N-BEATS for Heart Dysfunction Classification

Bartosz Puszkarski[1], Krzysztof Hryniów[2], Grzegorz Sarwas[2]

[1] Faculty of Electrical Engineering, Warsaw University of Technology, Warsaw, Poland
[2] Institute of Control and Industrial Technology, Warsaw University of Technology, Warsaw, Poland

Abstract

Introduction: Recurrent Neural Networks are useful tools for the prediction and classification of ECG problems. The most commonly used network for such a solution is Long Short-Term Memory (LSTM) architecture. This study aims to assess if another state-of-the-art solution, Neural Basis Expansion Analysis for Interpretable Time Series (N-BEATS) can be adopted to diagnose the same cardiac problems. In addition, a comparison is conducted for a different number of electrocardiogram leads.

Methods: Two architectures were tested for performance and dimension reduction problem, both in variants consisting of blended branches, that allow retaining accuracy while reducing the computational capacity needed.

Results: Due to a flaw in the challenge metric function, only results from cross-validation are presented. LSTM outperforms N-BEAST in terms of multi-label classification, data set resilience and obtained challenge metrics. Still, N-BEATS can obtain acceptable results and clearly outperforms LSTM in terms of complexity and speed.

Conclusions: This paper features a novel approach of using the N-BEATS, which was previously used only for forecasting, to classify ECG signals with success. While N-BEATS multi-label classification capacity is lower than LSTM, its speed allows it to be used on lower classes of wearable devices.

1. Introduction

Cardiovascular diseases (CVDs) are still one of the most common causes of death globally. An electrocardiogram (ECG) signal representing the heart's electrical activity is widely used to detect and classify cardiac arrhythmias. Detection can be done by visual inspection of ECG waveform, but as at early stages, they occasionally occur, continuous monitoring by wearable devices is used.

Three leading solutions are used: offline processing of stored signals, remote processing on cloud servers, and local execution on a wearable device. The first solution is the oldest and allows easy classification using the whole dataset but cannot be done for real-time detection. The second solution is now increasingly common, allowing to use of the power of cloud computing, but raises privacy issues. The final solution allows non-stop operation and real-time detection, regardless of network coverage. The main problem with such a solution is that the automatic ECG classification algorithm needs to be lightweight while retaining accuracy. This paper, pursues a solution that works both on server processors, GPU units, and low-powered processors on wearable devices.

Before, algorithms based on morphological features and classical signal processing techniques were used. However, fixed solutions proved insufficient due to the ECG waveform's morphological characteristics variance between patients and circumstances of measurement [1] Due to that, deep-learning-based algorithms using recurrent neural networks (RNNs) and convolutional neural networks (CNNs) were introduced [1–4]. ECG prediction and classification are based on the problem of analysis of time-series, where the most commonly used classifier is the Long Storm-term Memory (LSTM) network, the same network was used as a basis for this study. 2020 PhysioNet/ Computing in Cardiology Challenge also showed that the most common solutions used for 12-lead classification were also CNNs and RNNs [5]. Neural Basis Expansion Analysis for Interpretable Time Series (N-BEATS)[6, 7] is one of the newest RNNs, used for forecasting time series.

The article features a version of the LSTM algorithm based on [8], consisting of additional wavelet analysis and merged predictions from smaller models to lower the computational cost. N-BEATS network was modified in the same way, with added two blended sub-networks, wavelet analysis. Furthermore, it was adapted into a multi-label classifier.

This study aims to assess if N-BEATS can be utilized to diagnose cardiac problems as commonly used LSTM. In addition, a performance comparison is conducted for a different number of electrocardiogram leads, as obtaining the same accuracy with a reduced number of leads allows for arrhythmias detection and classification while using off-the-shelf wearable devices (Holter monitors, sport bands, etc.).

