The Accuracy of Beat-interval based Algorithms for Detecting Atrial Fibrillation

Alan Kennedy, Dewar D Finlay, Daniel Guldenring, Raymond Bond, James McLaughlin

NIBEC, University of Ulster, Jordanstown, United Kingdom

Abstract

Automated detection of Atrial Fibrillation (AF) from the surface electrocardiogram (ECG) remains a challenge. Some have suggested that a major source of false positives from R-R interval based AF algorithms are ectopic beats and/or other supraventricular arrhythmias. However, this has not been thoroughly investigated. This study aims to evaluate the accuracy of four commonly implemented R-R Interval based AF algorithms (1) The coefficient of variance, (2) Root Mean Square of the Successive Differences, (3) Turning Point Ratio (TPR) and (4) Shannon Entropy. All four algorithms were tested on R-R interval data from patients in normal sinus rhythm, during atrial fibrillation, with ectopic beats and with supraventricular tachycardia (SVT). Receiver operating characteristic analysis was used to determine the performance of each algorithm over different analysis segment lengths ranging from 30 to 120 beats. When comparing algorithm results, a clear reduction in algorithm performance was found in patients with ectopic beats and SVT. This must be taken into consideration when designing and evaluating algorithms for automated AF detection.

1. Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia affecting approximately 2% of the general population [1]. Within the European Union it is projected that the number of people at 55 years and over with AF will more than double between 2010 and 2060 [2]. AF is characterized by chaotic electrical activity in the atrial myocardium. When left untreated, it is a major risk factor for stroke or transient ischemic attack. Patients with AF are 5-7 times more likely to suffer a stroke [3] when compared to the normal population. However, appropriate intervention can substantially decrease stroke risk [4].

Some individuals suffering from AF are unaware of their condition due to the asymptomatic and sporadic (paroxysmal AF) nature of the AF. The characteristics of short paroxysmal AF events make them difficult to detect using the clinical standard for identification of cardiac arrhythmias, the 12-lead electrocardiogram (ECG). This is due to the relatively short duration of the recording (10 seconds) [5]. Usually patients with suspected paroxysmal AF are required to undergo 24-72 hour ambulatory ECG monitoring. As the relationship between AF and much more adverse thromboembolic events emerges [3] the clinical need for an accurate ambulatory monitoring solution for AF grows. Currently, ambulatory ECG monitoring systems perform poorly in detecting AF due to high false positive (FP) and high false negative (FN) rates [6]. This can force cardiologists or other training staff to manually review long-term ECG recordings, which is time-consuming and at times unreliable [7]. AF commonly presents on the ECG as the absence of a P-wave and the presence of a chaotic fibrillatory wave on the ECG baseline. The chaotic electrical activity in the atria triggers irregular ventricular repolarisation as impulses generated from the atria exceed the conduction capability of the atrioventricular node. Ectopic beats (EB) have a ventricular response that is similar to what is observed during AF. Such EB occur when fibers or groups of fibers, not associated with the sinoatrial node, stimulate contraction of the heart. However, ectopic beats are mostly harmless and usually do not require any clinical intervention. Many ambulatory systems rely on R-R interval based algorithms alone for AF detection. Previous studies have suggested that a major source of false positives, when using RR-interval based AF algorithms, is from patients suffering EB and/or supraventricular tachycardia (SVT) [8]. In fact, a recent study by *Petrénas et al.* [9] outlined the two major issues to be resolved in long-term AF monitoring are: (1) the inaccurate detection of short AF events (2) the number of false positives due to ectopic beats and other irregular cardiac rhythms.

In this study, we investigate four previously described R-R interval based AF detection algorithms [10] in order to assess their performance in detecting AF in the presence of EB and SVT.

2. Method

2.1. ECG data

Dataset1 consisted of R-R interval data from the MIT-BIH atrial fibrillation database (AFDB) [11]. This database contained 25 long-term (10-hour) recordings taken from patients with mostly paroxysmal atrial fibrillation captured at 250 samples per second. *Dataset2* consisted of the same MIT-BIH afdb however added to this were data from a further 22 patients from the MIT-BIH arrhythmia database [12] that were experiencing EB recorded at 360 samples per seconds along with the MIT-BIH SVT database [13] consisting of 30-minute recordings from 78 patients captured at 128 samples per second.

