Detection and Classification of Cardiac Arrhythmia from 12-Lead ECG Using Deep Neural Network

Doğukan Korkut
Department of Mechanical Engineering
Ozyegin University, Istanbul, Turkey
https://orcid.org/0000-0001-6914-7404

***Abstract*—aa**

***Index Terms*—arrhythmia classification, cardiac arrhythmia, computer-aided diagnosis, deep neural networks (DNNs), electrocardiogram (ECG).**

I. Introduction

The electrocardiogram (ECG) reflects the heart's electrical activity, and the interpretation of this recording can reveal numerous heart pathologies. An ECG is recorded using an electrocardiograph, whereas modern clinical devices usually contain automatic interpretation software that interprets the ECGs directly after recording. Although automatic ECG interpretation started in the 1950s, some limitations remain. Because of the errors they make, doctors have to read over the ECGs. This is time-consuming for the doctors and requires a high degree of expertise. There is a need for better ECG interpretation algorithms.

The recent years have shown a rapid improvement in machine learning. A sub-field of machine learning is called deep learning, where more complex architectures of neural networks can scale better with the amount of data in terms of performance. This type of machine learning has shown promising performance in many fields, including medicine.

This study aims to utilize open-source Electrocardiography (ECG) datasets to detect and classify arrhythmias (heart rhythm disorders) using deep learning methods. The ultimate goal is to contribute to the improvement of automatic diagnosis methods in the healthcare sector and advance medical systems.

II. Related Works

[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [15] [20] [21] [22] [23] [24] [25] [26] [27] [28] [29] [30] [31] [32] [33] [34] [35]

III. Problem Statement

While traditional methods of ECG analysis are valuable, they have certain limitations. Manually analyzing ECGs by cardiologists is time-consuming, leading to delays in diagnosis and treatment. Additionally, human experts are prone to errors, especially when dealing with subtle arrhythmias or variations in ECG patterns. Furthermore, accurate ECG interpretation requires extensive training and experience, limiting the accessibility of this expertise in specific settings.

This study proposes a deep neural network (DNN) based solution for the automated detection and interpretation of cardiac arrhythmias from 12-lead electrocardiograms (ECGs). This approach aims to address the limitations of traditional methods and improve the overall efficiency and accuracy of arrhythmia diagnosis. This proposed DNN-based solution has the potential to revolutionize the diagnosis of cardiac arrhythmias by offering a faster, more accurate, and standardized approach. With further development and validation, this system can become a valuable tool for cardiologists and other healthcare professionals, ultimately improving patient care and outcomes.

DNNs can potentially outperform traditional methods in arrhythmia detection by learning subtle patterns in large ECG datasets. DNN-based analysis can significantly reduce the time required for ECG interpretation compared to manual analysis. The system can be integrated into existing clinical workflows, facilitating automated arrhythmia screening and prioritization of patients requiring further evaluation. DNNs can provide consistent and standardized interpretations, minimizing inter-observer variability among healthcare professionals.

IV. Materials

A. Data Source

The most essential material for conducting an academic study is to obtain data sets. Therefore, open-source databases were investigated to provide input for the study. As a result of the literature search, access to 10 different open-source ECG databases was provided. These databases have been published in journals with high impact factors. These datasets are Chapman University and Shaoxing People's Hospital Database [3], Ningbo First Hospital Database (NFH) [5], Shandong Provincial Hospital Database (SPH) [6], European ST-T Database [36], MIT-BIH Arrhythmia Database [37], Georgia 12-Lead ECG Challenge Database (G12EC) [38], Physikalisch-Technische Bundesanstalt (PTB) Diagnostic ECG Database [39], Physikalisch-Technische Bundesanstalt (PTB-XL) Diagnostic ECG Database [40], St. Petersburg Institute of Cardiological Technics (INCART) 12-Lead Arrhythmia Database [41] and China Physiological Signal Challenge in 2018 (CPSC) [42], respectively. Their standard features are that they keep the numbers of male and female subjects as close as possible numerically, keep the age range wide, and create databases of subjects from various regions.

