

Combining 12-Lead ECG Information for a Beat Detection Algorithm

Victor Mondelo
University of Vigo, Spain
vmondelo@uvigo.es

María José Lado
University of Vigo, Spain
mrpepa@uvigo.es

Arturo José Méndez
University of Vigo, Spain
mrrarthur@uvigo.es

Xosé Antón Vila
University of Vigo, Spain
anton@uvigo.es

Leandro Rodríguez-Liñares
University of Vigo, Spain
leandro@uvigo.es

Abstract— This paper proposes a beat detection algorithm specially tailored to be used with 12 channel ECG records. The algorithm first obtains beat positions on each channel, and then combines this information to get an improved estimate. The detection process involves two stages: 1) single-channel detection: implemented by improving one of the most popular methods (Pan-Tompkins) developed to detect beat positions; and 2) multichannel detection: an algorithm that combines the information of the beat positions obtained in each of the 12 channels. In this way, our results clearly improve those obtained with the single-channel detection method, discarding detection errors, false positives, and duplicated beats. Besides, our single-channel method significantly reduces the temporal error when estimating the positions of QRS complexes. In the multichannel detection, the assessment of our algorithm against one-channel based approaches shows a significant improvement in detection outcome ($Se = 99.86\%$, $P+ = 99.98\%$, $RMS\ RR\ Interval\ Error = 25.98\ ms$, $F-Score = 0.9992$), making it a good starting point for automatic diagnosis of heart conditions.

Index Terms — 12-leads, beat position, computerized detection, ECG.

I. INTRODUCTION

ACCORDING to the World Health Organization, heart attack is still the main cause of death on a global basis [1]. Many of these deaths could be prevented through a more precise diagnosis and a more exhaustive monitoring of patients at risk. Early detection of cardiac conditions could help to establish a therapy that avoids potentially dangerous situations [2] [3] [4], and to contribute to reduce health care expenditure [5] [6].

Electrocardiogram (ECG or EKG) is one of the most used techniques to obtain information about the cardiac activity and to detect heart conduction defects [7] [8]. Whereas obtaining the ECG is nowadays a resolved issue from a technical point of view, its interpretation involves the participation of a specialist physician in order to get a precise diagnosis. A complete analysis of an ECG consists of, first, detecting beat positions, and, then, delineating the waves that form each beat. The information about the morphology, length and presence/absence of P and T waves and QRS complexes is essential to detect some affections and provide an accurate diagnosis. Analyzing ECG records is a tedious and time-consuming task, which justifies the necessity of having automatic systems to support the ECG diagnosis. Presently, these systems allow to automatize the most basic tasks: beats detection, acquisition of rhythm information, and delineation of the waves composing the beats.

In the literature, many proposals can be found that yield good results, both in beat detection and waves delineation [9]

[10] [11]. There are algorithms that achieve a rate of success greater than 99.5% in beat detection [12] [13] [14]. Besides, recent technical advances have made possible to achieve low execution times (even real-time) or to eliminate the need for specific hardware [15] [16]. Most of these proposals are based on single-channel detection, in some cases applied independently to two channels of the ECG. This can make these systems sensitive to interference, noise or bad contact between the electrodes and the patient's skin, since signal degradation can cause detection problems. Such problems can increase the number of false positives (artifacts erroneously detected as beats) and temporal errors (mismatches between detected and actual positions of beats), thus worsening the performance given by ECG delineation algorithms.

Although multichannel approaches could alleviate these issues and also help to improve the final performance of detection algorithms, only a few proposals can be found in the literature. The most interesting ones are based on a two-phase procedure: first, beat positions are estimated for each of the channels and, then, some decision is reached by combining this information. Whereas [17] considers the most frequent position across all channels, [18] makes use of a set of 14 thresholds to reach a decision. On the other hand, in [19] [20] and [21] support vector machines are used as a fusion technique to estimate beat positions.

This paper presents a new algorithm to detect beat positions in 12-lead electrocardiograms, which takes into account the specifics of each of the channels of the ECG to give a reliable and robust estimate of the beat positions. The rest of the paper is organized as follows. Next Section presents the experimental setup, whereas Section III describes in depth the proposed algorithms and the validation procedure. Then, Section IV gives the results obtained with our method compared against some widely used alternatives. Finally, in Section V we present the conclusions and ideas for future work.

