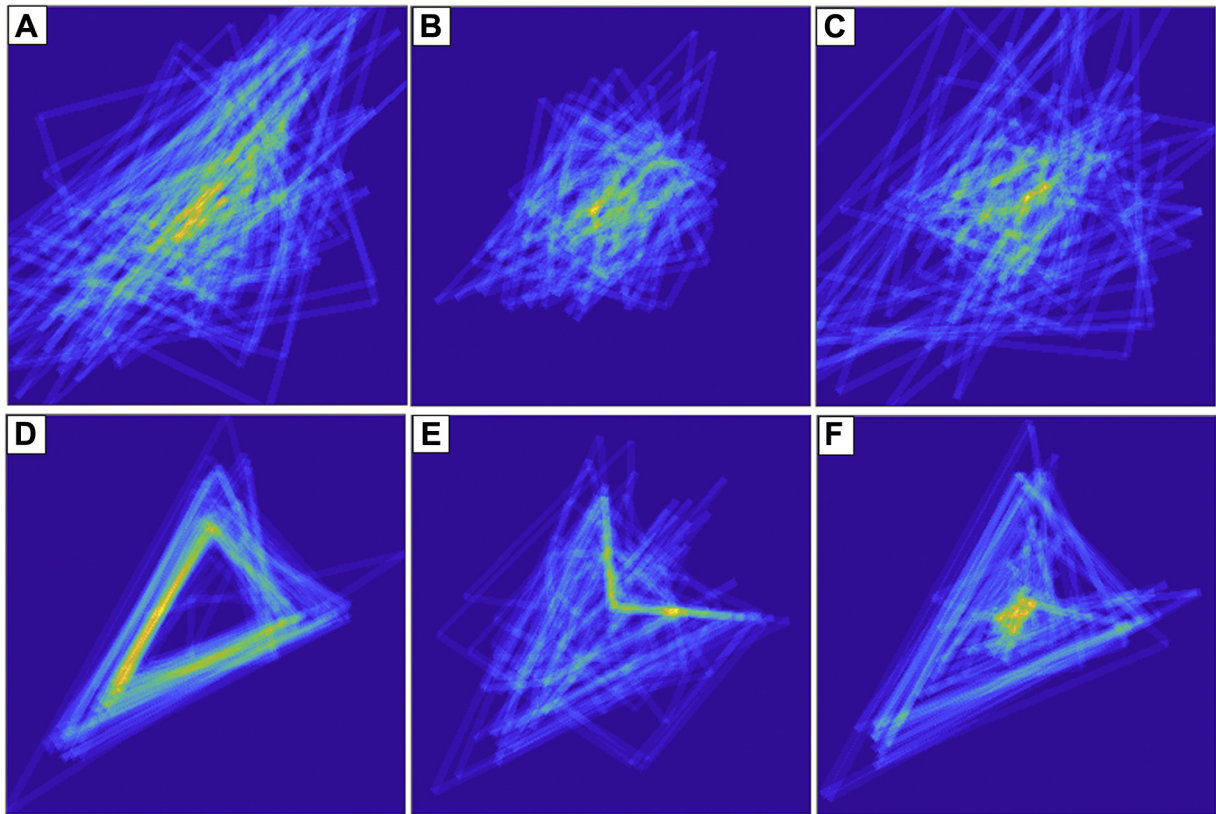


Figure 7 A–C: Three density Poincaré image templates for atrial fibrillation. D–F: Three density Poincaré image templates for premature atrial contractions.

uncertainty. The underlying hypothesis is that by analyzing the entropy and energy from different DWT coefficients, useful features can be obtained to discriminate between AF and PAC. Moreover, after normalizing the DWT coefficient over the $[0, 1]$ range, the entropy is calculated, which is referred to as normalized entropy. As a result, from 2 levels of decomposition, 8 energy features and 16 entropy features are obtained.

Feature Selection

Since we have extracted many features using the density Poincaré plot, we examined down-selection of features. Feature selection is very important in machine learning, as different features can describe different characteristics of the data distribution. Among many available feature selection methods, the infinite latent feature selection (ILFS) method was used in this study.²⁹ ILFS is a recently

