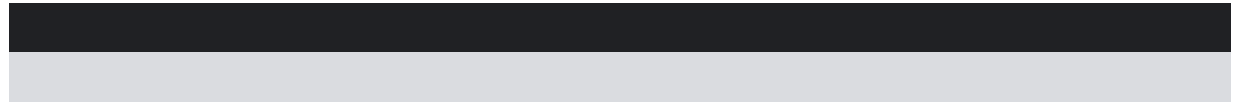


ECG Heartbeat Categorization Dataset

Segmented and Preprocessed ECG Signals for Heartbeat Classification



ECG Heartbeat Categorization Dataset



About Dataset Context

ECG Heartbeat Categorization Dataset

Abstract

This dataset is composed of two collections of heartbeat signals derived from two famous datasets in heartbeat classification, the MIT-BIH Arrhythmia Dataset and The PTB Diagnostic ECG Database. The number of samples in both collections is large enough for training a deep neural network.

This dataset has been used in exploring heartbeat classification using deep neural network architectures, and observing some of the capabilities of transfer learning on it. The signals correspond to electrocardiogram (ECG)

shapes of heartbeats for the normal case and the cases affected by different arrhythmias and myocardial infarction. These signals are preprocessed and segmented, with each segment corresponding to a heartbeat.

Content

Arrhythmia Dataset

- Number of Samples: 109446
- Number of Categories: 5
- Sampling Frequency: 125Hz
- Data Source: Physionet's MIT-BIH Arrhythmia Dataset
- Classes: ['N': 0, 'S': 1, 'V': 2, 'F': 3, 'Q': 4]

The PTB Diagnostic ECG Database

- Number of Samples: 14552
- Number of Categories: 2
- Sampling Frequency: 125Hz
- Data Source: Physionet's PTB Diagnostic Database

Remark: *All the samples are cropped, downsampled and padded with zeroes if necessary to the fixed dimension of 188.*

Data Files

This dataset consists of a series of CSV files. Each of these CSV files contain a matrix, with each row representing an example in that portion of the dataset. The final element of each row denotes the class to which that example belongs.

Acknowledgements

Mohammad Kachuee, Shayan Fazeli, and Majid Sarrafzadeh. "ECG Heartbeat Classification: A Deep Transferable Representation." *arXiv*

preprint arXiv:1805.00794 (2018).

Inspiration

Can you identify myocardial infarction?

Additional:

PTB Diagnostic ECG Database

<https://www.physionet.org/content/ptbdb/1.0.0/>

Data Description

The ECGs in this collection were obtained using a non-commercial, PTB prototype recorder with the following specifications:

- 16 input channels, (14 for ECGs, 1 for respiration, 1 for line voltage)
- Input voltage: ± 16 mV, compensated offset voltage up to ± 300 mV
- Input resistance: 100 Ω (DC)
- Resolution: 16 bit with 0.5 μV /LSB (2000 A/D units per mV)
- Bandwidth: 0 - 1 kHz (synchronous sampling of all channels)
- Noise voltage: max. 10 μV (pp), respectively 3 μV (RMS) with input short circuit
- Online recording of skin resistance
- Noise level recording during signal collection

The database contains 549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6; ages were not recorded for 1 female and 14 male subjects). Each subject is represented by one to five records.

There are no subjects numbered 124, 132, 134, or 161. Each record includes 15 simultaneously measured signals: the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank lead ECGs (vx, vy, vz). Each signal is digitized at 1000 samples per second, with 16 bit resolution over a range of ± 16.384 mV. On special request to the contributors of the database, recordings may be available at sampling rates up to 10 KHz.

Within the header (.hea) file of most of these ECG records is a detailed clinical summary, including age, gender, diagnosis, and where applicable, data on medical history, medication and interventions, coronary artery pathology, ventriculography, echocardiography, and hemodynamics. The clinical summary is not available for 22 subjects. The diagnostic classes of the remaining 268 subjects are summarized below:

Diagnostic class	Number of subjects
Myocardial infarction	148
Cardiomyopathy/Heart failure	18
Bundle branch block	15
Dysrhythmia	14
Myocardial hypertrophy	7
Valvular heart disease	6
Myocarditis	4
Miscellaneous	4
Healthy controls	52

Contributors

[Physikalisch-Technische Bundesanstalt \(PTB\)](#), the National Metrology Institute of Germany, has provided this compilation of digitized ECGs for research, algorithmic benchmarking or teaching purposes to the users of PhysioNet. The ECGs were collected from healthy volunteers and patients with different heart diseases by Professor Michael Oeff, M.D., at the

Department of Cardiology of University Clinic Benjamin Franklin
in Berlin, Germany. His current address is:

Department of Cardiology
Klinikum Brandenburg
14770 Brandenburg, Germany

The database was prepared for PhysioNet by:

