Detection of Chagas Disease Using Digital Electrocardiogram by Deep Transfer Learning of the InceptionTime Model

Arnaud Champetier[1], Ivo Strebel[1], Christian Mueller[1]

[1]Cardiovascular Research Institute of Basel, Basel, Switzerland

Abstract

This paper introduces the solution of team cardio_Basel25 to the PhysioNet Challenge 2025.

Chagas disease, caused by Trypanosoma cruzi, remains underdiagnosed in low-resource settings where serological testing is limited. Given the widespread availability of electrocardiography (ECG) and the conduction abnormalities characteristic of Chagas cardiomyopathy, ECG-based artificial intelligence offers a scalable alternative for early detection. A deep transfer learning model was developed based on a pretrained InceptionTime architecture, and fine-tuned on Brazilian (CODE-15%, Sami-Trop) and European (PTB-XL) ECG datasets. Recordings were preprocessed with filtering, downsampling, normalization, and 5-second segment extraction, with data augmentation applied during training. The training loss consisted of binary cross-entropy with a penalization term to emphasize the challenge metric. Inference combined predictions across multiple ECG segments and models. The approach achieved a cross-validation score of 0.42 and a score of 0.382 on the validation set, ranking 27th on the official leaderboard. These findings demonstrate the feasibility of deep transfer learning for ECG-based Chagas screening and its potential to expand diagnostic access in underserved regions.

1. Introduction

Chagas disease, caused by Trypanosoma cruzi, affects 6–8 million people in Latin America and causes an estimated 10,000 deaths annually [1]. Despite its significant burden, diagnosis still relies on multiple serological assays, which are costly and less accessible in low-resource regions [1], [2]. Machine learning approaches have been applied to serological data for detecting the parasite in blood samples, yet these methods leave the diagnostic gap in underserved areas unaddressed [3], [4].

The electrocardiogram (ECG) offers a low-cost, widely available alternative. Chagas cardiomyopathy is often marked by characteristic conduction disturbances—such as right bundle branch block, left anterior hemiblock, first-degree atrioventricular block, atrial fibrillation, and ventricular ectopy—which not only define its clinical profile but have also been integrated into established prognostic scores [5], [6], [7], [8], [9]. Yet, the direct use of ECG for Chagas disease screening remains underexplored, representing a critical opportunity to improve early detection and expand diagnostic reach.[10]

The 2025 PhysioNet Challenge seeks to address this research gap by leveraging two Brazilian datasets—CODE-15% and Sami-Trop—and one European dataset, PTB-XL [11], [12], [13], [14]. Our work aims to use a large pre-trained ECG model and fine tune it to better detect the Chagas disease.

2. Method

Our methodology was inspired by the preprocessing framework of the PhysioNet 2021 Challenge winner and by the binary outcome prediction model of Buscher et al., which relies on the InceptionTime backbone architecture [1], [2].

The architecture comprises two residual blocks, each containing three inception modules. The first inception module processes the raw ECG signal by applying a convolution across all channels to generate a bottleneck layer with 32 filters. This representation is subsequently passed through three parallel convolutional layers, each with 32 filters but distinct kernel sizes. The outputs of these convolutions are concatenated with the result of a max-pooling operation applied directly to the raw ECG (Figure 1). Each subsequent inception module receives as input the output of the preceding module.

At the final stage, an adaptive average pooling layer reduces the temporal dimension to a single value, producing a feature vector of size 128 (32 × 4). A fully connected layer (128, 1) then transformed this into the ECG feature vector. To leverage demographic information, patient age and sex were concatenated with the ECG feature before classification. The combined vector was passed through two fully connected layers of sizes (3, 3) and (3, 1), respectively, yielding the final prediction. The InceptionTime backbone was initialized with pretrained weights from Buscher et al., derived from a large emergency department ECG cohort [1].