These databases with various number of subjects have been used for various purposes in many deep learning algorithm

TABLE I
ECG DATABASES COMPARISON

Database	Participants	Recording	Standart	Classes	Duration	Sampling Rate	Age	Gender (male/female)	Lead
SPH	24,666	25,770	AHA	44	10-60 s	500 Hz	18 - 100	55% / 45%	12
European ST-T	78	90	N/A	N/A	120 min	250 Hz	30 - 84	89% / 11%	2
MIT-BIH	47	48	N/A	N/A	30 min	360 Hz	23 - 89	52% / 48%	2
G12EC	15,742	10,344	Snomed-CT	24	10 s	500 Hz	N/A	54% / 46%	12
PTB	262	516	SCP-ECG	71	10 s	1000 Hz	2 - 95	73% / 27%	12
PTB-XL	18,885	21,837	SCP-ECG	71	10 s	500 Hz	2 - 95	52% / 48%	12
INCART	32	74	N/A	10	30 min	257 Hz	18 -80	54% / 46%	12
CPSC	13,256	13,256	N/A	23	6-60 s	500 Hz	0 - 100	53% / 47%	12
Chapman-Shaoxing	**10,646**	**10,646**	**Snomed-CT**	**78**	**10 s**	**500 Hz**	**4 - 98**	**56% / 44%**	**12**
Ningho	**39,905**	**39,905**	**Snomed-CT**	**78**	**10 s**	**500 Hz**	**4 - 89**	**56% / 44%**	**12**

development studies. In this way, valuable outputs have been obtained by the international academic environment. It has been decided to use the Chapman University and Shaoxing People's Hospital Database [3] within the project, which is under development. It was influential in the selection that these databases remain up-to-date and that there is not as much academic research on them as other databases. They have the desired amount of data and distribution for the study.

Researchers created this database [3] to allow further studies on arrhythmias (irregular heartbeats) and other cardiovascular conditions. By providing a large collection of electrocardiogram (ECG) signals with corresponding diagnoses, the database aims to support new research in this field.

A new research database for 12-lead electrocardiogram signals was created under the auspices of Chapman University and Shaoxing People's Hospital. The database contains 12-lead ECGs collected from 10,646 participants. Each ECG has a 500 Hz sampling rate, features 11 common rhythms, and 67 additional cardiovascular conditions labeled by experts. The dataset consists of 10-second, 12-lead ECGs and labels for rhythms and other conditions for each subject.

This new database will enable the research community to conduct novel studies on arrhythmias and other cardiovascular conditions. The dataset can be used to design, compare, and fine-tune new and classical statistical and machine-learning techniques in studies focused on arrhythmia and other cardiovascular diseases.

B. Diagnostic Data

Snomed CT Code tablosu ekle (Diagnoses, SNOMED CT codes and abbreviations in the posted training databases for diagnoses that were scored for the Challenge.)

V. METHODOLOGY

A. Noise Reduction

aa

TABLE II
RHYTHMS, SNOMED CT CODES AND ABBREVIATIONS

Acronym Name	Rhythm Name	Snomed-CT Code
SB	Sinus Bradycardia	426177001
SR	Sinus Rhythm	426783006
AF	Atrial Fibrillation	164889003
ST	Sinus Tachycardia	427084000
AFL	Atrial Flutter	164890007
SI	Sinus Irregularity	427393009
SVT	Supraventricular Tachycardia	426761007
AT	Atrial Tachycardia	713422000
AVNRT	Atrioventricular Nodal Reentrant Tachycardia	251166008
AVRT	Atrioventricular Reentrant Tachycardia	233897008
WAP	Wandering Atrial Pacemaker	195101003

B. Feature Extraction

ECG Morphology

C. Feature Selection

aa

VI. MODEL AND METHOD OVERVIEW

A. Training and Validation Set Annotation

aaa

B. Test Set Annotation

aa

C. Deep Neural Network Architecture And Training

aa

TABLE III
RHYTHM INFORMATION AND BASELINE CHARACTERISTICS OF PARTICIPANTS

Acronym Name	Rhythm Name	Frequency, n(%)	Age, Mean ± SD	Male, n(%)
SB	Sinus Bradycardia	15,807 (36.93%)	58.58 ± 13.85	10,016 (63.36%)
AFL	Atrial Flutter	7,936 (18.54%)	72.73 ± 11.50	4,682 (59.00%)
SR	Sinus Rhythm	7,729 (18.06%)	53.61 ± 16.99	3,409 (44.11%)
ST	Sinus Tachycardia	7,126 (16.65%)	50.59 ± 24.08	3,673 (51.54%)
AF	Atrial Fibrillation	1,780 (4.16%)	73.24 ± 10.95	1,041 (58.48%)
SI	Sinus Irregularity	1,619 (3.78%)	31.83 ± 24.17	911 (56.27%)
SVT	Supraventricular Tachycardia	623 (1.46%)	55.69 ± 19.09	296 (47.51%)
AT	Atrial Tachycardia	141 (0.33%)	65.14 ± 19.58	75 (53.19%)
AVRT	Atrioventricular Reentrant Tachycardia	23 (0.05%)	62.39 ± 16.53	9 (39.13%)
AVNRT	Atrioventricular Nodal Reentrant Tachycardia	16 (0.04%)	57.88 ± 17.34	4 (25.00%)
WAP	Wandering Atrial Pacemaker	2 (0.01%)	52.50 ± 30.41	2 (100.00%)
ALL	**ALL**	**42,802**	**58.55 ± 19,14**	**24,118 (56,35%)**

