Learning from Leads: A 1D Dilated ResNet for ECG Chagas Disease Screening

Somesh Saini[1], Matheus Araujo[2]

[1] Computer and Data Sciences, Case Western Reserve University, Cleveland, United States of America
[2] Biomedical Engineering, Case Western Reserve University, Cleveland, United States of America

Abstract

Chagas disease, caused by the parasite Trypanosoma cruzi, presents significant diagnostic challenges due to asymptomatic impacts on heart health and limited resources in the affected areas. We leverage a deep-learning-based artificial neural network model to automate Chagas disease screening using 12-lead electrocardiograms from multiple datasets, including CODE-15, PTB-XL, and SamiTrop. We incorporate an automatic hyperparameter optimization and class imbalance solution combined with 1D dilated residual neural network with squeeze-and-excitation blocks. This approach is applied to the PhysioNet 2025 Computing in Cardiology Challenge. Our model, achieved a challenge score of 0.271, demonstrating a ~4-fold improvement over traditional machine learning approaches, which achieved a baseline of 0.062. Our findings support the use of deep-learning based approaches as a viable tool for scalable, automated Chagas disease screening, particularly in low-resource clinical environments where traditional diagnostics are unavailable.

1. Introduction

Chagas disease, termed *American trypanosomiasis*, is a chronic illness caused by the parasite *Trypanosoma cruzi*, predominantly transmitted through triatomine insects (1). Approximately 8 million people worldwide are infected annually, mainly concentrated in Latin America. Migration patterns have begun to introduce Chagas disease into non-endemic regions, including approximately 300,000 cases in the United States (2). Furthermore, around 13% of the population in Latin America is at risk of infection, and Chagas disease accounts for 670,000 disability-adjusted life-years annually (1). Environmental and socioeconomic factors significantly impact the epidemiology and transmission of Chagas disease. Subtropical climates with high humidity, warmer temperatures, and substandard housing conditions proliferate the activity of *Trypanosoma cruzi* parasites and human infection (3; 4).

The diagnosis of Chagas disease is a formidable challenge due to the asymptomatic or nonspecific clinical presentation observed in acute and chronic indeterminate phases. Once in the chronic phase, persistent inflammation and damage to myocardial cells can occur in heart tissue with the potential to lead to heart failure. Although serological diagnostic methods exist, such as enzyme-linked immunosorbent assays (ELISA) and polymerase chain reaction (PCR) testing, they are resource-intensive, often leading to delayed detection and ineffective disease management for affected patients (2; 3).

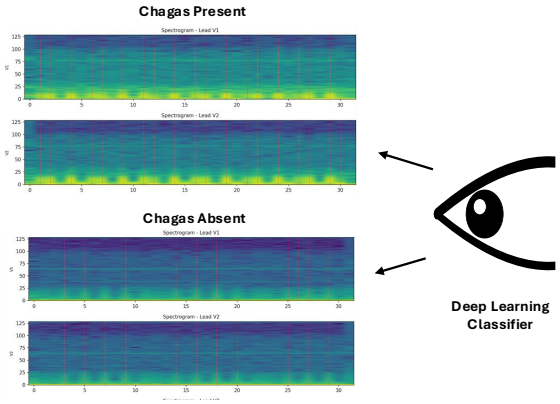

Figure 1. Spectrograms of ECG Data for Leads V1 and V2 with examples from 2 subjects, one present with Chagas and the other with Chagas absent. The brighter colors indicate higher signal energy at the time–frequency, specially present in the example with Chagas.

Electrocardiograms (ECGs) are crucial for detecting chronic Chagas cardiomyopathy. In Figure 1, we show how distinguishable can be the ECG spectrograms from a Chagas present and Chagas Absent patients. However, manual ECG interpretation is prone to variability among healthcare providers and costly time for interpretation. To address these challenges, the 2025 George B. Moody PhysioNet Challenge aims to advance research by providing large-scale, annotated 12-lead ECG datasets and a standardized framework for developing automated methods for Chagas disease detection (5; 6).

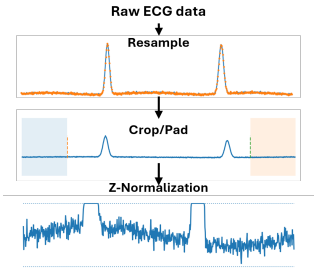

Figure 2. During preprocessing, ECG signals are resampled, cropped/padded, and z-normalized, then passed as input to the model.

In this work, we build on recent advances in Convolutional Neural Networks (CNNs) and residual network (ResNet) variants, which have achieved state-of-the-art performance in ECG classification by extracting complex temporal and spatial features from large datasets(6; 7). Using the 2025 PhysioNet dataset, we apply a ResNet-inspired deep-learning approach for automated Chagas disease detection from 12-lead ECGs. Our objective is to enhance screening availability, improve diagnostic consistency and precision, and ultimately support better patient outcomes.