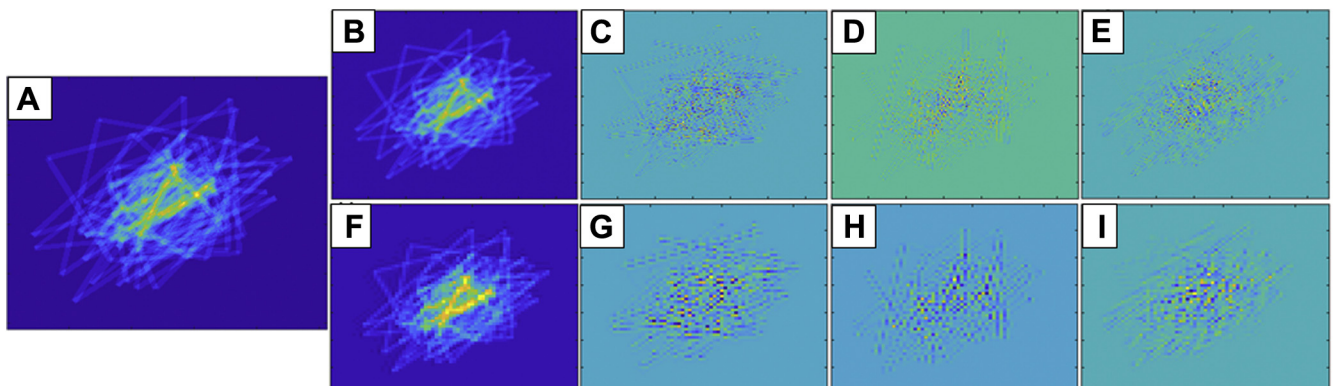


Figure 8 A: The input atrial fibrillation density Poincaré image. B–E: Four coefficient matrices obtained from level 1 of the discrete wavelet transform (DWT). F–I: Four coefficient matrices obtained from level 2 of the DWT. The 4 coefficients are approximation, horizontal, vertical, and diagonal, respectively.

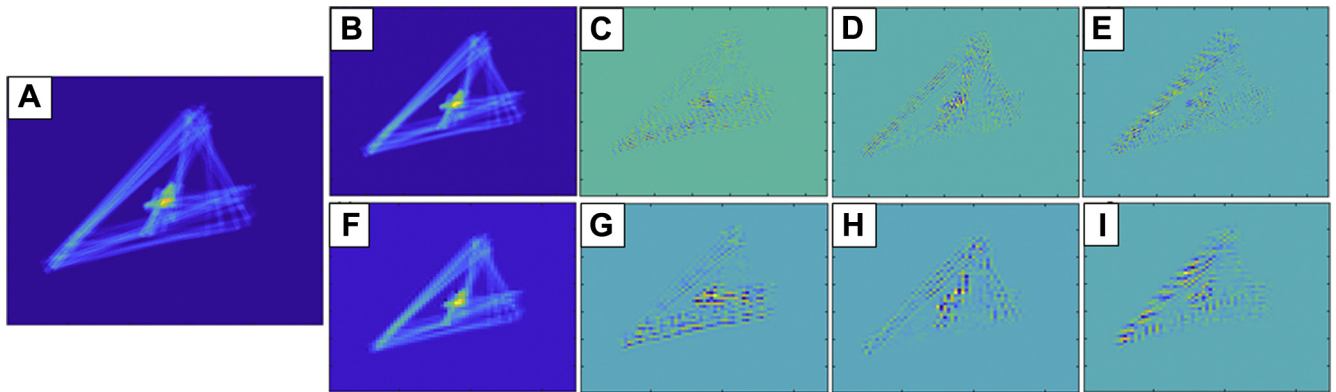


Figure 9 A: The input premature atrial contraction density Poincaré image. B–E: Four coefficient matrices obtained from level 1 of the discrete wavelet transform (DWT). F–I: Four coefficient matrices obtained from level 2 of the DWT. The 4 coefficients are approximation, horizontal, vertical, and diagonal, respectively.

developed method that is based on a probabilistic latent graph-based model, and we chose this approach as it has been shown to be effective. Here, feature ranking is performed by considering all the possible subsets of features where relevancy is modeled as a latent variable. Details of the ILFS algorithm are provided in Roffo and colleagues.²⁹

AF vs non-AF ECG segment classification

By combining the above-mentioned methods, the discrimination of AF vs non-AF ECG segments was performed in 2 steps. Each step was associated with a binary classification and for that purpose, the support vector machine (SVM) classifier was used.

SVM is a popular and widely used tool for solving binary classification problems owing to its excellent generalization properties. SVM works by finding a maximum margin between the training data and the decision boundary.³⁰ The training samples that are closest to the decision boundary are called support vectors.

The first step includes classifying each input ECG segment as either NSR or possible AF classes. For this stage, we have used RMSSD and SampEn as features; linear SVM was implemented as the classifier. This “possible AF” category includes PACs and AF.

cascaded SVM provides AF and non-AF decisions for an input ECG segment. Figure 10 shows the overall flowchart of the 2-step AF vs non-AF detection algorithm.

Results

In this section, first we describe the results from the MIMIC-III subset, which includes feature selection, AF vs PAC classification, and AF vs non-AF classification. Next, we present the test results using the wearable armband ECG data.