TABLE IV
PARTICIPANTS WITH AND WITHOUT ADDITIONAL CARDIAC CONDITIONS

Merged Rhythm	Participants With Additional Cardiac Conditions, n(%)	Participants Without Additional Cardiac Conditions, n(%)
SB	15,807 (36.93%)	8,895 (41.37%)
AFIB	9,716 (22.70%)	1,900 (8.84%)
SR	9,348 (21.84%)	7,055 (32.82%)
GSVT	7,931 (18.53%)	3,649 (16.97%)
ALL	**42,802**	**21,499**

TABLE V
THE QUANTITY OF PARTICIPANTS WITH ADDITIONAL CARDIAC CONDITIONS DATA AFTER MERGED CLASSES

Merged Rhythm	Total	Train Data Size (70%)	Validation Data Size (15%)	Test Data Size (15%)
SB	15,807	11,065	2,371	2,371
AFIB	9,716	6,802	1,457	1,457
SR	9,348	6,544	1,402	1,402
GSVT	7,931	5,551	1,190	1,190
ALL	**42,802**	**29,962**	**6,420**	**6,420**

D. *Hyperparameter Tuning*

aa

TABLE VI
THE QUANTITY OF PARTICIPANTS WITHOUT ADDITIONAL CARDIAC CONDITIONS DATA AFTER MERGED CLASSES

Merged Rhythm	Total	Train Data Size (70%)	Validation Data Size (15%)	Test Data Size (15%)
SB	8,895	6,227	1,334	1,334
AFIB	1,900	1,330	285	285
SR	7,055	4,939	1,058	1,058
GSVT	3,649	2,555	547	547
ALL	**21,499**	**15,051**	**3,224**	**3,224**

VII. RESULTS & DISCUSSION

A. *Testing And Performance Evaluation*

aa

B. *Performance Investigation*

aaa

VIII. CONCLUSION

aa

ACKNOWLEDGMENT

aa

REFERENCES

[1] Konstantinos C Siontis, Peter A Noseworthy, Zachi I Attia, and Paul A Friedman. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nature Reviews Cardiology*, 18(7):465–478, 2021.

[2] Antônio H Ribeiro, Manoel Horta Ribeiro, Gabriela MM Paixão, Derick M Oliveira, Paulo R Gomes, Jéssica A Canazart, Milton PS Ferreira, Carl R Andersson, Peter W Macfarlane, Wagner Meira Jr, et al. Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nature communications*, 11(1):1760, 2020.

[3] Jianwei Zheng, Jianming Zhang, Sidy Danioko, Hai Yao, Hangyuan Guo, and Cyril Rakovski. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific data*, 7(1):48, 2020.

[4] Yong-Soo Baek, Sang-Chul Lee, Wonik Choi, and Dae-Hyeok Kim. A new deep learning algorithm of 12-lead electrocardiogram for identifying atrial fibrillation during sinus rhythm. *Scientific reports*, 11(1):12818, 2021.

[5] Jianwei Zheng, Huimin Chu, Daniele Struppa, Jianming Zhang, Sir Magdi Yacoub, Hesham El-Askary, Anthony Chang, Louis Ehwerhemuepha, Islam Abudayyeh, Alexander Barrett, et al. Optimal multi-stage arrhythmia classification approach. *Scientific reports*, 10(1):2898, 2020.

[6] Hui Liu, Dan Chen, Da Chen, Xiyu Zhang, Huijie Li, Lipan Bian, Minglei Shu, and Yinglong Wang. A large-scale multi-label 12-lead electrocardiogram database with standardized diagnostic statements. *Scientific data*, 9(1):272, 2022.