II. EXPERIMENTAL SETUP

In order to implement and evaluate the algorithm, we used two databases: MIT-BIH Arrhythmia and St. Petersburg.

The MIT-BIH Arrhythmia database [22] [23] was used to test the single-channel implementation of our algorithm, since the fact that it is widely used makes it a useful resource for comparing our results with other researchers'. It contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained by the Boston's Beth Israel Hospital (now the Beth Israel Deaconess Medical Center). The records were digitized at 360 samples per second per channel. Most records contain II

and V1 channels, but channels V2, V4 and V5 were also used in some records. The database was manually annotated by, at least, two cardiologists working independently.

St. Petersburg database [23], also known as INCARTDB (St. Petersburg Institute of Cardiological Technics 12-lead Arrhythmia database), contains 75 half-hour recordings collected from 32 patients (17 men and 15 women) aged between 18 and 80 (mean age 58), suffering from diverse heart conditions. ECGs were obtained using a 12-lead device and digitized at 257 samples per second. Beat positions were estimated automatically, and annotations were later supervised and corrected manually by experts. This database was used with two purposes: (1) to study how to adapt a single-channel beat detection algorithm to the signals coming from a 12-lead ECG, and (2) to train and validate the multichannel detection algorithm proposed in this paper. In order to get non-biased results, for the second purpose the database was divided into two parts: 39 records (I01-I39) were used for training, whereas the other 36 records (I40-I75) form the validation set.

III. SYSTEM IMPLEMENTATION

The algorithm proposed in this paper is divided into two phases: first, it estimates beat positions on each of the channels of the 12-lead ECGs, and second, it combines this information to get a final estimate of beat positions. Due to differences across channels, using a standard single-channel method in the first phase would discard information that could be useful to increase both the precision and the robustness of the second one. Thus, an optimized single-channel detection method was devised to use its results as input to the multichannel detection phase. Both methods are explained in this Section.

A. Single-channel detection

The single-channel detection algorithm is based on Pan-Tompkins method [9], and it was deeply described in [24]. Briefly, it consists in adapting the single-channel Pan-Tompkins detection algorithm to the different leads of a multichannel ECG recording, to overcome the limitations when detecting beats due to high level of noise existing in the recordings, the missing of some part of the ECG signals, or the impossibility to use adequate preprocessing methods that improve the quality of the original signal.

The developed process involves: (1) adaptation of threshold values to each specific channel; and (2) filtering of the sequences of detected QRS complexes, rejecting those whose distance to the previous one is lower than a certain value.

By adapting the thresholds, we were able to improve the sensitivity Se (1) and the positive predictivity $P+$ (2) of the beat sequences, specially on channels where the standard Pan-Tompkins algorithm failed to detect many QRS complexes. Using the training set of the St. Petersburg database, we obtained a set of values for T that optimize sensitivity and positive predictivity [24].

With this threshold adaptation, the detection of QRS complexes is improved by using adapted thresholds. However, some false positives appear due to the decrease in the decision levels. This was easily solved by keeping only those beats whose distance to the previously detected one was greater than a certain value [24].

The effect of including these improvements in the Pan-Tompkins algorithm is two-fold: it produces a set of more homogeneous values for Se and $P+$ across channels, and it significantly reduces the temporal error when estimating the positions of QRS complexes.

B. Multichannel detection

As stated before, once the optimized single-channel

algorithm is applied to all leads, the multichannel detection phase is used on the whole set of beat sequences. The algorithm, represented in Fig. 1, is as follows:

1) Three lists, $auxBeats$, $Vbeat$ and $beats$, are created. The list $auxBeats$ will contain the positions of the detected QRS complexes of all channels in temporal order. $Vbeat$ and $beats$ are initialized as empty lists: $Vbeat$ will be used to store the positions that may belong to the beat under consideration, and $beats$ will store the estimated positions of beats considered as detected.