From the MIMIC-III subset, we had 10 AF, 10 NSR, and 10 PAC subjects. From these 30 subjects, we had a total of 505 AF, 551 NSR, and 468 PAC segments of 2 minutes each. After the 12 templates were discarded from the dataset, the final number of 2-minute ECG segments was 1512.

Statistical analysis

To evaluate the performance, the following binary classification accuracy measures were calculated:

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{TN} + \text{FP})$$

$$\text{F1 score} = 2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall}) = \text{TP} / (\text{TP} + 0.5 \times (\text{FP} + \text{FN}))$$

As the second step, we take only the “possible AF” segments (ie, AF and PACs) and discriminate between AF and non-AF, where non-AFs include PACs. For this discrimination, the density image-based features along with RMSSD and SampEn were used; linear SVM was used as the classifier. Details about the choice of the classifier are described in the Results section. As a result, after the 2 steps, the

where TP, TN, FP, and FN indicate true-positive, true-negative, false-positive, and false-negative, respectively.

Feature selection results

We used the density Poincaré plot-based approach for the AF vs PAC classification (step 2 of our proposed method), and subsequently extracted 36 features, from which we

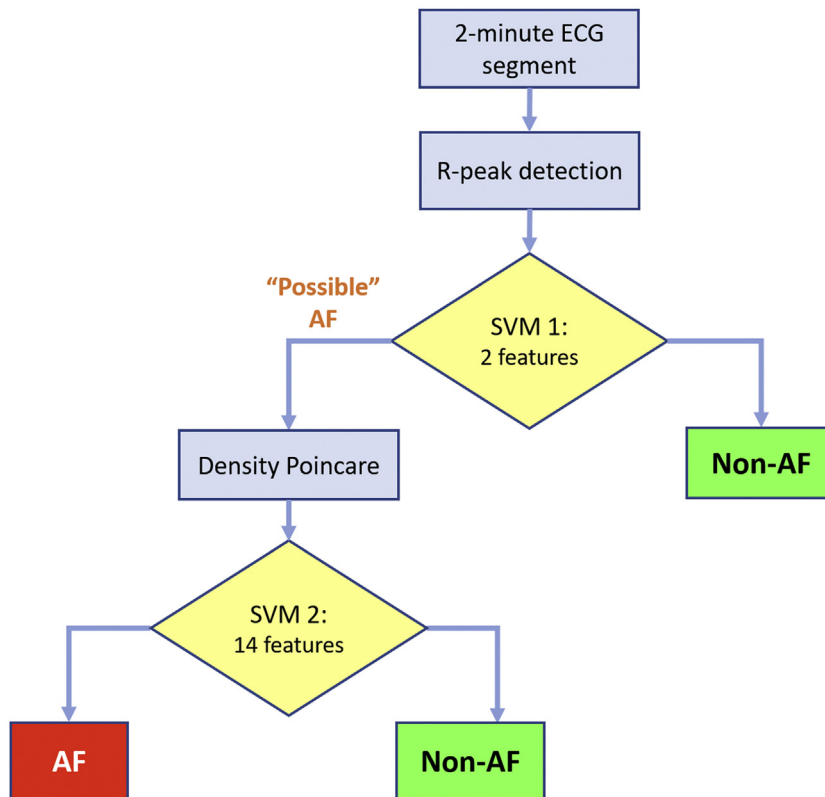


Figure 10 Flowchart of the proposed two-step atrial fibrillation (AF) vs non-AF discrimination algorithm. ECG = electrocardiogram; SVM = support vector machine.

reduced the features using the ILFS algorithm. For this purpose, 10-fold cross-validation was performed using the 10 AF and 10 PAC subjects and the ILFS ranking was calculated for each fold. In k-fold step, data were split into k disjoint segments; k-1 folds were used for training and the remaining fold was used for testing. At each fold, the ILFS feature ranking was obtained using only

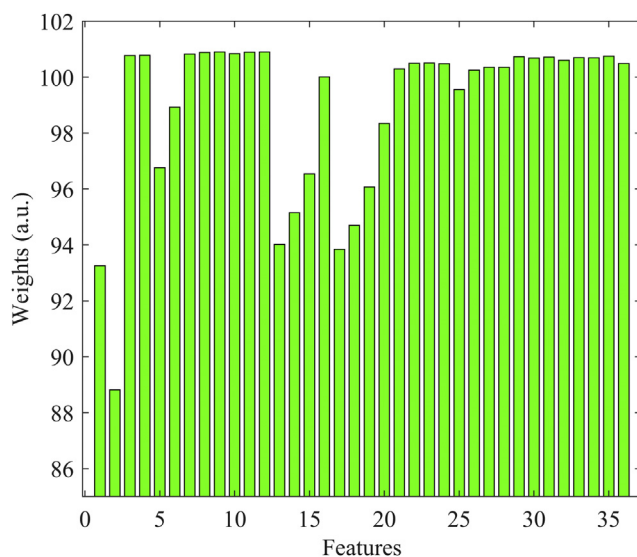


Figure 11 Bar chart of all the image-based feature weights obtained from the infinite latent feature selection algorithm.

