Lazaros Iliadis
Antonios Papaleonidas
Plamen Angelov
Chrisina Jayne (Eds.)

LNCS 14257

Artificial Neural Networks and Machine Learning – ICANN 2023

32nd International Conference on Artificial Neural Networks
Heraklion, Crete, Greece, September 26–29, 2023
Proceedings, Part IV

Springer

Analysis and Interpretation of ECG Time Series Through Convolutional Neural Networks in Brugada Syndrome Diagnosis

Alessio Micheli[1], Marco Natali[1], Luca Pedrelli[1(✉)], Lorenzo Simone[1], Maria-Aurora Morales[2], Marcello Piacenti[3], and Federico Vozzi[2]

[1] Department of Computer Science, University of Pisa, Pisa, Italy
luca.pedrelli@di.unipi.it
[2] Institute of Clinical Physiology, IFC-CNR, Pisa, Italy
[3] Fondazione Toscana Gabriele Monasterio, Pisa, Italy

Abstract. In this research, we present a novel approach to evaluate and interpret Convolutional Neural Networks (CNNs) for the diagnosis of Brugada Syndrome (BrS), a rare heart rhythm disease, from the electrocardiogram (ECG) time series. First, the model is assessed on the ECG classification of type-1 BrS. Then, we define a method to interpret the BrS prediction through Gradient-weighted Class Activation Mapping (Grad-CAM) applied to continuous time series. Finally, the proposed approach provides a tool to analyze the main areas of the ECG time series responsible for the BrS diagnosis through CNNs. In experimental assessments we use an original dataset of 306 ECGs collected from several clinical centers within the BrAID (Brugada syndrome and Artificial Intelligence applications to Diagnosis) project.

Keywords: Time series analysis · Health Informatics · Brugada Syndrome · Convolutional Neural Networks

1 Introduction

The Brugada Syndrome (BrS) is a rare heart rhythm disorder first reported in 1992 that can lead to sudden cardiac death, accounting for 4% to 12% of such events [3]. BrS is a channelopathy that arises from a modification of transmembrane currents of the cardiomyocyte, which are essential for cardiac action potential to function correctly [13]. Research has confirmed that around 20% of Brugada syndrome cases are connected to mutations in the SCN5A gene that encodes for the sodium ion channel in the cardiac cells. Individuals suffering from this syndrome are more likely to develop ventricular tachyarrhythmias, which can lead to fainting, cardiac arrest, or sudden cardiac death. The clinical onset of the disease usually occurs in adulthood, with a mean age of onset of 40 years. However, it can also affect children and adolescents.

Three BrS patterns [13] have been identified as: Type-1, characterized by ST-segment elevation (>2 mm) followed by a negative T wave in leads V1 and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
L. Iliadis et al. (Eds.): ICANN 2023, LNCS 14257, pp. 26–36, 2023.
https://doi.org/10.1007/978-3-031-44216-2_3

V2; Type 2, showing a similar ST-segment elevation but followed by a positive T wave; and Type 3, presenting a right precordial ST-segment elevation ($\leq$1 mm) followed by a saddleback morphology. However, identifying these patterns is time-consuming and challenging for the clinicians, since electrocardiograms (ECGs) may show dynamic changes thus impairing a definite diagnosis. In doubtful situations, to unmask the BrS pattern, patients were subjected to infusion with Sodium channel blockers, such as ajmaline or flecainide, able to provoke type-1 BrS ECG patterns. In conclusion, Brugada Syndrome is a rare but potentially life-threatening genetic disorder. Its early diagnosis, risk stratification, and appropriate management are crucial for preventing sudden cardiac death in affected individuals and their families.

Since ECG interpretation depends on the individual experience of the clinician, the use of machine learning (ML) approaches may represent a supportive tool for the ECG analysis in these patients. Several ML applications focus on detecting cardiac arrhythmias [2]. In the literature, few works are focused on BrS classification with ML approaches. Preliminary works that use recurrent neural networks to classify BrS were presented in [4,15]. Other works are focused on the use of CNNs [9,10]. However, to the best of our knowledge, no approaches face the interpretability of the model in the classification of BrS ECG patterns.

