

# Two Might Do: A Beat-by-Beat Classification of Cardiac Abnormalities using Deep Learning and Domain-Specific Features

Berken Utku Demirel<sup>1</sup>, Adnan Harun Dogan<sup>2</sup>, Mohammad Abdullah Al Faruque<sup>1</sup>

<sup>1</sup> Department of Electrical Engineering and Computer Science, University of California, Irvine ,USA

<sup>2</sup> Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

## Abstract

*This paper proposes an efficient convolutional neural network to detect 26 different classes of cardiac activities from different numbers of leads in the Physionet/Computing data in the Cardiology Challenge 2021. The proposed CNN architecture is designed to utilize heart rate variation features from ECG recordings and waveform morphologies of heartbeats simultaneously. Also, the designed architecture is flexible for the implementation of a different number of leads with a varied length of ECG recordings. The proposed algorithm achieved a score of 0.57 using only 2 channels of ECG on all recordings for the hidden validation set of the challenge, placing us 81, 64, 72, 71, 63rd (Team name: METU\_19) out of 252 entries for 12, 6, 4, 3, and 2-leads respectively. These results show the potential of an efficient, flexible novel neural network for beat-by-beat classification of raw ECG recordings to a complex multi-class multi-label classification problem.*

## 1. Introduction

Cardiovascular diseases are one of the leading causes of death worldwide [1]. According to American Heart Association, [2] the early diagnosis of these diseases is crucial for patients' health. Some works recently indicate that millions of deaths could have been prevented if the cardiovascular disease was detected earlier [3]. The electrocardiogram (ECG) signals are widely used for detecting different cardiovascular diseases, especially the clinical ECG setups are the main tools for detecting abnormal cardiac activity. However, the clinical ECG setup uses 12-lead recordings, making the cardiac monitoring limited outside the clinical hours. To solve this issue, people propose to use wearable devices with limited leads by utilizing the recent advancements in Internet-of-Things (IoT). However, this brings the question of how many leads are sufficient to diagnose a much larger number of cardiac problems? The Physionet/Computing in Cardiology Challenge 2021 [4] focus to identify clinical diagnoses from twelve-lead,

six-lead (I, II, III, aVR, aVL, aVF), four-lead (I, II, II, V2), three-lead (I, II, V2), and two-lead (I and II) ECG recordings. The differences in performances of the algorithms on the large-scale databases from multi-centers reveal the utility of reduced-lead ECGs compared to standard twelve-lead ECGs and answer the question.

This work proposed a novel, flexible neural network architecture to implement a different number of leads with a varied length of recordings to diagnose cardiac abnormalities using beat morphology and domain-specific features.

## 2. Methods

The objective is to create a model that can efficiently and accurately classify ECG recordings with different numbers of leads into multiple classes of 30 clinical diagnoses. Four pairs of classes were scored identically. Therefore, we considered these pairs the same and combined identical pairs as one class, so only 26 classes were considered. Although the datasets include varying length ECG recordings and our proposed model is designed to take beats as input and predict 26 different classes from a heartbeat, the overall decision is made using all heartbeats in a recording.

### 2.1. Preprocessing

The overall challenge dataset contained over 88,000 ECG records with varying lengths for training. Since our method is based on beat-by-beat classification, we have segmented each ECG recordings into beats using the Pan&Tompkins algorithm [5]. We take  $F_s/4$  samples before and  $F_s/3$  samples after the detected R-peak for segmentation. Additionally, the segmentation window is changed to  $F_s/5$  samples before and after when the heart rate is higher than 180 beats per minute to prevent overlapping two heartbeats in the same window.