2) The list of positions $auxBeats$ is run through, and for each position, $auxBeats[j]$, the following procedure is applied:

a) If the distance between positions $auxBeats[j]$ and $auxBeats[j-1]$ is equal or less than a value $limit1$, $auxBeats[j]$ is considered a candidate to belong to the beat under consideration. Then, $auxBeats[j]$ is added to $Vbeat$.

b) If the distance between positions $auxBeats[j]$ and $auxBeats[j-1]$ is greater than $limit1$, $auxBeats[j]$ will correspond to a new beat. Then, the values in $Vbeat$ will be a set of positions that may belong to the previous beat in different channels. If $Vbeat$ contains at least four values, it means that the beat was detected in, at least, four channels. Then, its position is estimated as the median of all the values in $Vbeat$, and added to $beats$. If $Vbeat$ contains three values or less, it is not considered as a valid beat. Either way, $Vbeat$ is reinitialized to an empty list.

3) From the list of tentative beats $beats$, a new list $finalBeats$ is generated, by removing false positives. To this end, every beat in $beats$ whose distance to the previous one is less than a value $limit2$ is not included in $finalBeats$. Finally, the algorithm returns $finalBeats$.

Thresholds $limit1$ and $limit2$ are obtained empirically, maximizing Se and $P+$ over the training set of the St. Petersburg database. Best results were obtained by setting $limit1$ and $limit2$ to 103 ms and 322 ms, respectively. Threshold $limit1$ must be high enough to take into account different positions of the same beat in different channels, but low enough to reject as many false positives as possible. However, most false positives will be removed later, because they have to be detected in at least four channels, and also because $limit2$ sets a minimum for the inter-beat distance (subjects are considered to be at rest).

This approach exploits the strengths of each channel of the multilead ECG, while avoiding their weaknesses. By combining all the beats in one structure, the problems caused by undetected beats in some channels can be successfully avoided.

C. Validation

As said before, the single-channel algorithm was validated using the MIT-BIH Arrhythmia and the St. Petersburg databases. The multichannel algorithm was validated using only the St. Petersburg database.

Performance of our algorithms was evaluated using the following indexes: defining the RR interval as the distance between two consecutive beats.

TP is the number of true positives (correctly detected beats), FN is the number of false negatives (undetected beats) and FP is the number of false positives (incorrectly detected beats).

These indexes were obtained using the tool bxb , that provides sensitivity, positive predictivity and RMS RR Interval Error, calculated by comparing two annotation files for ECGs. This tool is included in the PhysioToolkit library [23], which can be freely downloaded from the PhysioNet Website. In this case, one of the annotation files was obtained from the database, and the other was generated with our detection method.