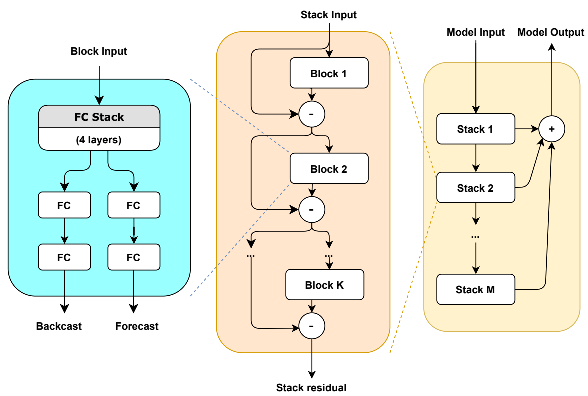

Figure 1. Model of N-BEATS network.

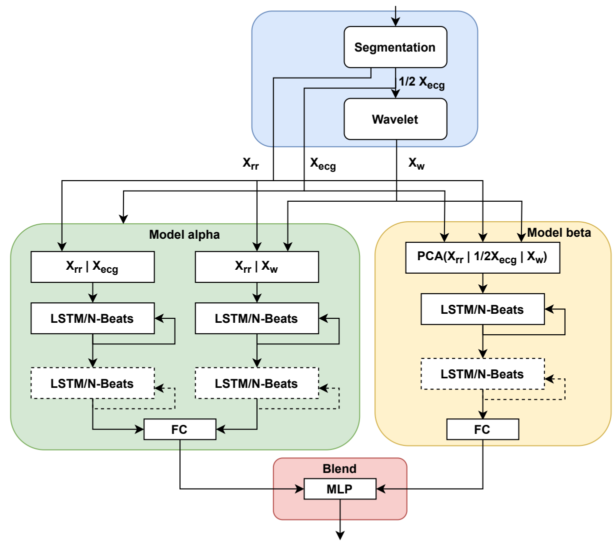

Figure 2. Model of classifier algorithm using LSTM / N-BEATS network.

2. Methods

Time series forecasting is an important problem in the machine learning field. Despite the rapid development of machine learning algorithms, classical statistical methods are used to solve the problems of this class. Increasingly, the use of hybrid solutions using neural residual/attention extended LSTM stack with the classic Holt-Winters statistical model [9] with learnable parameters. In 2020 Oreshkin et al. propose N-Beats architecture for interpretable time series forecasting [6]. The model of this architecture is presented in the Figure 1. The basic building block of this architecture is a multi-layer fully connected (FC) network with nonlinearities provided by activation function ReLU. It predicts basis expansion coefficients both forward (forecast) and backward (backcast). Blocks are organized into stacks using the doubly residual stacking principle. A stack may have layers with shared backcast and forecast blocks. Forecasts are aggregated in a hierarchical fashion, which enables building a very deep neural network with interpretable outputs.

From the point of ECG signal analysis, the essential part of the signal for CVDs diagnosis is part around R peak [8, 10], so R peak detection algorithm is used for signal segmentation. Pan-Tompkin's algorithm [11] is used for the R peak detection and segmentation process. Based on R peak detection, digitized ECG samples are segmented into a sequence of heartbeats containing precisely 0.7 seconds of an input signal (0.25 seconds before R peak and 0.45 seconds after R peak). This signal is denoted as X_{ecg} in Fig. 2. Following the algorithm presented in [8], as part of Pan-Tompkin's algorithm, additional information is obtained in the form of X_{rr} feature vector containing three features of heartbeat: $X_{rr}[1]$ is last peak's interval, $X_{rr}[2]$ is next peak's interval and $X_{rr}[3]$ is the average duration of five past and four next intervals. First-in-first-out (FIFO) memory is used for buffering ten ECG signals, giving minimal delay allowing the use of the algorithm in real-time.

Type 2 Daubechies wavelet transform (db2) with four levels of decomposition is applied to downsampled (by a factor of two) digitalized ECG samples to capture time and frequency domain information, resulting in second input vector $X_w = (A4, D4, D3, D2, D1)$.