All ECG recordings were preprocessed prior to model input. A zero-phase third-order Butterworth band-pass filter (1–47 Hz) was applied, followed by downsampling

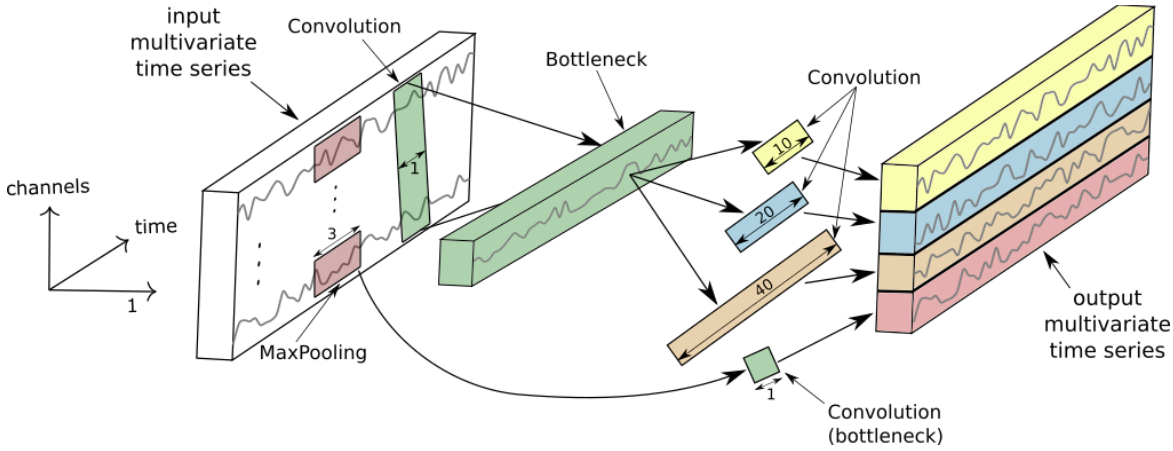

Figure 1: Inception module for time series classification (from the original paper). Bottleneck is visually described with a single dimension versus 32 in implementation.

to 200 Hz and z-score normalization. From each recording, a random 5-second segment was extracted. To improve model robustness, data augmentation was applied during training: random lead masking (10% of leads set to zero), two temporal masks covering approximately 6% of the signal window, and additive Gaussian noise sampled from $\mathcal{N}(0, 0.02)$.

The optimization objective was a penalized binary cross-entropy (pBCE) loss, designed to align with the challenge metric requirements. For a training batch with labels $y \in \{0,1\}$ and predictions $p \in [0,1]$, the loss is defined as:

$$L_{pBCE} = \alpha L_{BCE} + (1-\alpha) L_{FN} \quad (1)$$

Where α is a hyperparameter that was set at 0.66. The first term is the standard binary cross-entropy, and the second term penalizes false negatives above the 95th percentile of predicted probability:

$$L_{FN} = \sum_{i \in \mathfrak{B}} \mathbb{1}_{FN_i} \quad (2)$$

With:

$$\mathbb{1}_{FN_i} = \begin{cases} 1 & if\ y_i = 1\ and\ \hat{y}_i < \tau_{0.95} \\ 0 & otherwise \end{cases} \quad (3)$$

Where $\mathfrak{B}$ is the set of all indices in the batch, y_i the outcome, $\hat{y}_i$ the predicted outcomes, $\tau_{0.95}$ the 95th percentile of predicted outcomes in the batch.

A 5-fold cross-validation (CV) strategy was performed in which five models were trained independently with different data splits. Model training was conducted using the fastai framework (built on the PyTorch library) with the Adam optimizer. Training proceeded in two phases. In the first phase, all layers except the final linear classifier were frozen, and the model was pretrained for 5 epochs using the One-Cycle learning rate policy, with the learning rate decreasing from 1×10^{-2} 1×10^{-4}. In the second phase, all layers, including the Inception modules, were unfrozen and trained for 25 epochs with the One-Cycle policy, with learning rates ranging from 5×10^{-3} to 5×10^{-6}. A learning rate scheduler (ReduceLROnPlateau, patience = 2, reduction factor = 0.1) was applied, and early stopping was employed based on validation loss with a patience of 10 epochs.