[7] Saman Parvaneh, Jonathan Rubin, Saeed Babaeizadeh, and Minnan Xu-Wilson. Cardiac arrhythmia detection using deep learning: A review. *Journal of electrocardiology*, 57:S70–S74, 2019.

[8] Xinwen Liu, Huan Wang, Zongjin Li, and Lang Qin. Deep learning in ecg diagnosis: A review. *Knowledge-Based Systems*, 227:107187, 2021.

[9] Erick A Perez Alday, Annie Gu, Amit J Shah, Chad Robichaux, An-Kwok Ian Wong, Chengyu Liu, Feifei Liu, Ali Bahrami Rad, Andoni Elola, Salman Seyedi, et al. Classification of 12-lead ecgs: the physionet/computing in cardiology challenge 2020. *Physiological measurement*, 41(12):124003, 2020.

[10] Manuel Martínez-Sellés and Manuel Marina-Breysse. Current and future use of artificial intelligence in electrocardiography. *Journal of Cardiovascular Development and Disease*, 10(4):175, 2023.

[11] Kuan-Cheng Chang, Po-Hsin Hsieh, Mei-Yao Wu, Yu-Chen Wang, Jan-Yow Chen, Fuu-Jen Tsai, Edward SC Shih, Ming-Jing Hwang, and Tzung-Chi Huang. Usefulness of machine learning-based detection and classification of cardiac arrhythmias with 12-lead electrocardiograms. *Canadian Journal of Cardiology*, 37(1):94–104, 2021.

[12] Zahra Ebrahimi, Mohammad Loni, Masoud Daneshtalab, and Arash Gharehbaghi. A review on deep learning methods for ecg arrhythmia classification. *Expert Systems with Applications: X*, 7:100033, 2020.

[13] Fatma Murat, Ozal Yildirim, Muhammed Talo, Ulas Baran Baloglu, Yakup Demir, and U Rajendra Acharya. Application of deep learning techniques for heartbeats detection using ecg signals-analysis and review. *Computers in biology and medicine*, 120:103726, 2020.

[14] Nils Strodthoff, Patrick Wagner, Tobias Schaeffter, and Wojciech Samek. Deep learning for ecg analysis: Benchmarks and insights from ptb-xl. *IEEE journal of biomedical and health informatics*, 25(5):1519–1528, 2020.

[15] Özal Yıldırım, Paweł Pławiak, Ru-San Tan, and U Rajendra Acharya. Arrhythmia detection using deep convolutional neural network with long duration ecg signals. *Computers in biology and medicine*, 102:411–420, 2018.

[16] Wenjuan Cai, Yundai Chen, Jun Guo, Baoshi Han, Yajun Shi, Lei Ji, Jinliang Wang, Guanglei Zhang, and Jianwen Luo. Accurate detection of atrial fibrillation from 12-lead ecg using deep neural network. *Computers in biology and medicine*, 116:103378, 2020.

[17] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1):65–69, 2019.

[18] Sean Shensheng Xu, Man-Wai Mak, and Chi-Chung Cheung. Towards end-to-end ecg classification with raw signal extraction and deep neural networks. *IEEE journal of biomedical and health informatics*, 23(4):1574–1584, 2018.

[19] U Rajendra Acharya, Hamido Fujita, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Muhammad Adam. Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. *Information Sciences*, 415:190–198, 2017.

[20] Yong Xia, Naren Wulan, Kuanquan Wang, and Henggui Zhang. Detecting atrial fibrillation by deep convolutional neural networks. *Computers in biology and medicine*, 93:84–92, 2018.

[21] Yanrui Jin, Zhiyuan Li, Mengxiao Wang, Jinlei Liu, Yuanyuan Tian, Yunqing Liu, Xiaoyang Wei, Liqun Zhao, and Chengliang Liu. Cardiologist-level interpretable knowledge-fused deep neural network for automatic arrhythmia diagnosis. *Communications Medicine*, 4(1):31, 2024.

[22] Tsai-Min Chen, Chih-Han Huang, Edward SC Shih, Yu-Feng Hu, and Ming-Jing Hwang. Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model. *Iscience*, 23(3), 2020.

[23] Runnan He, Yang Liu, Kuanquan Wang, Na Zhao, Yongfeng Yuan, Qince Li, and Henggui Zhang. Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional lstm. *IEEE access*, 7:102119–102135, 2019.

[24] Xiaolong Zhai and Chung Tin. Automated ecg classification using dual heartbeat coupling based on convolutional neural network. *IEEE Access*, 6:27465–27472, 2018.