2.1.1. R-R interval data

To extract R-R interval based data from the ECG recordings an open source QRS detection algorithm, GQRS [14, 15], was used. This algorithm was chosen based on a number of studies describing its improved performance over other available QRS detection algorithms [16]. Table 1 shows the partitioning of each dataset in terms of AF and non-AF annotated beats from manual annotation of the ECG.

Table 1. The number of ECG beats in each dataset

	AF beats	Non-AF beats
Dataset1	486658	635795
Dataset2	486658	876002

2.2. AF detection algorithms

The irregular ventricular response during AF creates an irregular R-R interval on the ECG. Automated detection of AF can therefore be performed through statistical analysis of consecutive R-R intervals. The R-R interval based AF algorithms investigated in this study were implemented as detailed below.

(1) The coefficient of variance [17]

$$CV = \frac{RR_{\sigma}}{RR_{\mu}} \tag{1}$$

Where σ is the standard deviation of a series of R-R intervals and μ is the mean R-R interval. When the CV exceeds a predefined threshold the analysis segment is classified as AF.

(2) Root Mean Square Successive Differences [10]

$$RMSSD = \sqrt{\frac{1}{N-1}(\sum_{I=1}^{N-1}(RR_{i+1} - RR_i)^2} \tag{2}$$

Where i is the R-R interval and N is the length of the analysis segment. If the RMSSD of the analysis segment is found to be above a predefined threshold the analysis segment is classified as AF.

(3) Shannon Entropy (SE) [18] is a calculation used to measure the level of uncertainty in a data series by measuring and comparing the probability of patterns. SE was calculated from a histogram of R-R intervals in a analysis segment of length *l* using 16 equally spaced bins as follows.

$$SE = \frac{-\sum_{i=1}^{16} prob(i) \cdot \log(prob(i))}{\log(\frac{1}{16})} \tag{3a}$$

The probability histograms were calculated as:

$$prob(i) = \frac{N_{bin(i)}}{l} \tag{3b}$$

Where, $prob(i)$ is the probability distribution for the RR interval analysis segment and $N_{bin(i)}$ is the number of beats in the *i*th bin of the histogram.

(4) Turning Point Ratio [10] for AF detection is based on the nonparametric *runs test*. It is used to measure the degree of randomness in a particular time-series of data. Each beat-interval in the RR analysis segment is compared to its two nearest neighbors and is designated a Turning Point (TP) if it is greater than or less than both. For example, in a time series of R-R intervals when:

$$(RR(n) - RR(n-1)) * (RR(n) - RR(n+1)) > 0 \tag{4}$$

A turning point is assigned. The number of turning points in a analysis segment are then normalised by the analysis segment length. R-R intervals were classified as AF if the number of turning points exceeded the predefined threshold.

2.2.1 Analysis segment length

To determine the optimal analysis segment length for AF detection, the four algorithms investigated in this study were trialed on analysis segment lengths ranging from 30-beats in increments of 30 to 120-beats. This is of interest as shorter analysis segment lengths are highly desirable for new body-worn detection devices due to reduced memory consumption [9].

2.3.Receiver operator characteristic analysis

The receiver operating characteristic (ROC) curve shows the relationship between specificity and sensitivity of a classifier over a range of detection thresholds. An ROC curve was produced for each algorithm over each analysis segment length for each dataset giving multiple ROC curves. ROC area was then calculated from each ROC curve allowing for comparison between algorithms and between datasets.

3. Results

3.1. Algorithm performance

The maximum sensitivity and specificity that could be achieved from the CV on *dataset1* was 90% and 86% respectfully. This was reduced to Sensitivity = 85% and Specificity = 80% on *dataset2* (Fig. 1).

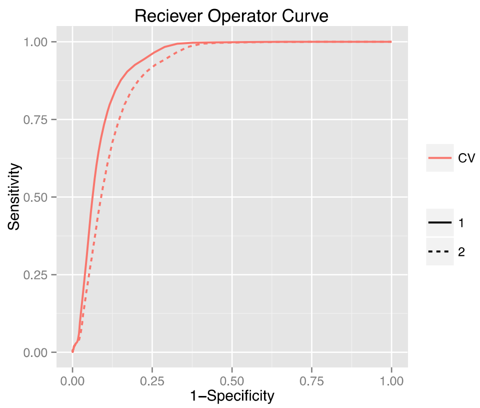

Figure 1. An example receiver operator curve for the coefficient of variance on *dataset1* and *dataset2* at analysis segment length of 120 beats. A clear reduction in algorithm performance can be seen on *dataset2*.