2. Methodology

The following subsections present the model architecture approach for Chagas classification. Our approach consists of a ResNet Network for prediction. Training utilizes an imbalance-aware strategy and precision-recall curve calibration.

For comparison purposes, we experimented with alternative machine learning methods, including a Random Forest model combined with features extracted via `tsfresh`.

2.1. Data

This study uses three data sets from the PhysioNet 2025 Challenge, each offering different demographic, geographic, and clinical characteristics relevant to the detection of Chagas disease. The CODE-15 data set comprises 345,779 ECG recordings from 233,770 patients, collected through the Minas Gerais Telehealth Network in Brazil. These recordings include self-reported cases of Chagas disease and are available in 7.3-second or 10.2-second durations, sampled at 400 Hz. 1.91% of the records were labeled as Chagas-positive and 98.09% as Chagas-negative (8). The PTB-XL dataset provides 21,799 ECG records from 18,869 European patients, each 10.0 seconds in length and sampled at 500 Hz. It contains only Chagas-negative cases, making it valuable for training balanced models and generalization (9). Finally, the SamiTrop dataset contains 1,631 ECGs from 1959 patients, specifically validated for chronic Chagas cardiomyopathy, where all cases are Chagas-positive. Specifically, it includes women aged 50 to 74 years of age from lower socioeconomic backgrounds, with recordings of 7.3 or 10.2-second durations and a sampling frequency of 400 Hz, offering high clinical relevance for Chagas-specific modeling (10).

2.2. Preprocessing

All raw 12-lead ECGs were standardized with a three-stage preprocessing pipeline as shown in Figure 2. First, signals were resampled to 500 Hz so that recordings from different sources shared the same frequency content. Second, to present a uniform input size, each record was converted to a fixed 5120-length window. Traces shorter than the set length were zero-padded, and longer traces were centered and trimmed accordingly. Third, we normalized each lead by median-centering and median absolute deviation (MAD) scaling, removing offset and placing amplitudes on a common scale, robust to ECG outliers.

For lead ℓ with samples $x_\ell(t)$,

$$z_\ell(t) = \mathrm{clip}_{[-5,5]}\left(\frac{\mathrm{x}_\ell(\mathrm{t}) - \mathrm{median}(\mathrm{x}_\ell)}{\mathrm{MAD}(\mathrm{x}_\ell) + 10^{-6}}\right)$$

$$\mathrm{MAD}(x) = \mathrm{median}\big(|x - \mathrm{median}(x)|\big)$$

Here, the first equation highlights the z-normalization process applied to all ECG signals. The $\mathrm{clip}_{[-5,5]}$ interval is applied on the z-score, capping standardized values at ± 5 to suppress extreme outliers. The second equation defines the median absolute deviation (MAD), which measures variability by taking the median of absolute deviations from the median.

2.3. Model Architecture

We use a compact dilated residual 1-dimensional (1D) ResNet CNN with squeeze-and-excitation (SE) and Group Normalization (11; 12). The network processes the preprocessed 12-lead ECG segment through three residual convolutional blocks. Each block contains a 1D convolution (*Conv1d*) with a 3x1 kernel and increasing dilation rates (1, 2, and 4), enabling the model to capture progressively wider temporal contexts within the receptive field while preserving signal resolution. This is followed by GroupNorm layer to rescale ECG feature maps across leads, and ReLU activation. When moving to the next residual block, we use a 1x1 *Conv1d*. The ECG sequence length is gradually reduced between blocks via max pooling.

While training, data was processed as tensors below (Length, Channels):

$$(5120, 12) \xrightarrow[\text{Block 1 } (d=1)]{\text{MaxPool}} (2560, 64) \xrightarrow[\text{Block 2 } (d=2)]{\text{MaxPool}} (1280, 128) \xrightarrow[\text{Block 3 } (d=4)]{\text{GlobalMaxPool}} (1, 256)$$

Ultimately, the classifier head operates on the pooled features by flattening them into a single vector and passing them through a fully connected neural network with a

