Novel N-BEATS Architecture for Classification of Chagas Disease

Bartosz Puszkarski, Krzysztof Hryniów, Marcin Iwanowski

Warsaw University of Technology, Warsaw, Poland

Abstract

Chagas disease is a tropical parasitic disease endemic to South America. Widespread serological testing is impossible due to limited capacities. One of the solutions is to make an initial screening based on electrocardiogram (ECG) results. To classify Chagas disease, we propose a multi-branch architecture for Neural Basis Expansion Analysis for Interpretable Time Series (N-BEATS), previously used for fast classification of cardiovascular diseases. N-BEATS is rarely used for classification purposes, but we want to test for what cases it can be a valid solution and an alternative to other state-of-the-art networks.

Multi-branch architecture allows for better use of the metadata and signal modifications (like wavelet transformations or drift removal) while not losing the essential elements of pure ECG signal, making it suited for 'Detection of Chagas Disease from the ECG: The George B. Moody PhysioNet Challenge 2025'.

For performance comparison, we show the effects of the exact modification on two other state-of-the-art networks: long-short-term memory (LSTM) and convolutional neural network (CNN).

Our team's (WEAIT) submissions for NBEATS, LSTM, and CNN scored 0.068, 0.195, and 0.237, respectively, on the validation set.

1. Introduction

Chagas disease, caused by the protozoan parasite Trypanosoma cruzi and transmitted primarily by triatomine insects, remains one of the most neglected tropical diseases, affecting an estimated 7 million people and causing nearly 10,000 deaths yearly in endemic regions. In the advanced stages of infection, Chagas disease can lead to the development of cardiomyopathy, which may result in heart failure, arrhythmias, and thromboembolic events, significantly increasing the risk of mortality. While serological testing remains the standard method for diagnosing Chagas disease and has revealed high prevalence in certain regions, access to such testing is often limited. Also, in non-endemic, high-income countries, where awareness and laboratory-based serological testing are limited, migrant populations often go untested and unmonitored, allowing chronic infection to progress silently [1, 2]. As Chagas-related cardiac abnormalities–like right bundle branch block or left anterior fascicular block, which correlate with positive diagnosis – have frequently appeared in electrocardiograms (ECGs) over the past decade, electrocardiography has emerged as a promising, low-cost, widely available tool for screening chronic Chagas cardiomyopathy [3]. Exploring the opportunities and challenges related to the classification of ChD based on the ECG signal was the subject of the PhysioNet Challenge 2025 [4] to which the research described in this paper was submitted. Tackling the problem from an ECG analysis perspective, residual neural networks (ResNets) have been employed in order to recognize major signal abnormalities related to chronic Chagas cardiomyopathy [5]. In this paper, we aim to explore the possibility of employing a smaller and simpler Neural Basis Expansion Analysis for Time Series - N-BEATS [6], embedded within an architecture of multiple independent branches [7], for the task of automated Chagas disease binary classification based on a 12-lead ECG signal.

2. Methods

Neural networks are widely employed for the automated classification of cardiovascular diseases from ECG signals. Architectures are often based on ResNets, CNNs, or RNNs, which tend to be relatively complex. To balance simplicity with interpretability, the N-BEATS network was selected as the core architecture, as in previous studies, it has demonstrated performance comparable to widely used RNNs such as LSTM and GRU [8, 9]. To further examine the contribution of different signal features and to provide multiple perspectives on the ECG data, a multi-branch design was adopted, which has been shown to improve performance in multiclass CVD classification [7]. The N-BEATS, tested during the unofficial phase of the challenge, yielded very promising results–it was scored 17th (out of more than 173 entries) with a 0.605 challenge score. The LSTM model in the same multi-branch architecture achieved a challenge score of 0.294 – which justified continuing to work with this architecture for the Chagas diagnosis problem.

2.1. The N-BEATS model

Neural Basis Expansion Analysis for Interpretable Time Series (N-BEATS), introduced by Oreshkin et al. in 2019 [6], is a forecasting architecture built on a deep stack of fully connected layers that use ReLU activations. Each block predicts basis-expansion coefficients in both the forward (forecast) and backward (backcast) directions, and the blocks are linked through a doubly residual stacking scheme; some layers handle backcast and forecast roles simultaneously. After a block subtracts its backcast component from the input, the residual passes to the next block, allowing subsequent blocks to model portions of the signal not yet explained. The architecture presented in Figure 1 shows how layers, blocks, and stacks form the doubly residual connections. In [8], researchers repurposed N-BEATS, previously used only for forecasting, for cardiac-diagnosis tasks and compared it with LSTM and GRU models. They found comparable performance and particularly strong results when only a few ECG leads were available.