For inference, the three models with the highest validation performance were selected. As the input length was restricted to 5 seconds, five random 5-second segments were extracted from each ECG recording. Each model generated probability estimates for all segments, which were averaged to obtain a segment-level prediction per model. The final output probability was then computed as the mean of the three model-level predictions.

Computations were performed using a NVIDIA GeForce RTX 3090 (24 GB), computational resources were allocated to hyperparameter optimization using the Optuna framework, followed by iterative refinement in subsequent training runs.

3. Results

Cross-validation training scores were not directly computed, as this would have substantially increased training time. Instead, for each fold, the training loss and validation score corresponding to the epoch with the best validation loss were extracted. Table 1 reports the mean ± standard deviation of these values across the 5 folds.

Table 1:Challenge scores and losses across datasets

Dataset	Score	Custom Loss
CV training	-	1.10±0.06
CV validation	0.42±0.01	1.14±0.03
Validation	0.382	-
Test	-	-

The validation score was the 27th/365 of the official phase leaderboard.

4. Alternatives and future work

Several alternative methodologies were explored during the official phase, though each presented limitation. Training either the classification head or the full ECGFounder model on ECG signals upsampled to 500 Hz resulted in unstable training dynamics (gradient explosion). Similarly, we evaluated an approach in which the 27 cardiac outcome predictions produced by ECGFounder were used as input features for an XGBoost classifier. However, when assessed using 5-nested-fold cross-validation, this strategy achieved lower performance than our deep learning model. A student–teacher framework based on the ECGFM model, which incorporated 150 predictions into a distillation loss, was also investigated; however, this approach suffered from either gradient explosion or severe overfitting.

These limitations highlight the challenges of adapting large, pretrained ECG models within the competition's time constraints. Future work could examine generalization performance of ECGFounder and ECGFM under specific training strategies and explore advanced regularization schemes for distillation.

5. Conclusion

In this study, we developed and evaluated a deep learning model for Chagas disease detection from digital ECG signals, motivated by the urgent need for accessible diagnostic tools in low-resource regions. Leveraging a pretrained InceptionTime backbone and fine-tuning it on Brazilian and European datasets, we achieved competitive performance in the 2025 PhysioNet Challenge. Our results confirm the potential of ECG-based AI to complement traditional serology and expand diagnostic reach.

Acknowledgments

This work was supported by the university hospital Basel, the university of Basel, the Swiss Heart Foundation and the Swiss National Science Foundation.

References

[1] A. S. de Sousa, D. Vermeij, A. N. Ramos, and A. O. Luquetti, 'Chagas disease', The Lancet, vol. 403, no. 10422, pp. 203–218, Jan. 2024, doi: 10.1016/S0140-6736(23)01787-7/ASSET/C5C4D911-3A1E-471F-829C-B11F930861EF/MAIN.ASSETS/GR2.JPG.

[2] J. A. Pérez-Molina and I. Molina, 'Chagas disease', The Lancet, vol. 391, no. 10115, pp. 82–94, Jan. 2018, doi: 10.1016/S0140-6736(17)31612-4/ASSET/3FEACF50-4570-48E8-9A37-6C08C25A088E/MAIN.ASSETS/GR3.JPG.

[3] A. Ojeda-Pat, A. Martin-Gonzalez, · Carlos Brito-Loeza, · Hugo Ruiz-Piña, and · Daniel Ruz-Suarez, 'Effective residual convolutional neural network for Chagas disease parasite segmentation', Med Biol Eng Comput, vol. 1, p. 3, 2022, doi: 10.1007/s11517-022-02537-9.

