

# Combining ResNet Model with Handcrafted Temporal Features for ECG Classification with varying number of leads

Stefano Magni<sup>1</sup>, Andrea Sansonetti<sup>1</sup>, Chiara Salvi<sup>1</sup>, Tiziana Tabiaddon<sup>1</sup>, Guadalupe García-Isla<sup>1</sup>

<sup>1</sup>Politecnico di Milano, Milano, Italy

## Abstract

*This study presents PhysioNauts Team’s contribution to the PhysioNet/CinC Challenge 2021 on ECG classification for variable leads. Three types of labels were identified: those affecting cardiac rhythm, ECG morphology or both. The full model integrated handcrafted rhythm features and deep learning features into a residual neural network (ResNet) with a squeeze and excitation module and a wide 10-neuron single-layer fully connected (FC) branch to leverage the learning of both feature types. The ResNet inputs were ECG segments of 4096 samples downsampled to 257 Hz. The FC inputs were standard rhythm features extracted from the RR-series. Class imbalance was mitigated by selecting only a third of normal sinus rhythm and sinus bradycardia recordings. Moreover, threshold optimization was performed based on a grid search and the Nelder-Mead method to maximize the Challenge metric (CM). The complete model obtained as official CM score 0.613, 0.585, 0.603, 0.594, and 0.582 on 12-lead, 6-lead, 4-lead, 3-lead, 2-lead (ranked 43rd, 37th, 38th, 46th and 50th out of 256 classified entries).*

## 1. Introduction

The electrocardiogram (ECG) records the heart’s electrical activity. It is used for the identification of cardiac electrical abnormalities. ECGs may be acquired with a variable number of leads depending on the device used, the duration and scope of the registration. While the standard clinical ECG consists on a 12-lead 30s ECG, Holter monitoring intended for a long-term recording (24-48 h) typically uses a reduced number of leads.

ECG manual annotation is highly time consuming and requires of trained professionals. Although many studies in the past have attempted to develop automatic algorithms for ECG analysis [1–3], still no method has proven to detect pathologies using variable ECG leads from signals acquired using different devices and clinical centers. The PhysioNet/Computing in Cardiology Challenge 2021 addresses this problem by providing a multi center ECG

database [2–9] with variable leads and signal lengths to develop an algorithm capable of detecting 26 different pathologies.

In this work we present the PhysioNauts Team’s contribution to the CinC Challenge. The presented study focused on the nature of the detected pathologies: some affect the *morphology* of the ECG, others its *rhythm*, or both. The combination of a modified ResNet with another branch fed with temporal parameters was exploited to leverage both morphological and rhythmic properties of the pathologies, and improve the classification capabilities of the neural networks.

## 2. Material and Methods

Eight different datasets were made available by the Challenge (for further details the reader is referred to [2–9]) with a total of 88253 subjects associated with 133 labels, representing different pathologies. Only 30 out of 133 were of interest for the challenge. The unscored signals were discarded, and only 81966 were considered. The recordings were relabeled, keeping only 26 out of the 30 scored labels, as 8 of them were considered equivalent [10]. Recordings were multi-labeled, and had variable sampling frequencies, between 257 Hz up to 1000 Hz, and duration, with signals of 5 s, 10 s and 120 s long, only 74 recordings were 30 minutes long.

### 2.1. ResNet data processing

Recordings were resampled to 257 Hz, then a common length of 4096 samples was fixed: if the signal was shorter it was zero padded, if it was longer, windowing in a random part of the signal was performed, as done by Zhao et al. [11]. These processed data were fed to the ResNet, as shown in Figure 1.

### 2.2. Temporal Features extraction

Lead 'II' was chosen to obtain the RR time series since, together with lead 'I', it was the only one common to all