the training data; the algorithm ranked every feature from the total pool of 36 features and a weight was assigned to each feature. Figure 11 shows a bar chart of the ILFS weights for all of the density image-based features for a single fold. We analyzed the feature rankings for each fold and then selected the top 12 features, which were common for most of the 10 folds, thus resulting in a reduction of 67% of the number of features. These features include 8 from the template correlation and 4 from the DWT. The selected density Poincaré plot features were then used for the leave-one-subject-out cross-validation.

Results on MIMIC data

First, we describe the AF vs PAC classification results using the image-based features. To analyze which classifier model works best with the ILFS selected features, we analyzed the 10 AF and 10 PAC subjects from the MIMIC-III subset. From this dataset, after discarding the templates, we have 499 AF and 462 PAC segments, where each segment consists of 2 minutes of ECG.

To study this AF vs PAC classification performance and find out the model parameters, leave-one-subject-out cross-validation was performed using the 20 subjects. Table 1 shows the performance of several machine learning classifiers along with different kernels/parameters where the 12 ILFS-selected density Poincaré image-based features were used. The classifiers compared include SVM with both linear and radial basis function kernels, k-nearest neighbors (KNN)

Table 1 Atrial fibrillation vs premature atrial contraction classification performance for different classifiers

Classifier name	Parameter/kernel	Overall accuracy (%)	AF accuracy (%)	PAC accuracy (%)
SVM	Linear	97.56	97.68	97.45
	RBF	97.27	97.24	97.3
DA	Quadratic	93.76	89.81	97.7
	Linear	88.86	94.89	82.83
	Diaglinear	88.75	99.79	77.71
	Diagquadratic	92.24	91.75	92.72
	Mahalanobis	85.76	71.73	99.8
KNN	Euclidean, k = 3	90.2	99	81.4
	Manhattan, k = 3	92.78	98.98	86.58

AF = atrial fibrillation; DA = discriminant analysis; KNN = k-nearest neighbors; PAC = premature atrial contraction; RBF = radial basis function; SVM = support vector machine.

classifier with 2 distance parameters, and discriminant analysis (DA) classifier with different variants. From the table it can be seen that both SVMs give higher overall accuracy than the others; however, the linear SVM resulted in the highest AF, PAC, and overall accuracies of 97.68%, 97.45%, and 97.56%, respectively. As a result, the linear SVM was selected to classify AF and PAC in the second step of our proposed algorithm.

Next, we calculated the performance of the SVM classifiers when all 36 of the density Poincaré plot–based features were used. With all 36 features, the linear SVM achieved 98.37% overall accuracy, 98.72% sensitivity, and 98.01% specificity. The radial basis function SVM had 98.36% accuracy, 98.11% sensitivity, and 98.61% specificity. This shows that using all 36 image-based features increases overall accuracy by only ~1% over using only the features selected by ILFS, which had a 67% decrease of feature space (from 36 to 12 features). As a result, this set of 12 features was the best selection, instead of all the 36 density image–based features.

Finally, we calculated the overall AF vs non-AF discrimination results using the 30 subjects from the MIMIC-III subset. For this purpose, we performed leave-one-subject-out cross-validation. This ensures that we do not overfit for individual subjects and allows us to examine how the algorithm performs on a new unseen subject. The flowchart provided in Figure 10 describes the process. We used 2 different SVMs for the 2 steps: the first step used RMSSD and SampEn with linear SVM to identify “possible AF” vs NSR; the “possible AF” class included AF and PACs. Then the second SVM used the previously mentioned 12 image-based features as well as RMSSD and SampEn, and used as its input only the “possible AF” detected segments from the first step. This second SVM is also binary and outputs the final AF or non-AF labels.

Table 2 shows the confusion matrix obtained from the 30 MIMIC subjects. Since we have 3 classes (AF, PAC, and NSR) in the MIMIC training data, we first show the 3-by-3

confusion matrix. From the table it can be seen that our proposed model achieves high performance for the 3-class detection. Cross-diagonally we have only 4 samples, which shows that only 4 samples were misdetected from all of the 1512 samples. The sensitivity values for AF, PAC, and NSR were 99.40%, 99.78%, and 100%, respectively, whereas the overall accuracy was 99.74%.

