Database Open Access

Brno University of Technology ECG Quality Database (BUT QDB)

Andrea Nemcova, Radovan Smisek, Kamila Opravilová, Martin Vitek, Lukas Smital, Lucie Maršánová

Published: July 2 020. Version: 1.0.0

When using this resource, please cite: (show more options)
Nemcova, A., Smisek, R., Opravilová, K., Vitek, M., Smital, L., & Maršánová, L. (2020). Brno University of Technology ECG Quality Database (BUT QDB) (version 1.0.0). *PhysioNet*. https://doi.org/10.13026/kah4-0w24.

Please include the standard citation for PhysioNet: (show more options)
Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.

Abstract

BUT QDB is a database created by the cardiology team at the Department of Biomedical Engineering, Brno University of Technology, for the purpose of evaluating ECG quality. The data comprises 18 long-term recordings of single-lead ECGs and associated 3-axis accelerometer data, collected from 15 subjects (9 female, 6 male) aged between 21 to 83 years. Recordings were carried out between August 2018 and October 2019 while the subjects were undertaking ordinary everyday activities ("free-living conditions"). Data was collected using a mobile ECG and accelerometer (Bittium Faros 180) with a sampling frequency of 1,000 Hz for ECG signals and 100 Hz for accelerometer signals. The minimum length of recording is 24 hours.

Three signals were fully annotated in terms of ECG signal quality. The remaining 15 signals were annotated in two selected segments, each of 20 minutes in duration. Furthermore, five additional segments of poor signal quality were also annotated. Signal quality was classified as follows: Class 1 indicates that all ECG significant waveforms (P wave, T wave and QRS complex) are clearly visible and their onsets and offsets can be detected reliably; Class 2 indicates that the noise level is increased and ECG significant points cannot be reliably detected, but the signal enables reliable QRS detection; Class 3 indicates that QRS complexes cannot be detected reliably and the signal is unsuitable for any further analysis.

Background

This database was created for the purpose of evaluating algorithms that are designed to assess the quality of ECG records. Evaluation of ECG quality has become a popular research topic, driven in part by the increased use of mobile monitoring devices such as AliveCor, Apple Watch, and Withings Move ECG Watch. Due to the nature of these devices, it is often necessary to adapt analyses to account for varying signal quality [1-6]. For example, in a very noisy ECG signal it may be highly challenging to reliably detect P waves and to measure PR intervals. Although the recordings are short in the case of self-recorded signals, algorithms for quality estimation are nevertheless useful to filter out possible false positives and to request updated measurements if poor ECG signal quality is detected.

Methods

We recorded 18 signals longer than 24 hours using the Bittium Faros 180 device (ECG recorder with integrated accelerometer) under free-living conditions. By "free-living conditions", we mean that the subjects were carrying out ordinary everyday activities. The only limitation was that water activities such as bathing, showering, and swimming were avoided to protect the Bittium Faros 180 recorder. The database is broadly balanced in terms of gender (10 female records and 8 male records) and age (21 to 83 years, mean 41 years, median 37 years). The signals were recorded from 15 subjects (9 women and 6 men) with 13 subjects being monitored only once, one woman being monitored twice, and one man being monitored three times.

The database contains signal-quality labels provided by three ECG experts, as well as the consensus of these experts, who grouped the signals into three quality classes:

- Class 1: All significant waveforms (P wave, T wave, and QRS complex) are clearly visible and the onsets and offsets of these

waveforms can be detected reliably.
- Class 2: The noise level is increased and significant points in the ECG are unclear (for example, PR interval and/or QRS duration cannot be measured reliably), but QRS complexes are clearly visible and the signal enables reliable QRS detection.
- Class 3: QRS complexes cannot be detected reliably and the signal is unsuitable for any analysis.

Three signals were annotated fully and the other 15 signals were annotated in two segments, each of 20 minutes duration. Five more segments were subjectively selected and annotated to increase the proportion of noisy segments. The duration of four of these five segments is 20 minutes, and the remaining segment is two minutes.

All human studies were approved by the Institutional Review Board of DBME Faculty of Electrical engineering and communication, Brno University of Technology, on July 27, 2018 (IRB Protocol EC:EK:05b/2018) . Informed written consent was obtained from all subjects prior to the studies.

Data Description

Each record contains one-lead ECG recorded with a sampling frequency of 1,000 Hz and 3-axis accelerometer data recorded with a sampling frequency of 100 Hz. All data are provided in the WaveForm Database (WFDB) format. The names of the recordings are six-digit numbers where the first three numbers are unique subject identifiers and the next three numbers indicate the measurement number of this subject. The ECG signal and ACC signals are in two separate files: `*_ECG.dat`, `*_ECG.hea` and `*_ACC.dat`, `*_ACC.hea`, respectively.

