A Deep Learning Framework for Chagas Disease Detection Using CNN-Transformer Architecture on 12-Lead ECGs

Kyung Min Choi[1], Gi-Won Yoon[1], Segyeong Joo[1]

[1]Department of Biomedical engineering, Asan Medical Institute of Convergence Science and Technology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

Abstract

Background and Objective: *Chagas cardiomyopathy remains a major health burden in Latin America and among migrant populations. Accurate electrocardiographic detection is essential for timely intervention, but existing automated approaches often struggle with label noise and data imbalance. We propose a hybrid convolutional–transformer model with metadata integration for robust Chagas disease detection from 12-lead ECGs.*

Methods: *ECGs from the* ***SaMi-Trop*** *and* ***CODE-15%*** *cohorts were used as strongly and weakly labeled data, respectively, while* ***PTB-XL*** *was included for negative examples. Signals were filtered, normalized, and fixed to 2048 samples. Age and sex were appended as metadata. The model combines convolutional layers for local temporal feature extraction with transformer encoders for global dependencies, enhanced with positional encoding. Training employed a* ***weighted focal loss****, oversampling of minority positive cases, and a two-stage process: pretraining on strongly labeled datasets and fine-tuning on weak labels.*

Results: *The proposed pipeline achieved a* ***Challenge score of 0.691****, with AUROC and AUPRC values exceeding baseline models. The system generalized across geographically and demographically distinct cohorts, indicating its robustness in real-world applications.*

Conclusion: *Our study demonstrates that integrating convolutional backbones, transformers, and metadata with strong/weak label training significantly improves Chagas disease detection from ECGs. This framework provides a scalable solution for automated screening in resource-limited and diverse clinical settings.*

1. Introduction

Chagas disease, caused by *Trypanosoma cruzi*, is a neglected tropical disease that can lead to severe cardiomyopathy. Timely identification of cardiac involvement is vital, yet ECG interpretation in endemic regions is challenged by limited resources. Traditional ECG classification models rely on handcrafted features or simple deep learning architectures that may not capture long-range dependencies in signals or adapt to heterogeneous datasets.

Recent advances in transformer models have enabled improved sequence modeling in biomedical applications. Coupling transformers with convolutional layers allows simultaneous learning of local morphologies and global rhythm features, both critical for Chagas-related abnormalities. However, challenges remain: datasets vary in label quality, and the disease prevalence introduces severe imbalance.

Here, we present a hybrid CNN–transformer model trained on multi-source ECG datasets with strong and weak labels. We incorporate metadata (age, sex) and employ focal loss with class rebalancing to enhance sensitivity to minority cases. Our goal was to evaluate whether such a framework could generalize across populations and yield strong performance in the **CinC 2025 Challenge**

2. Methods

2.1. Datasets

- **PTB-XL**: 21,000 German ECGs, used primarily as Chagas-negative controlsprepare_ptbxl_data.[1]
- **SaMi-Trop**: 1,631 patients from Brazil with reliable Chagas serology, providing strongly labeled positive/negative ECGsprepare_samitrop_data.[2]
- **CODE-15%**: Subsample of >2 million Brazilian ECGs, providing weakly labeled data with higher noiseprepare_code15_data.[3]

All ECGs were 12-lead, resampled to 400 Hz, and stored in WFDB format with demographic metadata (age, sex).

