Teaching a Machine to Diagnose a Heart Disease

Beginning from digitizing scanned ECGs to detecting the Brugada Syndrome (BrS)

Simon Jaxy[1, 2] Isel Grau[2] Nico Potyka[1] Gudrun Pappaert[3] Catharina Olsen[3] Ann Nowé [2]

[1]University of Osnabrück [2]Vrije Universiteit Brussels [3]Universitair Ziekenhuis Brussel

Introduction

The Brugada Syndrome (BrS) is a severe cardiovascular disease that can lead to a sudden cardiac death following a swift loss of any cardiac function and therefore, posing an imminent threat to the human life. Particularly, this can appear in patients with structurally normal hearts [2].

Ever since its first description [2], only one clear diagnostic case is acknowledged, characterized by an anomaly in the electrocardiogram (ECG); an accentuation of the J wave found in the right precordial leads (V1, V2), which results in an ST-segment elevation that is often followed by a negative T-wave [2], see Figure 1.

Since the only diagnostic case of the BrS is known by the ECG, we concentrate ourselves on it while investigating the BrS from the aspect of computational means.

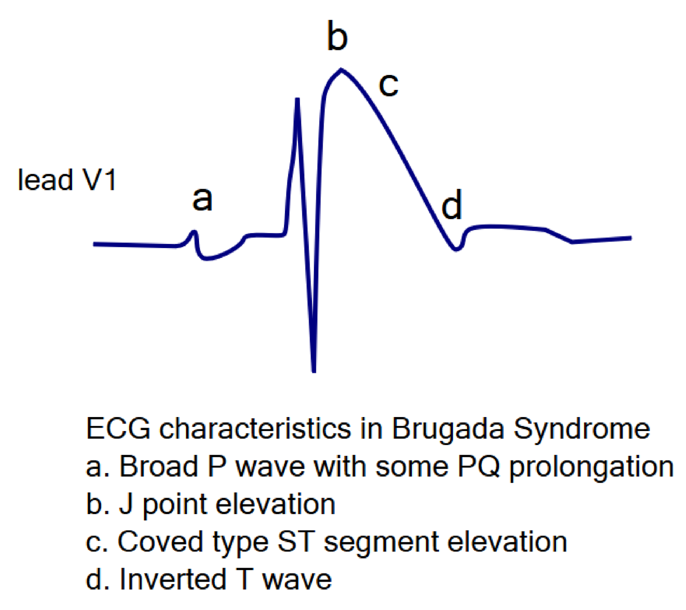

Figure 1: The BrS positive pattern [8].

In the following, we present an automated pipeline that transformed scanned images of ECGs to time-voltage data, which is then used as basis for our long short-term memory (LSTM) [6, 4] classifier able to differentiate BrS positive ECGs from negative ones.

The Brugada Syndrome

- Cardiovascular diseases are the number one cause of death world-wide, causing an estimated fatality of 17.9 million people or 31% globally [9].
- BrS is associated with 4% of the total count of cardiac death for all patients.
- BrS is the cause of 20% of the deaths in patients with structurally normal hearts [3].
- The disease was first described in 1992 by [2].
- A characterized by a **coved ST-segment elevation followed by a negative T wave** in one of the precordial leads [3].

The IMAGica Project

- An Integrative personalized Medical Approach for Genetic diseases.
- An hollistic approach towards analysing genetic diseases.
- It combines physiological, psycho-social and environmental data and their complex interactions to investigate cardiac arrhythmias, particularly, the Brugada Syndrome.
- The underlying thesis provides the blueprint towards a classifier that is specialized on the Brugada Syndrome able to make the diagnosis.
- In future research, the classifier will be refined and different methods of interpretability will be presented.

