

Robust, real-time generic detector based on a multi-feature probabilistic method

Matthieu Doyen, Di Ge, Alain Beuchée, Guy Carrault, Alfredo I Hernandez

▶ **To cite this version:**

Matthieu Doyen, Di Ge, Alain Beuchée, Guy Carrault, Alfredo I Hernandez. Robust, real-time generic detector based on a multi-feature probabilistic method. PLoS ONE, 2019, 14 (10), pp.e0223785. 10.1371/journal.pone.0223785 . inserm-02419006

HAL Id: inserm-02419006

https://inserm.hal.science/inserm-02419006

Submitted on 19 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PLOS ONE

RESEARCH ARTICLE

Robust, real-time generic detector based on a multi-feature probabilistic method

Matthieu Doyen, Di Ge*, Alain Beuchée, Guy Carrault, Alfredo I. Hernández

Univ Rennes, CHU Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

These authors contributed equally to this work.
* di.ge@univ-rennes1.fr

Check for updates

OPEN ACCESS

Citation: Doyen M, Ge D, Beuchée A, Carrault G, I. Hernández A (2019) Robust, real-time generic detector based on a multi-feature probabilistic method. PLoS ONE 14(10): e0223785. https://doi.org/10.1371/journal.pone.0223785

Editor: Esther Pueyo, Universidad de Zaragoza, SPAIN

Received: May 13, 2019

Accepted: September 27, 2019

Published: October 29, 2019

Copyright: © 2019 Doyen et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data of the MIT-BIH Arrhythmia Database - PhysioNet used in this paper are available at: https://physionet.org/physiobank/database/mitdb/. The noisy benchmark database is available at: https://github.com/dge996/MIT_NoiseStress/ and can also be constructed using the method described at: https://physionet.org/physiobank/database/nstdb/".

Funding: This work was mainly funded by the French Ministry of Higher Education, Research and Innovation (MESRI). It also received support from the European Union's Horizon 2020 research and

Abstract

Robust, real-time event detection from physiological signals acquired during long-term ambulatory monitoring still represents a major challenge for highly-artifacted signals. In this paper, we propose an original and generic multi-feature probabilistic detector (MFPD) and apply it to real-time QRS complex detection under noisy conditions. The MFPD method calculates a binary Bayesian probability for each derived feature and makes a centralized fusion, using the Kullback-Leibler divergence. The method is evaluated on two ECG databases: 1) the MIT-BIH arrhythmia database from Physionet containing clean ECG signals, 2) a benchmark noisy database created by adding noise recordings of the MIT-BIH noise stress test database, also from Physionet, to the MIT-BIH arrhythmia database. Results are compared with a well-known wavelet-based detector, and two recently published detectors: one based on spatiotemporal characteristic of the QRS complex and the second, as the MFDP, based on feature calculations from the University of New South Wales detector (UNSW). For both benchmark Physionet databases, the proposed MFPD method achieves the lowest standard deviation in sensitivity and positive predictivity (+P) despite its online algorithm architecture. While the statistics are comparable for low-to mildly artifactual ECG signals, the MFPD outperforms reference methods for artifacted ECG with low SNR levels reaching 87.48 ± 14.21% in sensitivity and 89.39 ± 14.67% in +P as compared to 88.30 ± 17.66% and 86.06 ± 19.67% respectively from UNSW, the best performing reference method. With demonstrations on the extensively studied QRS detection problem, we consider that the proposed generic structure of the multi-feature probabilistic detector should offer promising perspectives for long-term monitoring applications for highly-artifacted signals.

1 Introduction

Event detections from physiological signals are often faced with important noise perturbations, especially in clinical monitoring context. Main strategies are focused on finding an efficient feature reliable in most cases. Generally, these methods such as [1] get interesting results under low- to mid-level noise conditions, but performances decrease significantly with the

innovation program under grant agreement no. 689260 (Digi-NewB project).

Competing interests: The authors have declared that no competing interests exist.

signal-to-noise ratio (SNR) diminution or with a change in the noise type since all features have vulnerabilities to specific distortions. To circumvent this weakness, multi-feature detectors were proposed [2] but the decentralized fusion method does not fully exploit feature informations using statistical learning strategies. The main objective of this paper is to propose a generic event detection method with centralized fusion, in which the final decision is made by a weighted sum of posterior detection probabilities derived from each feature's statistical properties. The power of the method is illustrated by its application to real-time QRS complex detection from electrocardiogram (ECG) signals.

QRS complex is the most prominent deflection in ECG signal and corresponds to the electrical depolarization of ventricles. The detection is often the first analysis performed on ECG signal processing, in order to estimate basic cardiac markers, such as heart rate or to perform further ECG segmentation and analysis. The QRS complex detection has been investigated for many decades [1] and yet remains a challenge [3] as an event detection problem from physiological signals. Many different methods have been proposed and a number of review publications have been dedicated to this subject [4] [5]. The main proposed methods are based on filtering and non linear transformations [1], fuzzy hybrid neural networks [6], S-Transform [7] or wavelet analysis [8–11]. Although these QRS detection methods perform well in low- to mid-level noise conditions, their applications on long-term periods of ECG recordings under noisy conditions such as in ambulatory care and intensive care units still pose a significant challenge as evidenced by the large number of RR correction methods reported in [12]. Indeed, these ECG recordings are often prone to episodes of strong signal non-stationarity, sudden modifications of beat morphologies and most importantly the presence of several types of noise (baseline drift, saturation, power-line pickup, muscular contractions and motion artifacts [13]). Recently in [14], the authors explained that QRS detection has not been completely assessed in terms of robustness to noise. As a consequence, recent publications [15] [16] [17] and a recent PhysioNet challenge [3] have been focused on the specific problem of robust QRS detection. Furthermore, the emergence of wearable cardiac monitors, with a limited number of leads [18] [19] for long-term daily-life recordings [20] further revives the research interests on robust QRS detection for low-quality electrocardiograms.

In our previous works, we have proposed different methods to improve the robustness of QRS detection, through multisensor fusion [2], adaptive selection of QRS detectors as a function of the signal context [21] or through optimal detector parameter configuration, using evolutionary methods [22] [23]. More recently, we revisited this optimization process in order to identify optimal parameter configurations with respect to changes in signal noise [24].

In this paper, we propose and evaluate a novel, generic event detector, that provides improved robustness through the probabilistic combination of a set of signal features. Section 2 presents the general architecture of the proposed Multi-Feature Probabilistic Detector (MFPD) and a specific implementation adapted to robust QRS detection. Section 4 evaluates its performances on two ECG databases: 1) the benchmark MIT-BIH [25] to validate the proposed method on clean ECG signals with consensus annotations, 2) the benchmark noisy database (created by adding noise recordings to the MIT-BIH) with known artifact types and levels, to validate the proposed method on highly artifactual ECG signals including various artifacts.

2 Methods

2.1 General architecture of the detector

The general architecture of the MFPD depicted in Fig 1 is based on the following steps:

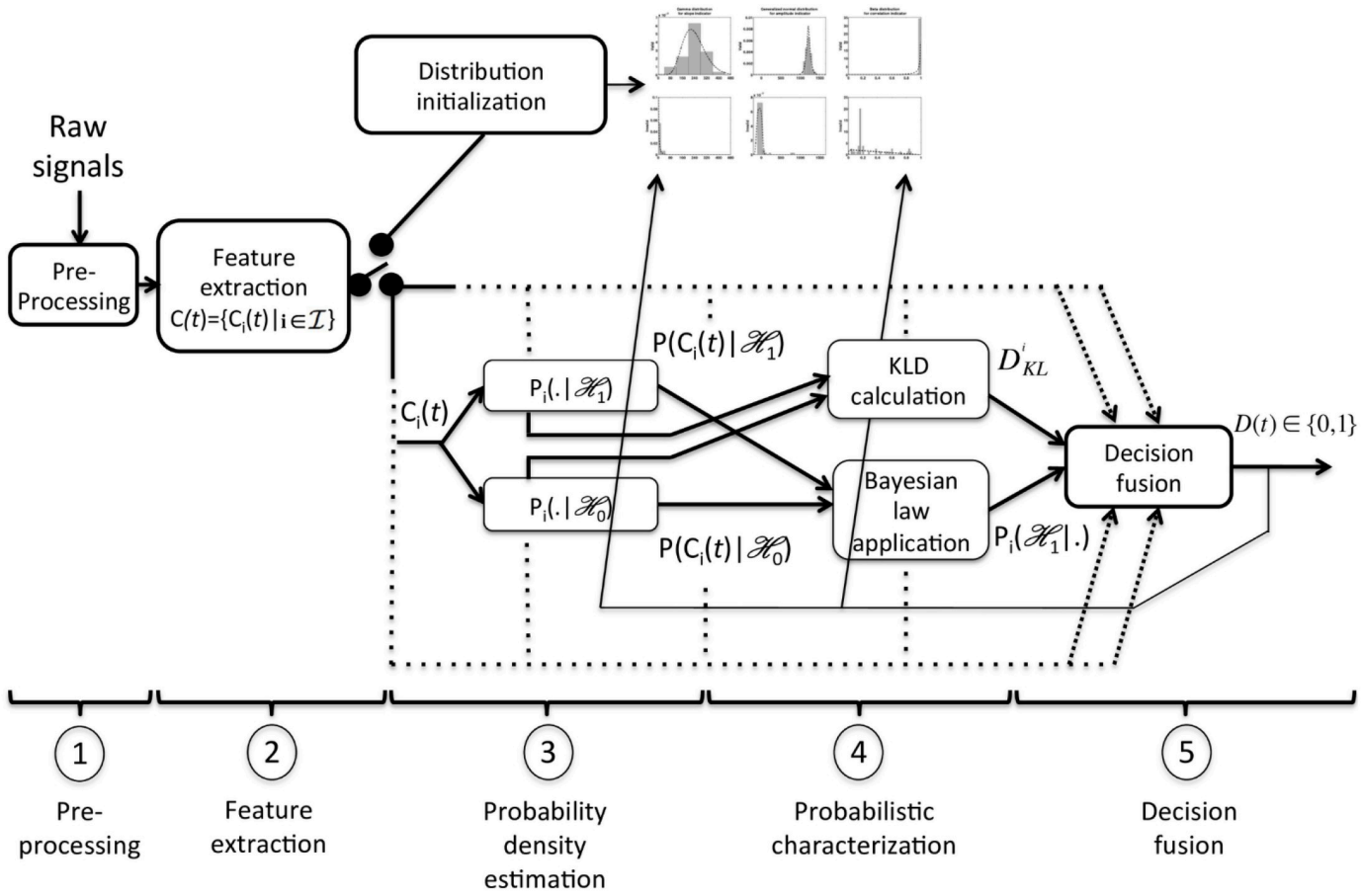


Fig 1. Global architecture of the MFPD. In the first two steps, data coming from ECG are filtered then converted into features $C_i(t)$. $P_i(.|\mathscr{H}_1)$ and $P_i(.|\mathscr{H}_0)$ are parametric probability models of feature i, representing valid and invalid detections.