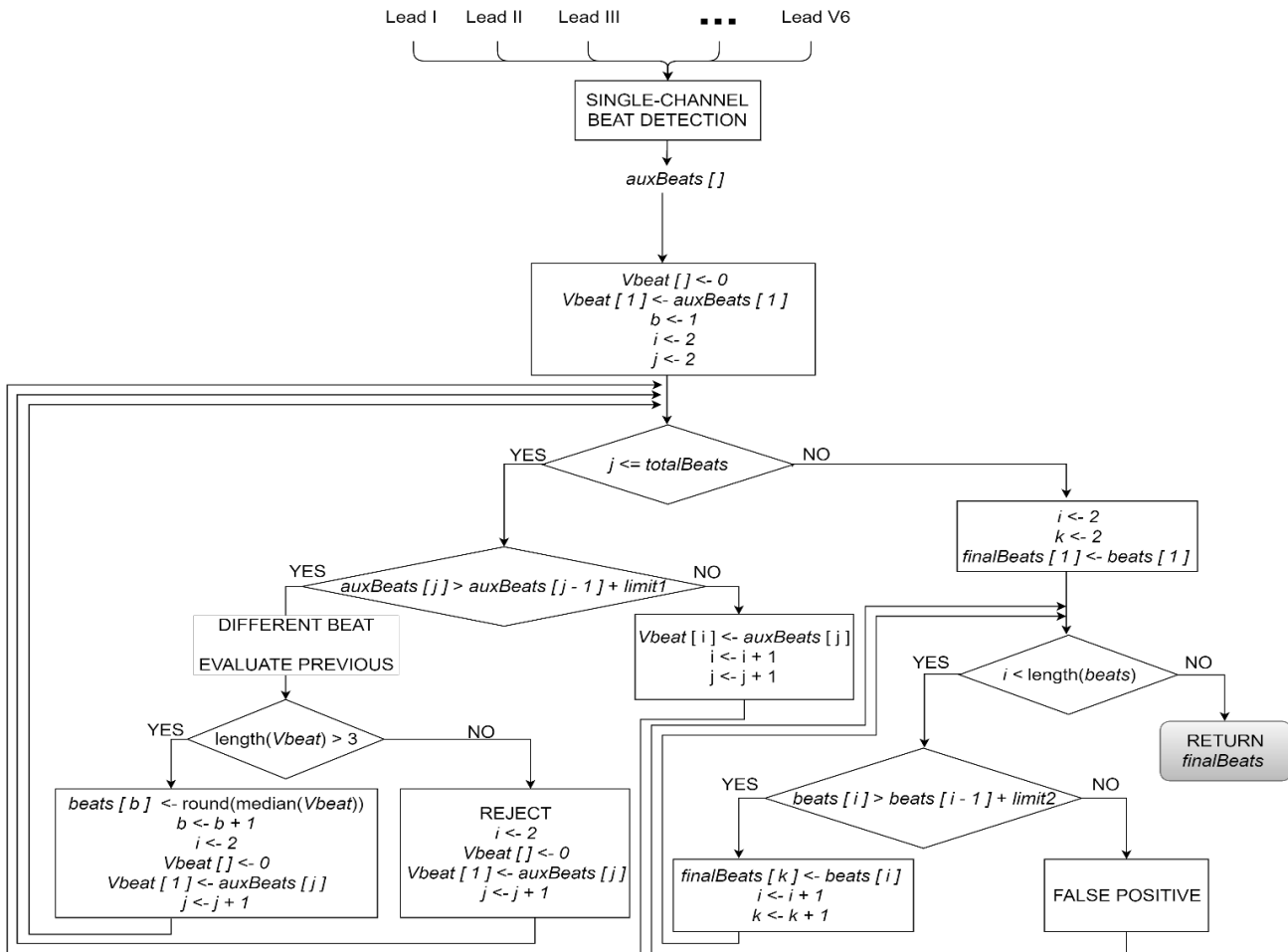


Fig. 1: Multichannel detection algorithm flow-chart

- Sensitivity (Se):

$$Se = \frac{TP}{(TP + FN)}$$

(1)

- Positive predictivity (P⁺):

$$P^+ = \frac{TP}{(TP + FP)}$$

(2)

- F-Score:

$$F - Score = 2 \times \left(\frac{P^+ \times Se}{P^+ + Se} \right)$$

(3)

- RMS (Root Mean Square) RR Interval Error,

IV. RESULTS AND DISCUSSION

In this Section, we present the results obtained with both the single-channel and the multichannel algorithms proposed in this paper.

A. Single-channel detection

As stated before, we implemented first the standard Pan-Tompkins algorithm to perform single-channel detection, and we validated this algorithm with the MIT-BIH Arrhythmia database. Results obtained with our method were similar to other published results for P⁺, and slightly better for Se [24].

The standard Pan-Tompkins method was also modified and adapted to the different leads of a multichannel ECG. This adaptation was tested using the St. Petersburg database, giving a significant improvement over the unmodified Pan-Tompkins method [24].

B. Multichannel detection

Comparing our multichannel results against similar approaches is not an easy task, since there are very few works dealing with beat detection using all channels of a 12-lead ECG. For instance, [17] follows this idea, although validation is performed against a set of manually reviewed beat positions. These are not publicly available, making it not adequate to use it for comparing results. The only paper that presents results on the St. Petersburg database comparable to ours is [25], and performance of their Pan-Tompkins implementation on St. Petersburg database is worse than ours (Se = 99.03%, P⁺ = 97.09%).

Our multichannel detection algorithm was tuned using the training set of the St. Petersburg database, and using this set, we obtained Se = 99.83% and P⁺ = 99.94% on average. Then, using the validation set of the database, our algorithm gave the results that are included in Fig. 2. As it can be seen, performance on the validation set (Se = 99.86%, P⁺ = 99.98%, RMS RR Interval Error = 25.98 ms, F-Score = 0.9992) is similar to the one obtained using the training set. Whereas there are some records where performance is noticeable lower (record I71, for example), in 14 out of 36 cases we obtained both Se and P⁺ equal to 100%. There are more records achieving these values in the validation set than in the training one, probably due to the presence of better quality records in the former.