The sampling rate and the gain value of ECG recordings vary in different datasets. Due to this difference, we have resampled and normalized each extracted beat before feeding them to a classifier. All beats have been resampled to 200 Hz using Fourier resampling. The segmented beats

were normalized by their maximum value to have the same signal amplitudes. Also, the Pearson cross-correlation of each extracted beat from the 10-second record is calculated by their linear dependence as Equation 1.

$$\rho_{xy} = \frac{1}{N} \sum_{i=1}^N \frac{x_i - \mu_x}{\sigma_x} \frac{y_i - \mu_y}{\sigma_y} \quad (1)$$

where each variable (segmented heartbeats as  $x$  and  $y$ ) has  $N$  scalar observations (length of each heartbeat), and  $\mu_x$  and  $\sigma_x$  are the mean and standard deviation of  $x$ , respectively, and  $\mu_y$  and  $\sigma_y$  are the mean and standard deviation of  $y$ . Then using these correlations, the correlation matrix is formed as shown in Equation 2, where the  $K$  is the number of segmented heartbeats in a 10-second record.

$$P = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1K} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2K} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{K1} & \rho_{K2} & \cdots & \rho_{KK} \end{bmatrix} \quad (2)$$

Finally, each column, which refers to a heartbeat in  $P$ , is summed and normalized with the maximum value and compared with a threshold value found empirically and set to 0.8 value. If the normalized correlation coefficient of a segment is lower than 0.8, this segmented section is removed from the recording without further procession. This last step, based on cross-correlation of segmented beats, in preprocessing acts as a control mechanism to avoid a wrongly segmented heartbeat feeding as input to the classifier. Although this approach assumes that the performance of the peak detection algorithm is highly accurate and detects most of the correct peaks in the recording, it is observed that this preprocessing step reduces the number of wrong segments in a recording.

## 2.2. Domain-Specific Feature Extraction

When we segment the ECG records into heartbeats and feed them to a classifier, it is observed that the heart rate variability information is lost, resulting in a significant performance decrease for some cardiovascular diseases, which are characterized by heart rhythm such as atrial fibrillation, bradycardia. To prevent this performance decrease for some classes, we have calculated the heart rate for each beat in all recordings using the following equation:

$$HR_{(i)} = \frac{60}{(R_{(i)} - R_{(i-1)})/F_s} \quad (3)$$

Where  $HR_{(i)}$  is the heart rate of the  $i^{th}$  heartbeat segment in beat per minute unit.  $R_{(i)}$  is the location of the R-peak in the  $i^{th}$  heartbeat segment, and  $F_s$  is the sampling rate of this record. Then, using the calculated  $HR_{(i)}$

values, an HR array ( $\overrightarrow{HR}$ ) is formed by adding them one after another, as shown in Equation 4.

$$\overrightarrow{HR} = [HR_0, HR_1, HR_2, \dots, HR_i] \quad (4)$$

Finally, the heart rate variability (HRV) features are calculated from the recordings and integrated into the designed deep learning model. The following HRV features are calculated from a 10-second ECG recording from lead I,

- Mean of  $\overrightarrow{HR}$
- Median of  $\overrightarrow{HR}$
- Range of  $\overrightarrow{HR}$
- Standard deviation of  $\overrightarrow{HR}$  values

$$\sigma = \sqrt{\frac{1}{I-1} \sum_{n=1}^I |HR[n] - \mu_i|^2} \quad (5)$$

- Root mean square of  $\overrightarrow{HR}$  values

$$RMS = \sqrt{\frac{1}{I} \sum_{n=1}^I HR_n^2} \quad (6)$$

- pNN20  
(percentage of HR intervals differing more than 20 ms)

In total, six different heart rate variability features are extracted from a 10-second recording to be used for classification. Since these features are obtained using the complete recording instead of a single heartbeat and our proposed model is based on the beat-by-beat classification, we have used the extracted feature with all heartbeats simultaneously at the neural network. Although, at first glance, it seems a problem due to overfitting as the same features are used with the different heartbeats in the same recording. It is observed that it works as a regularizer instead of resulting in overfitting.

## 2.3. Classifier

The proposed convolutional neural network classifies the segmented heartbeat into 26 different classes using the extracted HRV features from a 10-second recording. The proposed CNN architecture is designed to be flexible with the different number of leads. The CNN architecture for the heartbeat classification is designed both utilizing the residual connections introduced by He et al. in [6] and the dropout [7] that has been shown to achieve good performance for increasing the generalization ability of the neural networks. Especially, it is noted that the CNN architecture classification performance increases approximately 3 – 5% when the dropout with 0.1 is used at the lead ways as shown in Figure 1.