As shown in Figure 2, both algorithms use two models (donated as α model and β model) consisting of multiple parallel RNNs that are blended using a multi-layer perceptron (MLP) network with two hidden layers. The α model processes X_w and X_{ecg} inputs in separate branches consisting of parallel RNNs. Outputs of those branches are concatenated and used as input for a fully connected neural network layer (FC) to produce probabilities of all output classes. In the β model, the principal component analysis (PCA) is used on a downsampled version of X_{ecg} that is concatenated with X_w. The result is used as input for RNNs, and extracted features form the input for FC neural network layer. Results from both models are blended using MLP to obtain the final probabilities of output classes.

3. Experiments

In the conducted experiments, we used the N-Beats and LSTM architectures in the configurations presented in Table 4. The experiments were performed for signals consisting of 2, 3, 4, 6, and 12 ECG leads. For tests were taken only 26 classes, defined by the challenge organizer, which described 30 heart diseases. The main goals of the research were to check the possibility of using the N-Beats model in the classification process, comparing the capacities of this architecture with the LSTM architecture, determining the required number of leads for the correct diagnosis of as many diseases as possible, and determining the computational complexity of the analyzed algorithms and their inference times. It was done to determine a solution that will allow the diagnosis of CVDs while retaining the speed and low computational capacity needed to use it for real-time detection on an off-the-shelf wearable device.

Our experiments were conducted on the training data sets include annotated twelve-lead ECG over 88,000 ECGs shared publicly. For results comparison, we used cross-validation techniques. The whole dataset was divided using five folds for the training and testing datasets. We randomized the validation dataset for each training dataset to select a number of epoch for the training process. For this purpose, we initialized network weights using frozen seed. We searched for the best number of epochs for a given training set based on the early stopping technique and validation dataset. Then, we fixed obtained number of epochs and trained the model for all training data in this fold using network weights initialization for frozen seed. The achieved model was tested by calculating average accuracy, F-measure, and challenge metric.

Table 1. Average scores (and score standard deviation) obtained by LSTM and N-BEATS.

Leads	LSTM			N-BEATS		
	avg	max	stdDev	avg	max	stdDev
12	0.411	0.426	0.012	0.247	0.365	0.066
6	0.360	0.402	0.058	0.307	0.352	0.061
4	0.361	0.430	0.066	0.300	0.358	0.055
3	0.331	0.433	0.099	0.265	0.326	0.039
2	0.359	0.416	0.058	0.326	0.365	0.047

As can be seen in Table 1, LSTM achieved the best results for 12 leads, with a degrading challenge metric score while decreasing the number of leads. For all tests, obtained results were with minor variation, as shown by low standard deviation. That means the LSTM model is resistant to the selection of training set. N-BEATS fared much worse, with a high standard deviation and lower challenge score for all numbers of leads. Best results (0.326-0.365) were on par with the average obtained by LSTM for 2-6 leads and below worst result obtained by 12 lead LSTM. Still, even the worst result achieved with use of N-BEATS has challenge score above 0.21.

Remark 1: An interesting observation is that adding the third lead (the chest V2 lead) obtained a lower challenge metric for nine of the ten test sets while obtaining the second-best result for the tenth test set.

Remark 2: There are no published official challenge scores for both tested architectures as in the organizer's challenge metric function; an error ignored float values and returned the lowest possible score. We found it after the end of an official phase, and as our submitted networks were not tested again and the hackathon was canceled, there is no way to determine how the networks will fare against the hidden set.

The main advantage of the N-BEATS solution compared to LSTM is low complexity and higher speed, as can be seen in Table 2. N-BEATS needed on average 20 times less time to classify given 1-peak sample.

Table 2. Average challenge scores and time classification times for one peak.