The annotations are recorded in a CSV file with 12 columns. Each annotator receives three columns and the last three columns are for consensus (3 columns x 3 annotators + consensus). The first column for each annotator is the first sample of the annotated segment; the second column is the end sample of the annotated segment; and the third column is quality class (1, 2 or 3; 0 means the quality was not annotated in this segment). Patient demographics (gender, age, height, weight, and smoking status) are provided in the file named `subject-info.csv`.

Usage Notes

The database is intended for the development and objective comparison of algorithms designed to assess the quality of ECG records. One of the unique features of this database is that the quality of the ECG signals is annotated sample-by-sample (the signals have varying quality during time). Besides ECG signals, the dataset includes information about motion of the subjects - data from 3-axis accelerometer. Thus, quality assessment algorithms can take into consideration ACC data besides ECG signal or they can be based only on ACC data. For MatLab users, the `ann_reader.m` function file can be used to read the data files and convert them to a point-by-point matrix.

Acknowledgements

This work has been funded by the United States Office of Naval Research (ONR) Global, award number N62909-19-1-2006.

Conflicts of Interest

The authors have no conflict of interest.

References

1. Smital, L., Felton, C. L., Gilbert, B., Holmes, D., Haider, C. R., Vitek, M., … Provaznik, I. (2020). Real-time quality assessment of long-term ECG signals recorded by wearables in free-living conditions. IEEE Transactions on Biomedical Engineering. doi: 10.1109/tbme.2020.2969719.
2. Satija, U., Ramkumar, B., & Manikandan, M. S. (2019). A New Automated Signal Quality-Aware ECG Beat Classification Method for Unsupervised ECG Diagnosis Environments. IEEE Sensors Journal, 19(1), 277–286. doi: 10.1109/jsen.2018.2877055.
3. Shahriari, Y., Fidler, R., Pelter, M. M., Bai, Y., Villaroman, A., & Hu, X. (2018). Electrocardiogram Signal Quality Assessment Based on Structural Image Similarity Metric. IEEE Transactions on Biomedical Engineering, 65(4), 745–753. doi: 10.1109/tbme.2017.2717876.
4. Bai, Y., Do, D., Ding, Q., Palacios, J. A., Shahriari, Y., Pelter, M. M., … Hu, X. (2017). Is the Sequence of SuperAlarm Triggers More Predictive Than Sequence of the Currently Utilized Patient Monitor Alarms? IEEE Transactions on Biomedical Engineering, 64(5), 1023–1032. doi: 10.1109/tbme.2016.2586443.
5. Tobón, D. P., Falk, T. H., & Maier, M. (2016). MS-QI: A Modulation Spectrum-Based ECG Quality Index for Telehealth Applications. IEEE Transactions on Biomedical Engineering, 63(8), 1613–1622. doi: 10.1109/tbme.2014.2355135.
6. Satija, U., Ramkumar, B., & Manikandan, M. S. (2018). A Review of Signal Processing Techniques for Electrocardiogram Signal

Quality Assessment. IEEE Reviews in Biomedical Engineering, 11, 36–52. doi: 10.1109/rbme.2018.2810957.

Contents

Share

Access

Access Policy:
Anyone can access the files, as long as they conform to the terms of the specified license.

License (for files):
Creative Commons Attribution 4.0 International Public License

Discovery

DOI (version 1.0.0):
https://doi.org/10.13026/kah4-0w24

DOI (latest version):
https://doi.org/10.13026/qcnm-a155

Corresponding Author

You must be logged in to view the contact information.

Files

Total uncompressed size: 4.2 GB.

Access the files

- Download the ZIP file (2.2 GB)
- Access the files using the Google Cloud Storage Browser here. Login with a Google account is required.
- Access the data using the Google Cloud command line tools (please refer to the gsutil documentation for guidance):
  `gsutil -m -u YOUR_PROJECT_ID cp -r gs://butqdb-1.0.0.physionet.org DESTINATION`
- Download the files using your terminal: `wget -r -N -c -np https://physionet.org/files/butqdb/1.0.0/`
- Download the files using AWS command line tools: `aws s3 sync s3://physionet-open/butqdb/1.0.0/ DESTINATION`

Visualize waveforms

Folder Navigation: <base>

Name	Size	Modified
100001		
100002		

Name		Size	Modified
103001			
103002			
103003			
104001			
105001			
111001			
113001			
114001			
115001			
118001			
121001			
122001			
123001			
124001			
125001			
126001			
ANNOTATORS		30 B	2020-02-24
LICENSE.txt		14.5 KB	2020-07-20
RECORDS		648 B	2020-07-20
SHA256SUMS.txt		8.0 KB	2020-07-22
ann_reader.m		1.2 KB	2020-07-16
subject-info.csv		432 B	2020-02-24