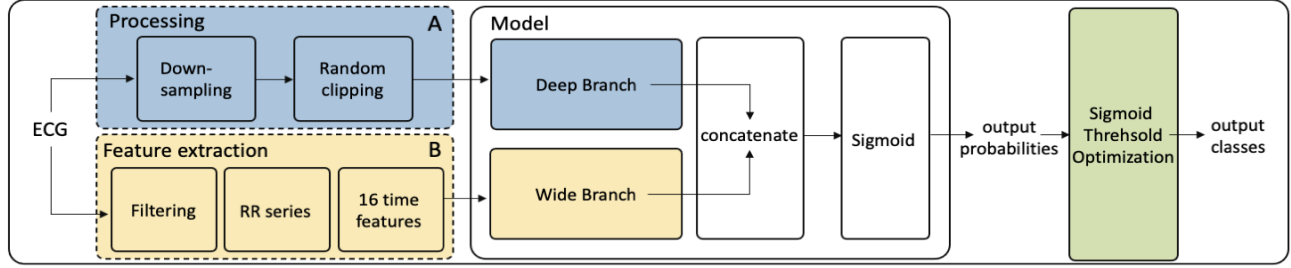


Figure 1: The two flows of information. A. processing of ECG before giving it to the ResNet; B. extraction of temporal features for the wide branch. The two branches are then concatenated to produce the output probabilities and classes.

the lead subsets, and better suited for R peak extraction. Raw signals were processed as follows:

1. Application of 2<sup>nd</sup> order Butterworth band-pass filter in the range 5-15 Hz.
2. Normalization of filtered signals in the range [-1,1].
3. A combination of Pan Tompkins and amplitude and time (equations (1) and (2), respectively) thresholds were applied for R peak detection.

$$Thr_{amp} = \mu_{signal} + 2 * \sigma_{signal}, \quad (1)$$

$$Thr_{time} = 150ms, \quad (2)$$

where  $\mu_{signal}$  is the mean value of the filtered and normalized ECG, and  $\sigma_{signal}$  is its standard deviation. If less than three peaks or less than 30 beats/min were found with the thresholds, Pan Tompkins was applied.

4. Extraction of the RR event series in milliseconds [ms].
5. Outlier removal procedure using as threshold the relationship of equation (3):

$$Thr_{outliers} = \mu_{rrseries} + 5 * \sigma_{rrseries}, \quad (3)$$

where  $\mu_{rrseries}$  and  $\sigma_{rrseries}$  are the mean and standard deviation of the RR series respectively [12].

6. Extraction from RR series the 16 temporal parameters typically used in HRV analysis [13, 14], shown in Table 1.
7. Normalization of the parameters with MinMax method. These parameters were fed to the wide branch of the network, as shown in Figure 1.

### 2.3. Model Architecture

The network consisted in two branches, as shown in Figure 2. The first one was a modified ResNet [11], composed by two types of blocks: the Residual Block (ResBs) and the Squeeze-and-Excitation block (SE). The standardized-windowed signals (flow A of Figure 1) were fed to a Convolutional Layer, and then given to the eight successive 'ResBs's. In each ResBs there were two Convolutional Layers followed by the SE. The SE block encompassed

Time Parameter	Description
mean_nni	Mean RR interval
sdnn	RR interval std <sup>a</sup>
sdsd	Std of adjacent RR-intervals differences
nni_50	Number of RR exceeding 50ms
pnni_50	Percentual of nni50 over total RR intervals
nni_20	Number of RR exceeding 20ms
pnni_20	Percentual of nni20 over total RR intervals
rmsd	RMS <sup>b</sup> of RR differences
median_nni	Median RR interval
range_nni	Max-Min difference
cvsd	CV <sup>c</sup> of successive differences
cvnni	Coefficient of Variation
mean_hr	Mean Heart Rate
max_hr	Max Heart Rate
min_hr	Min Heart Rate
std_hr	std of Heart Rate

Table 1: List of temporal properties extracted.

<sup>a</sup> Standard Deviation, <sup>b</sup> Root Mean Square, <sup>c</sup> Coefficient of Variation

two Fully Connected (FC) layers and ended with a sigmoid function, which transferred the input to the successive block. This block detected morphological correlations among leads. The Convolutional Layers increased the number of characteristics extracted by the network while reducing the input dimensionality, thanks to the doubling of filter dimension every two 'ResBs's. The dropout layer placed in between the Convolutional Layers of the ResBs increased the network generalization capabilities. The wide branch took as input 16 temporal features extracted from the ECGs (flow B of Figure 1) and it consisted of a FC layer of 10 neurons. This limited number of neurons forced the model to learn an embedding of the temporal features useful for the classification.

The ResNet capability to recognize the implicit morphological patterns and the wide branch explicit temporal features were put together into a concatenate layer, without distinction between the two. Since both parts had different number of parameters, a two-step training procedure was followed to guarantee correct error back-propagation to the deep branch (further explained in section 2.4).