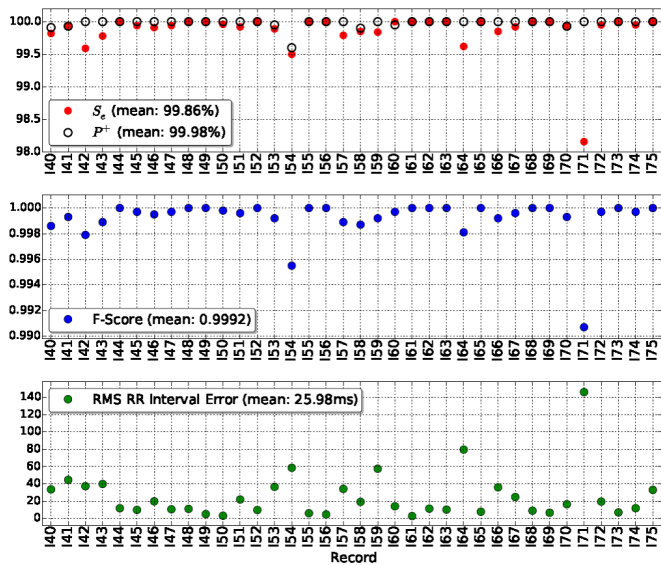


Fig. 2: Se, P+, F-Score and RMS RR Interval Error for each record of the validation set of St. Petersburg database with the multichannel algorithm.

One of the strongest points of our method is that it provides robustness against situations similar to the one shown in Fig. 3: combining the information of all the channels, it is possible to correct both the undetected and the erroneously detected beats.

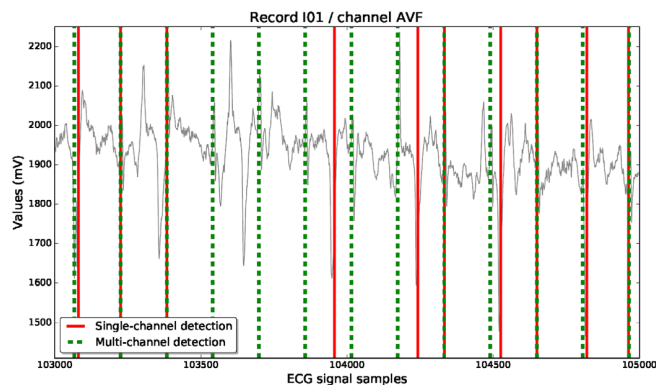


Fig. 3: Beats detected by the single-channel (red lines) and multichannel (dashed green lines) algorithms in a segment of a problematic channel. Whereas several beats are missing and some false positives appear when using the single-channel algorithm, all the beats are correctly detected by the multichannel algorithm.

Fig. 4 shows all leads on the same segment. It can be seen that our method offers a simple solution for beats undetected, duplicated or erroneously detected in some channels and for possible combinations of these three situations. A potential unresolved situation would be when a beat appears duplicated on a certain channel. In this case, the estimated position of this beat would be biased due to the influence of this channel on the estimation. However, this is not a very frequent situation and, besides, the error would be smoothed by the rest of the channels.

From the very beginning, we thought of using some procedure to discard channels prone to yield false positives, due to the presence of noise or to a bad contact of its electrode.

Although our first idea was to require a beat to be detected in 50% of the available channels, our experiments showed that the algorithm best performs when just four channels are required to detect a beat. This value allows the algorithm to discard enough defective channels, while requiring a minimum of detections to ensure the actual presence of the beat under study.