https://doi.org/10.1371/journal.pone.0223785.g001

1. Pre-processing: Raw signals are processed in order to improve the SNR and to pre-select potential candidates (to be validated by the detector) of the events of interest at instants t.
2. Feature extraction: For every event candidate selected at instant t, a vector $C(t) = \{C_i(t)|i \in \mathcal{I}\}$ is created, where $\mathcal{I}$ is a set of complementary features extracted from the preprocessed signals.
3. Probability density estimation: The probability density functions (pdf), noted as $P_i(C_i(t); \Theta_{i0/1}, \mathscr{H}_{0/1})$ are used to model feature i on the observed candidate $C(t)$, with the two hypotheses:

$$\mathscr{H}_0 : D(t) = 0,$$
$$\mathscr{H}_1 : D(t) = 1,$$

where $D(t)$ is the final detection decision ($D(t) = 1$ for detection and $D(t) = 0$ otherwise) and $\Theta_{i0/1}$ the parameter set for each hypothesis. Note that for each feature, the two pdf belong to the same distribution family, whose parameters $\Theta_{i0/1}$ are initialized at the beginning of the recording and updated throughout the detection process (see Section 2.2.3 for more details).

4. Probabilistic characterization: The posterior probability $P_i(\mathscr{H}_1|C_i(t))$ is calculated by applying the Bayes law. Moreover, the Kullback-Leibler divergence (KLD) between each pdf pair characterizing feature i, D^i_{KL}, is calculated.

5. Decision fusion: The posterior probabilities, weighted by their respective KLD, are combined to build a binary decision $D(t)$ on whether candidate $C(t)$ is a valid event ($D(t) = 1$) or not ($D(t) = 0$). According to the decision for the current candidate, distribution parameters are updated to complete the real-time learning process.

2.2 Real-time QRS detection implementation

In this section, we detail the realization of the above-mentioned MFPD, adapted for real-time detection of QRS complexes. From the generic approach of Fig 1, the specific adaptions to QRS complex detection include mainly the pre-processing (step 1) and the feature extraction (step 2), they are depicted in Fig 2.

2.2.1 Pre-processing. As in many other QRS detection methods, the first step consists in applying to the raw ECG signal different transformations. Typically, a band-pass filter, a derivative filter, a non-linear transformation and a final smoothing filter are applied. We adopt here the pre-processing method used in [1, 26]: Fig 2 represents a diagram of these steps. In the following, the band-pass filtered ECG signal, computed using finite impulse response (FIR) low pass (f_{cutoff} = 19 Hz, order = 256) and high pass (f_{cutoff} = 8 Hz, order = 256) filters, will be denoted SAecg; and the output of a squared transformation followed by a derivative FIR filter (f_{cutoff} = 30 Hz, order = 129) then a smoothing filter (length = 101 ms) on signal SAecg will be denoted SFecg (cf Fig 2). FIR filters were designed using Remez exchange algorithm. Each local maximum detected at an instant t on SFecg is considered as a potential QRS candidate.

2.2.2 Feature extraction. In this paper, a set of 3 features $C(t) = \{C_i(t), i \in \mathcal{I}\}, \mathcal{I} = \{s, a, c\}$ is extracted for each QRS candidate. These features are the input to the probability

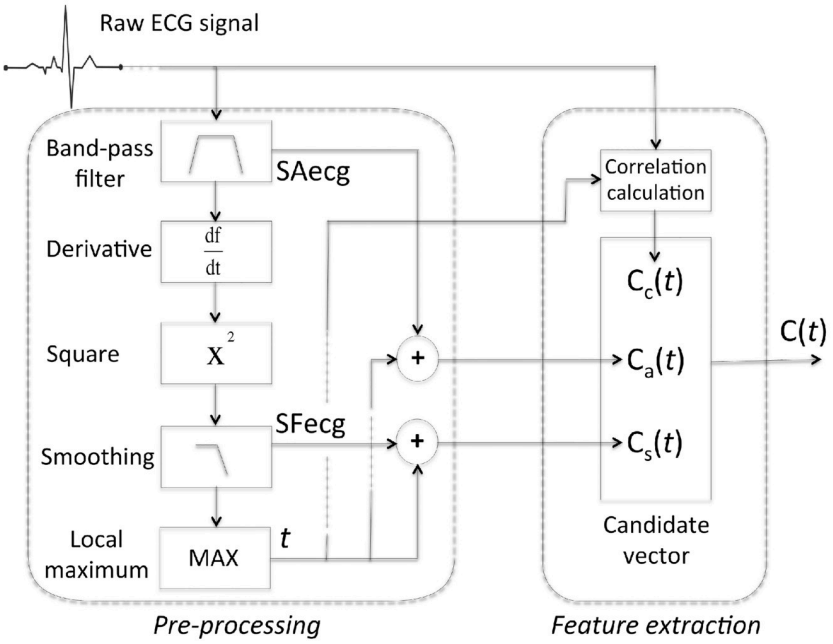


Fig 2. Specification of signal pre-processing and feature extraction (steps 1 and 2 in Fig 1). For the application of robust QRS detection, a set of 3 features $C(t) = \{C_i(t), i \in \mathcal{I}\}, \mathcal{I} = \{s, a, c\}$ is extracted.

https://doi.org/10.1371/journal.pone.0223785.g002

density estimation step (step 3 in Section 2.1 and Fig 1). The squared slope (s), the raw amplitude (a) and the absolute correlation with a beat template (c) constitute the most common features used in the literature for QRS detection. In the current study, we prove the concept of decision fusion with adaptive weights to track the relative importance of each feature while the optimization of feature selection lies beyond the scope of our investigation.

2.2.3 Feature probability density.

1. Feature probability models: in the proposed QRS detection application, each candidate is characterized by a set of 3 features $\mathcal{I} = \{s, a, c\}$. The probability distribution function (pdf) of features are chosen with several conditions: 1) distributions should have the same support as the calculated feature, e.g. $(0, \infty)$ for the squared slope, $(-\infty, +\infty)$ for the amplitude and [0, 1] for the absolute correlation, 2) pdf parameter estimations should be easy to calculate, with low complexity in the updating methods, 3) the distance measures on these pdfs should also be easy to compute and tractable, without numerical integrations. In this paper, we implemented three different pdfs: (i) a Beta distribution with closed support, (ii) a Gamma distribution with one-sided open support, and (iii) a generalized normal distribution (GND) with two-side open support. Thus, pdf parameter update and distance calculation methods in this paper can be generalized to include most continuous features in the future.

- The squared slope of the peak (s) is the value of SFecg signal at instant t. This feature is represented with the Gamma distribution with two degrees of freedom:

$$P_s(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k} \mathbb{1}_{x>0}, \tag{1}$$

  where $k \in \mathbb{R}^+$ is the shape parameter, and $\theta \in \mathbb{R}^+$ the scale parameter. The indicator function $\mathbb{1}_{x>0}$ typically limits the function support to $\mathbb{R}^+$.

- The peak amplitude (a) is the value of SAecg signal at instant t. We characterized it using the GND defined as:

$$P_a(x; \alpha; \beta; \mu) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \; e^{-\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}} \tag{2}$$

  where $\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} \mathrm{d}x$ is the gamma function and $\mu \in \mathbb{R}$ the position parameter, $\alpha \in \mathbb{R}^+$ the scale parameter and $\beta \in \mathbb{R}^+$ the shape parameter. Note that both positive and negative peak values can be fitted with the GND model.

- The absolute Bravais-Pearson correlation (c) is calculated between the candidate peak (represented by 50 ms of raw ECG signal centered 20 ms before the peak) at instant t and an adaptive template. The template duration was chosen in order to extract mainly the information around the peak, where the information is most characteristic of the QRS complex in our opinion. With a longer template duration, QRS complexes affected by noise can obtain a lower correlation, this weakness is less present with a short template duration. With an even shorter duration, high-frequency artifacts can be very similar to QRS complexes. In order to model this feature, we have chosen the Beta distribution, defined as:

$$P_c(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \; x^{\alpha-1}(1-x)^{\beta-1} \; \mathbb{1}_{[0,1]}(x). \tag{3}$$

  This can also be considered as a special case of the Dirichlet distribution, with two positive shape parameters α and β. Parameters are estimated using maximum a posteriori (MAP) method, further details are reported in section 4.2.

Note again that the same pdf family is proposed for both $\mathscr{H}_0$ and $\mathscr{H}_1$ considering the numerical tractability of the KLD calculations (see discussion in 2.2.4).

2. Initialization: As in step 3) in section 2.1, the model parameters—$\{k, \theta\}$ for the Gamma distribution, $\{\alpha, \beta, \mu\}$ for the GND, and $\{\alpha, \beta\}$ for the Beta distribution—are initialized during the heating-up period (cf block *distribution initialization* in Fig 1) using the Pan-Tompkins detector [1] for the 40 first validated QRS detections (note that these detections are considered in the performance evaluation step). At the end of the heating-up period, all QRS candidates are labelled as either validated ($\mathscr{H}_1$) or invalidated ($\mathscr{H}_0$) and the model parameters are estimated for each distribution, using the maximum likelihood estimator (MLE) for the *s* and *a* feature, and MAP estimator for the *c* feature. Beat templates (for both $\mathscr{H}_1$ and $\mathscr{H}_0$) are also initialized (aligned and averaged) using the validated and invalidated candidates of the heating-up.