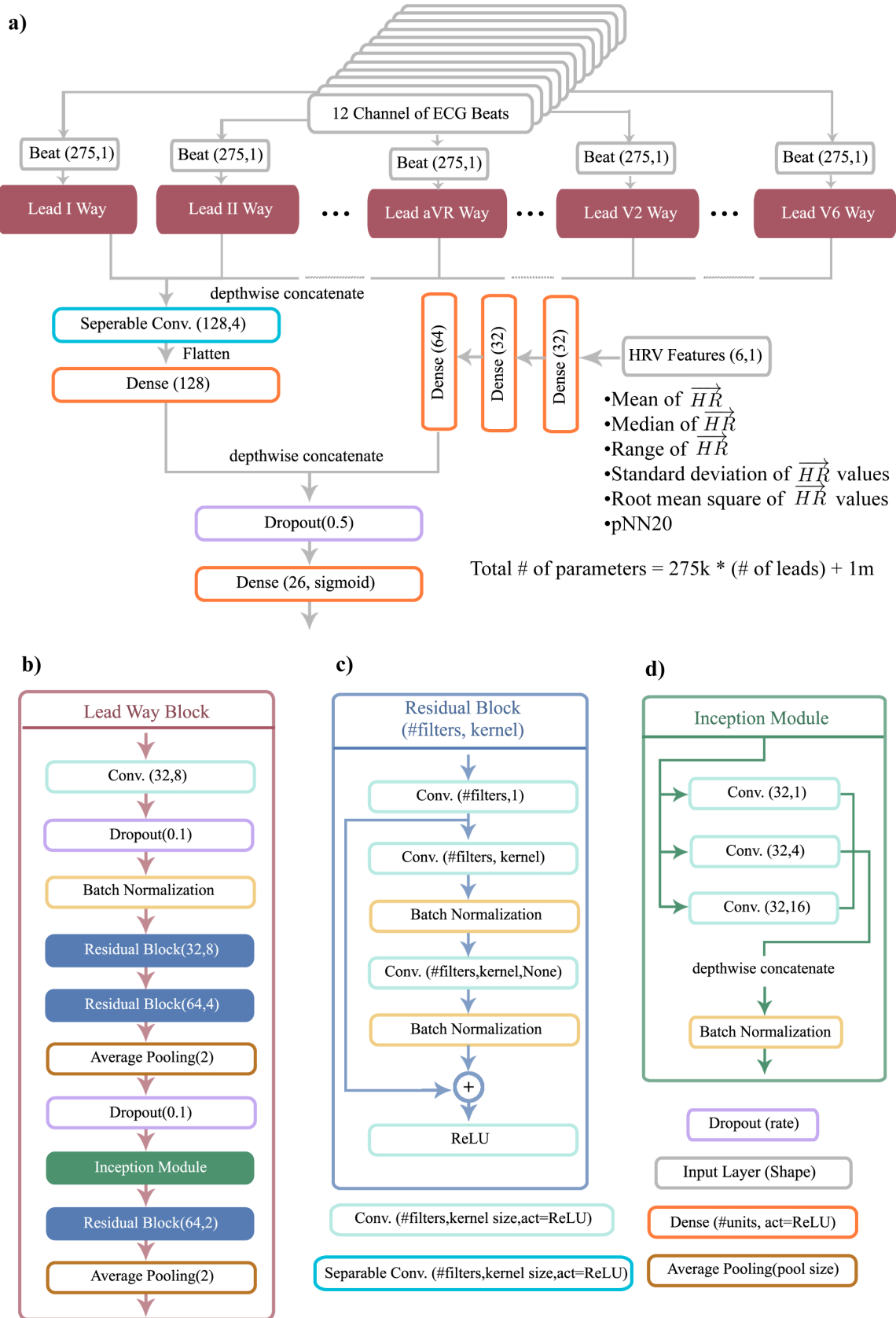


Figure 1: a) Complete model architecture, b) Channel block, c) Residual block, d) Inception block

## 2.4. Network Training

The network was trained with Glorot initialization of the weights [8]. L2-regularization with 0.0002 is applied for each convolution layer. We used the Adam optimizer [9] with the default parameters  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ , and a mini-batch size of 90. The learning rate is initialized to 0.001 and reduced by a factor of 10 when the validation F1 score stopped improving for 75 consecutive epochs. The training continues until 100 successive epochs without performance improvements with a maximum of 500 epochs. 10% of training records were used for validation to implement early-stopping based on the F1 score.