Digitization

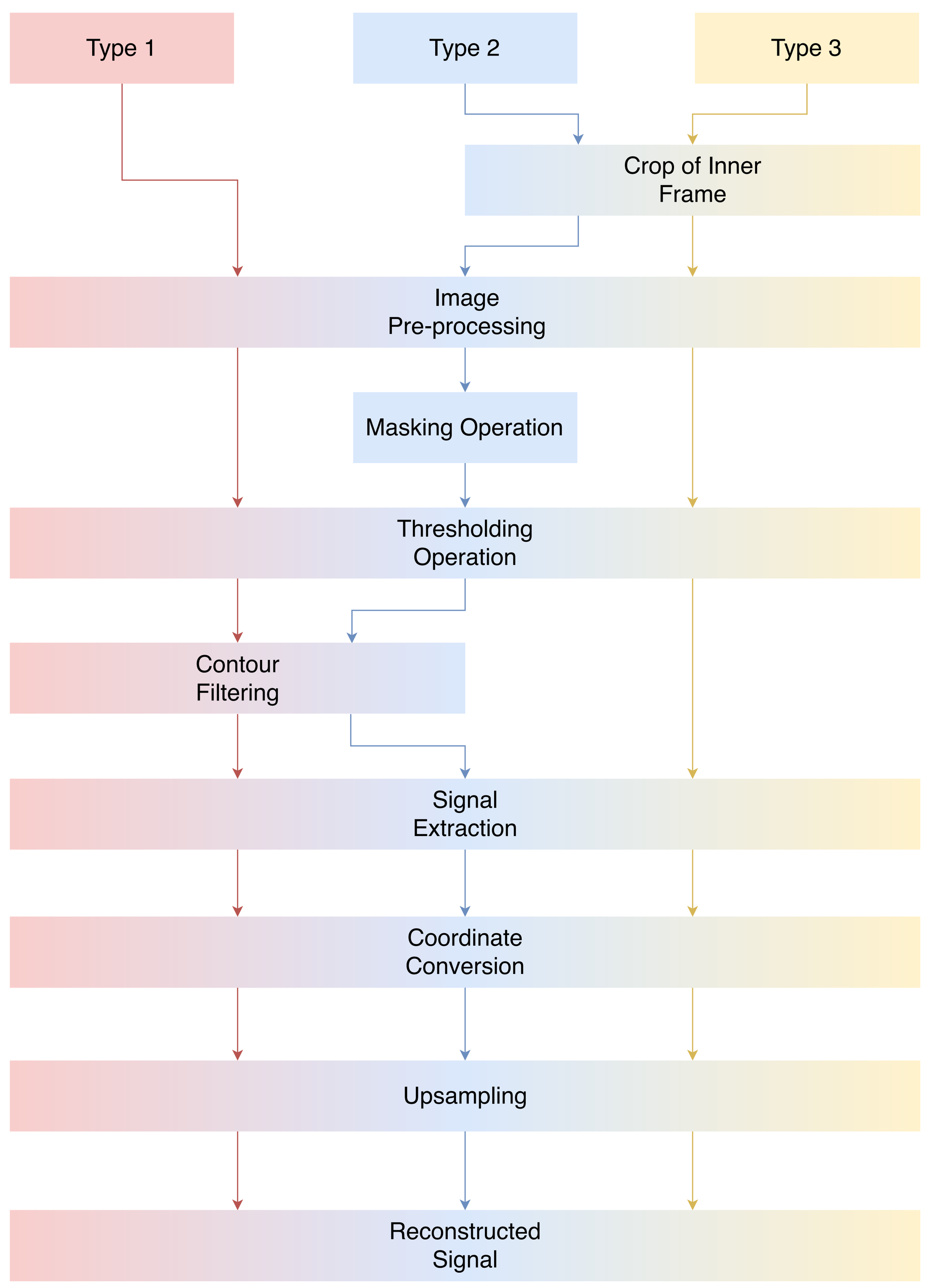

Figure 3: Digitization Pipeline

The digitization process follows an automatic pipeline that transforms scans of ECG images, comprising three distinctive image types (e.g., background and foreground color). A full description of the process can be found in the original thesis [7], the pipeline is displayed in Figure 3.

First, the images are gray-scaled and rotated if needed. Then, obstacles, such as a black frame surrounding the signals or the background grid, are removed.

Thirdly, we split the sequences into distinctive images by summing over the pixels in its columns. We use the minima in between peaks of pixels, each representing a single signal, as cut off points. Finally, every signal is upsampled and then mapped to time-voltage coordinates.

For most of the ECG leads, the pipeline preserves the signals, see Figure 4. Yet, some sequences cannot be separated, leading to distortion.

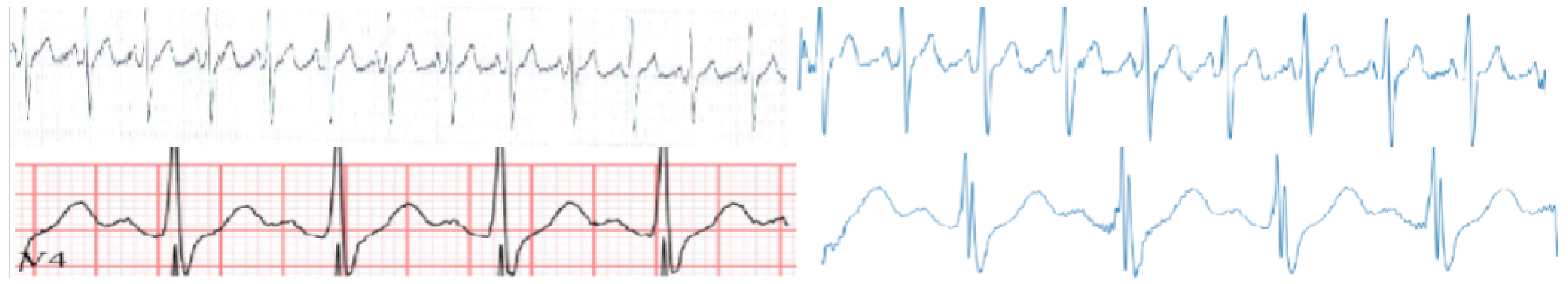

Figure 4: The outcome of the digitization process for two ECG image types. The top image displays a good result the bottom image a distorted one.

