1. Motivation and significance

As the amount of recorded data increases in all fields, automatic processing is inevitably needed. This is the case for health data, too. Especially the electrocardiogram (ECG) as a cheap and available standard heart activity monitoring device [2] is used in medicine around the world. Automatic ECG processing can be applied to batch process many ECGs in a short amount of time. In this way, manual annotation work of physicians can be avoided. Furthermore, retrospective analysis especially plays an important role in clinical research to reveal effects of new drugs on the
heart, optimize diagnosis and treatment strategies of cardiac diseases [3–8]. The automatic analysis of the ECG usually starts with determining wave types (P waves, T waves and QRS complexes), followed by finding the peaks and boundaries, i.e., wave onsets and offsets. Starting from that, further temporal parameters, like QT intervals, RR intervals, etc. can be derived. As well, these annotations are needed to calculate further features like amplitudes, slopes, and parameters like heart rate variability [9].

In this work, we present a toolbox for MATLAB comprising algorithms for ECG pre-processing and ECG wave delineation for research use, freely available on github. We call this open source toolbox ECGdeli. The aim of this work is not to present new approaches for wave delineation as the algorithms or slightly modified versions were already used in several studies [11–15]. Rather, this work should be seen as accompanying documentation of the open software resource. We will describe shortly the main ideas behind the algorithms and compare them to two existing open implementations on PhysioNet: ecg-kit [16] and ecgpuwave [17]. The algorithms for comparison, especially ecgkit, offer more options for evaluation and data import/export,
however, at the cost of a more complicated non-modular framework or perform worse. The advantages of ECGdeli in comparison with the existing implementations are a simple and modular design concentrating on the essential algorithms needed for ECG evaluation yielding an easy to use software; ECGdeli focuses on one concrete purpose: ECG wave detection and delineation. File input/output, post processing like feature calculation and a graphical user interface are intentionally not included. Thus, users can easily integrate the single algorithms or the whole toolbox in their evaluation frameworks, replace single algorithms, all by using few clear and standardized interfaces.

In the following, we present a short description of the single algorithms provided with the toolbox and discuss the main advantages of ECGdeli.

2. Software description

2.1. ECG preprocessing

The ECG preprocessing functionalities are shown in Fig. 5, lines one to four and include a baseline removal technique [15] ECG_Baseline Removal.m, a bandpass filter.
Fig. 3. A: Flowchart of the T wave detection algorithm. B: Flow chart of the P wave detection algorithm. WT: wavelet transform, AUC: area under the curve, min: minimum, max: maximum, fid: fiducial.

ECG_High_Low_Filter.m and a notch filter Notch_Filter.m, as well as a method for correcting the electrical isoline Iso-line_Correction.m. All preprocessing steps are performed lead by lead.

2.2. QRS detection

In the QRS detection algorithm QRS_detection.m (an overview is shown in Fig. 1), the input signal is first bandpass filtered (5 Hz to 100 Hz) in order to reduce high frequency noise, baseline wander, to attenuate P and T waves and to obtain QRS complexes with pronounced Q and S waves. Then, a phase-free stationary wavelet transformation (SWT) using the Haar wavelet is applied [12]. The detail coefficients $SWD_n$ of the SWT are calculated at the level including frequencies around 45 Hz, which is the frequency range where spectral components of the QRS complex being relevant for delineation can be found [18].

To detect intervals containing the QRS complexes, a threshold based approach is chosen. This threshold is adapted dependent on a changing time windows which is applied to $SWD_n$ and a signal dependent statistical measure. Among all sets of solutions calculated with different thresholds, a voting algorithm yields the most consistent solution. The R peak is annotated at the position with the highest amplitude within the bounds of the QRS interval. The Q- and S peak are subsequently marked at the minimum amplitude in the regions from QRS onset to R peak and from R peak to QRS offset, respectively. The onset and the offset of the QRS complex are afterwards marked 20 ms prior the Q-peak and 20 ms after the S-peak, respectively. In the end, all detected peaks and boundaries are stored in a fiducial point table (FPT) in columns five to eight. The structure is shown in Table 1.
2.3. T wave detection

The implemented T wave detection method T_Detection.m relies on QRS detection and is performed for each lead separately. Fig. 3A shows an overview of the algorithm. To detect the T waves, the following pre-processing steps are needed: QRS replacement, filtering and SWT. First, a fixed part of each RR interval is replaced by a sigmoid function, such that QRS complexes and P waves are blanked (see Fig. 2A, second line). Subsequently, phase-free Butterworth band-pass filtering (passband from 0.3 Hz to 20Hz) of order 4 is performed. This QRS-free signal is fed into a phase-free SWT [12] with the bior3.3 wavelet yielding a wavelet signal hardly showing activity in the QRS segment (see Fig. 2A, third line). The level $N$ of the detail coefficients to be analyzed is determined based on the central frequency of the T wave (7 Hz) [18].