Leads	LSTM		N-BEATS	
	score	time(ms)	score	time(ms)
12	0.411	3.44	0.247	0.191
6	0.360	3.37	0.307	0.167
4	0.361	2.82	0.300	0.165
3	0.331	2.46	0.265	0.162
2	0.359	2.22	0.326	0.156

As can be seen in Table 3 both accuracy and F-score of tested models are low. It is because to maximize challenge metrics, only diseases belonging to 26 classes that were part of the metric were learned by the networks. Networks that were learned for all existing classes in the dataset were penalized by how the challenge metric was calculated and not presented in the paper.

Table 3. Average accuracy and average F-score for LSTM and N-BEATS models.

Parameters	LSTM		N-BEATS	
	Accuracy	F-score	Accuracy	F-score
12	0.020	0.270	0.002	0.116
6	0.016	0.243	0.005	0.184
4	0.013	0.224	0.002	0.178
3	0.011	0.209	0.001	0.160
2	0.012	0.237	0.003	0.194

4. Conclusions

In the paper, we analyzed the ability to utilization N-Beats architecture for multi-label classification task. Conducted research was focused on the architecture performance and influence of ECG leads reduction for classification results. Achieved results were compared with the results obtained from the LSTM. Due to a flaw in the challenge metric function, only results from cross-validation are presented. Based on them, we can say that while LSTM outperforms N-BEATS in terms of multi-label classification, data set resilience, and challenge results, N-BEATS,

Table 4. Configuration of LSTM and N-BEATS architectures.

Param.	LSTM_{α_1}	LSTM_{α_2}	LSTM_{β}	$\text{N-Beats}_{\alpha_1}$	$\text{N-BEATS}_{\alpha_2}$	N-BEATS_{β}
input_{size}	12	12	1	12	12	1
hidden_{size}	17	17	1	17	17	17
num_{layers}	4	4	1	2	1	1
FC	[9197, 128]		[2256, 26]	[6492, 128]		[2256,26]
FC_2	[128, 26]			[128, 26]		
BLEND FC	[52, 26]			[52, 26]		

due to its greater speed and lower complexity, is a viable solution to be used on off-the-shelf wearable devices for real-time CVDs detection.

References

[1] Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1d convolutional neural networks. IEEE transactions on bio medical engineering August 2015; Vol. 63.

[2] Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine 2019; (25):65–69.

[3] Jun T, Nguyen HM, Kang D, Kim D, Kim D, Kim YH. Ecg arrhythmia classification using a 2-d convolutional neural network, 2018. URL `https://arxiv.org/abs/1804.06812v1`.

[4] Ribeiro AH, Ribeiro MH, Paix/˜ao GM, Oliveira DM, Gomes PR, Canazart JA, Ferreira MP, Andersson CR, Macfarlane PW, Jr MW, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. Nature Communications 2020;Vol. 11(1):1—9.

[5] Computing in Cardiology Challenge 2020 PN. Classification of 12-lead ECGs: the physionet/computing in cardiology challenge 2020, 2020. URL `https://physionetchallenges.org/2020/papers/2020ChallengePaper.pdf`.

[6] Oreshkin BN, Carpov D, Chapados N, Bengio Y. N-beats: Neural basis expansion analysis for interpretable time series forecasting, 2019. URL `https://arxiv.org/abs/1905.10437v4`.

[7] Oreshkin BN, Carpov D, Chapados N, Bengio Y. Meta-learning framework with applications to zero-shot time-series forecasting, 2020. URL `https://arxiv.org/abs/2002.02887v3`.

[8] 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.

[9] Smyl S. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting 2020;36(1):75–85. ISSN 0169-2070. M4 Competition.

[10] How the heart works. United States Department of Health & Human Services. URL `https://www.nhlbi.nih.gov/health-topics/how-heart-works`.

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

Address for correspondence:

Bartosz Puszkarski

Plac Politechniki 1, 00-661 Warsaw, Poland

puszkarski.bartosz@gmail.com