Hans Koch, Prof., Ph.D (email: Hans.Koch@ptb.de)
Ralf Bousseljot, Dr. Ing.
Dieter Kreiseler, Dr. Ing.
Physikalisch-Technische Bundesanstalt, Abbestrasse
2-12, 10587 Berlin, Germany

and by

Lothar Schmitz, M.D.
Charité Medical Center, Campus Virchow-Klinikum
and German Heart Institute, Augustenburger Platz 1,
13353 Berlin, Germany

References

Bousseljot, R.; Kreiseler, D.; Schnabel, A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik*, Band 40, Ergänzungsband 1 (1995) S 317

Kreiseler, D.; Bousseljot, R. Automatisierte EKG-Auswertung mit Hilfe der EKG-Signaldatenbank CARDIODAT der PTB. *Biomedizinische Technik*, Band 40, Ergänzungsband 1 (1995) S 319

Comments on the experience with this dataset from the users of this service would be highly appreciated; in particular, a preprint of any publication relying on some of these ECGs would be very welcome.

Share

Access

Access Policy:

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

License (for files):

[Open Data Commons Attribution License v1.0](#)

Discovery

DOI (version 1.0.0):

<https://doi.org/10.13026/C28C71>

Topics:

[ecg](#)

Corresponding Author

You must be logged in to view the contact information.

Files

Total uncompressed size: 1.7 GB.

Access the files

- [Download the ZIP file](#) (1.7 GB)
- Download the files using your terminal: `wget -r -N -c -np https://physionet.org/files/ptbdb/1.0.0/`
- Download the files using AWS command line tools: `aws s3 sync --no-sign-request s3://physionet-open/ptbdb/1.0.0/ DESTINATION`

MIT-BIH Arrhythmia Database

<https://www.physionet.org/content/mitdb/1.0.0/>

Background

Since 1975, our laboratories at Boston's Beth Israel Hospital (now the Beth Israel Deaconess Medical Center) and at MIT have supported our own research into arrhythmia analysis and related subjects. One of the first major products of that effort was the MIT-BIH Arrhythmia Database, which we completed and began distributing in 1980. The database was the first generally available set of standard test material for evaluation of arrhythmia detectors, and has been used for that purpose as well as for basic research into cardiac dynamics at more than 500 sites worldwide. Originally, we distributed the database on 9-track half-inch digital tape at 800 and 1600 bpi, and on quarter-inch IRIG-format FM analog tape. In August, 1989, we produced a CD-ROM version of the database.

Data Description

The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%) at Boston's Beth Israel Hospital; the remaining 25 recordings were selected from the same set to include less common but clinically significant arrhythmias that would not be well-represented in a small random sample.

The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database.

This directory contains the entire MIT-BIH Arrhythmia Database. About half (25 of 48 complete records, and reference annotation files for all 48 records) of this database has been freely available here since PhysioNet's inception in September 1999. The 23 remaining signal files, which had been available only on the MIT-BIH Arrhythmia Database CD-ROM, were posted here in February 2005.

Much more information about this database may be found in the [MIT-BIH Arrhythmia Database Directory](#).

Release Info

- 07/06/2018: File 102.atr has been edited. Annotation number 1991 (0 indexed) has been shifted from sample 590296 to 590262. The original file is here: [102-0.atr](#)

Related Databases

- [MIT-BIH Noise Stress Test Database](#). Twelve half-hour ECG recordings and 3 half-hour recordings of noise typical in ambulatory ECG recordings. The ECG recordings were created by adding calibrated amounts of noise to clean ECG recordings from the MIT-BIH Arrhythmia Database.
- [MIT-BIH P-wave Annotations](#) This database contains reference p-wave annotations for twelve signals from the MIT-BIH arrhythmia database.

Additional references

- 1 Mark RG, Schluter PS, Moody GB, Devlin, PH, Chernoff, D. An annotated ECG database for evaluating arrhythmia detectors. *IEEE Transactions on Biomedical Engineering* 29(8):600 (1982).
- 2 Moody GB, Mark RG. The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it. *Computers in Cardiology* 17:185-188 (1990).

Share

Access

Access Policy:

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

License (for files):

[Open Data Commons Attribution License v1.0](#)

Discovery

DOI (version 1.0.0):

<https://doi.org/10.13026/C2F305>

Topics:

[arrhythmia](#) [ecg](#)

Corresponding Author

You must be logged in to view the contact information.

Files

Total uncompressed size: 104.3 MB.

Access the files

- [Download the ZIP file](#) (73.5 MB)
- Download the files using your terminal: `wget -r -N -c -np https://physionet.org/files/mitdb/1.0.0/`
- Download the files using AWS command line tools: `aws s3 sync --no-sign-request s3://physionet-open/mitdb/1.0.0/ DESTINATION`