3. Learning: Once initialized, the MFPD detector shifts to the decision fusion mode (step 3-5 in section 2.1 and Fig 1) whose final detection results (of step 5) feed the pdf parameter updating (step 3) in the same manner as during the initialization. KLDs are updated as a direct consequence (step 4). In a similar manner, the beat templates are updated after each decision using 80% of the previous template and 20% of the current candidate. If no detection occurs during 3.5 sec, all parameters are reset and a new initialization starts.

2.2.4 Probabilistic characterization. Based on the pdf model for each feature in 2.2.3, two probabilistic markers are calculated for each feature: the posterior probability and the KLD. The posterior probability of validating $\mathscr{H}_1$ for feature $C_i(t)$ is given by

$$P_i(\mathscr{H}_1|C_i(t)) = \frac{P_i(C_i(t)|\mathscr{H}_1)\,P_i(\mathscr{H}_1)}{P_i(C_i(t)|\mathscr{H}_1)P_i(\mathscr{H}_1) + P_i(C_i(t)|\mathscr{H}_0)P_i(\mathscr{H}_0)} \tag{4}$$

using the Bayes rule. A binary decision can typically be made by thresholding this probability as in most single feature-based QRS detectors.

We propose in this paper a centralized fusion by making a weighted sum of these posterior probabilities. The key here is to update dynamically a metric to measure the pertinence of each feature, or the power of separating two classes in the detection context by measuring the distance between the two antagonist pdfs. The KLD is such a non-negative measure defined by:

$$D_{\mathrm{KL}}(p\|q) = \int_{-\infty}^{\infty} p(x)\log\frac{p(x)}{q(x)}\,\mathrm{d}x. \tag{5}$$

It is particularly well-suited to assess the distance between the distribution pair $P_i(C_i(t); \Theta_{i0/1}; \mathscr{H}_{0/1})$. Analytic expressions can be found in the literature in the case of Beta [27] and Gamma distributions [28]. However for the GND case and to our best knowledge, no analytic expression can be found in the general case when $\mu_p \neq \mu_q$. One theoretical contribution of this paper is to efficiently calculate the KLD between two GND distributions in the case of general settings. While detailed derivation is reported in the appendix B, here we give a summary of the main results:

1. Analytic expressions for $\beta_q \in \mathbb{N}^+ \cup \{0\}$ have been derived;
2. Eq (5) is monotonously increasing with β_q;
3. The computational complexity requires $2 \times (\beta_q + 1)$ gamma function evaluations.

Thus without numerical integration of Eq (5), we are able to obtain a close approximation of the KLD value for all $\beta_q \in \mathbb{R}^+$. Note that the other parameters (α_p, α_q, β_p, μ_p, μ_q) have no effect on the calculation complexity.

Finally, and without loss of generality, the KLD is one of the *f*-divergence functions that measures the difference between two probability distributions. They are non-negative, monotonous and jointly convex. For example, the reverse KLD $D^R_{\mathrm{KL}}(p\|q) = D_{\mathrm{KL}}(q\|p)$ is another *f*-divergence and can be easily calculated by reversing the role of the two distributions. However, it is less evident to derive tractable numerical methods for the Hellinger distance, or the χ^2-divergence for the exponential pdf family.

2.2.5 Decision fusion. Based on the probabilistic markers, the following decision rule is applied

$$D(t) = 1 \text{ if and only if } \sum_{i\in\mathcal{I}} \frac{\bar{D}^i_{\mathrm{KL}} . P_i(\mathscr{H}_1|C_i(t))}{\sum_{j\in\mathcal{I}} \bar{D}^j_{\mathrm{KL}}} > \lambda \tag{6}$$

where λ is the decision threshold. The modified $\bar{D}^i_{\mathrm{KL}}$ is calculated by letting $i^* = argmax\{D^i_{KL}\}$ and:

$$\bar{D}^i_{KL} = \begin{cases} min\left\{ D^i_{KL}, 2\sum_{h\neq i^*} D^h_{KL} \right\}, & i = i^* \\ D^i_{KL}, & i \neq i^* \end{cases} \tag{7}$$

Such that the most significant KLD should not exceed 2/3 after normalization. Intuitively, the decision rule represents the sum of all posterior probabilities, weighted by their normalized KLD, such that features that are better separated in distributions (between $\mathscr{H}_0$ and $\mathscr{H}_1$) have more weight in the final decision making.

3 Performance evaluation for QRS detection

3.1 Databases

Two databases were used:

- Benchmark database: we first performed a comparative study using the first lead of the MIT-BIH Arrhythmia Database with consensus annotations to show the detection performances on clean ECG signals even though the main objective of the paper is to evaluate detector's robustness under artifact conditions.
- Benchmark noisy database: we then created a benchmark simulated database by adding to the MIT-BIH Arrhythmia Database three noise sources (baseline wander, muscle and electrode motion artifact) extracted from the MIT-BIH noise stress test database. Noises were recorded from physically active volunteers [25] and with SNR levels from −6dB to 24dB by 6dB increments. It is composed of 864 noisy signals for which the reference annotations are simply copies of those for the MIT-BIH Arrhythmia Database. The purpose of this database test is to provide a ground truth for detection performance comparison with different levels and types of noises. A quick access to this benchmark database is available at: https://github.com/dge996/MIT_NoiseStress/ though it can also be constructed using the method described in [25].

We choose to report results on two databases to prove the design concept of the MFPD robustness for clean and noisy databases with consensus annotations. For the second one, both artifact type and level classification annotations are provided, they allow differentiated comparisons.

3.2 Comparison methods

As in recent publications [29] [15], performance of the proposed MFPD was compared with the following state of the art QRS detection methods:

- University of New South Wales detector (UNSW): a feature-based (ECG amplitude and derivative) detector, with adaptive thresholding, suitable for both clinical and poorer quality tele-health ECG [15].
- Spatiotemporal Characteristics Detector (SCD): a simple and robust realtime QRS detection algorithm based on spatiotemporal characteristics [30], with state-of-the-art performances reported for noisy signals.
- Wavelet-Based Detector (WBD): a wavelet-based QRS detector [31], with implementation provided in the ECG-toolkit [32] [33].

WBD was chosen because it is a reference in QRS detection. The other two methods were selected because their objectives are the same as MFDP, which is not the case for many detectors from the literature focused on clean ECG signals. SCD is described by their authors as robust and real-time, UNSW detector is described as suitable for poorer quality ECG signals. Moreover, both were recently published (2016) with open access toolbox. We underline however that contrary to UNSW and WBD methods, the MFPD is an online algorithm designed for monitoring applications which forbids typically bi-directional filtering and beat by beat corrections in post-treatment. We thus further included detection delay statistics, that cannot be obtained for classical detectors.

3.3 Performance criterion

In this study, we used the `bxb` program (as part of the WFDB applications available on physionet) to obtain the beat-by-beat performance statistics for QRS complex detection according to the American national standard for ambulatory ECG analyzers (ANSI/AAMI EC38). We report for the two databases the total number of TP, FP and FN as well as the sensitivity (Se) and positive predictivity (+P) metrics calculated per record and a mean and standard deviation alongside the overall result. As in recent publications [24] [34], we chose a match window of 50 ms. Only one detection is considered as true positive (TP) inside the match window centered at each QRS annotation while the rest are considered as false positive (FP) detections. A false negative (FN) is counted when no detection is found inside the match window. Due to their direct influences on heart rate variability analysis, we also report the mean and standard deviation of *jitters*: the difference in time between a TP and its associated annotation. To easily observe the increase in detection performance as a whole, the detection error rate DER = (FN+FP)/(TP+FN) is also reported. Furthermore, computational complexity and detection delay of the MFPD are also given since they are key implementation issues in realtime monitoring applications.

4 Results

In addition to the global performance evaluation and comparison summarized in 4.4, we also provide here some intermediate results to illustrate the multi-feature complementarity in 4.1,

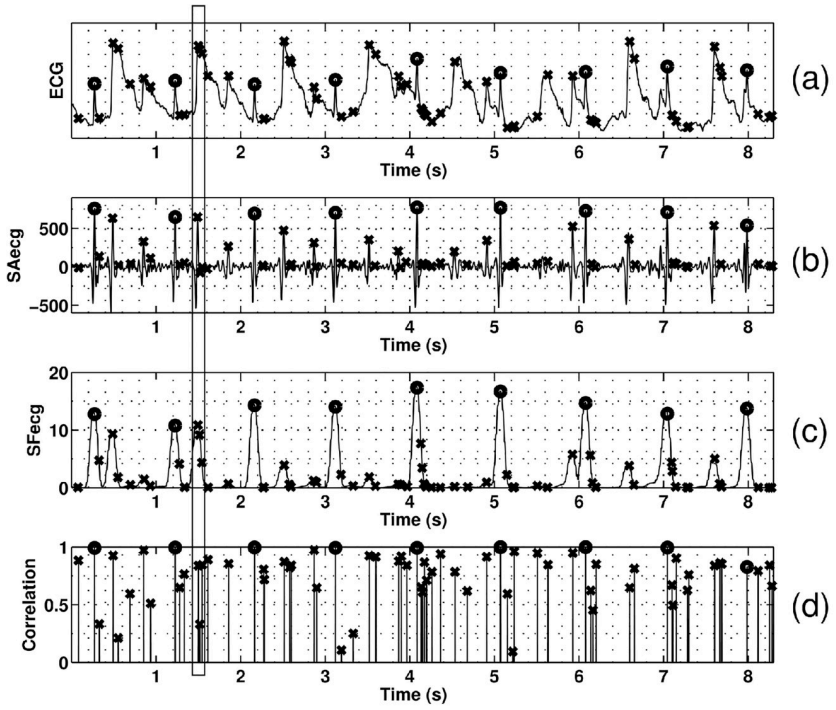


Fig 3. Illustration of the multi-feature complementarity in fusion decision. a) Raw ECG segment from record MIT-101 (SNR = 6 dB with electrode noise) b), c), d) represents the feature *a*, *s*, *c* respectively. The × and ∘ indicate invalidated and validated QRS respectively by the fusion.

https://doi.org/10.1371/journal.pone.0223785.g003

their distribution estimation results in 4.2, and the importance of KLD weighting in the centralized decision making in 4.3.