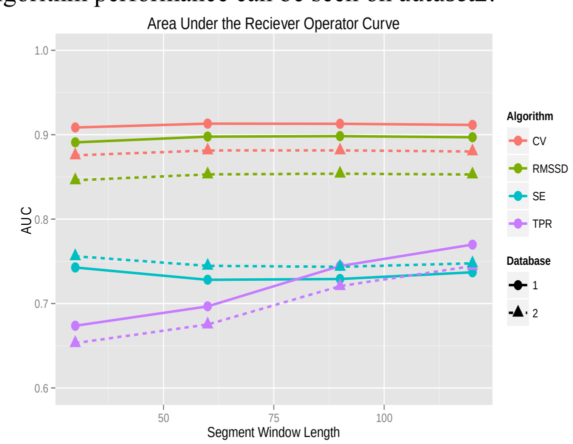

Figure 2. The area under the curve (AUC) for the receiver operator curves of the four beat interval algorithms at analysis segment length of 30, 60, 90 and 120. Again, reduced algorithm accuracy can be seen in *dataset2* from CV, RMSSD and TPR.

From the four analysis segments an average AUC ($\overline{AUC}$) was calculated and used to compare performance on the two datasets. On *dataset1*, the coefficient of variance performed best ($\overline{AUC}$ = 91) followed by RMSSD ($\overline{AUC}$ = 90), Shannon Entropy ($\overline{AUC}$ = 73) then turning point ratio ($\overline{AUC}$ = 72). A reduction in algorithm performance can be seen for three of the four algorithms tested on *dataset2* (Fig. 1) with CV ($\overline{AUC}$ = 88) followed by RMSSD ($\overline{AUC}$ = 85) then turning point ratio ($\overline{AUC}$ = 70). Demonstrating that the data contained in *dataset2* resulted in an increased number of false positives from these algorithms.

3.2. Detection thresholds

Through creation of the ROC curves by varying the detection threshold optimal detection thresholds were established (Fig. 1).

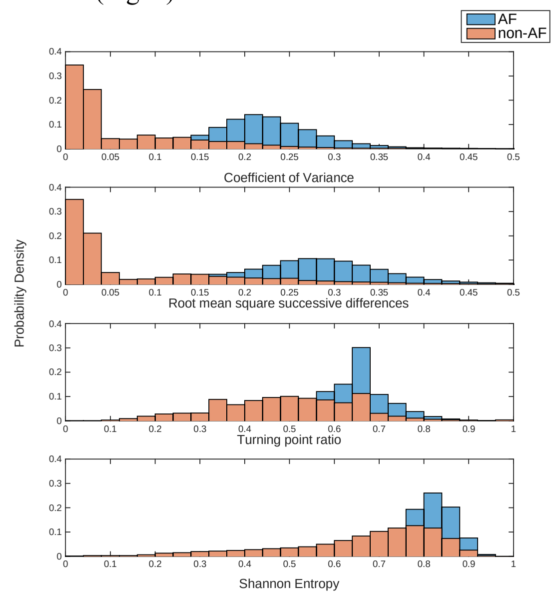

Figure 3. Histograms of the distribution of AF and non-AF beats for all four algorithms over the analysis segment length of 30 beats.

Figure 3 shows probability distributions of AF and non-AF beats from all four algorithms included in this study, it can be seen that CV and RMSSD can better differentiate between AF and non-AF beats compared to TPR and SE.

Table 2. Average optimal detection thresholds across all window lengths.

Algorithm	*Dataset1*	*Dataset2*
CV	0.16	0.16
RMSSD	0.19	0.20
TPR	0.02	0.02
SE	0.76	0.76

As can be seen in table 2, detection thresholds are consistent between databases. which may be something to consider when tuning algorithms on the MIT-BIH AF database for clinical application.

4. Discussion

The majority of studies describe the performance of automated AF detection methods on the MIT-BIH database alone. However, as described in this study a significant challenge for R-R interval based AF algorithms is differentiating between true AF episodes, sinus rhythm, episodes of EB and other irregular heart rhythms not present in the MIT-BIH AF database. When assessing each algorithm individually the segment analysis window had a marginal influence on the CV, RMSSD and SE However, TPR appears to perform more effectively at longer analysis segment lengths (Fig. 2). When comparing results from *Dataset1* to *Dataset2* a clear reduction in algorithm performance is evident from three of the four tested algorithms. SE did not have a reduced performance on *dataset2*, however, its low overall accuracy would not make it an appropriate choice for AF detection. Dash *et al.* [10] successfully combined RMSSD, Shannon Entropy and Turning point ratio which proved to be highly effective on the MIT-BIH AFDB. However, it was also noted by the authors than this R-R based combination algorithm had a reduced performance when applied to the MIT-BIH arrhythmia database demonstrating its inability to differentiate between AF and other rhythm disturbances.