Subsequently, we merged the PAC and NSR classes into the combined “non-AF” label and calculated the 2-class confusion matrix, shown in Table 3. This was calculated because eventually we will test the binary classification performance of AF vs non-AF using the novel wearable armband data. From the table, it can be seen that for the binary classification, the sensitivity, specificity, and accuracy were found to be 99.40%, 100%, and 99.80%, respectively; the F1 score was 99.70%.

Results on the Armband Data

To study the feasibility of AF detection using the novel armband device, we applied the trained model on the armband data. In other words, we trained our algorithm using the MIMIC-III ICU data and tested the model directly on the armband ECG data without changing or tuning the model further. This resulted in a blind test scenario.

The armband data had a total of 12 subjects; among them, 2 had AF and the rest had either NSR or PACs. We have 7560 2-minute ECG segments from the non-AF subjects (including NSR and PAC) and 936 2-minute segments from the AF subjects, resulting in a total of 8496 2-minute segments for testing. Table 4 shows the results of AF detection using the wearable armband data. Subjects 1 and 2 are the AF subjects, whereas the remaining 10 are non-AF (NSR and PAC). From the table, it can be seen that the algorithm achieved high detection accuracy for all 12 subjects; the detection accuracy was 100% for 10 subjects. It is to be noted that the number of segments per subjects differs because the subjects had different amounts of clean ECG signals, as determined by the algorithm of Lazaro and colleagues.¹²

Table 5 shows the confusion matrix from armband data results. From the table, it can be seen that only 2 segments were incorrectly labeled out of the total 8496 segments, resulting

Table 2 Confusion matrix of atrial fibrillation, premature atrial contraction, and normal sinus rhythm classification using the MIMIC data

	Predicted labels		
	AF	PAC	NSR
True labels			
AF	496	3	0
PAC	0	461	1
NSR	0	0	551

AF = atrial fibrillation; NSR = normal sinus rhythm; PAC = premature atrial contraction.

Table 3 Atrial fibrillation vs non-atrial fibrillation classification on MIMIC data

	Predicted labels	
	AF	Non-AF
True labels		
AF	496	3
Non-AF	0	1013

AF = atrial fibrillation.

in sensitivity, specificity, accuracy, and an F1 score of 99.89%, 99.99%, 99.98%, and 0.9989, respectively.

Finally, in [Table 6](#), the performance of our AF detection algorithm is compared with several existing methods for the 12 subjects of the armband database; all necessary training was performed using the MIMIC data for fair comparison. From the table, it can be seen that our method achieved the best accuracy and F1 score of all the methods, which demonstrates the efficacy of our method.

Discussion

We presented the feasibility of AF detection using data from the novel wearable armband device. After several R-R intervals and density Poincaré plot-based features were extracted, the proposed model was trained using MIMIC-III ICU data. Finally, the model was tested using the armband ECG data. The high accuracy suggests that the novel wearable armband device can be used for continuous AF monitoring, once our algorithm is embedded in it.

In the first step of the algorithm, RMSSD and SampEn were used with an SVM classifier. This step was relatively straightforward, as its goal was to identify NSR from possible AF. Since the possible AF category includes AF and PACs and both had higher R-R variability than did NSR, this classification/discrimination was less challenging. SampEn was shown to classify the NSR from possible AF with a fixed threshold in Bashar and colleagues.¹⁸ However, to make this approach generalizable and independent of predefined

Table 4 Classification performance using the wearable armband electrocardiogram for individual subjects

Subject ID	Number of segments	TP	TN	Accuracy (%)
1	14	14	0	100
2	922	921	0	99.89
3	919	0	919	100
4	272	0	272	100
5	970	0	970	100
6	543	0	543	100
7	9	0	9	100
8	548	0	548	100
9	725	0	725	100
10	1756	0	1756	100
11	421	0	420	99.76
12	1397	0	1397	100
Total	8496	935	7559	Mean = 99.97

TN = true-negative; TP = true-positive.

Table 5 Confusion matrix using the armband data (blind testing)

	Predicted labels	
	AF	Non-AF
True labels		
AF	935	1
Non-AF	1	7559

AF = atrial fibrillation.

thresholds, an SVM classifier was used with RMSSD and SampEn features.

Ectopic beat is one of the most common reasons for false-positive detection of AF. Although in literature the absence of a P wave was commonly used to reduce false-positives,¹⁸ this was not used in our case. The armband ECG signal often has severe muscle noise, which can make P-wave detection difficult. Since the armband was shown to capture accurate R-R intervals when compared with Holter ECG,¹² only the R-R interval-derived features were used to detect the PAC beats from the possible AF detected segments.