4.1 Multi-feature complementarity

Fig 3 shows an example of the processed ECG signals and the features extracted in this paper. Panel (a) shows an ECG segment from record MIT-101, with added electrode movement noise at 6 dB. Panels (b) (c) and (d) represent, respectively, the SAecg (feature *a*), the SFecg (feature *s*) signals and the Bravais-Pearson correlation (feature *c*) related to each QRS complex candidates in panel (a). Candidates with the × symbol are not validated, while those with the ∘ symbol are validated as QRS by the MFPD method. In this example, all validated detections were TP and all invalidated candidates were true negatives (TN), not belonging to any annotation's match window. The vertical box pinpoints a segment that, if analyzed individually with thresholding, would have produced a false positive. The weighted fusion by Eq (6) however takes the opposite decision. This example shows the power of the multi-feature complementarity in such complex signal context.

4.2 Distribution estimation

Estimated parametric distributions (in dashed line) and normalized histogram (in vertical boxes) for each feature are illustrated in Fig 4. The x-axis stands for the dimensionless feature values while y-axis the probability density (also dimensionless). Generally speaking, the pdf type associated with updated parameters can reasonably fit the feature histograms. Kolmogorov–Smirnov tests were realized at each pdf parameter update, with the current candidates

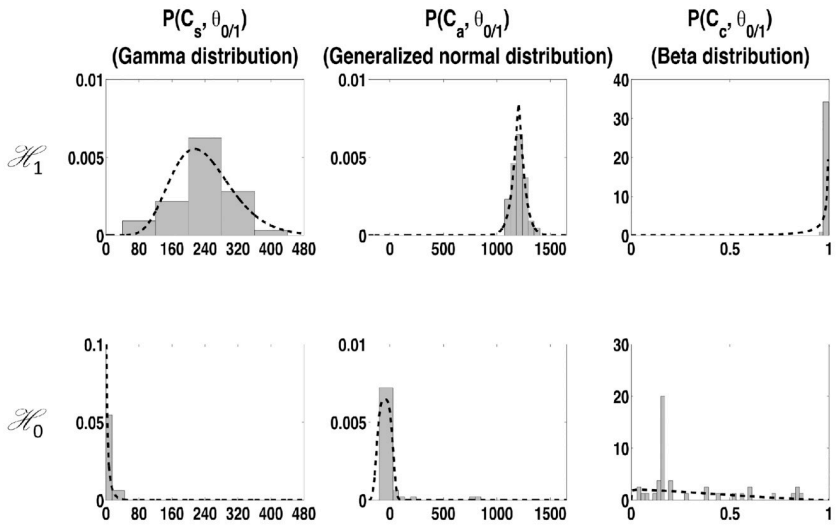


Fig 4. Estimated distributions (dashed line) vs normalized histograms (vertical boxes) for the three features and for both $\mathscr{H}_0$ and $\mathscr{H}_1$. Record MIT-101 is used with added baseline noise (SNR = 12dB).

https://doi.org/10.1371/journal.pone.0223785.g004

history. Of the 10678 tests performed for the MIT-101 signal, all were positive at the 0.05 significance level.

For mildly-artifacted ECG signals, due to the stability of the QRS complex morphology, the correlation features $C_c(t)$ can be close to 1 for most valid candidates. Thus large α values are estimated for the beta distribution $P_c(C_c(t); \Theta_{c1}, \mathscr{H}_1)$ yielding a rapid convergence towards a dirac-like distribution around 1, and a large D_{KL}^c value (cf the net separation in the two distributions of the third column in Fig 4). As a consequence, future candidates must have a very high correlation to be validated, giving a decision with a high +P, but with a low sensitivity in the case of sudden noise artifacts or even mild morphology change. In order to obtain a better trade-off between +P and sensitivity, we propose to limit the estimated α parameter of the beta distribution by imposing a conjugate prior law:

$$P(\alpha, \beta) \propto B(\alpha, \beta)^K e^{-a\alpha} e^{-b\beta}$$

for $K, a, b \in \mathbb{N}^+$. The $\exp^{-a\alpha}$ term indeed forbids large values in estimating α to control the distribution shape of $\mathscr{H}_1$. Numerical implementation is detailed in appendix A for the MAP estimator.

4.3 KLD weighting

In this section, we show the importance of the KLD measure in comparison with both Single Featured Probabilistic Detector (SFPD) and reference methods in Fig 5. SFPD is implemented as MFPD but with only one feature (*s*, *a* or *c*). Note that the evolution of KLD values during the successive (in)-validation process is a direct consequence of the parameter-fitting process since Eq (5) is a function of the updated pdf parameters (cf [27, 28], Appendix B). Indeed, the relative importance of the KLD of the *a*-feature during artifactual period suggests a better separability between the two antagonist distributions ($\mathscr{H}_0$ vs $\mathscr{H}_1$) and consequently highlights the decision from the *a*-feature in the centralized fusion (see Eq (6)). Interestingly, the KLD values tend to approach each other once the artifact period ends (near 50s), from which point all feature decisions participate more equally in the fusion. On the other hand, that the SFPD using the *s*-feature performs the worst (with multiple FP detections) is inline with its lowest KLD

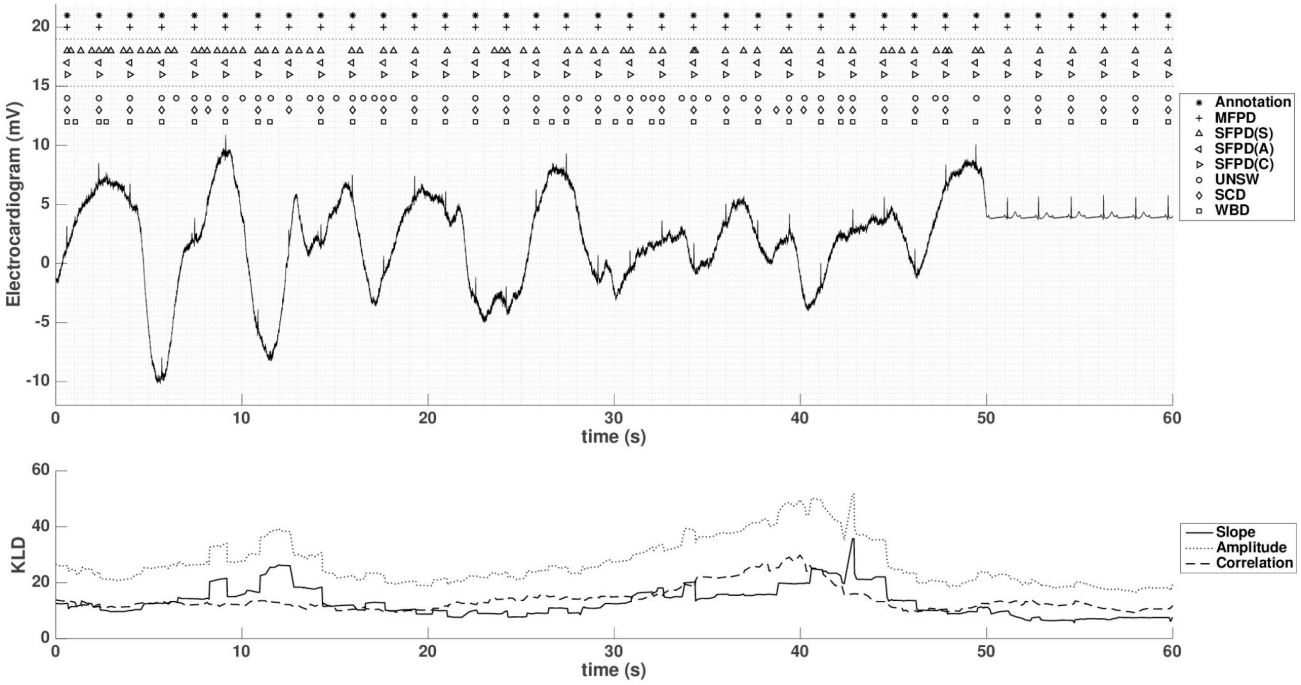


Fig 5. KLD variations for successive QRS candidates in 60 s from record MIT-231 with baseline noise and -6dB SNR. Upper panel: raw ECG signal with annotations (marked by *), MFPD fusion results (+), single feature (SFPD) results (△, ◁ and ▷), and results of the reference methods (○ for UNSW, ♢ for SCD and □ for WBD). Lower panel: KLD evolution for three features in the same period.

https://doi.org/10.1371/journal.pone.0223785.g005

measure, or the smallest distances between the related $\mathscr{H}_0$ and $\mathscr{H}_1$ pdfs. Therefore, depending on the noise onset, the KLD measure of the distribution distances evolve dynamically to quantify the pertinence of each feature and to adapt their relative weights in the decision fusion. The power of the proposed method lies on this engineered flexibility since KLD are not learned from a particular database, but from the past observations of the features labelled by the detection results.

Furthermore, the detection statistics of the whole record MIT-231 are given in Fig 5: the MFPD outperforms the SCD method in both sensitivity and +P, has lower sensitivity than UNSW and WBD but highest in +P (see Table 1), which is consistant with the illustration of the 60s segment detection results in Fig 5.

4.4 Performance evaluation

4.4.1 Benchmark MIT-BIH arrhythmia database. In Table 2, we first present the performance comparison on the MIT-BIH database, containing mainly clean ECG recordings, and high Se and +P results for all tested methods. We simply note that the MFPD has the lowest

Table 1. Performances of the whole record MIT-231 of Fig 5.

record MIT-231 with baseline noise (−6dB)						
	# TP	#FN	#FP	Se (%)	+P(%)	DER(%)
MFPD	1468	103	63	93.44	95.89	10.57
UNSW	1524	47	170	97.01	89.96	12.81
SCD	1038	533	74	66.07	93.34	38.63
WBD	1510	61	223	96.12	87.13	18.08

https://doi.org/10.1371/journal.pone.0223785.t001

Table 2. Performance comparison on the benchmark MIT-BIH arrhythmia database.