This paper presents a novel approach for ECG time series to interpret temporal data analysis in BrS Type 1 diagnosis based on Convolutional Neural Networks (CNNs). In particular, (i) we use an original dataset composed of 306 ECGs (123 of positive class and 183 of negative class) collected from several clinical centers within the BrAID project [1], a funded project by the Health Research Fundings established as part of the Tuscany Region initiatives. Then, (ii) in order to address low data volume we experimentally assess the CNN model in classifying BrS through a double cross-validation (DCV) approach. Finally, (iii) we implement an interpretation of the ECG time-series analysis through Grad-CAM [12] adapted to time series. Overall, the approach provides a support tool to discover and analyze the main areas of the ECG time series that are important to make the BrS diagnosis. Differently from previous works [4,15], in this research we propose a 1-dimensional CNN for temporal data, a novel dataset composed by more data with respect to [4] and single heart beat with respect to [15]. Moreover, we introduce a novel approach to analyze the CNN prediction in temporal data through Grad-CAM adapted to time series. Other works exploit CNN but with a different dataset, validation approach and without interpretation [9,10].

The paper is organized as follows. In Sect. 2, the BrAID project is presented and the dataset used in the study is described. In Sect. 3, the CNN model for time-series classification is introduced. Section 4 describes the experimental assessment for the classification of BrS. Section 5 presents our approach to interpreting the model predictions in BrS analysis. Section 6 shows the quantitative and qualitative results obtained from our approach. Finally, in Sect. 7, conclusions are drawn.

2 BrAID Project and BrS Dataset

The BrAID project [1] was initiated in September 2020 with the primary objective of creating a clinical tool supporting clinicians in the diagnosis of type 1 BrS. The main objectives of BrAID focused on: developing a novel approach to disease management through the implementation of standardized procedures across clinical centers; an Electronic Health Record (EHR) system for data management and the construction of a supervised Machine Learning (ML) model for ECG analysis.

The project consists of three incremental phases: (1) The first phase involves setting standard clinical procedures and data collection protocols for BrS patients. (2) The second phase focuses on organizing a retrospective study with the development of ML models for ECG analysis, which is being tested in the collected population. (3) In the final phase, a prototype platform based on an electronic health record (EHR) system integrated with the proposed model is being released, and later validated by clinical centers.

Before conducting experiments, the process of creating the dataset was characterized by a multitude of technical hurdles to address for having an input suitable for a machine learning model following a pre-processing pipeline defined in [4]. Among these issues we noticed inconsistencies in ECG recordings from various clinical centers, differences in the type of data collected (electronic ECGs or scanned paper ECGs), and the presence of noise or signal artifacts. We addressed these challenges through the following pre-processing steps:

- Paper format ECGs were scanned by the clinical center before analysis. Then, several image processing filters were applied to reduce the source of noise from manual procedures.
- To limit the loss of information during the extraction process, a second-order spline interpolation was utilized in combination with a 50 Hz Butterworth filter of order 2.
- Each recording has been resampled to a frequency of 500 Hz with a single beat selected as suggested by clinicians to maintain a time window within a 1-second duration.

The original dataset is composed of 306 time series pre-processed from ECGs. Among these recordings 123 were positive class (87 Spontaneous BrS and 36 Positive *provocative test*) and 183 were negative class (169 Controls and 14 Negative *provocative test*). In particular, *provocative test* samples are related to the administration of sodium blocker channels under medical supervision in the Electrophysiology Lab. In Spontaneous BrS and Positive provocative cases the BrS is diagnosed, while in Control and Negative provocative cases the BrS is not diagnosed. Table 1 presents the distribution of samples across the four original cases. These four original cases are combined into two classes, considering *Spontaneous BrS* and *Positive provocative* as the positive class, while *Control* and *Negative provocative* as the negative class.

Figures 1a) and 1b) show two time series pre-processed from ECG lead V2. Figure a) shows an example of a positive class while Figure b) shows an example of a negative class.

3 Convolutional Neural Networks

A convolutional neural network (CNN) is a type of neural network characterized by the usage of convolutional operators, which apply filters to an input image or signal to extract features, and pool layers reducing the size of the feature maps [8,16]. The model optimizes its weights through backpropagation in order to minimize a predefined supervised loss function. Recently, CNNs have been successfully adapted to processing temporal data, such as audio signals or time series. They have shown great potential in the classification of heart arrhythmias due to their ability to extract local patterns or subtle changes in the morphology of ECGs, which is crucial for an accurate diagnosis [6]. In Fig. 2, we provide a summary of our CNN-based workflow specifically designed for Brugada syndrome prediction from ECG recordings.