After preparing the signal, the polarity of the T wave is determined considering all positive and negative extrema in the SWT signal. In a search window dependent on the RR interval and QRS offset, we determine the peak of the T wave ($T_{peak}$) in each RR interval by choosing the time instant of maximum amplitude (for positive polarity) or of the minimum amplitude (for negative polarity). According to the polarity, a correction of this peak is made in the ECG signal (time domain). Afterwards, T wave boundaries are determined. Around every T peak in each RR interval, we calculate the area under the curve of the SWT at level $N-1$ within a search window dependent on the respective RR interval. With two preset thresholds for the area distribution and the positions of the peaks in the interval, we estimate the position of T wave onset ($T_{on}$) and T wave offset ($T_{off}$). Finally, all detected peaks and boundaries are stored in the FPT (columns ten to twelve as shown in Table 1).

2.4. P wave detection

P wave detection P_Detection.m also relies on the phase-free SWT as well as on QRS detection and T wave detection. Fig. 3B provides an overview of the algorithm which was presented in [12]. The single lead input signal is band-pass filtered with a Gaussian filter (passband 1Hz to 15Hz). QRS complexes and T waves are replaced by sigmoid functions (see Fig. 2B, second line). The SWT is performed with the quadratic spline wavelet at the detail coefficients of level $N$ corresponding to the P wave center frequency of 7 Hz [18] (see Fig. 2B, third line).

The peaks of the SWT in each RR interval are determined by annotating the absolute maximum of the SWT as P wave peak in a search window dependent on each RR interval and the Q peak location. The user can also choose to set a nearby extremum in the filtered ECG signal as peak.

After peak detection, P wave boundaries are determined. Since P waves are low in amplitude, a template of the SWT in a window of 50 waves around the current P wave peak is built to estimate the boundaries. Hotelling’s T-squared method is applied to obtain a clean template. P wave onset and offset are estimated by initial guesses from the template considering the nearest wavelet extremum as the respective wave boundary.

The found peaks and wave boundaries are stored in the FPT, columns one to three as shown in Table 1.

2.5. Multilead processing

As already mentioned, P wave, QRS complex, and T wave detection are generally performed lead by lead. To achieve a multilead delineation, a voting algorithm (Sync_Beats.m) discards possible beats in the FPT if they are not visible in the majority of all leads and averages over the found positions to deliver a overall delineation result. Innately, a reasonable result can only be achieved with more than two leads.

2.6. Software architecture and functionality

A high-level view on the pipeline for the ECG preprocessing and annotation is shown in Fig. 4. First, baseline removal, filtering and isoline correction are performed (see Fig. 5, lines one to four). Following that, the signal can be annotated by launching the script Annotate_ECG_Multi.m. Here, QRS detection, T and P wave detection are run (see Fig. 5, line five). Afterwards, in the case of a multilead signal, the voting algorithm (Sync_Beats.m) would discard possible beats in the FPT if they are not visible in the majority of all leads.
3. Performance evaluation

The algorithms provided in the toolbox were evaluated with the QT database, an ECG database comprising 105 signals that were either recorded during normal sinus rhythm or represent one of six selected cardiac diseases with fiducial points annotated by clinicians [10]. We extracted annotations from the *.q1c and *.q2c files of the database. To obtain the detection errors, we analyzed both of the two leads separately. For each annotation, we subsequently took the best result among the two leads meaning choosing the annotation being closest to the annotation. To get the closest annotation, every annotation available in one beat was considered finding the respective annotation beat. We carried out the exact same evaluation for ECGdeli, ecg-kit [16] available on github (version 1.0, Commit c8e3de4) and the ecgp-wave [17] algorithm as provided in PhysioNet. With ecg-kit, we used the wavedetect delineation algorithm which is the default selection for ECG delineation. Detection errors for all types of annotations for all algorithms were calculated, including median errors and the interquartile ranges, mean errors and standard deviations of the signed and unsigned (absolute) errors.

3.1. Results

Fig. 6 shows the differences between the calculated fiducial points with the ECGdeli and the manually annotated points by clinicians for all annotated beats in the QT database. Especially with the detection of QRS and T wave annotations, outliers were
Clinicians clearly found and annotated the respective wave in the event, although they should ideally be detected due to the fact that last lines in Table 2). ecg-kit and ecgpuwave discarded waves. This number was highest in the case of ECGdeli (see ranges, ecgpuwave was always outperformed by ecg-kit. Performing the other. Regarding only median and interquartile performance measure and the annotation type, one was out-
visible. Detection errors for all three algorithms are given in Table 2. ECGdeli and ecg-kit performed comparably, dependent on the performance measure and the annotation type, one was out-performing the other. Regarding only median and interquartile ranges, ecgpuwave was always outperformed by ecg-kit. A particular difference was visible in the number of detected waveforms. This number was highest in the case of ECGdeli (see last lines in Table 2). ecg-kit and ecgpuwave discarded waves even though they should ideally be detected due to the fact that clinicians clearly found and annotated the respective wave in the signal. ECGdeli did not so.