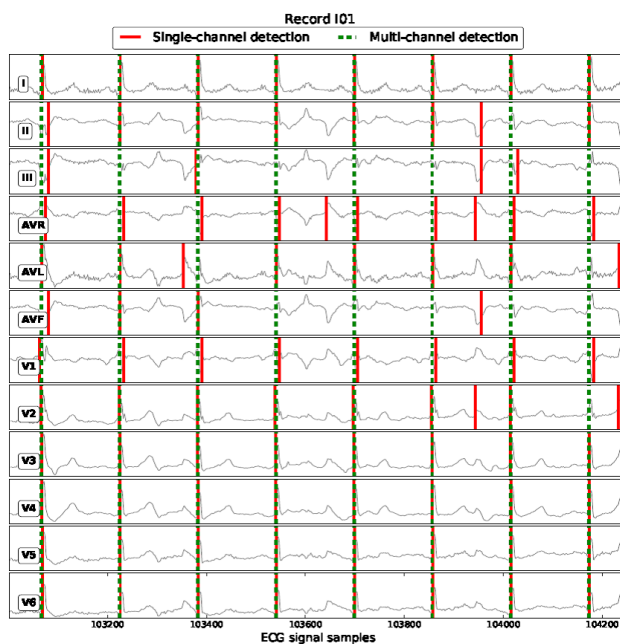


Fig. 4: Detected beats employing the single-channel (red lines) and multichannel (dashed green lines) algorithms in a segment of record I01. Those beats detected in at least four channels will be kept and their positions will be store as beat positions. Otherwise, they will be rejected.

Another possibility we considered was to implement some weighted voting system, similar to the approaches proposed in [26] and [27]. Each channel would have a weight estimated from its RMS RR Interval Error, thus conveying the reliability of the channel for beat detection. We tried several alternatives for obtaining and combining the weights, but none gave better results than the algorithm finally implemented. Best results obtained with the weighted channels approach on the training set of the database ($S_e = 99.58\%$, $P^+ = 99.57\%$, RMS RR Interval Error = 80.05 ms, F-Score = 0.9957) were clearly worse than the ones obtained when using the unweighted channels. This suggests that adapting the channels partially equalizes their performances, making it pointless either to penalize or to favor some leads over the rest. Weighting the channels could be a good alternative for systems without channel adaptation, although more difficult to implement and to train.

V. CONCLUSION

During the last decades, many efforts were devoted to the task of beat detection on ECG signals. There are different approaches that give good results, and many papers propose modifications over these approaches to improve the success rate. However, works based on multichannel detection are scarce, although combining the information present in the different channels to improve the results seems a good alternative.

In this paper, we have analyzed how to adapt a single-channel algorithm to a multichannel environment, combining the results of all the leads present in records obtained from a 12-channel ECG. Our approach increases the results of Se, P+ and F-Score, while reducing the RMS RR Interval Error significantly. The use of adapted thresholds brings out the best in problematic channels, going, in some cases, from unusable results to values similar to the ones obtained from the best channels. All this is achieved without preprocessing or modification of the signals, while keeping the runtime within acceptable values, even when working with complete databases.

Given the performance of our beat detection algorithm,

we think it is a good basis for another projects in which we are working at the moment, namely wave detection and ECG-based diagnosis.