[25] Jingshan Huang, Binqiang Chen, Bin Yao, and Wangpeng He. Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network. *IEEE access*, 7:92871–92880, 2019.

[26] Chen Chen, Zhengchun Hua, Ruiqi Zhang, Guangyuan Liu, and Wanhui Wen. Automated arrhythmia classification based on a combination network of cnn and lstm. *Biomedical Signal Processing and Control*, 57:101819, 2020.

[27] Manu Thomas, Manab Kr Das, and Samit Ari. Automatic ecg arrhythmia classification using dual tree complex wavelet based features. *AEU-International Journal of Electronics and Communications*, 69(4):715–721, 2015.

[28] Jing Zhang, Aiping Liu, Min Gao, Xiang Chen, Xu Zhang, and Xun Chen. Ecg-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network. *Artificial Intelligence in Medicine*, 106:101856, 2020.

[29] Zhaohan Xiong, Martyn P Nash, Elizabeth Cheng, Vadim V Fedorov, Martin K Stiles, and Jichao Zhao. Ecg signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiological measurement*, 39(9):094006, 2018.

[30] Wenliang Zhu, Xiaohe Chen, Yan Wang, and Lirong Wang. Arrhythmia recognition and classification using ecg morphology and segment feature analysis. *IEEE/ACM transactions on computational biology and bioinformatics*, 16(1):131–138, 2018.

[31] Saroj Kumar Pandey and Rekh Ram Janghel. Automatic detection of arrhythmia from imbalanced ecg database using cnn model with smote. *Australasian physical & engineering sciences in medicine*, 42(4):1129–1139, 2019.

[32] Santanu Sahoo, Asit Subudhi, Manasa Dash, and Sukanta Sabut. Automatic classification of cardiac arrhythmias based on hybrid features and decision tree algorithm. *International Journal of Automation and Computing*, 17(4):551–561, 2020.

[33] Jagdeep Rahul and Lakhan Dev Sharma. Automatic cardiac arrhythmia classification based on hybrid 1-d cnn and bi-lstm model. *Biocybernetics and Biomedical Engineering*, 42(1):312–324, 2022.

[34] Saroj Kumar Pandey and Rekh Ram Janghel. Automatic arrhythmia recognition from electrocardiogram signals using different feature methods with long short-term memory network model. *Signal, Image and Video Processing*, 14(6):1255–1263, 2020.

[35] Siti Nurmaini, Alexander Edo Tondas, Annisa Darmawahyuni, Muhammad Naufal Rachmatullah, Jannes Effendi, Firdaus Firdaus, and Bambang Tutuko. Electrocardiogram signal classification for automated delineation using bidirectional long short-term memory. *Informatics in Medicine Unlocked*, 22:100507, 2021.

[36] Alessandro Taddei, G Distante, M Emdin, P Pisani, GB Moody, C Zeelenberg, and C Marchesi. The european st-t database: standard for evaluating systems for the analysis of st-t changes in ambulatory electrocardiography. *European heart journal*, 13(9):1164–1172, 1992.

[37] George B Moody and Roger G Mark. The impact of the mit-bih arrhythmia database. *IEEE engineering in medicine and biology magazine*, 20(3):45–50, 2001.

[38] Emory University. Georgia 12-lead ecg challenge database. 2020.

[39] Ralf Bousseljot, Dieter Kreiseler, and Allard Schnabel. Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet. *Biomedical engineering / biomedizinische technik*, 40(s1):317–318, 1995.

[40] Patrick Wagner, Nils Strodthoff, Ralf-Dieter Bousseljot, Dieter Kreiseler, Fatima I Lunze, Wojciech Samek, and Tobias Schaeffter. Ptb-xl, a large publicly available electrocardiography dataset. *Scientific data*, 7(1):1–15, 2020.

[41] Vikto Tihonenko, A Khaustov, S Ivanov, A Rivin, and E Yakushenko. St petersburg incart 12-lead arrhythmia database. *PhysioBank PhysioToolkit and PhysioNet*, 2008.

[42] Feifei Liu, Chengyu Liu, Lina Zhao, Xiangyu Zhang, Xiaoling Wu, Xiaoyan Xu, Yulin Liu, Caiyun Ma, Shoushui Wei, Zhiqiang He, et al. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. *Journal of Medical Imaging and Health Informatics*, 8(7):1368–1373, 2018.