3.2. Discussion

In this section, three main points connected with the results from Section 3.1 will be discussed.

First, we decided to evaluate the two leads separately and then take the lowest error per annotation as the annotator had also both leads at hands during annotation and we did not know which lead was chosen. We therefore did not evaluate a possible improvement by using a multilead approach which would have been hardly possible for ECGdeli since a voting between two different annotations does not offer an advantage.

Second, we want to refer to further publications comparing more closed source implementations [19,20]. We intentionally did not repeat the results of the closed source implementations here for two main reasons: first, we wanted to compare our results to the two most visible algorithms for ECG delineation offering P wave, T wave and QRS detection. Second, as there is no standard procedure to generate the final evaluation results, it is hardly possible to guarantee the comparability.

Third, we want to highlight the possibility of a postprocessing step for ECGdeli and a possible improvement of the average detection errors. With the current implementation, and as already stated, the algorithm is forced to detect a P and T wave in each RR interval. On the one hand, this implies that every wave is detected. Nevertheless, an adequate detection of the outliers (visible e.g. for the T wave in Fig. 6) and a subsequent correction or dropping of those, could have lowered the average errors of ECGdeli.

4. Impact

ECG delineation algorithms are important in clinical and research practice. As stated in Section 1, many parameters depend on the result of the wave detection. With ECGdeli we offer one of the few open toolboxes to solve this relevant problem. By comparing the toolbox with two alternative implementations for wave delineation, we showed that ECGdeli is already at its current state delivering results en par with existing open approaches. However, several partly unique features of ECGdeli should be highlighted:

The simple input/output and modular structure make the toolbox functions easy and intuitive to use. In this way, the annotation functions can also be executed separately if for example only a P wave detection is necessary. This goes hand in hand with the fact that by making ECGdeli freely available under the GPLv3, single algorithms can be extracted and incorporated into existing projects to extend them.

Furthermore, others can easily apply ECGdeli as a benchmark for new algorithms as we did with ecg-kit and ecgpuwave. Usually, the evaluation of the algorithms with standard databases,
like the QT database, allows to compare new work to already existing. However, manual annotations in these databases can be prone to error and different ways of calculating performance parameters impair comparability (as discussed in Section 3). Moreover, there might be pathologies not represented in a freely available database that can however be relevant for the intended application of a new algorithm. The best way to compare the own work to existing algorithms is to have benchmark algorithms available.

At last, ECGdeli is capable of detecting all P and T waves in a provided input ECG signal if the respective QRS complex was detected. Thus, for example a P wave gets detected in any case between two subsequent R peaks. In this way, no false negative detected waves are possible and false positive waves can still be eliminated in a postprocessing step that is designed for the specific problem.

5. Conclusions

In this work, we present the ECGdeli MATLAB toolbox for ECG wave delineation. The modularity is the major advantage as single algorithms can be used and adapted independently from the others (e.g., use another QRS detector as long as the standardized FPT is used). By publishing ECGdeli under GPLv3, the toolbox can be freely used for wave detection, serve as a benchmark method for future works on non-standard datasets or as a basis for extended feature extraction algorithms.

CRediT authorship contribution statement

N. Pilia: Methodology, Software, Conceptualization, Investigation, Validation, Visualization, Writing - original draft, Writing - review & editing. C. Nagel: Methodology, Conceptualization, Validation, Visualization, Writing - original draft, Writing - review & editing. G. Lenis: Methodology, Software, Conceptualization, Investigation, Validation, Writing - original draft. O. Dössel: Funding acquisition, Project administration, Supervision, Conceptualization. A. Loewe: Funding acquisition, Project administration, Data curation, Supervision, Conceptualization, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank several colleagues, collaborators and students at our institute for using the software, reporting bugs and proposing improvements. Furthermore, we acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology, Germany.

We gratefully acknowledge financial support by the European Metrology Programme for Innovation and Research, Germany [18HLT07_BiomedRay]. We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology, Germany.

We also gratefully acknowledge support by Deutsche Forschungsgemeinschaft (DFG), Germany [project ID 391128822, LO2093/1-1].

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j sofxt.2020.100639.