REFERENCES

- [1] World Health Organization. (2011) Fact sheet no. 317. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs317/en/index.html>
- [2] D. M. Lloyd-Jones, E. P. Leip, M. G. Larson, R. B. d'Agostino, A. Beiser, P. W. Wilson, P. A. Wolf, and D. Levy, "Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age," *Circulation*, vol. 113, no. 6, pp. 791–798, 2006.
- [3] M. L. Dansinger, J. A. Gleason, J. L. Griffith, H. P. Selker, and E. J. Schaefer, "Comparison of the Atkins, Ornish, Weight Watchers, and Zone diets for weight loss and heart disease risk reduction: a randomized trial," *J Amer. Med. Assoc.*, vol. 293, no. 1, pp. 43–53, 2005.
- [4] S. C. Smith, E. J. Benjamin, R. O. Bonow, L. T. Braun, M. A. Creager, B. A. Franklin, R. J. Gibbons, S. M. Grundy, L. F. Hiratzka, D. W. Jones et al., "AHA/ACCF secondary prevention and risk reduction therapy for patients with coronary and other atherosclerotic vascular disease: 2011 update: a guideline from the American Heart Association and American College of Cardiology Foundation endorsed by the World Heart Federation and the Preventive Cardiovascular Nurses Association," *J Am. Coll. Cardiol.*, vol. 58, no. 23, pp. 2432–2446, 2011.
- [5] S. S. Lim, T. A. Gaziano, E. Gakidou, K. S. Reddy, F. Farzadfar, R. Lozano, and A. Rodgers, "Prevention of cardiovascular disease in high-risk individuals in low-income and middle-income countries: health effects and costs," *Lancet*, vol. 370, no. 9604, pp. 2054–2062, 2007.
- [6] R. Kahn, R. M. Robertson, R. Smith, and D. Eddy, "The impact of prevention on reducing the burden of cardiovascular disease," *Diabetes care*, vol. 31, no. 8, pp. 1686–1696, 2008.
- [7] G. D. Clifford, F. Azuaje, and P. McSharry, *Advanced methods and tools for ECG data analysis*. Norwood, MA, USA: Artech House, Inc., 2006.
- [8] M. S. Thaler, *The only EKG book you'll ever need*, ser. Board Review Series. Lippincott Williams & Wilkins, 2010.
- [9] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE T. Bio-med. Eng.*, no. 3, pp. 230–236, 1985.
- [10] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ECG delineator: evaluation on standard databases," *IEEE T. Bio-med. Eng.*, vol. 51, no. 4, pp. 570–581, 2004.
- [11] P. Hamilton, "Open source ECG analysis," in *Comput. Cardiol.* 2002, vol. 29. IEEE, 2002, pp. 101–104.
- [12] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE T. Bio-med. Eng.*, vol. 46, no. 2, pp. 192–202, 1999.
- [13] M. Adnane, Z. Jiang, and S. Choi, "Development of QRS detection algorithm designed for wearable cardiorespiratory system," *Comput. Meth. Prog. Bio.*, vol. 93, no. 1, pp. 20–31, 2009.
- [14] J. Lee, K. Jeong, J. Yoon, and M. Lee, "A simple real-time QRS detection algorithm," in *Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE*, vol. 4. IEEE, Oct 1996, pp. 1396–1398.
- [15] J. Quero, M. Elena, J. Segovia, C. Tarrida, J. Santana, and C. Santana, "CardioSmart: Cardiological monitoring intelligent system using GPRS," *IEEE Lat. Am. T.*, vol. 2, no. 3, pp. 152–158, 2005.
- [16] E. B. Mazomenos, D. Biswas, A. Acharyya, T. Chen, K. Maharatna, J. Rosengarten, J. Morgan, and N. Curzen, "A low-complexity ECG feature extraction algorithm for mobile healthcare applications," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 2, pp. 459–469, 2013.
- [17] M. Yochum, C. Renaud, and S. Jacquir, "Automatic detection of P, QRS and T patterns in 12-leads ECG signal based on CWT," *Biomed. Signal Proces.*, vol. 25, pp. 46–52, 2016.
- [18] V. Chouhan and S. Mehta, "Detection of QRS complexes in 12-lead ECG using adaptive quantized threshold," *IJCSNS*, vol. 8, no. 1, pp. 155–163, 2008.
- [19] S. S. Mehta and N. S. Lingayat, "Development of entropy based algorithm for cardiac beat detection in 12-lead electrocardiogram," *Signal Process.*, vol. 87, no. 12, pp. 3190–3201, 2007.
- [20] —, "Combined entropy based method for detection of QRS complexes in 12-lead electrocardiogram using SVM," *Comput. Biol. Med.*, vol. 38, no. 1, pp. 138–145, 2008.
- [21] —, "Application of support vector machine for the detection of P and T-waves in 12-lead electrocardiogram," *Comput. Meth. Prog. Bio.*, vol. 93, no. 1, pp. 46–60, 2009.
- [22] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol.*, vol. 20, no. 3, pp. 45–50, 2001.
- [23] 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, no. 23, pp. e215–e220, 2000.
- [24] V. Mondelo, M. J. Lado, A. J. Méndez, X. A. Vila Liñares, "A Channel-Dependent Algorithm for Heart Beats Detection in ECG Recordings", in *Information Systems and Technologies (CISTI), 2017 12th Iberian Conference*, Jun 2017, pp. 2225–2229.
- [25] M. Elgendi, "Fast QRS detection with an optimized knowledge-based method: Evaluation on 11 standard ECG databases," *PloS One*, vol. 8, no. 9, p. e73557, 2013.
- [26] C. A. Ledezma and M. Altuve, "Fusión de datos para detectar complejos QRS en registros electrocardiográficos multicanal," in *V Congreso Venezolano de Bioingeniería*, Mérida, Venezuela, 2015.
- [27] S. Torbey, S. G. Akl, and D. P. Redfearn, "Multi-lead QRS detection using window pairs," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Aug 2012, pp. 3143–3146.