The density Poincaré plot-based features along with SVM classifier showed high accuracy in discriminating PAC segments from AF segments, reducing false-positives. Leave-one-subject-out validation ensured that the model was not over-fitted to subjects. As a result, when we applied the model trained with the MIMIC-III dataset directly on the wearable armband data, high AF vs non-AF classification accuracy was obtained. Despite having PACs in the armband data, our method obtained an F1 score of 0.9989 and 99.98% accuracy, which suggests potential clinical utility.

There have been numerous studies on AF detection using the ECG signal. Most of the works are based on analysis of either atrial activities or ventricular responses.³³ Since the armband records the ECG signal from the upper left arm, the ECG is corrupted with extreme muscle noise and the P wave is not always detectable. As a result, we focus on detecting AF solely based on ventricular response.

Ventricular response-based AF detection methods analyze different properties of RR interval irregularities. In Duverney and colleagues,³⁴ discrete wavelet transform of R-R intervals was first used to identify periods of high heart rate variability, which was followed by fractal analysis to classify these high-variability periods into AF or sinus

Table 6 Performance comparison with existing methods using the wearable armband data

Method	Sensitivity (%)	Specificity (%)	Accuracy (%)	F1 score
AR model ³¹	87.18	87.71	87.65	0.6087
COSEn ²⁰	100	92.9	93.68	0.7771
RR features + KNN ³²	85.47	99.93	98.34	0.919
RR features + RF ³²	92.52	99.97	99.15	0.9601
Proposed method	99.89	99.99	99.98	0.9989

AR = autoregressive; COSEn = coefficient of sample entropy; KNN = k-nearest neighbors; RF = random forest; RR = R-R interval.

rhythm. Autoregressive model order–based percentage of predicted power was used to detect AF episodes in Cerutti and colleagues.³¹ In Tateno and Glass,³⁵ standard-density histograms of R-R intervals and the difference between successive R-R intervals were first calculated. Next, the Kolmogorov-Smirnov test was used to compare the similarity between the standard- and test-density histograms to determine the AF status. In Dash and colleagues,¹⁷ RMSSD, turning point ratio, and Shannon entropy, calculated from R-R interval series after ectopic beat filtering, were used to detect AF. In Zhou and colleagues,³⁶ symbolic dynamics were obtained from the preprocessed R-R interval series, followed by calculating Shannon entropy from the symbolic sequence to determine the presence of AF. In Lake and Moorman,²⁰ the coefficient of sample entropy (COSEn) calculated from the R-R interval series was proposed to detect AF from short ECG recordings. In Huang and colleagues,³⁷ the density histogram of delta R-R intervals (defined as the difference between 2 successive RR intervals) was used to detect an AF event and later to determine the boundary of AF. The Lorentz plot obtained from the R-R time series was used to detect AF as well as atrial tachycardia in Sarkar and colleagues.³⁸ RMSSD, coefficient of variance, median absolute deviation, and COSEn calculated from R-R intervals were used as features and KNN along with random forests were used as classifiers in Kennedy and colleagues³² to detect AF.

When we compared the performance of several existing AF detection methods on the wearable armband data, our proposed method outperformed the others. Both the autoregressive model–based³¹ and COSEn-based²⁰ methods resulted in low F1 scores. The R-R features described in Kennedy and colleagues³² resulted in high accuracies of 98.34% and 99.15% when used with KNN and RF, respectively. However, the F1 scores were low, as these models missed many true AF segments. Our presented approach achieved higher accuracy and F1 score than all of the compared methods, which shows the effectiveness of our algorithm for the wearable armband data. Moreover, in terms of readable ECG, previously in another study from our group,¹² it is shown that 95% of data recorded from this novel armband device during night-time (bedtime) were readable when compared to the standard chest ECG recordings, whereas for the daytime (non-bedtime), it was 53.85% owing to different motion artifacts. The readable ECG data recorded from this novel armband device as well as the high AF detection performance in this pilot study establishes the feasibility of continuous AF monitoring using this novel armband device.

Participants wore the devices for 14 days and the usability of the armband was briefly assessed in several domains, including general impressions of the device, how easy the device was to use, and how comfortable the device was. The mean responses to these questions were all found to be greater than 3, where 1 is very negative and 5 is very positive.

Finally, our results should be considered in light of the study limitations. Although the algorithm achieved high ac-

curacy during testing the armband data, we had only 2 AF subjects. In future, more expansive clinical trials will need to be conducted to include more AF subjects and analyze continuous AF monitoring on a large scale. Moreover, some subjects had many fewer ECG segments than other subjects did, owing to noise artifacts as well as not wearing the device correctly, which will be improved in future studies. However, since the entire algorithm is trained on the ICU ECG data and it still achieves high performance on the wearable armband ECG data in blind testing, we hope to obtain similar performance for the new unseen data.