References

[1] Pilia N, Nagel C. KIT-IB (ecgdeli): Ecgdeli v1.0.2. Zenodo; 2020, http://dx.doi.org/10.5281/zenodo.3979791.

[2] Fye W. A History of the origin, evolution, and impact of electrocardiography. Amer J Cardiol 1994;73(13):937–49. http://dx.doi.org/10.1016/0002-9149(94)00135-X.

[3] Lyon A, Minchölé A, Martínez JP, Laguna P, Rodriguez B. Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. J R Soc Interface 2018;15(138). http://dx.doi.org/10.1098/rsif.2017.0821.

[4] Majumder S, Chen L, Marinov O, Chen C-H, Mondal T, Deen MJ. Noncontact wearable wireless ECG systems for long-term monitoring. IEEE Rev Biomed Eng 2018;11:306–21. http://dx.doi.org/10.1109/RBME.2018.2840336.

[5] Nielsen JB, Kuhl JT, Pietersen A, Graff C, Lind B, Struijk JJ, et al. P-wave duration and the risk of atrial fibrillation: Results from the Copenhagen ECG study. Heart Rhythm ; Off J Heart Rhythm Soc 2015;12(9):1887–95. http://dx.doi.org/10.1016/j.hrthm.2015.04.026.

[6] Loewe A, Krueger MW, Holmqvist F, Dössel O, Seemann G, Platovon PG. Influence of the earliest right atrial activation site and its proximity to interatrial connections on P-wave morphology. Europace 2016;18(Suppl. 4):iv35–43. http://dx.doi.org/10.1093/europace/euw348.

[7] Andlauer R, Seemann G, Baron L, Dössel O, Kohl P, Platovon P, et al. Influence of left atrial size on P-wave morphology: differential effects of dilation and hypertrophy. Europace 2018;20(5):i33–44. http://dx.doi.org/10.1093/europace/euy231.

[8] Jahmounah V, Oh SL, Wei JKE, Ciaccio EJ, Chua K, San TR, et al. Computer-aided diagnosis of congestive heart failure using ECG signals - A review. Phys Med Off J Ital Assoc Biomed Phys 2019;62:95–104. http://dx.doi.org/10.1016/j.lejmip.2019.05.004.

[9] Clifford GD, Auzaje F, McSharry P. Advanced methods and tools for ECG data analysis. Boston, Mass.[u.a.]: Artect House; 2006.

[10] Laguna P, Mark RG, Goldberg A, Moody GB. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In: Computers in cardiology 1997; 1997. p. 673–76.

[11] Lenis G, Dössel O. T wave morphology during heart rate turbulence in patients with chronic heart failure. Biomed Tech Biomed Eng 2013;58(1). http://dx.doi.org/10.1515/bmt-2013-4177.

[12] Lenis G, Pilia N, Oesterlein T, Luik A, Schmitt C, Dössel O. P wave detection and delineation in the ECG based on the phase free stationary wavelet transform and using intracardiac atrial electrograms as reference. Biomed Tech Biomed Eng 2016;61(1):37–56. http://dx.doi.org/10.1515/bmt-2014-0161.

[13] Lenis G, Reichensperger P, Sommer D, Heinze C, Colz M, Dössel O. Detection of microsleep events in a car driving simulation study using electrocardiographic features. Curr Direct Biomed Eng 2016;2(1):283–7.

[14] Lenis G, Kircher M, Lazoro J, Bailon R, Gil E, Dössel O. Separating the effect of respiration on the heart rate variability using Granger’s causality and linear filtering. Biomed Signal Process Control 2017;31:272–8. http://dx.doi.org/10.1016/j.bspc.2017.05.014.

[15] Lenis G, Pilia N, Loewe A, Schulze WHW, Dössel O. Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: A simulation study. Comput Math Methods Med 2017;2017:13. http://dx.doi.org/10.1155/2017/9295029, (Article ID 9295029).

[16] Demski A, Lamedo Soria M. Ecg-kit a Matlab toolbox for cardiovascular signal processing. J Open Res Softw 2016;4. http://dx.doi.org/10.5334/jors. 90.

[17] Laguna P, Jané R, Caminal P. Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE database. Comput Biomed Eng 1994;27(1):45–60.

[18] Sörnmo L, Laguna P. Bioelectrical signal processing in cardiac and neurological applications. Amsterdam: Elsevier; 2005.

[19] Di Marco LY, Chiari L. A wavelet-based ECG delineation algorithm for extended feature extraction algorithms. In: Computers in cardiology 1997; 1997. p. 673–76.

[20] Spicher N, Kukuk M. Delineation of electrocardiograms using multiscale parameter estimation. IEEE J Biomed Health Inf 2020;24(8):2216–29. http://dx.doi.org/10.1109/JBHI.2019.2963786.