A sigmoid layer was used as final classifier to return the recording's probability values of belonging to each of the

classes independently in the [0,1] range. The classes were assigned by comparing this output probability with optimal learnt thresholds (see par. 2.4).

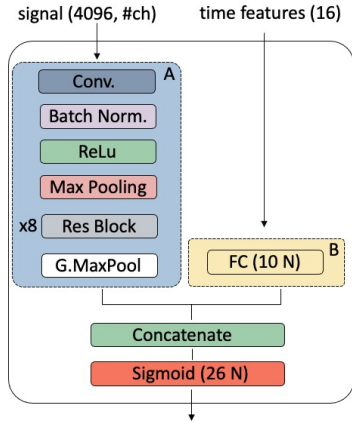


Figure 2: Network architecture: on the left (block A) the deep branch, consisting of the modified ResNet; on the right (block B) the wide branch, composed of the FC layer. Their outputs are concatenated to perform the final predictions.

## 2.4. Model Training

The available full data-set was unbalanced, especially towards Normal Sinus Rhythm (NSR) and Sinus Bradycardia (SB) classes. To avoid bias and enhance the generalization capability of the models, the original distribution was modified, keeping all data except for the NSR and SB, which were both randomly reduced to a third.

For each leads subset a model was trained using the same network structure but changing the number of input channels. Models were trained in a two-stage fashion for 12 epochs. During the first 9 only the deep network was trained, freezing the wide layers, while in the remaining 3 only the wide branch parameters were updated. The chosen batch size was 64 and the initial Learning Rate was 0.003, reduced tenfold each 10 epochs. The training error used was binary cross-entropy, optimized using the Adam optimizer.

To maximize the Challenge Metric (CM) and handle class unbalance a threshold optimization [11] was applied to the 26 output probabilities. This was done in two steps:

1. Maximization of the CM score with a step of 0.01 and maximum value of 0.4 through grid search of threshold value.
2. Nelder-Mead downhill simplex minimization method over the negative value of the CM score, initialized with the output of step 1.

## 2.5. Model Evaluation

Signals were processed as in the training phase. However, instead of performing the prediction on a single time window, signals were segmented into windows of 4096 samples with a 256 samples overlap. The final prediction probability output was the average across all independently classified segments. Output classes were assigned according to the optimal thresholds found.

The model performances were assessed holding out a stratified local test set with 20% of training data. The actual learning of the model was conducted using a stratified 5-fold cross validation.

## 3. Results

The results obtained after a 20% hold out 5-fold stratified cross-validation are presented in column "Training" of Table 2, while columns "Validation", "Test" and "Ranking" were provided by the challenge organizers.

Leads	Training	Validation	Test *	Ranking *
12	0.689 ± 0.004	0.613	-	-
6	0.663 ± 0.008	0.585	-	-
4	0.673 ± 0.006	0.603	-	-
3	0.672 ± 0.005	0.594	-	-
2	0.656 ± 0.007	0.582	-	-

Table 2: Challenge scores for final entry: 5-fold cross validation on the public training set, repeated scoring on the hidden validation set, and one-time scoring on the hidden test set as well as the ranking on the hidden test set. \* Waiting for the final Official scores.

## 4. Discussion and Conclusions

The model presented encompassed two branches: a deep ResNet SE neural network and a wide single layer handcrafted feature classifier. Both branches concatenated to obtain an embedding of the morphological and temporal features and to provide with the recording's probability of belonging to each class. The choice of adding a wide branch was driven by the reported poor capability of deep models to effectively entail temporal patterns of the ECG signal [15]. The combination of HRV handcrafted features and deep features aimed to take advantage of the implicit complex morphological information obtained by the ResNet model together with the explicit rhythmic data of the HRV features. A similar approach was followed in [16], but in contrast with that procedure which directly combined handcrafted features with deep features, in this work the learning of the wide branch was performed through 16 handcrafted HRV features (Table 1)

fed to a FC 10 neurons layer. This was combined with the ResNet model [11], which had already proven in the past Challenge its efficacy for 12-leads ECGs classification.

In addition, to mitigate the overfitting and avoid giving excessive importance to the wide branch a learning routine in two phases was implemented, as described in section 2.4. The network was firstly allowed to focus on extracting morphological features by freezing the wide branch and the end sigmoid classifier and only updating the ResNet weights. On a second stage the wide branch was then allowed to learn by freezing the deep network and preserving the embedding. This approach reduced overfitting, even if it was not completely removed, as shown in Table 2.