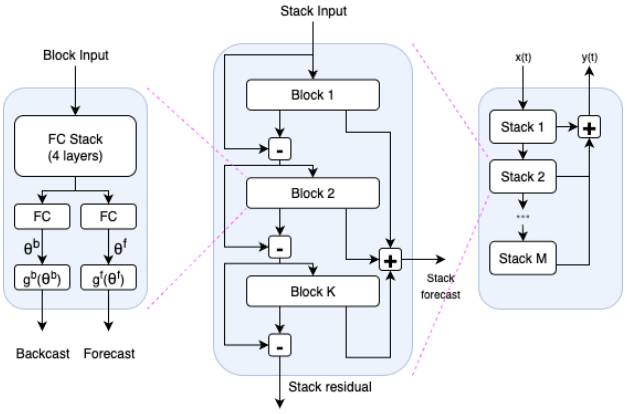

Figure 1. N-BEATS architecture diagram.

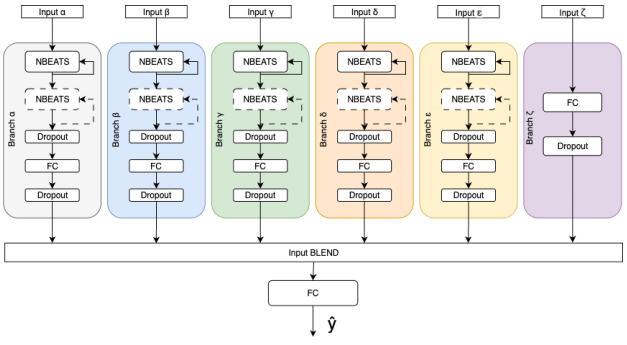

Figure 2. Multi-branch N-BEATS architecture diagram.

2.2. Multi-branch architecture

Building on the multi-branch strategy described in [7], which integrates six parallel branches, we introduce a modification to this architecture by simplifying branch ζ. In this branch, the core network was replaced with a single fully connected layer. This design choice reflects the nature of the inputs: branch ζ processes patient metadata and general recording statistics, which are less directly related to the ECG signal than the windowed input data used in the other branches. Instead of merging domain knowledge, patient information, auxiliary transformations, and raw ECG signals into a single input stream, each source is provided to the network independently (see Fig.3). The resulting model consists of six distinct branches, each with its own structure and dedicated input (Fig.2). Throughout this paper, these branches are denoted as α, β, γ, δ, ϵ, and ζ.

2.3. Preprocessing

The difference in the number of representatives between the positive and negative classes was vital to the challenge, as negative cases heavily outnumbered positive ones. Because disk space was limited to 100GB, we had to refrain from using all records available, as the size of the created HDF5 database exceeded the limit. In order to build a training and test dataset, a random pre-selection of records was performed to reduce their number. All records from the SaMi-Trop dataset were incorporated, from CODE-15% and PTB-XL, only 5% and 10%, respectively; randomly selected records were taken into consideration. All data was split in a 2:1 ratio between the training and test datasets, ensuring the positive label ratio between the datasets is also 2:1. Established approaches from recent state-of-the-art studies were incorporated as inputs to the individual architectural branches. In each branch, the raw ECG signal was subjected to a denoising stage implemented through wavelet thresholding. For this purpose, the Daubechies wavelet was employed, as it remains the most widely adopted basis for ECG signal denoising [10]. The *wavedec* method from the `PyWT` library was employed to extract the corresponding wavelet coefficients. Subsequently, a robust median estimator of the standard Gaussian noise level was computed and utilised to determine the Bayesian Shrink threshold [11]. This thresholding procedure was applied to the detail coefficients obtained in the initial wavelet decomposition, after which the signal was reconstructed from the modified coefficients. All signals were resampled to a unified target frequency of 400 Hz to ensure consistency across datasets, as most data comes with said sampling rate. For recordings sampled at frequencies other than 400 Hz, resampling was performed using the *resample* function from the `scipy.signal`