5. Conclusions

This study demonstrates that when performing ECG monitoring for AF using R-R interval algorithms alone an increased number of false positives is observed. This increased number of false positives is from patients with ectopic beats and supraventricular tachycardia. This must be considered when testing and training automated algorithms for AF detection.

Acknowledgements

This work has been supported by the Health Informatics Society of Ireland, the Institution of Engineering and Technology and the Department of Employment and Learning Northern Ireland.

References

[1] J. Mant, D. Wade and S. Winner, "Health care needs assessment: stroke," Health Care Needs Assessment: The Epidemiologically Based Needs Assessment Reviews, vol. 2, 2004.

[2] A. Alonso and L. G. Bengtson, "A rising tide: the global epidemic of atrial fibrillation," Circulation, vol. 129, pp. 829-830, Feb 25, 2014.

[3] G. Saposnik, D. Gladstone, R. Raptis, L. Zhou, R. G. Hart and Investigators of the Registry of the Canadian Stroke Network (RCSN) and the Stroke Outcomes Research Canada (SORCan) Working Group, "Atrial fibrillation in ischemic stroke: predicting response to thrombolysis and clinical outcomes," Stroke, vol. 44, pp. 99-104, Jan, 2013.

[4] R. G. Hart, L. A. Pearce and M. I. Aguilar, "Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation," Ann. Intern. Med., vol. 146, pp. 857-867, 2007.

[5] J.S. Taggar, T. Coleman, S. Lewis, C. Heneghan and M. Jones, "Accuracy of methods for diagnosing atrial fibrillation using 12-lead ECG: A systematic review and meta-analysis," Int. J. Cardiol., vol. 184, pp. 175-183, 2015.

[6] M. F. Márquez, L. Colín, M. Guevara, P. Iturralde and A. G. Hermosillo, "Common electrocardiographic artifacts mimicking arrhythmias in ambulatory monitoring," Am. Heart J., vol. 144, pp. 187-197, 2002.

[7] K. Harris, D. Edwards and J. Mant, "How can we best detect atrial fibrillation?" J. R. Coll. Physicians. Edinb., vol. 42 Suppl 18, pp. 5-22, 2012.

[8] G. A. Ng, "Treating patients with ventricular ectopic beats," Heart, vol. 92, pp. 1707-1712, 11, 2006.

[9] A. Petrėnas, L. Sörnmo, A.Lukoševičius, & V. Marozas."Detection of occult paroxysmal atrial fibrillation," Med. Biol. Eng. Comput., vol. 53, pp. 287-297, 2015.

[10] S. Dash, K. H. Chon, S. Lu and E. A. Raeder, "Automatic Real Time Detection of Atrial Fibrillation," Ann. Biomed. Eng., vol. 37, pp. 1701-1709, 09/01, 2009.

[11] G. B. Moody and R. G. Mark, "A new method for detecting atrial fibrillation using RR intervals," Comput. Cardiol., 1983;10: 227-230.

[12] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," Engineering in Medicine and Biology Magazine, IEEE, vol. 20, pp. 45-50, 2001.

[13] S. D. Greenwald, R. S. Patil and R. G. Mark, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," Computers in Cardiology 1990:. 461-464.

[14] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," Circulation, vol. 101, pp. E215-20, Jun 13, 2000.

[15] G. Moody, B. Moody and I. Silva, "Robust detection of heart beats in multimodal data: The PhysioNet/computing in cardiology challenge 2014", Computing in Cardiology 2014;41: 549-552.

[16] M. Llamedo and J. P. Martinez, "QRS detectors performance comparison in public databases," Computing in Cardiology 2014;41: 357-360.

[17] K. Tateno and L. Glass, "Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and ΔRR intervals," Medical and Biological Engineering and Computing, vol. 39, pp. 664-671, 2001.

[18] X. Zhou, H. Ding, B. Ung, E. Pickwell-MacPherson and Y. Zhang, "Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy," Biomed. Eng. Online, vol. 13, pp. 18-925X-13-18, 2014.

Address for correspondence:
Alan Kennedy a.kennedy@ulster.ac.uk
NIBEC Building, University of Ulster, Shore Road
Newtownabbey, Co. Antrim