The performances obtained showed that the models were capable of classifying ECGs for every lead subset. Some small differences were present, presumably due to the loss of information in reduced lead subsets, affecting the overall ECG morphological patterns.

The main limitation of the proposed approach is overfitting, probably caused by the high importance given to the handcrafted temporal features and by the differences in the hidden validation and test datasets labelling. Future works should aim to expand the number of features to be fed to the wide branch, encompassing possibly explicit morphological features of the signal, and to perform regularization on the wide branch weights.

In conclusion, the proposed algorithm tackles some of the complexities and challenges posed by the dataset. By working on the unbalancing of the dataset results were improved. However, the lack of an unique approach in the labelling process of the training data-set poses a challenge that should be overcome to increase the robustness of classification.

## References

- [1] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 2000;101(23):e215–e220.
- [2] Perez Alday EA, Gu A, Shah A, Robichaux C, Wong AKI, Liu C, et al. Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020. *Physiological Measurement* 2020;41.
- [3] Reyna MA, Sadr N, Perez Alday EA, Gu A, Shah A, Robichaux C, et al. Will Two Do? Varying Dimensions in Electrocardiography: the PhysioNet/Computing in Cardiology Challenge 2021. *Computing in Cardiology* 2021;48:1–4.
- [4] Liu F, Liu C, Zhao L, Zhang X, Wu X, Xu X, et al. An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection. *Journal of Medical Imaging and Health Informatics* 2018;8(7):1368—1373.
- [5] Tihonenko V, Khaustov A, Ivanov S, Rivin A, Yakushenko E. St Petersburg INCART 12-lead Arrhythmia Database. *PhysioBank PhysioToolkit and PhysioNet* 2008;Doi: 10.13026/C2V88N.
- [6] Zheng J, Cui H, Struppa D, Zhang J, Yacoub SM, El-Askary H, et al. Optimal Multi-Stage Arrhythmia Classification Approach. *Scientific Data* 2020;10(2898):1–17.
- [7] Wagner P, Strodthoff N, Boussejot RD, Kreiseler D, Lunze FI, Samek W, et al. PTB-XL, a Large Publicly Available Electrocardiography Dataset. *Scientific Data* 2020;7(1):1–15.
- [8] Boussejot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik* 1995;40(S1):317–318.
- [9] 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* 2020;7(48):1–8.
- [10] PhysioNet/Computing in Cardiology Challenge 2020. <https://physionetchallenges.org/2020/>. Accessed: 2021-06-08.
- [11] Zhao Z, Fang H, Relton S, Yan R, Liu Y, Li Z, et al. Adaptive lead weighted ResNet trained with different duration signals for classifying 12-lead ECGs. In *2020 Computing in Cardiology Conference (CinC)*. *Computing in Cardiology*, dec 2020; URL <https://doi.org/10.22489/cinc.2020.112>.
- [12] Cabiddu R, Cerutti S, Werner S, Viardot G, Bianchi A. Modulation of the sympatho-vagal balance during sleep: Frequency domain study of heart rate variability and respiration. *Frontiers in Physiology* 2012;3:45. ISSN 1664-042X. URL <https://www.frontiersin.org/article/10.3389/fphys.2012.00045>.
- [13] Hrv analysis parameters extraction from rr series. <https://aura-healthcare.github.io/hrv-analysis/>. Accessed: 2021-08-27.
- [14] of the European Society of Cardiology TF, the North American Society of Pacing, Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 1996;93(5):1043–1065. URL <https://pubmed.ncbi.nlm.nih.gov/8598068>.
- [15] Yan G, Liang S, Zhang Y, Liu F. Fusing transformer model with temporal features for ecg heartbeat classification. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2019; 898–905.
- [16] Natarajan A, Chang Y, Mariani S, Rahman A, Boverman G, Vij S, et al. A wide and deep transformer neural network for 12-lead ecg classification. In *2020 Computing in Cardiology*. 2020; 1–4.

Address for correspondence:

Stefano Magni	Andrea Sansonetti
Pizza Leonardo Da Vinci, 32	Pizza Leonardo Da Vinci, 32
Milano(MI), Italy	Milano(MI), Italy
stefano4.magni@mail.polimi.it	andrea.sansonetti@mail.polimi.it