library, which applies Fourier-based interpolation to obtain uniformly spaced samples at the desired target frequency. Given the diagnostic importance of the temporal region surrounding the R-peak in identifying cardiovascular diseases (CVDs) [12], an R-peak detection algorithm was employed to guide signal segmentation. In the present study, the widely adopted Pan-Tompkins algorithm [13] was implemented to perform both R-peak detection and subsequent segmentation. Following detection, the ECG recordings were partitioned into windows of 3s in duration, each comprising 1 s of signal preceding the R-peak and 2 s following it. As illustrated in Figure 3, the inputs to the β, γ, δ, and ϵ branches were derived by first eliminating baseline drift from the denoised ECG signal. The baseline component was estimated using the SNIP algorithm, and the signals were subsequently re-centred by subtracting this estimate, thereby normalising them to a zero-line baseline. ζ input differs from others as it expresses basic patient-related features based on the information from the header file; we extract age and sex, which are later one-hot-encoded. Additionally, the signal mean and standard deviation are extracted for all leads.

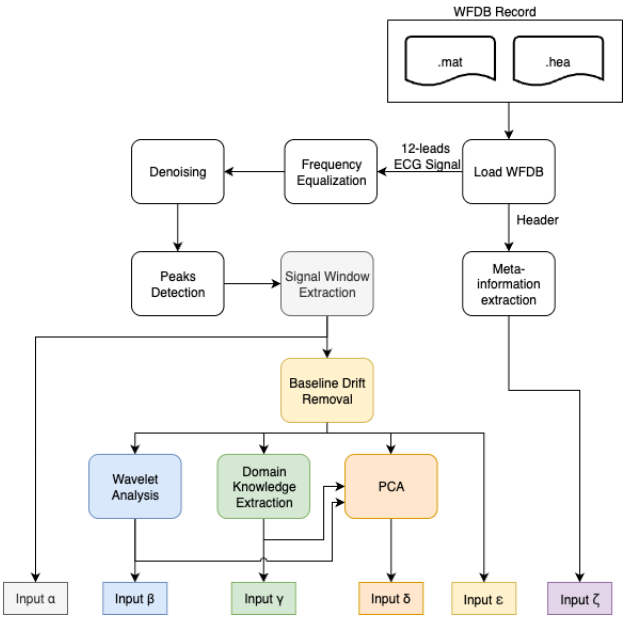

Figure 3. Diagram of ECG Signal preprocessing flow from data read to forming branch inputs.

2.4. Training and Hyper-Parameters

Throughout the training phase, a range of network hyperparameters, as summarized in Table 1, was systematically evaluated. The configurations highlighted in bold demonstrated superior performance across the conducted experiments and were therefore identified as the most effective settings. These selected hyperparameters were subsequently adopted as the parameters of choice for the experimental runs discussed in Section 3. It is also crucial to note that the training hyperparameters and network configuration were run with the following hardware limitations imposed by the challenge: 16 vCPU, 64GB RAM, 16 GB NVIDIA T4 GPU, and 100GB of disc storage space.

Table 1. Network hyperparameters. Bolded values indicate the configurations that achieved the best performance.

Parameter	Values
Optimizers	Adam
Learning rates	0.1, 0.01, **0.001**, 0.0001
Number of epochs	10, **30**, 50, 100
Early stop	5, **10**, 20
Loss Function	Binary Cross Entropy with Logistic Loss
Positive weight	**yes**, no
Batch size	100, **200**, 300, 400, 500
Dropouts	No, 0.1, 0.2, 0.3, **0.4**

3. Results

During the official phase of the challenge, we successfully ran multiple experiments testing N-BEATS, LSTM, and CNN-based architectures. Best results for each network are presented in Tab. 2.

Table 2. Scores on the validation set of the official phase.

Submission ID	Network	Score
2064	N-BEATS	0.068
2095	LSTM	0.195
2248	CNN	0.237

4. Discussion and Conclusions

Classifier performance varied depending on the network used, with scores ranging from 0.068 to 0.237 on the hidden validation set. Comparing to widely used LSTM and CNN architectures, the N-BEATS multi-branch architecture maintains a simpler design due to its reliance solely on fully connected layers. Comparing with our previous approach using multi-branch architecture [8] where performance differences also narrowed when the number of input leads was reduced, problem complexity plays a vital role. We expected, that in binary classification, such input reduction is unnecessary, and N-BEATS performs competitively when trained on the full input data. Although our fine-tuned submissions failed to run successfully in the challenge environment, we ran additional experiments locally. Local cross-validation results for multi-branch architecture is visible as first row of Table 3 and it shows, that

N-BEATS can obtain comparable results to other widely used networks.