	Benchmark MIT-BIH DB						
	# TP	#FN	#FP	Se (%)	+P(%)	DER(%)	Jitter (ms)
MFPD	101799	7695	3195	93.16±11.46	96.81±9.03	9.86 ± 18.96	7.91±4.62
UNSW	103010	6484	6439	94.01±16.44	94.01±16.45	11.97 ± 32.86	7.04±4.77
SCD	104942	4552	4631	95.88±13.62	95.79±13.60	8.32 ± 27.20	5.46±3.69
WBD	103125	6369	5684	93.88±13.77	94.33±13.77	11.77 ± 27.45	3.08±3.33

Performances on the benchmark MIT-BIH arrhythmia database including: total number of TP, FN, FP, Se and +P (mean and standard deviation in %) and detection jitter (mean and std in ms) for 48 recordings.

https://doi.org/10.1371/journal.pone.0223785.t002

standard deviation for both Se and +P statistics, a phenomenon also observed in the next tests to illustrate its stability. On the whole, the results indicate a trade-off in favor of the +P as compared with the reference methods while the overall score DER also shows high detection accuracy in both mean and standard deviation terms. To illustrate the method's inter-subject stability, 2.08% of signals had a DER higher than 60% with MFPD versus 6.25% for UNSW, 2.08% for SCD and 8.33% for WBD. This test validates the MFPD on a clean ECG database with consensus annotations and peer methods.

4.4.2 MIT-BIH noise stress database. A second performance analysis was done using the MIT-BIH noise stress database. Fig 6 compares Se and +P (in %) for different noise types. Firstly, for +P the MFPD outperforms almost all reference methods except in the case of muscle artifact with SNR lower than 0dB. As for Se, the MFPD has slightly lower performances in

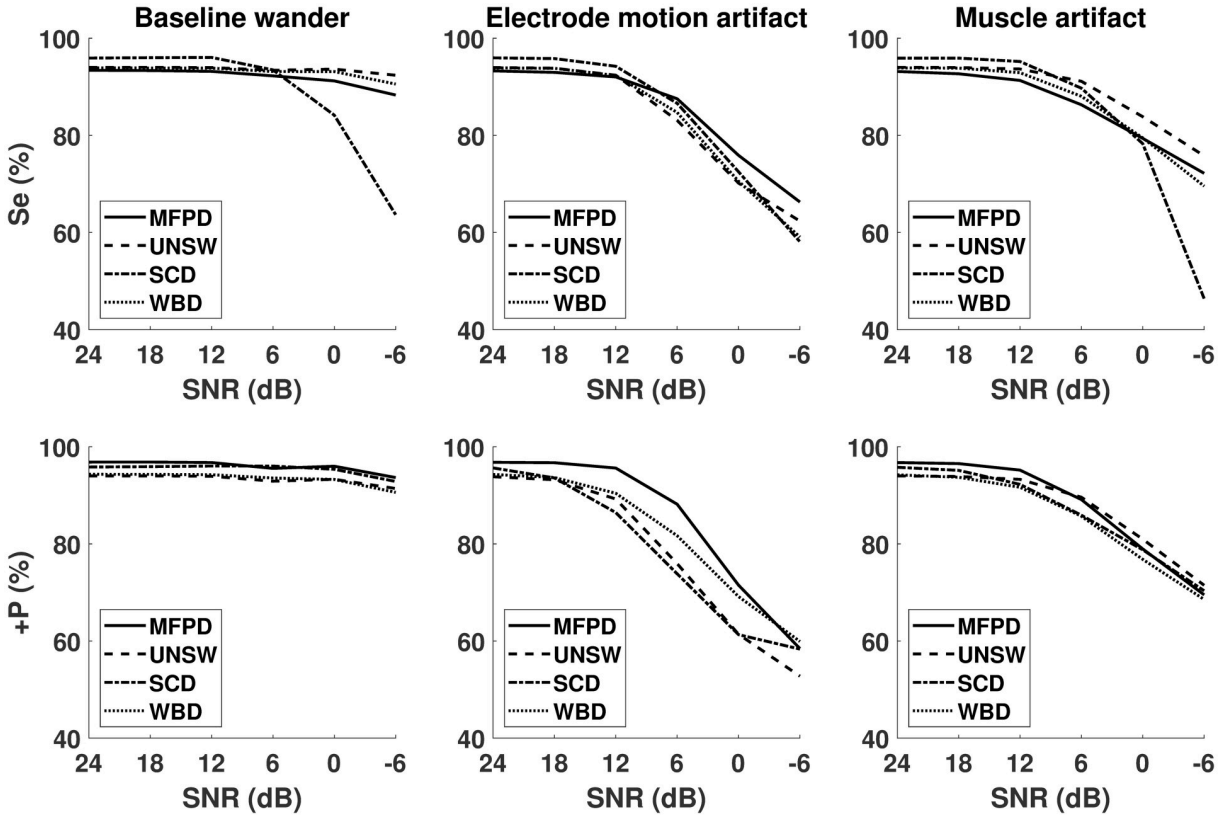


Fig 6. Se and +P (in %) of MFPD, UNSW, SCD and WBD on the benchmark noisy database, with different noise types and SNR levels.

https://doi.org/10.1371/journal.pone.0223785.g006

Table 3. Performance comparison on the benchmark noisy database.

Benchmark Noisy DB							
	# TP	#FN	#FP	Se (%)	+P (%)	DER (%)	Jitter (ms)
MFPD	1723	247	208	87.48±14.21	89.39±14.67	23.45 ± 28.46	8.15±4.24
UNSW	1743	227	287	88.30±17.66	86.06±19.27	26.82 ± 38.47	7.23±4.38
SCD	1679	291	257	85.21±18.63	86.64±16.71	28.43 ± 33.54	5.60±3.24
WBD	1724	246	250	87.25±16.41	86.68±16.67	26.38 ± 33.09	3.32±3.25

Performances on the benchmark noisy database including: total number of TP, FN, FP (in $\times 10^3$), Se and +P (mean and standard deviation in %) and detection jitter (mean and std in ms) for all 864 recordings.

https://doi.org/10.1371/journal.pone.0223785.t003

the high SNR region for all artifact types, but higher performances in the low SNR regions. In general the MFPD has better Se scores in comparison with the SCD and WBD method as reported by Table 3. The UNSW, on the other hand, has shown comparable performances in most cases, but suffers from FP detections in the presence of electrode motion artifacts. The MFPD also achieves the lowest overall DER score in terms of mean and standard deviation. To illustrate the method's inter-subject stability, 11.81% of signals had a DER higher than 60% with MFPD versus 16.78% for UNSW, 17.82% for SCD and 16.90% for WBD. These results demonstrate the MFPD viability and highlight its efficiency in noisy context.

Finally, we can notice that performances with baseline wander artifacts are generally higher than with other noise types for all tested methods, which can be attributed to the preprocessing filtering steps that attenuate baseline wanders.

4.5 Complexity and realtime detection delay

For the computation complexity, the MFPD algorithm implemented in the C++ language costs 19.32 ± 4.75 s to analyze an hour of ECG signal resampled at 1000 Hz using a MacBook Pro laptop (2017, 3,1 GHz Intel Core i7) without multi-threading or graphic parallel computing. The resampling at 1000 Hz was done because, according to us, this frequency is more representative of actual ECG acquisition devices and easier to interpret for readers.

In realtime monitoring applications, event detection delay is also part of the performance that can be used to trade-off for more accuracy for example. It is different from the jitter reported in Tables 2 and 3 that measures the time distance between annotation and detection. The delay reported in Table 4 measures the time distance between the annotation and the sample time at which the MFPD fusion decision validates the corresponding QRS and is typically influenced by lengths of the filters and the templates to calculate the correlation feature. Table 4 shows that this delay in the monitoring is rather stable across the two databases. These results confirm the feasibility of the MFPD in realtime monitoring of ECG signals, even for multi-lead recordings.

5 Discussion

In this paper, a novel, generic and robust detector combining different features extracted from the signal of interest has been proposed. The original aspects of this method concern

Table 4. MFPD detection delay (mean and std in ms).

Detection delay in ms	
MIT-BIH arrhythmia	MIT-BIH noise stress
612±84	648±78

https://doi.org/10.1371/journal.pone.0223785.t004

particularly i) the probabilistic approach with online learning ii) the multi-feature design iii) the centralized fusion method based on KLD.

In the proposed method, the pdf of each feature is patient, device and even experience specific. Parametric probability models are designed with regard to the real-time constraints of our application to avoid the tuning of the number of bins and widths as in histogram based approaches, and the increasing evaluation costs, inherent to variable-bandwith kernel density estimation approaches [35]. Our proposed online learning method requires a small data sample (40 validated candidates) to initialize the probabilistic model for each recording. No manual annotations are needed since the Pan-Tompkins detector [1] results are used for feature extraction and classification labeling ($\mathscr{H}_0/\mathscr{H}_1$). Prior laws of the Beta distribution for the correlation feature are fixed and not database dependent. Among the wide variety of pdf families, the GND seems well-suited for uncentered features with long tail distributions, but to our best knowledge, no analytic expression of the KLD can be found in the general case. Existing solutions [36] are limited to cases of equal means (see 2.2.4). We proposed in this paper an innovative estimation method of the KLD in the general case. With a reasonable computational cost, it can be used in real-time context.

The proposed MFPD QRS detector has been evaluated using two different databases and compared with two most recent state of the art detectors and one reference detector from the literature. Results show that the features provide complementary information to improve detection performance compared with single featured based QRS detectors (see Fig 5), particularly under noisy conditions. This is essentially due to the fact that different types of noise or uncommon pathological artifacts might influence features in different manners while MFPD makes a centralized decision using the KLD as weight to measure the relative pertinence of each feature. In a way, the MFPD is capable of neglecting features (that yield low KLD between the antagonist distributions) that are mostly corrupted by artefacts. Previously proposed methods based on decentralized fusion [2] and algorithm-switching [21] also prove the relevance of multi-feature approaches. To our knowledge, this is the first real-time QRS detection method integrating such an adaptive, multi-feature, centralized decision fusion. Furthermore, MFPD method is more compact and easy to implement than [2] or [21]. Quantitative comparisons results are particularly encouraging for challenging monitoring situations, in which the heterogeneity and levels of noise may be particularly high. For the two databases under evaluation, the ms-level jitter for all tested methods should be acceptable for further HRV parameter analysis and is thus not regarded as a distinguishing factor for comparison.