Classification

The classifier's task is to read-in ECG images and make a binary decision of whether it is BrS positive or negative. We gathered positive ECGs (30 in total) with our pipeline while extracting negative examples (80 in total) from the PTB Database of Physionet [1], [5]. Our model is composed of a single LSTM-layer followed by a dropout layer and a sigmoid activation function for classification.

		True		
		BrS+	BrS−	Total
Predicted	BrS+	22	45	67
	BrS−	0	24	24
	Total	22	69	91

Table 1: Classification Matrix of the LSTM-based model

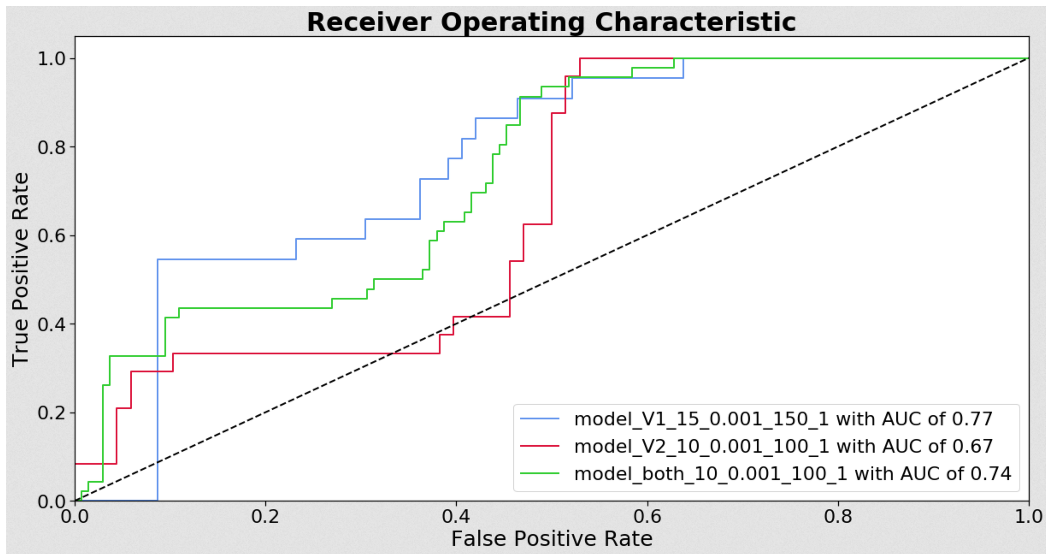

Figure 5: ROC curve of tested models. We selected the blue one based on V1 data.

Given the limited amount of data, the classifier scores a high amount of false positives, while avoiding any false negatives, see Table 1 and Figure 5.

In a medical context this might be favorable, as one would prefer to perform additional test rather to miss a diagnosis, however, before it becomes relevant for any practical purpose, further improvements and testing has to be conducted.

Conclusion

We presented an automated pipeline capable of transforming scanned ECGs to time-voltage data. Furthermore, we explore the capabilities of the LSTM-based classifier on differentiating BrS positive ECGs from negative ones. Further research and experimentation are needed for obtaining a classifier achieving better performance. Another aspect is to investigate the role of different segments of the ECGs in the classification as positive for BrS.

References

[1] Bousseljot, R., Kreiseler, D., Schnabel, A.: Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet. Biomedical Engineering / Biomedizinische Technik **40**(s1), 317--318 (1995)

[2] Brugada, P., Brugada, J.: Right bundle branch block, persistent st segment elevation and sudden cardiac death: A distinct clinical and electrocardiographic syndrome. Journal of the American College of Cardiology **20**(6), 1391--1396 (1992). https://doi.org/10.1016/0735-1097(92)90253-J, `http://www.onlinejacc.org/content/20/6/1391`

[3] Charles, A., Pedro, B., Martin, B., Josep, B., Ramon, B., Domenico, C., Ihor, G., Herve, L., Koonlawee, N., Ricardo, P.R.A., Wataru, S., Eric, S.B., Hanno, T., Arthur, W.: Brugada syndrome: Report of the second consensus conference. Circulation **111**(5), 659--670 (Feb 2005). https://doi.org/10.1161/01.CIR.0000152479.54298.51, `https://doi.org/10.1161/01.CIR.0000152479.54298.51`

[4] Gers, F., Schmidhuber, J., Cummins, F.: Learning to forget: continual prediction with LSTM. In: 9th International Conference on Artificial Neural Networks: ICANN '99. pp. 850--855. Institution of Engineering and Technology (1999)

[5] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215--e220 (2000)

[6] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735--1780 (1997)

[7] Jaxy, S.: Teaching a machine to diagnose a heart disease; beginning from digitizing scanned ecgs to detecting the brugada syndrome (brs) (2020)

[8] de Jong, J.: Brugada ecg characteristics (2013), `https://en.ecgpedia.org/images/f/f1/Brugada_ecg_characteristics.svg`, last accessed 27 Aug 2020

[9] Cardiovascular diseases (cvds). `https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)`, accessed: 2020-10-07

ARTIFICIAL INTELLIGENCE RESEARCH GROUP

Universitair Ziekenhuis Brussel

UNIVERSITÄT OSNABRÜCK