To assess if the multi-branch design improves N-BEATS performance, a series of additional local studies was conducted, where cross-validation scores for the multi-branch architecture were compared to results obtained with single-branch variants. Cross-validation scores are summarized in Table 3. Local evaluations were obtained using the *evaluate_model.py* script provided as part of the Challenge code base [14].

Table 3. Comparison between multi-branch and single-branch architecture utilising N-BEATS network.

Branch	Score	AUROC	Acc	F2
All	0.251	0.742	0.024	0.005
α	0.040	0.636	0.023	0.004
β	0.259	0.698	0.092	0.135
γ	0.138	0.591	0.023	0.004
δ	0.061	0.505	0.023	0.004
ϵ	0.236	0.709	0.029	0.017
ζ	0.05	0.500	0.979	0.968

Based on the extended metrics collected during cross-validation, as shown in Table 3, it is evident that the N-BEATS multi-branch architecture yields significantly better results than applying the same network solely to the raw signal input, represented by branch α. Notably, when compared against the performance of all individual branches, branch β, which corresponds to the wavelet analysis coefficients, achieved the highest challenge score. However, relative to the multi-branch approach, the performance of branch β exhibited a markedly lower AUROC. Branch ζ performed the worst when evaluated independently; nevertheless, it achieved the highest accuracy and F-measure scores. The model utilizing only branch ζ effectively learned to predict only negative responses, which—given the statistical distribution of the positive class—resulted in the highest accuracy.

Further investigation into the more extensive use of wavelet analysis outputs would be required. Moreover, the blending fully connected layer could potentially be replaced with a multi-head attention layer to examine whether the network could mitigate the negative influence of underperforming branches.

References

[1] Rassi A, Marin-Neto JA. Chagas disease. The Lancet 2010; 375:1388–1402.

[2] Bern C. Chagas' disease. N Engl J Med 2015;373(5):456–466.

[3] Brito B, Ribeiro A. Electrocardiogram in chagas disease. Rev Soc Bras Med Trop 2018;51:570–577.

[4] Reyna MA, Koscova Z, Pavlus J, Weigle J, Saghafi S, Gomes P, Elola A, Hassannia MS, Campbell K, Bahrami Rad A, Ribeiro AH, Ribeiro AL, Sameni R, Clifford GD. Detection of Chagas Disease from the ECG: The George B. Moody PhysioNet Challenge 2025. Computing in Cardiology 2025;52:1–4.

[5] Jidling C, Gedon D, Schön TB, Oliveira CDL, Cardoso CS, Ferreira AM, Giatti L, Barreto SM, Sabino EC, Ribeiro ALP, Ribeiro AH. Screening for chagas disease from the electrocardiogram using a deep neural network. PLOS Neglected Tropical Diseases July 2023;17(7):e0011118. ISSN 1935-2735.

[6] Oreshkin BN, Carpov D, Chapados N, Bengio Y. N-beats: Neural basis expansion analysis for interpretable time series forecasting. CoRR 2019;abs/1905.10437.

[7] Hryniów K, Puszkarski B, Iwanowski M. The multi-branch deep-learning-based approach to heart dysfunction classification. Applied Sciences 2025;15(15). ISSN 2076-3417.

[8] Puszkarski B, Hryniów K, Sarwas G. N-beats for heart dysfunction classification. In 2021 Computing in Cardiology (CinC), volume 48. 2021; 1–4.

[9] Puszkarski B, Hryniów K, Sarwas G. Comparison of neural basis expansion analysis for interpretable time series (n-beats) and recurrent neural networks for heart dysfunction classification. Physiological Measurement jun 2022; 43(6):064006.

[10] Zaynidinov H, Juraev U, Tishlikov S, Modullayev J. Application of Daubechies Wavelets in Digital Processing of Biomedical Signals and Images. Springer Nature Switzerland. ISBN 9783031538278, 2024; 194–206.

[11] Donoho DL, Johnstone IM. Ideal spatial adaptation by wavelet shrinkage. Biometrika 09 1994;81(3):425–455. ISSN 0006-3444.

[12] Saadatnejad S, Oveisi M, Hashemi M. LSTM-based ECG classification for continuous monitoring on personal wearable devices. IEEE Journal of Biomedical and Health Informatics JBHI Feb. 2019;2:515–523.

[13] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering 1985;BME-32(3):230–236.

[14] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet. Circulation 2000;101(23):e215–e220.

Address for correspondence:

Bartosz Puszkarski
Plac Politechniki 1, Warszawa, 00-661
bartosz.puszkarski.dokt@pw.edu.pl