Even though this method was implemented for single-lead ECG signals, it can be extended to the multi-lead and multi-source cases. Indeed, further improvements in detection robustness are expected by combining multiple ECG leads, but also by integrating other physiological signals (pulse oximetry, phonocardiography, etc) or other sensors sensitive to noise (accelerometers for movement noise, etc). Future works will be directed towards the extension and evaluation of the method in these multi-channel, multi-source contexts. Finally, in addition to the qualitative results of the selected features' relevance shown in Fig 3 and in Fig 5 through KLD evolution, integrating the amplitude and derivative features as used by UNSW [15] into the MFPD architecture should lead to higher sensitivity in QRS detection.

6 Conclusion

We proposed an original multi-feature probabilistic detector working in real-time and applicable to different physiological signal applications. The method, illustrated on QRS complex detection, has been compared to two latest detectors in the literature, using the MIT-BIH arrhythmia benchmark database and a benchmark noisy database by adding noise recordings

[25] to the MIT-BIH database. The proposed MFPD has achieved significant performance improvements on artifacted signals and comparable (in mean) but more stable (in std) performances for low-to mildly artifacted signals. These performance improvements are mainly due to the multi-feature probabilistic model and the KLD-based decision fusion that adaptively adjust the relative contribution of each feature's decision in real-time.

The following contributions and originalities can be highlighted. The MFPD uses a probabilistic approach with online learning and an original adaptive, multi-feature, centralized decision fusion based on KLD. Its application on QRS detection achieved notable performance gains on artifacted ECG signals in comparison with one classical and two recent methods. The proposed method can also be easily extended to the multi-lead and multi-source cases, or to other physiological event detection applications. Besides the theoretical contribution and experimental validation, this new approach boasts a reasonable computational cost and thus can be embedded into low-power devices offering interesting possibilities in the current context of connected health applications.

A Application of the Karush-Kuhn-Tucker on the MAP of Beta distribution

A.1 About the KKT

Consider the non-linear optimization problem: maximize $f(x), \mathbb{R}^n \rightarrow \mathbb{R}$ (the cost function) subject to m inequality and l equality constraints:

$$g_i(x) \leq 0, i = 1, \ldots, m, \text{ and } h_j(x) = 0, j = 1, \ldots, l.$$

Suppose further that both $f(x)$ and the constraint functions $g_i(x)$, $h_j(x)$ are continuously differentiable at a point $\tilde{x}$. If $\tilde{x}$ is a local maximum of $f(x)$ satisfying some regularity conditions, then there exist the KKT multipliers: $\mu_i \in \mathbb{R}, i = 1, \ldots, m$ and $\lambda_j \in \mathbb{R}, j = 1, \ldots, l$ such that:

$$\nabla f(\tilde{x}) = \sum \mu_i \nabla g_i(\tilde{x}) + \sum \lambda_j \nabla h_j(\tilde{x}); \tag{8}$$

$$g_i(\tilde{x}) \leq 0, \ \mu_i \geq 0, \ \mu_i g_i(\tilde{x}) = 0 \text{ for all } i = 1, \ldots, m \tag{9}$$

$$h_j(\tilde{x}) = 0 \text{ for all } j = 1, \ldots, l \tag{10}$$

We note that in the particular case of $m = 0$, the KKT conditions are reduced to the Lagrange conditions. If g_i and h_j are affine functions (MAP of beta distribution), then no regularity condition is needed.

A.2 MAP for the Beta distribution

We derive here the numerical method to calculate the maximum-likelihood estimator (MLE) and MAP estimator given N independent samples of the Beta distribution. Recall the Beta density function:

$$P(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}_{0 \leq x \leq 1}$$

where $B(\alpha, \beta)$ is the normalization constant. For the MLE, the function to maximize is the joint log-likelihood function:

$$f(\alpha, \beta) = \sum_{i=1}^{N} \log P(x_i; \alpha, \beta) = (\alpha - 1)X + (\beta - 1)Y - N \log B(\alpha, \beta),$$

with ($X = \sum_{i=1}^{N} \log x_i$, $Y = \sum_{i=1}^{N} \log(1 - x_i)$). In a Bayesian setting typically to avoid the dirac-like shape of the beta distribution (see discussion in Sect. 4.2), a prior law can be added:

$$P(\alpha, \beta) \propto B(\alpha, \beta)^K e^{-a\alpha} e^{-b\beta} \text{ for } K < N, a, b \in \mathbb{R}^+.$$

The objective function of the MAP estimator becomes:

$$f^*(\alpha, \beta) = \alpha(X - a) + \beta(Y - b) - (N - K)\log B(\alpha, \beta) \tag{11}$$

with inequality constraints $g_1 = \epsilon - \alpha \leq 0$ and $g_2 = \epsilon - \beta \leq 0$. $\epsilon > 0$ is close to zero to form a closed space. Note that these constraints are affine functions and satisfy the regularity conditions. We thus search the solution of:

$$F_1(\boldsymbol{z}) = (X - a) - (N - K)(\psi(\tilde{\alpha}) - \psi(\tilde{\alpha} + \tilde{\beta})) + \tilde{\mu}_1 = 0,$$

$$F_2(\boldsymbol{z}) = (Y - b) - (N - K)(\psi(\tilde{\beta}) - \psi(\tilde{\alpha} + \tilde{\beta})) + \tilde{\mu}_2 = 0,$$

$$\epsilon - \tilde{\alpha} \leq 0,\ \tilde{\mu}_1 \geq 0,\ F_3(\boldsymbol{z}) = \tilde{\mu}_1(\epsilon - \tilde{\alpha}) = 0, \tag{12}$$

$$\epsilon - \tilde{\beta} \leq 0,\ \tilde{\mu}_2 \geq 0,\ F_4(\boldsymbol{z}) = \tilde{\mu}_2(\epsilon - \tilde{\beta}) = 0, \tag{13}$$

where $\psi(\cdot)$ represents the di-gamma function.

We apply the Newton Raphson method on $F(\boldsymbol{z}) = [F_1, F_2, F_3, F_4]^t$ with:

$$\boldsymbol{z}^{(n+1)} - \boldsymbol{z}^{(n)} = -\mathbf{J}(\boldsymbol{z}^{(n)})^{-1} F(\boldsymbol{z}^{(n)}), \tag{14}$$

and checking the inequality constraints in (12) and (13).

We next show that $\mathbf{J}(\boldsymbol{z})$ is always invertible given the inequality constraints. The $\mathbf{J}(\boldsymbol{z}) = [\frac{\partial F_i}{\partial z_j}]$ writes:

$$\mathbf{J}(\boldsymbol{z}) = \left[\begin{array}{cc:c} (N-K)\begin{bmatrix} \psi^{(1)}(\alpha+\beta) - \psi^{(1)}(\alpha) & \psi^{(1)}(\alpha+\beta) \\ \psi^{(1)}(\alpha+\beta) & \psi^{(1)}(\alpha+\beta) - \psi^{(1)}(\beta) \end{bmatrix} & & \mathbf{I}_2 \\ \hdashline \multicolumn{2}{c:}{\begin{matrix} -\mu_1 & 0 \\ 0 & -\mu_2 \end{matrix}} & \begin{matrix} -\alpha & 0 \\ 0 & -\beta \end{matrix} \end{array}\right]$$

where $\psi^{(1)}(\cdot)$ is the tri-gamma function (second derivative of the log-gamma). Its determinant is:

$$\begin{aligned} \det \mathbf{J}(\boldsymbol{z}) &= \det \left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right] = \det AD - BC = (N-K)^2 \alpha\beta \\ &\Big(\psi^{(1)}(\alpha)\psi^{(1)}(\beta) - (\psi^{(1)}(\alpha) + \psi^{(1)}(\beta))\psi^{(1)}(\alpha+\beta)\Big) \\ &+ (N-K)\mu_1\beta(\psi^{(1)}(\beta) - \psi^{(1)}(\alpha+\beta)) \\ &+ (N-K)\mu_2\alpha(\psi^{(1)}(\alpha) - \psi^{(1)}(\alpha+\beta)) \end{aligned}$$

The second equality is due to the fact that C and D commute (i.e. $CD = DC$). From the relation

$$\psi^{(1)}(\alpha)\psi^{(1)}(\beta) > (\psi^{(1)}(\alpha) + \psi^{(1)}(\beta))\psi^{(1)}(\alpha+\beta)$$

it can be verified that $\det \mathbf{J}(\boldsymbol{z}) > 0$. Thus the Jacobian matrix is always invertible.

B Kulback-Leiber divergence for generalized normal distributions

The probability density of the generalized normal distribution writes:

$$P(x;\alpha,\beta,\mu) = \frac{\beta}{2\alpha\Gamma(1/\beta)}\ e^{-(|x-\mu|/\alpha)^{\beta}} \text{ for } \alpha,\beta>0$$

Thus, the Kullback-Leibler divergence is:

$$D_{\mathrm{KL}}(P\|Q) = \int_{\mathbb{R}} \frac{\beta_p}{2\alpha_p\Gamma(1/\beta_p)}\ e^{-(|x-\mu_p|/\alpha_p)^{\beta_p}} \times \log\left(\frac{\frac{\beta_p}{2\alpha_p\Gamma(1/\beta_p)}\ e^{-(|x-\mu_p|/\alpha_p)^{\beta_p}}}{\frac{\beta_q}{2\alpha_q\Gamma(1/\beta_q)}\ e^{-(|x-\mu_q|/\alpha_q)^{\beta_q}}}\right)\mathrm{d}x$$

$$= \log\left(\frac{\beta_p\alpha_q\Gamma(1/\beta_q)}{\beta_q\alpha_p\Gamma(1/\beta_p)}\right) + \int_{\mathbb{R}} \frac{\beta_p}{2\Gamma(1/\beta_p)} e^{-(\frac{|x-\mu_p|}{\alpha_p})^{\beta_p}} \left(-(\frac{|x-\mu_p|}{\alpha_p})^{\beta_p} + (\frac{|x-\mu_q|}{\alpha_q})^{\beta_q}\right)\frac{\mathrm{d}x}{\alpha_p}.$$