[4] V. Uc-Cetina, C. Brito-Loeza, and H. Ruiz-Piña, 'Chagas Parasite Detection in Blood Images Using AdaBoost', 2015, doi: 10.1155/2015/139681.

[5] M. C. P. Nunes et al., 'Chagas Cardiomyopathy: An Update of Current Clinical Knowledge and Management: A Scientific Statement From the American Heart Association', Circulation, vol. 138, no. 12, pp. e169–e209, Sep. 2018, doi: 10.1161/CIR.0000000000000599/SUPPL_FILE/CIRC-D-18-00039_SUPPL1.PDF.

[6] L. Z. Rojas et al., 'Electrocardiographic abnormalities in Chagas disease in the general population: A systematic review and meta-analysis', PLoS Negl Trop Dis, vol. 12, no. 6, Jun. 2018, doi: 10.1371/JOURNAL.PNTD.0006567.

[7] M. S. Marcolino, D. M. Palhares, L. R. Ferreira, and A. L. Ribeiro, 'Electrocardiogram and Chagas Disease: A Large Population Database of Primary Care Patients', Glob Heart, vol. 10, no. 3, p. 167, Sep. 2015, doi: 10.1016/j.gheart.2015.07.001.

[8] C. Di Lorenzo Oliveira et al., 'Risk Score for Predicting 2-Year Mortality in Patients With Chagas Cardiomyopathy From Endemic Areas: SaMi-Trop Cohort Study', doi: 10.1161/JAHA.119.014176.

[9] A. Rassi et al., 'Development and Validation of a Risk Score for Predicting Death in Chagas' Heart Disease', New England Journal of Medicine, vol. 355, no. 8, pp. 799–808, Aug. 2006, doi: 10.1056/NEJMOA053241/ASSET/838B1395-3FDF-43CA-96F2-9F32668A6CCB/ASSETS/IMAGES/LARGE/NEJMOA053241_T3.JPG.

[10] C. Jidling et al., 'Screening for Chagas disease from the electrocardiogram using a deep neural network', PLoS Negl Trop Dis, vol. 17, no. 7, p. e0011118, Jul. 2023, doi: 10.1371/JOURNAL.PNTD.0011118.

[11] Reyna MA et al., 'Detection of Chagas Disease from the ECG: The George B. Moody PhysioNet Challenge 2025', vol. 52, pp. 1-4..

[12] A. L. Goldberger et al., 'PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.', Circulation, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.CIR.101.23.E215/ASSET/9716B30D-2E4D-4D06-8F9E-A21157DFE97C/ASSETS/GRAPHIC/HC2304183003.JPEG.

[13] A. H. Ribeiro et al., 'CODE-test: An annotated 12-lead ECG dataset', Jan. 2020, doi: 10.5281/ZENODO.3765780.

[14] A. L. P. Ribeiro et al., 'Tele-electrocardiography and bigdata: The CODE (Clinical Outcomes in Digital Electrocardiography) study', J Electrocardiol, vol. 57, pp. S75–S78, Nov. 2019, doi: 10.1016/J.JELECTROCARD.2019.09.008.

[15] A. L. P. Ribeiro et al., 'Sami-Trop: 12-lead ECG traces with age and mortality annotations', doi: 10.5281/ZENODO.4905618.

[16] P. Wagner et al., 'PTB-XL, a large publicly available electrocardiography dataset', Sci Data, vol. 7, no. 1, 2020, doi: 10.1038/s41597-020-0495-6.

[17] A. Büscher et al., 'Deep Learning Electrocardiogram Model for Risk Stratification of Coronary Revascularization Need in the Emergency Department', Eur Heart J, Mar. 2025, doi: 10.1093/eurheartj/ehaf254.

[18] P. Nejedly et al., 'Classification of ECG Using

Ensemble of Residual CNNs with Attention Mechanism', Comput Cardiol (2010), vol. 48, pp. 1–4, 2021, doi: 10.23919/CINC53138.2021.9662723.