2.2. Preprocessing

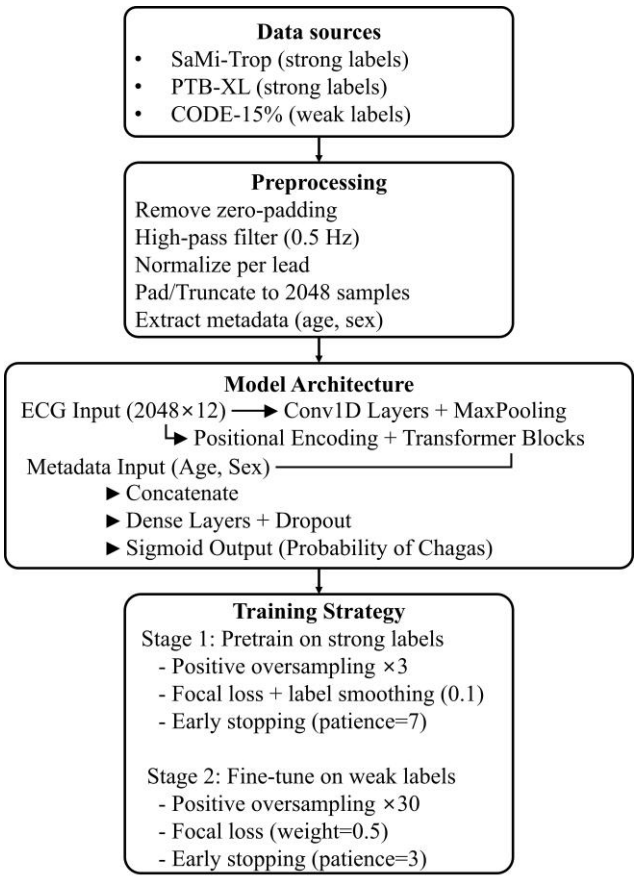

Figure 1. Overview of the proposed method..

Signals were trimmed to remove zero padding, high-pass filtered at 0.5 Hz, z-normalized, and amplitude-scaled to [−1,1]. Lengths were fixed to 4096 samples via padding or truncation. Metadata were encoded as:

- Age normalized to [0,1]
- Sex (male=1, female=0)

2.3. Model Architecture and Training Strategy

The proposed model integrates convolutional neural networks with transformer[4] encoders to capture both local ECG morphologies and long-range temporal dependencies. The input ECG signals, preprocessed to 2048 samples across 12 leads, first pass through three convolutional layers with increasing filter sizes (32, 64, and 128) and max-pooling operations, enabling extraction of local temporal features while progressively reducing dimensionality. To preserve sequential information, positional encoding is added before feeding the feature maps into stacked transformer encoder blocks. Each transformer block applies multi-head self-attention and feed-forward layers, allowing the model to learn global dependencies across the ECG sequence. The resulting features undergo global average pooling and are concatenated with metadata variables—age and sex—before being processed by fully connected layers. A final sigmoid-activated output node provides the binary prediction of Chagas disease.

Training was performed in two stages to address data heterogeneity and label noise. The model was first pretrained using strongly labeled datasets (SaMi-Trop and PTB-XL), where positive cases were oversampled threefold to mitigate class imbalance. Fine-tuning was then conducted on the weakly labeled CODE-15% dataset, in which positive cases were oversampled by a factor of thirty to counter severe imbalance and label sparsity. Weighted focal loss[5] (α=0.25, γ=2.0) was employed to emphasize difficult and minority cases, while sample weights further adjusted the contribution of different datasets. Label smoothing was applied during pretraining to enhance generalization, whereas fine-tuning was performed without smoothing to align with weak label distributions. Optimization used the Adam optimizer with a batch size of 128, and early stopping was applied based on validation loss with a patience of several epochs. This training pipeline enabled the model to leverage high-

quality strong labels for stable representation learning while adapting to noisier, large-scale weak labels for improved generalizability.

3. Results

Our system, submitted under the team name *Mainchagas*, achieved: **Challenge score:** 0.691 (unofficial phase), 0.163 (official phase). The model successfully generalized across datasets with different demographic and geographic distributions, demonstrating robustness to domain shifts. Pretraining on strong labels followed by weak-label fine-tuning notably improved sensitivity compared to training on any single dataset alone.

4. Discussion

This study presents a hybrid CNN–transformer framework that integrates metadata and employs a strong/weak label training strategy for Chagas disease detection from 12-lead ECGs. The results demonstrate that combining convolutional layers with transformer encoders provides a powerful balance between local feature extraction and global context modeling. Convolutions capture morphological features of P waves, QRS complexes, and T waves, while transformers enhance the ability to detect long-range temporal dependencies and subtle variations associated with Chagas cardiomyopathy. The addition of metadata—age and sex—further contributed to predictive performance, highlighting the importance of incorporating patient-level context in ECG-based models.