Let $t = \frac{x-\mu_p}{\alpha_p} \Leftrightarrow \mathrm{d}x = \alpha_p \mathrm{d}t$. Since $\alpha_p > 0$, we have:

$$D_{\mathrm{KL}}(P\|Q) = \log\left(\frac{\beta_p\alpha_q\Gamma(1/\beta_q)}{\beta_q\alpha_p\Gamma(1/\beta_p)}\right) - \int_{\mathbb{R}} \beta_p/2\Gamma(1/\beta_p)\cdot e^{-|t|^{\beta_p}}\ |t|^{\beta_p}\mathrm{d}t$$

$$+ \underbrace{\int_{\mathbb{R}} \beta_p/2\Gamma(1/\beta_p)\cdot e^{-|t|^{\beta_p}}\ (|t\alpha_p+\mu_p-\mu_q|/\alpha_q)^{\beta_q}\mathrm{d}t}_{(*)}$$

Since $\Gamma(z+1) = z\,\Gamma(z)$, the second term can be further simplified:

$$\int_0^{+\infty} \frac{\beta_p}{\Gamma(1/\beta_p)}\ e^{-t^{\beta_p}}\ t^{\beta_p}\mathrm{d}t = \frac{\Gamma(1/\beta_p+1)}{\Gamma(1/\beta_p)} = \frac{1}{\beta_p}$$

In the following, we treat the term in (*). First define $\tilde{u} = \frac{|\mu_p-\mu_q|}{\alpha_p}$ to replace (*) with:

$$\left(\frac{\alpha_p}{\alpha_q}\right)^{\beta_q} \frac{\beta_p}{2\Gamma(1/\beta_p)}\int_{\mathbb{R}} e^{-|t|^{\beta_p}}|t+\tilde{u}|^{\beta_q}\mathrm{d}t = k_1\int_{\mathbb{R}} e^{-|t|^{\beta_p}}|t+\tilde{u}|^{\beta_q}\mathrm{d}t$$

$$\overset{t'=t/\tilde{u}}{=} k_1\int_{\mathbb{R}} e^{-(\tilde{u}t')^{\beta_p}}|\tilde{u}(t'+1)|^{\beta_q}\tilde{u}\mathrm{d}t' = k_1\tilde{u}^{\beta_q+1}\int_{\mathbb{R}} e^{-\tilde{u}^{\beta_p}|t|^{\beta_p}}|t+1|^{\beta_q}\mathrm{d}t$$

$$= k_2\int_{\mathbb{R}} e^{-\xi|t|^{\beta_p}}|t+1|^{\beta_q}\mathrm{d}t$$

with $k_1 = \left(\frac{\alpha_p}{\alpha_q}\right)^{\beta_q}\frac{\beta_p}{2\Gamma(1/\beta_p)}$, $\xi = \tilde{u}^{\beta_p}$ and $k_2 = k_1\ \tilde{u}^{\beta_q+1} > 0$.

Since both $|t|$ and $|t+1|$ exist in the expression, we further decompose (*) into:

$$k_2\int_0^1 e^{-\xi t^{\beta_p}}[(t+1)^{\beta_q} + (1-t)^{\beta_q}]\mathrm{d}t + k_2\int_1^{+\infty} e^{-\xi t^{\beta_p}}[(t+1)^{\beta_q} + (t-1)^{\beta_q}]\mathrm{d}t$$

Notice that when β_q is even, the two functions inside the integrals are identical since $(t-1)^{\beta_q} = (1-t)^{\beta_q}$. For instance, if $\beta_q = 2$:

$$k_2\int_0^1 e^{-\xi t^{\beta_p}}2[t^2+1]\,\mathrm{d}t + k_2\int_1^{+\infty} e^{-\xi t^{\beta_p}}2[t^2+1]\,\mathrm{d}t$$

$$= 2k_2\int_0^{\infty} e^{-\xi t^{\beta_p}}[t^2+1]\,\mathrm{d}t = \frac{2k_2}{\beta_p}\left(\frac{\Gamma(3/\beta_p)}{\xi^{3/\beta_p}} + \frac{\Gamma(1/\beta_p)}{\xi^{1/\beta_p}}\right),$$

by exploiting the relation:

$$\int_0^{+\infty} e^{-\xi t^{\beta_p}} t^k \mathrm{d}t = \Gamma\left(\frac{k+1}{\beta_p}\right) \Big/ \left(\beta_p \, \xi^{\frac{k+1}{\beta_p}}\right). \tag{15}$$

This result can be generalized for all even-valued β_q including $\beta_q = 0$ though it is supposed to be strictly positive by definition. Let's then investigate the case of $\beta_q = 2n + 1$ with $n \in \mathbb{N}$. First, for $\beta_q = 1$,

$$\begin{aligned}(*) = \;& k_2 \int_0^1 e^{-\xi t^{\beta_p}} (2)\mathrm{d}t + k_2 \int_1^{+\infty} e^{-\xi t^{\beta_p}} (2t)\mathrm{d}t = 2k_2 \left[\int_0^1 e^{-\xi t^{\beta_p}} \mathrm{d}t + \int_1^{+\infty} e^{-\xi t^{\beta_p}} t\mathrm{d}t\right] \\ \overset{y=\xi t^{\beta_p}}{=} \;& \frac{2k_2}{\beta_p} \left[(1/\xi)^{\frac{1}{\beta_p}} \int_0^{\xi} e^{-y} y^{\frac{1}{\beta_p}-1} \mathrm{d}y + (1/\xi)^{\frac{2}{\beta_p}} \int_{\xi}^{\infty} e^{-y} y^{\frac{2}{\beta_p}-1} \mathrm{d}y\right] \\ = \;& (2k_2/\beta_p)\left[(1/\xi)^{\frac{1}{\beta_p}} \gamma(1/\beta_p, \xi) + (1/\xi)^{\frac{2}{\beta_p}} \Gamma(2/\beta_p, \xi)\right]\end{aligned}$$

in which

$$\gamma(s, x) = \int_0^x t^{s-1} e^{-t} \, \mathrm{d}t, \quad \Gamma(s, x) = \int_x^{\infty} t^{s-1} e^{-t} \, \mathrm{d}t$$

are the lower and upper incomplete gamma functions respectively. Similarly, for $\beta_q = 3$, we develop the $(1 + t)^3$, $(1 - t)^3$ and $(t - 1)^3$ terms:

$$(*) = k_2 \left(\int_0^1 e^{-\xi t^{\beta_p}} (2 + 6t^2)\mathrm{d}t + \int_1^{+\infty} e^{-\xi t^{\beta_p}} (6t + 2t^3)\mathrm{d}t\right),$$

to coerce $(*)$ into a sum of lower and upper incomplete gamma functions. As in (15), we use the following relations:

$$\int_0^1 e^{-\xi t^{\beta_p}} t^k \mathrm{d}t \;\; = \gamma\left(\frac{k+1}{\beta_p}, \xi\right) \Big/ \left(\beta_p \xi^{\frac{k+1}{\beta_p}}\right), \tag{16}$$

$$\int_1^{\infty} e^{-\xi t^{\beta_p}} t^k \mathrm{d}t \;\; = \Gamma\left(\frac{k+1}{\beta_p}, \xi\right) \Big/ \left(\beta_p \xi^{\frac{k+1}{\beta_p}}\right). \tag{17}$$

To generalize, $(*)$ is a sum of either weighted gamma functions using (15) if β_q is even or a sum of weighted upper and lower incomplete gamma functions using (16) and (17) if β_q is odd. β_q-degree binomial coefficients are used to calculate the weights. Next we show that $(*)$ is monotonously increasing:

$$\frac{\partial(*)}{\partial \beta_q} = k_2 \int_{\mathbb{R}} e^{-\xi |t|^{\beta_p}} |1 + t|^{\beta_q} \, \log|1 + t| \mathrm{d}t > 0$$

in which k_2 and $e^{-\xi |t|^{\beta_p}} |1 + t|^{\beta_q}$ are positive for $t \in \mathbb{R}$. $\log|1 + t|$ is negative for $t \in [-2, 0]$, and positive otherwise. A sufficient condition is:

$$\int_{-2}^{2} e^{-\xi t^{\beta_p}} |1 + t|^{\beta_q} \, \log|1 + t| \; \mathrm{d}t > 0 \tag{18}$$

By splitting the integral into 2 parts and letting $y = -t$ in the first part, the integral of the above

inequality becomes

$$\int_0^2 e^{-\xi y^{\beta_p}} |1-y|^{\beta_q} \log|1-y| \mathrm{d}y + \int_0^2 e^{-\xi y^{\beta_p}} (1+y)^{\beta_q} \log(1+y) \mathrm{d}y$$

$$= \int_0^2 e^{-\xi y^{\beta_p}} \underbrace{\left(|1-y|^{\beta_q} \log|1-y| + (1+y)^{\beta_q} \log(1+y)\right)}_{G_{\beta_q}(y)} \mathrm{d}y$$

Note that the function $G_{\beta_q}(y)$ is a continuous function of y, differentiable by piece and $G_{\beta_q}(y) \geq 0$ for all $y \in [0, 2]$, $\beta_q > 0$, which validates the condition in (18).