Conclusion

In this paper, we examined the feasibility of AF monitoring using a novel wearable armband ECG device. We extracted 2 R-R interval–based features as well as a novel density Poincaré plot–based feature to develop the AF algorithm. The 2-step algorithm uses the extracted features along with the SVM classifier. With the aid of density image–based features, we were able to detect the ectopic beats (eg, PACs/PVCs), which reduced false-positive AF detection. To train and validate the algorithm, we used long-term ECG recordings obtained from a subset of the MIMIC-III ICU database and achieved high performance. When we tested our proposed method using the wearable armband ECG data, high sensitivity, specificity, and accuracy were obtained. Moreover, we compared our method to several existing methods and it outperformed them. Although for this pilot study we used only a limited number of subjects, our study demonstrates the potential for continuous AF monitoring using the novel wearable armband device. Future study will incorporate a larger number of subjects to examine how the device and AF detection algorithm perform when data are collected in subjects' home environments.

Funding Sources

This work was supported by NSF SBIR Phase I (#1746589) and NIH R43 HL135961. This work was partly supported by MCIU, AEI and FEDER under project RTI2018-097723-B-I00, by Aragon Government and FEDER through BSICoS group (T39_20R) and LMP44-18, and by CIBER in Bioengineering, Biomaterials & Nanomedicine through Instituto de Salud Carlos III.

Disclosures

D.D.M. has received honoraria, speaking/consulting fee, or grants from Flexcon, Rose Consulting, Bristol-Myers Squibb, Pfizer, Boston Biomedical Associates, Samsung, Phillips, Mobile Sense, CareEvolution, Flexcon Boehringer Ingelheim, Biotronik, Otsuka Pharmaceuticals, and Sanofi. D.D.M. also declares financial support for serving on the Steering Committee for the GUARD-AF study (NCT04126486) and Advisory Committee for the Fitbit Heart Study (NCT04176926). The other authors have no disclosures.

Authorship

All authors attest they meet the current ICMJE criteria for authorship.

Patient Consent

All patients provided written informed consent.

Ethics Statement

The authors designed the study and gathered and analyzed the data according to the Helsinki Declaration guidelines on human research. The research protocol used in this study was reviewed and approved by the institutional review board.