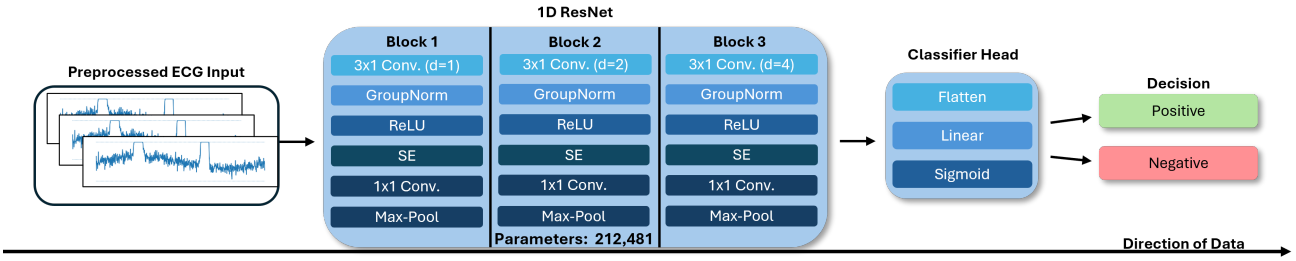

Figure 3. From left to right, this figure depicts model architecture. The 1D ResNet processes each ECG segment through three residual blocks with increasing dilation (1–4) to expand the receptive field. Each block applies GroupNorm, ReLU, squeeze-excitation, and max-pooling. The flattened features feed a linear layer with sigmoid to produce a binary (positive/negative) prediction.

sigmoid activation to output the probability of Chagas disease. The model was implemented in Python 3.13.0 using PyTorch 2.1.2.

2.4. Training and Inference

We trained the model using stratified $90/5/5$ respective train, validation, and test splits. To improve robustness, we applied data augmentations including Gaussian noise, scale jitter, and mixup. Training employed Focal Loss ($\alpha = 0.5,\ \gamma = 1.5$), which emphasizes rare positive cases and mitigates the effects of the 1:30 positive–negative class imbalance. We optimized the 1D ResNet model with *AdamW* under a *OneCycleLR* schedule, helping regularize data and avoid overfitting. OneCycleLR makes a single learning rate sweep from short to high peaks to learn robust detectors for Chagas. For stability on ECG sequences, we enable exponential moving average (EMA) of weights to reduce variance and yield better generalization. The final training uses early stopping (patience $= 4$) to limit overfitting.

In Table 1, we show the selected hyperparameters with Optuna (30 trials) over augmentation metrics (noise, scale jitter, mixup), optimizer settings (learning rate, weight decay), and model depth (number of residual blocks). After training, After training, we pick a single final decision threshold on the validation split by maximizing F_1 on the precision–recall curve, and we use this fixed threshold at inference to convert probabilities into class labels.

2.5. Evaluation

For consistency with the PhysioNet Challenge framework, we submitted all trained models to the official challenge organizers. The submitted model performance was evaluated on a hidden test set and returned the Challenge Score, which served as the primary criterion for ranking models. Since the Challenge setting included multiple competitors and limited each team to only ten submission attempts, careful model selection and tuning were necessary before submission. While additional internal metrics (AUROC, AUPRC, F1) were monitored during development, the Challenge Score ultimately determined our comparative evaluation.

Hyperparameter	Value / Description
Sequence length	5120
Number of residual blocks	3
Base channel width	64
Dropout rate	0.26
Mixup coefficient	0.04
Noise sigma	0.02
Input scaling range	[0.99, 1.01]
Stochastic drop probability	0.06
Classification threshold	0.53
Softmax temperature	0.58
EMA decay rate	0.99

Table 1. Final hyperparameter configuration for the 1D ResNet model, obtained via Optuna optimization in the local experimental environment.

3. Results

Results for our scoring submissions are in Table 2. The best performance was obtained with a 1D SE ResNet model, receiving a challenge score of 0.271 and AUROC of 0.671. We were able to increase the challenge score significantly from the baseline score performance of 0.062. Through the iteration of multiple models, including Random Forest classifiers with tsfresh-based feature extraction, we observed that traditional approaches reached substantially lower performance (e.g., 0.097 and 0.062 for the Random Forest-based models).

4. Discussion

Our results suggest that a dilated 1D SE ResNet can learn and discriminate features from 12-lead ECG data for automated Chagas screening. We also demonstrated how deep

Model (ID)	Architecture	HPO	Score	AUROC	AUPRC	Accuracy	F1
Model 9 (2684)	ResidualDilatedBlocks + SE	Optuna	0.271	N/A	N/A	N/A	N/A
Model 7 (2627)	ResidualDilatedBlocks + SE	Optuna	0.270	0.671	0.119	0.980	0.020
Model 8 (2617)	No Residual 3-Block CNN	Optuna	0.262	N/A	N/A	N/A	N/A
Model 6 (2564)	ResidualDilatedBlocks	Optuna	0.236	0.677	0.094	0.977	0.098
Model 1 (2061)	ResidualDilatedBlocks + SE	Optuna	0.230	0.653	0.058	0.971	0.076
Model 5 (2560)	ResidualDilatedBlocks	Optuna	0.215	0.667	0.073	0.979	0.038
Model 2 (2443)	Dilated Conv. Blocks	Optuna	0.194	0.612	0.061	0.978	0.094
Model 3 (2342)	RandomForestClassifier + tsfresh	GridSearchCV	0.097	0.526	0.025	0.979	0.000
Model 4 (2494)	ResidualDilatedBlocks + SE	Optuna	0.050	0.500	0.021	0.979	0.000
Baseline Model	RandomForestClassifier	N/A	0.062	N/A	N/A	N/A	N/A