C List of acronyms, abbreviations and symbols

a	Raw amplitude
c	Absolute correlation
$C(t)$	Candidate selected at instant t
$C_i(t)$	Feature i of the candidate selected at instant t
$D(t)$	Final detection decision at instant t
D_{KL}^i	Kullback-Leibler divergence of feature i
$\bar{D}_{\mathrm{KL}}^i$	Modified Kullback-Leibler divergence of feature i
$\mathcal{I}$	Set of complementary features
λ	Decision threshold
$P_i(C_i(t); \Theta_{i0}, \mathscr{H}_0)$	Probability density function of feature i corresponding to hypothesis $\mathscr{H}_0$ with Θ_{i0} parameter(s)
$P_i(\mathscr{H}_1 \mid C_i(t))$	Posterior probability of validating $\mathscr{H}_1$ for feature $C_i(t)$
s	Squared slope
Θ	Parameter(s) of the probability density function set for hypothesis $\mathscr{H}_0$
DER	Detection error rate
ECG	Electrocardiogram
FIR	Finite Impulse Response
FN	False Negative
FP	False Positive
GND	Generalized Normal Distribution
KLD	Kullback-Leibler Divergence
MAP	Maximum A Posterior
MFPD	Multi-Feature Probabilistic Detector
MLE	Maximum Likelihood Estimator
pdf	Probability Density Function
+P	Positive predictivity
SAecg	Signal use to extract amplitude feature
SCD	Spatiotemporal Characteristics Detector
Se	Sensitivity
SFecg	Signal use to extract slope feature
SFPD	Single Featured Probabilistic Detector
SNR	Signal-to-Noise Ratio
TN	True Negative
TP	True Positive
UNSW	University of New South Wales detector
WBD	Wavelet-Based Detector

Acknowledgments

This work was mainly funded by the French Ministry of Higher Education, Research and Innovation (MESRI). It also received support from the European Union's Horizon 2020 research and innovation program under grant agreement no. 689260 (Digi-NewB project).

Author Contributions

Conceptualization: Matthieu Doyen, Di Ge, Alfredo I. Hernández.

Formal analysis: Matthieu Doyen, Di Ge, Alfredo I. Hernández.

Funding acquisition: Guy Carrault, Alfredo I. Hernández.

Investigation: Di Ge.

Methodology: Matthieu Doyen, Di Ge, Guy Carrault, Alfredo I. Hernández.

Project administration: Di Ge, Guy Carrault, Alfredo I. Hernández.

Software: Matthieu Doyen, Alfredo I. Hernández.

Supervision: Di Ge, Guy Carrault, Alfredo I. Hernández.

Validation: Matthieu Doyen, Di Ge, Alain Beuchée, Alfredo I. Hernández.

Visualization: Matthieu Doyen, Alain Beuchée.

Writing – original draft: Matthieu Doyen, Di Ge, Guy Carrault, Alfredo I. Hernández.

Writing – review & editing: Matthieu Doyen, Di Ge, Guy Carrault, Alfredo I. Hernández.

References

1. Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm. IEEE Transactions on Biomedical Engineering. 1985; BME-32(3):230–236. https://doi.org/10.1109/TBME.1985.325532
2. Hernández AI, Carrault G, Mora F, Thoraval L, Passariello G, Schleich JM. Multisensor fusion for atrial and ventricular activity detection in coronary care monitoring. IEEE transactions on bio-medical engineering. 1999; 46(10):1186–1190. https://doi.org/10.1109/10.790494 PMID: 10513122
3. Hansen B. Robust Detection of Heart Beats in Multimodal Data: the PhysioNet/ CINC 2014. 2014.
4. Kohler B, Hennig C, Orglmeister R. The principles of software QRS detection. IEEE Engineering in Medicine and Biology Magazine. 2002; 21(1):42–57. https://doi.org/10.1109/51.993193 PMID: 11935987
5. Jain S, Ahirwal MK, Kumar A, Bajaj V, Singh GK. QRS detection using adaptive filters: A comparative study. ISA Transactions. 2017; 66:362—375. https://doi.org/10.1016/j.isatra.2016.09.023 PMID: 27745689
6. Osowski S, Tran L. ECG Beat Recognition Using Fuzzy Hybrid Neural Network. vol. 48; 2001.
7. Zidelmal Z, Amirou A, Ould-Abdeslam D, Moukadem A, Dieterlen A. QRS detection using S-Transform and Shannon energy. Computer Methods and Programs in Biomedicine. 2014; 116(1):1—9. https://doi.org/10.1016/j.cmpb.2014.04.008 PMID: 24856322
8. Li C, Zheng C, Tai C. Detection of ECG characteristic points using wavelet transform. vol. 42; 1995.
9. Benitez D, Gaydecki PA, Zaidi A, Fitzpatrick AP. The use of the Hilbert transform in ECG signal analysis. Computers in Biology and Medicine. 2001; 31(5):399—406. https://doi.org/10.1016/s0010-4825(01)00009-9 PMID: 11535204
10. Zidelmal Z, Amirou A, Adnane M, Belouchrani A. QRS detection based on wavelet coefficients. Computer Methods and Programs in Biomedicine. 2012; 107(3):490—496. https://doi.org/10.1016/j.cmpb.2011.12.004 PMID: 22296976
11. Rani R, VS C, HP S. Automated Detection of QRS Complex in ECG Signal using Wavelet Transform. International Journal of Computer Science and Network. 2015; 15.
12. Peltola MA. Role of editing of R-R intervals in the analysis of heart rate variability. Frontiers in physiology. 2012; 3:148–148. https://doi.org/10.3389/fphys.2012.00148 PMID: 22654764

13. Friesen GM, Jannett TC, Jadallah MA, Yates SL, Quint SR, Nagle HT. A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Transactions on Biomedical Engineering. 1990; 37(1):85–98. https://doi.org/10.1109/10.43620 PMID: 2303275

14. Elgendi M, Eskofier B, Dokos S, Abbott D. Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems. PloS one. 2014; 9(1):e84018–e84018. https://doi.org/10.1371/journal.pone.0084018 PMID: 24409290

15. Khamis H, Weiss R, Xie Y, Chang CW, Lovell NH, Redmond SJ. QRS Detection Algorithm for Telehealth Electrocardiogram Recordings. IEEE Transactions on Biomedical Engineering. 2016; 63:1377–1388. https://doi.org/10.1109/TBME.2016.2549060 PMID: 27046889

16. Elgendi M, Mohamed A, Ward RK. Efficient ECG Compression and QRS Detection for E-Health Applications. In: Scientific Reports; 2017. https://doi.org/10.1038/s41598-017-17101-x

17. Kim J, Shin H. Simple and Robust Realtime QRS Detection Algorithm Based on Spatiotemporal Characteristic of the QRS Complex. PloS one. 2016; 11 3:e0150144. https://doi.org/10.1371/journal.pone.0150144 PMID: 26943949

18. Baig MM, Gholamhosseini H, Connolly MJ. A comprehensive survey of wearable and wireless ECG monitoring systems for older adults. Medical & Biological Engineering & Computing. 2013; 51(5):485–495. https://doi.org/10.1007/s11517-012-1021-6

19. Deepu CJ, Xu X, Zou X, Yao L, Lian Y. An ECG-on-Chip for Wearable Cardiac Monitoring Devices. In: 2010 Fifth IEEE International Symposium on Electronic Design, Test Applications; 2010. p. 225–228.

20. Goldenholz DM, Kuhn A, Austermuehle A, Bachler M, Mayer CC, Wassertheurer S, et al. Long-term monitoring of cardiorespiratory patterns in drug-resistant epilepsy. Epilepsia. 2017; 58 1:77–84. https://doi.org/10.1111/epi.13606 PMID: 27864903

21. Portet F, Hernandez A, Carrault G. G.: Evaluation of real-time QRS detection algorithms in variable contexts. Medical & biological engineering & computing. 2005; 43:379–85. https://doi.org/10.1007/BF02345816

22. Dumont J, Hernandez A, Carrault G. Improving ECG Beats Delineation With an Evolutionary Optimization Process. Biomedical Engineering, IEEE Transactions on. 2010; 57:607—615. https://doi.org/10.1109/TBME.2008.2002157

23. Altuve M, Carrault G, Cruz J, Beuchae A, Pladys P, Hernandez A. Analysis of the QRS Complex for Apnea-Bradycardia Characterization in Preterm Infants. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference. 2009; 2009:946–9.

24. Hernando D, Bailón R, Almeida R, Hernández AI. QRS detection optimization in stress test recordings using evolutionary algorithms. Computing in Cardiology 2014. 2014; p. 737–740.

25. Moody GB, Muldrow WK, Mark RG. NOISE STRESS TEST FOR ARRHYTHMIA DETECTORS. Computers in Cardiology. 1984; 11:381–384.

26. Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using MIT/BIH Arrhythmia database. Biomedical Engineering, IEEE Transactions on. 1987; 33:1157—1165.

27. Gil M, Alajaji F, Linder T. Rényi divergence measures for commonly used univariate continuous distributions. Information Sciences. 2013; 249:124–131. https://doi.org/10.1016/j.ins.2013.06.018

28. Mathiassen JR, Skavhaug A, Bø K. Texture Similarity Measure Using Kullback-Leibler Divergence between Gamma Distributions. In: Heyden A, Sparr G, Nielsen M, Johansen P, editors. Computer Vision—ECCV 2002. Springer Berlin Heidelberg; 2002. p. 133–147.

29. Elgendi M. Fast QRS Detection with an Optimized Knowledge-Based Method: Evaluation on 11 Standard ECG Databases. PLOS ONE. 2013; 8(9):1–18. https://doi.org/10.1371/journal.pone.0073557

30. Kim J, Shin H. Simple and Robust Realtime QRS Detection Algorithm Based on Spatiotemporal Characteristic of the QRS Complex. PLOS ONE. 2016; 11(3):1–13.

31. Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A Wavelet-Based ECG Delineator: Evaluation on Standard Databases. IEEE transactions on bio-medical engineering. 2004; 51:570–81. https://doi.org/10.1109/TBME.2003.821031 PMID: 15072211

32. Llamedo M. ecg-kit a Matlab Toolbox for Cardiovascular Signal Processing. Journal of Open Research Software. 2016; 4:e8. https://doi.org/10.5334/jors.86

33. Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation. 2000; 101:E215–20. https://doi.org/10.1161/01.cir.101.23.e215 PMID: 10851218

34. Liu F, Liu C, Zhao L, Jiang X, Zhang Z, Li J, et al. Dynamic ECG Signal Quality Evaluation Based on the Generalized bSQI Index. IEEE Access. 2018; 6:41892–41902. https://doi.org/10.1109/ACCESS.2018.2860056

35. Hansen B. (2004, May). University of Wisconsin. Madison, US. Bandwidth selection for nonparametric distribution estimation [Online]. Available: http://www.ssc.wisc.edu/bhansen/papers/smooth.pdf

36. Do M, Vetterli M. Wavelet-based Texture Retrieval using Generalized Gaussian Density and Kullback-Leibler Distance. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society. 2002; 11:146–58. https://doi.org/10.1109/83.982822