References

- Murphy A, Banerjee A, Breithardt G, et al. The World Heart Federation roadmap for nonvalvular atrial fibrillation. *Global Heart* 2017;12:273–284.
- Danelich IM, Reed BN, Hollis IB, Cook AM, Rodgers JE. Clinical update on the management of atrial fibrillation. *Pharmacotherapy* 2013;33:422–446.
- Pereira T, Tran N, Gadhomi K, et al. Photoplethysmography based atrial fibrillation detection: a review. *npj Digital Medicine* 2020;3:1–12.
- European Heart Rhythm Association; European Association for Cardio-Thoracic Surgery, Camm AJ, et al. Guidelines for the management of atrial fibrillation The Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J* 2010;31:2369–2429.
- Naccarelli GV, Varker H, Lin J, Schulman KL. Increasing prevalence of atrial fibrillation and flutter in the United States. *Am J Cardiol* 2009;104:1534–1539.
- Camm AJ, Corbucci G, Padeletti L. Usefulness of continuous electrocardiographic monitoring for atrial fibrillation. *Am J Cardiol* 2012;110:270–276.
- Bashar SK, Han D, Hajeb-Mohammadalipour S, et al. Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches. *Sci Rep* 2019;9:1–10.
- Lee J, Reyes BA, McManus DD, Maitas O, Chon KH. Atrial fibrillation detection using an iPhone 4S. *IEEE Trans Biomed Eng* 2013;60:203–206.
- Couderc J-P, Kyal S, Mestha LK, et al. Detection of atrial fibrillation using contactless facial video monitoring. *Heart Rhythm* 2015;12:195–201.
- Jabaudon D, Sztajzel J, Sievert K, Landis T, Sztajzel R. Usefulness of ambulatory 7-day ECG monitoring for the detection of atrial fibrillation and flutter after acute stroke and transient ischemic attack. *Stroke* 2004;35:1647–1651.
- Jung H, Moon JH, Baek DH, et al. CNT/PDMS composite flexible dry electrodes for long-term ECG monitoring. *IEEE Trans Biomed Eng* 2012;59:1472–1479.
- Lázaro J, Reljin N, Hossain M-B, Noh Y, Laguna P, Chon KH. Wearable armband device for daily life electrocardiogram monitoring. *IEEE Trans Biomed Eng* 2020;67:3464–3473.
- Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000;101:E215–E220.
- Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Scientific Data* 2016;3:160035.
- Bashar SK, Ding E, Walkey AJ, McManus DD, Chon KH. Noise detection in electrocardiogram signals for intensive care unit patients. *IEEE Access* 2019;7:88357–88368.
- Bashar SK, Noh Y, Walkey AJ, McManus DD, Chon KH. VERB: VFCDM-based electrocardiogram reconstruction and beat detection algorithm. *IEEE Access* 2019;7:13856–13866.
- Dash S, Chon KH, Lu S, Raeder EA. Automatic real time detection of atrial fibrillation. *Ann Biomed Eng* 2009;37:1701–1709.
- Bashar SK, Hossain MB, Ding E, Walkey AJ, McManus DD, Chon KH. Atrial fibrillation detection during sepsis: study on MIMIC III ICU data. *IEEE J Biomed Health Inform* 2020;24:3124–3135.
- Alcaraz R, Rieta JJ. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. *Biomedical Signal Processing and Control* 2010;5:1–14.
- Lake DE, Moorman JR. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am J Physiol Heart Circ Physiol* 2011;300:H319–H325.
- Langley P, Dewhurst M, Di Marco LY, et al. Accuracy of algorithms for detection of atrial fibrillation from short duration beat interval recordings. *Med Eng Phys* 2012;34:1441–1447.
- Chong JW, Esa N, McManus DD, Chon KH. Arrhythmia discrimination using a smart phone. *IEEE J Biomed Health Inform* 2015;19:815–824.
- Han D, Bashar SK, Lazaro J, et al. Smartwatch PPG peak detection method for sinus rhythm and cardiac arrhythmia. *Annu Int Conf IEEE Eng Med Biol Soc* 2019;2019:4310–4313.
- Bashar SK, Han D, Zieneddin F, et al. Novel density Poincaré plot based machine learning method to detect atrial fibrillation from premature atrial/ventricular contractions. *IEEE Trans Biomed Eng* 2021;68:448–460.
- Bashar SK, Han D, Zieneddin F, et al. Preliminary results on density poincare plot based atrial fibrillation detection from premature atrial/ventricular contractions*. *Annu Int Conf IEEE Eng Med Biol Soc* 2020;2020:2594–2597.
- Tong M, Pan Y, Li Z, Lin W. Valid data based normalized cross-correlation (VDNCC) for topography identification. *Neurocomputing* 2018;308:184–193.
- Nayak DR, Dash R, Majhi B. Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. *Neurocomputing* 2016;177:188–197.
- Singh A, Dutta MK, ParthaSarathi M, Uher V, Burget R. Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image. *Comput Methods Programs Biomed* 2016;124:108–120.
- Roffo G, Melzi S, Castellani U, Vinciarelli A. Infinite latent feature selection: a probabilistic latent graph-based ranking approach. In: *Proceedings of the IEEE International Conference on Computer Vision* 2017;1398–1406.
- Yu S-N, Chou K-T. Selection of significant independent components for ECG beat classification. *Expert Systems with Applications* 2009;36:2088–2096.
- Cerutti S, Mainardi LT, Porta A, Bianchi AM. Analysis of the dynamics of RR interval series for the detection of atrial fibrillation episodes. in *Computers in Cardiology* 1997;77–80.
- Kennedy A, Finlay DD, Guldenring D, Bond RR, Moran K, McLaughlin J. Automated detection of atrial fibrillation using R-R intervals and multivariate-based classification. *J Electrocardiol* 2016;49:871–876.
- Mukherjee A, Dutta Choudhury A, Datta S, et al. Detection of atrial fibrillation and other abnormal rhythms from ECG using a multi-layer classifier architecture. *Physiol Meas* 2019;40:054006.
- Duverney D, Gaspoz JM, Pichot V, et al. High accuracy of automatic detection of atrial fibrillation using wavelet transform of heart rate intervals. *Pacing Clin Electrophysiol* 2002;25:457–462.
- Tateno K, Glass L. A method for detection of atrial fibrillation using RR intervals. *Comput Cardiol* 2000;27:391–394.
- Zhou X, Ding H, Ung B, Pickwell-MacPherson E, Zhang Y. Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy. *BioMedical Engineering OnLine* 2014;13:18.
- Huang C, Ye S, Chen H, Li D, He F, Tu Y. A novel method for detection of the transition between atrial fibrillation and sinus rhythm. *IEEE Trans Biomed Eng* 2011;58:1113–1119.
- Sarkar S, Ritscher D, Mehra R. A detector for a chronic implantable atrial tachyarrhythmia monitor. *IEEE Trans Biomed Eng* 2008;55:1219–1224.