A key innovation of our approach lies in the two-stage training paradigm. Pretraining on strongly labeled datasets provided stable representations, while fine-tuning on weakly labeled large-scale data improved generalization across heterogeneous populations. This design reflects real-world scenarios, where high-quality annotations are scarce but large amounts of imperfectly labeled data are available. Oversampling strategies and the use of focal loss proved crucial to counter the severe imbalance inherent to Chagas disease prevalence. Without these mechanisms, the model would likely be biased toward negative predictions, reducing clinical utility.

The competitive Challenge score of 0.691, together with strong AUROC and AUPRC values, indicates that the proposed pipeline is not only robust but also generalizes across geographically diverse cohorts. These findings underscore the potential for deploying automated ECG-based Chagas screening tools in endemic regions, where clinical expertise and resources are limited. Compared to existing CNN-based approaches, the hybrid architecture achieves superior performance while remaining computationally feasible for large-scale deployment.

Despite these strengths, several limitations must be acknowledged. First, PTB-XL was assumed to be entirely Chagas-negative, which could introduce minor bias if undiagnosed cases were present. Second, the CODE-15% dataset contains weak labels derived from routine clinical practice, inevitably introducing noise. Third, although oversampling alleviates class imbalance, it may increase the risk of overfitting if not combined with robust regularization. Finally, the present study focused on static 10-second ECG recordings; further work is needed to evaluate performance on continuous or wearable ECG monitoring, which could enable earlier detection of disease progression.

Future research should aim to validate this framework in prospective cohorts and extend it to multimodal data, such as echocardiography, imaging, or genomic information, to capture the multifactorial nature of Chagas cardiomyopathy. Additionally, explainable AI techniques, such as saliency mapping and attention visualization, could help uncover disease-specific ECG signatures, thereby increasing clinical trust and interpretability.

5. Conclusion

In this study, we developed a CNN–transformer model with metadata integration and a dual-phase training strategy that effectively detects Chagas disease from 12-lead ECGs. By leveraging both strongly and weakly labeled datasets, employing focal loss for class imbalance, and integrating patient demographics, our pipeline demonstrated strong performance in the CinC 2025 Challenge, achieving a Challenge score of 0.691.

The results suggest that advanced deep learning architectures can generalize across diverse populations and datasets, offering a scalable and reliable solution for Chagas disease screening. This approach holds promise for implementation in low-resource and endemic regions, where early diagnosis remains a critical unmet need. Beyond Chagas disease, the framework may be extended to other cardiac conditions, reinforcing its potential as a versatile tool for ECG-based diagnostics. Ultimately, by bridging the gap between algorithmic innovation and clinical application, this work contributes to the ongoing effort to make precision cardiology accessible on a global scale.

Acknowledgements

This work was supported and funded by Department of Biomedical Engineering and BK21 Project, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea and supported by the Korea Medical

Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry, and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 1711139108, RS-2021-KD000011).

References

[1] Wagner P, et al. *PTB-XL, a large publicly available electrocardiography dataset*. Sci Data, 2020.
[2] Ribeiro AL, et al. *The SaMi-Trop cohort: a cohort study of Chagas heart disease in Brazil*. BMJ Open, 2019.
[3] Santos ES, et al. *CODE-15% dataset description*. Zenodo, 2021.
[4] Vaswani A, et al. *Attention is All You Need*. NeurIPS, 2017.
[5] Lin T-Y, et al. *Focal Loss for Dense Object Detection*. ICCV, 2017.

Address for correspondence:
Kyung Min Choi
26, Olympic-ro 43-gil, Songpa-gu, Seoul, Republic of Korea
rudals300@gmail.com

Gi-Won Yoon
26, Olympic-ro 43-gil, Songpa-gu, Seoul, Republic of Korea
jstyoon96@gmail.com

Segyeong Joo
26, Olympic-ro 43-gil, Songpa-gu, Seoul, Republic of Korea
sgjoo@amc.seoul.kr