## 3. Results

After training, the model predicts scores  $p \in [0, 1]$  for each class. To binarize these scores, we have applied a threshold which is found empirically, and set it to 0.25 for all classes. As our model takes heartbeats as input instead of complete recording, the prediction outputs of beats in a record are combined into a single score for each class similar to our correlation matrix calculation. However, different from the correlation matrix, each row represents a heartbeat in a recording, and each column refers to the scored classes in the challenge. After creating this scoring matrix for each heartbeat, the columns are summed and normalized with the maximum value. Finally, the normalized scores are compared with the threshold value, which is indicated above and set to 1 if it is higher than the threshold.

Table 1: 10-fold cross-validation results

Database	AUROC	AUPRC	F-measure	Challenge Score
Georgia	0.83	0.42	0.46	0.51
Chapman-Shaoxing [10]	0.85	0.45	0.57	0.72
CPSC [11]	0.79	0.28	0.42	0.76

It is important to emphasize that cross-validation was always performed on a per-subject basis in all datasets using two leads. In other words, the subjects' recordings for testing the performance are not included the training.

## 4. Discussion and Conclusion

This paper proposed an efficient, flexible CNN architecture for classifying raw ECG recordings with varying lengths and different numbers of leads. The designed CNN is not only utilized the heartbeat morphology but also employed the heart rate variation features simultaneously to detect arrhythmias based on rhythms. Also, novel preprocessing and postprocessing methods are introduced in this paper to integrate beat-by-beat classification to long, varying length ECG recordings. Considering all of these proposed methods together contributes a significant performance improvement for identifying cardiac abnormalities. As part of the PhysioNet/Computing in Car-

diology Challenge 2021, the performance of the proposed method was evaluated using comprehensive heterogeneous datasets with a new scoring metric in a realistic clinical setting.

## References

- [1] Heart Disease and Stroke Statistics—2020 Update: A Report From the American Heart Association. *Circulation* March 2020;141(9):e139–e596.
- [2] Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* October 2016; 388(10053):1459–1544. ISSN 0140-6736, 1474-547X.
- [3] McGill HC, McMahan CA, Gidding SS. Preventing Heart Disease in the 21st Century. *Circulation* March 2008; 117(9):1216–1227.
- [4] Reyna M, Sadr N, Gu A, Alday P, A. E, Liu C, Seyedi S, Shah A, Clifford G. Will two do? Varying Dimensions in Electrocardiography the PhysioNet Computing in Cardiology Challenge 2021;2021. URL <https://doi.org/10.13026/jz9p-0m02>.
- [5] Pan J, Tompkins WJ. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering* 1985;BME-32(3):230–236.
- [6] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition, 2015.
- [7] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 2014;15(56):1929–1958.
- [8] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In Teh YW, Titterton M (eds.), *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*. Chia Laguna Resort, Sardinia, Italy: JMLR Workshop and Conference Proceedings, 13–15 May 2010; 249–256.
- [9] Kingma DP, Ba J. Adam: A method for stochastic optimization. *CoRR* 2015;abs/1412.6980.
- [10] Zheng J, Zhang J, Danioko S, Yao H, Guo H, Rakovski C. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific Data* February 2020;7(1):48.
- [11] Liu FF, Liu\* CY, Zhao LN, Zhang XY, Wu XL, Xu XY, Liu YL, Ma CY, Wei SS, He ZQ, Li JQ, Kwee NY. An open access database for evaluating the algorithms of ecg rhythm and morphology abnormal detection. *Journal of Medical Imaging and Health Informatics* 2018;8(7):1368–1373.

Address for correspondence:

Berken Utku Demirel  
Department of Electrical Engineering and Computer Science  
2200 Engineering Hall Irvine, CA, USA  
bdemirel@uci.edu