Table 2. Model comparison across architectures and hyperparameter optimization (HPO). Reported are the PhysioNet Official Phase Challenge Score and metrics (AUROC, AUPRC, Accuracy, F1).

learning architectures are more capable of capturing complex temporal dependencies in ECG signals compared to state-of-the-art machine learning models such as Random Forest.

We noticed two main challenges during this project. First, the 90/5/5 validation data split and its impact on heavy calibration efforts. With the 1:30 class imbalance, very few positive cases were in the 5% validation split, which introduces statistical uncertainty in the precision–recall curve and calibration performance. We also noticed how limited computing resources can impact model development. We explored Synthetic Minority Oversampling (SMOTE) to mitigate the 1:30 class imbalance, but applying it to thousands of 12-lead ECG sequences was prevented by memory limits. As a result, we adopted a lightweight sampling strategy with focal loss, a function that down-weights easy negatives and gives more weight to the rare positive cases, providing a reasonable improvement for model learning despite the imbalance. Future work, with robust computing resources, could revisit imbalance handling in stratified split adjustments and synthetic oversampling.

5. Conclusion

We demonstrate that a dilated 1D SE ResNet achieves approximately 4-fold gains over traditional machine learning baselines. We acknowledge that the current algorithm is not as accurate as experts in detecting Chagas disease. Still, we showed that learning discriminative features for Chagas disease directly from 12-lead ECGs is possible. By incorporating imbalance-aware objectives, the approach achieved competitive results in the PhysioNet 2025 Challenge under the constraint of limited submissions. Our findings highlight the significant potential for scalable Chagas disease screening in resource-constrained areas.

References

[1] World Health Organization, "Chagas disease (american trypanosomiasis)," 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis)

[2] Centers for Disease Control and Prevention, "About chagas disease," 2025. [Online]. Available: https://www.cdc.gov/chagas/about/index.html

[3] Pan American Health Organization, "Chagas disease," 2025. [Online]. Available: https://www.paho.org/en/topics/chagas-disease

[4] SciELO Brasil, "Chagas disease transmission and housing," 2022. [Online]. Available: https://www.scielo.br/j/mioc/a/TXvCrGp77QxBq9nqx7Scq6p/

[5] PhysioNet, "Physionet/computing in cardiology challenge 2025," 2025. [Online]. Available: https://physionetchallenges.org/2025/

[6] J. Li, S.-p. Pang, F. Xu, P. Ji, S. Zhou, and M. Shu, "Two-dimensional ecg-based cardiac arrhythmia classification using dseresnet," *Scientific Reports*, vol. 12, no. 1, p. 14485, 2022.

[7] A. Saglietto, D. Baccega, R. Esposito, M. Anselmino, V. Dusi, A. Fiandrotti, and G. M. De Ferrari, "Convolutional neural network (cnn)-enabled electrocardiogram (ecg) analysis: a comparison between standard twelve-lead and single-lead setups," *Frontiers in Cardiovascular Medicine*, vol. 11, p. 1327179, 2024.

[8] A. H. Ribeiro, G. M. Paixao, E. M. Lima, M. Horta Ribeiro, M. M. Pinto Filho, P. R. Gomes, D. M. Oliveira, W. Meira Jr., T. B. Schon, and A. L. P. Ribeiro, "CODE-15%: a large scale annotated dataset of 12-lead ECGs," 2021. [Online]. Available: https://zenodo.org/records/4916206

[9] P. Wagner, N. Strodthoff, R.-D. Bousseljot, W. Samek, and T. Schaeffter, "PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3)," 2022. [Online]. Available: https://physionet.org/content/ptb-xl/1.0.3/

[10] A. L. P. Ribeiro, A. H. Ribeiro, G. M. Paixao, E. M. Lima, M. Horta Ribeiro, M. M. Pinto Filho, P. R. Gomes, D. M. Oliveira, W. Meira Jr, T. B. Schn, and E. C. Sabino, "Sami-trop: 12-lead ecg traces with age and mortality annotations," 2021. [Online]. Available: https://zenodo.org/records/4905618

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[12] Y. Wu and K. He, "Group normalization," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3–19.

Address for correspondence:

Somesh Saini

3919 Magnolia Circle Maumee, Ohio, 43537

sbs166@case.edu