In this work, we consider a CNN composed of N_b convolutional blocks with varying temporal window sizes (also referred to as kernel) and number of filters to capture a hierarchical representation across the signal. Each convolutional block consists of a 1D convolutional layer followed by a batch normalization [7] and a rectified linear unit (ReLU) activation function. More formally, considering the input signal $\mathbf{A}^{(0)} = \mathbf{X}$, a convolutional layer at depth $m > 0$ receives as input $\mathbf{A}^{(m-1)}$ producing the output $\mathbf{A}_k^{(m)}$ also referred to as a feature map and computed as:

$$\mathbf{A}_k^{(m)} = ReLU(BN(\sum_{i=1}^{K} \mathbf{W}_{ki}^{(m)} * \mathbf{A}_i^{(m-1)} + b_k^{(m)})), \tag{1}$$

where $\mathbf{W}$ and b refer respectively to weight and bias matrices, BN denotes the batch normalization layer and $*$ the 1D convolutional operator.

A dropout regularization mask [14] is also applied in between each convolutional block, while on top of the architecture there is a fully connected output layer with sigmoid activation function. Lastly, the initial weights are set following Glorot uniform initialization [5] and the model is trained from scratch using stochastic gradient descent (SGD) optimizing the binary cross-entropy loss function $\mathcal{L}(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$, given a label y and the model prediction $\hat{y}$.

4 Experimental Assessment (BrS Classification)

First, we started with the classification task of evaluating the CNN model in BrS diagnosis. We perform a model selection considering CNN configurations with a number of convolutional blocks $N_b = 3, 4, 5, 6, 7$ with V1 lead (5 hyperparameter configurations) and V2 lead (5 hyperparameter configurations) in input. The

Table 1. Distribution of the original four cases, later merged into two distinctive classes for the binary classification task.

Total	Positive Class		Negative Class	
	Spontaneous BrS	Positive provocative	Control	Negative provocative
306	87	36	169	14

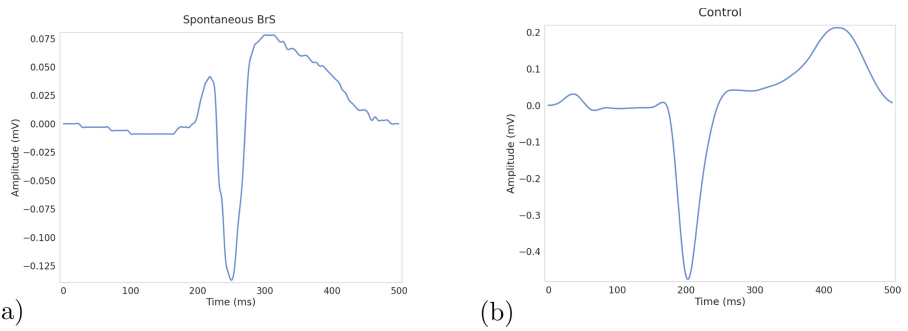

Fig. 1. An example of two time series of the dataset pre-processed from ECG lead V2. A positive class (a) and a negative class (b).

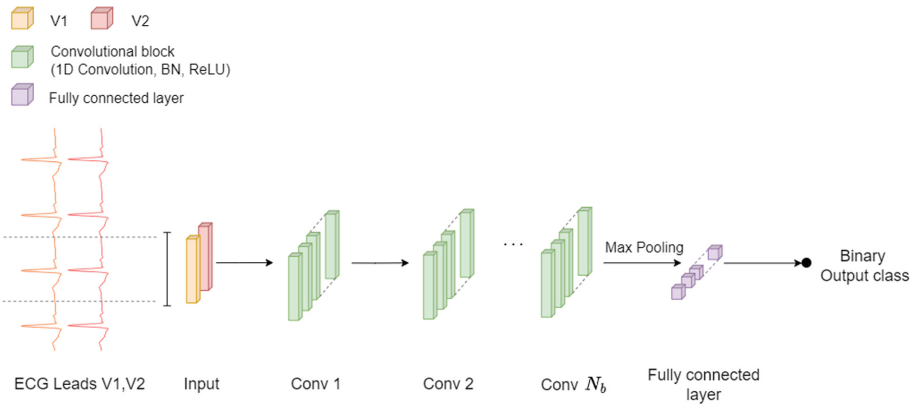

Fig. 2. Visual summary of our CNN-based classifier used in the experimental assessment for Brugada syndrome prediction from ECG recordings.

model selection is performed through a DCV approach with 5 external and 4 internal folds. Accordingly, the dataset is randomly shuffled with uniform distribution and then split into 5 test outer folds. The hyper-parameters of each convolutional block configuration were selected in validation folds through a random search by using a range of values shown in Table 2.

As metrics to evaluate the model we use accuracy (rate of correctly classified ECG), sensitivity (true positive rate) and specificity (true negative rate). The

Table 2. Ranges of hyperparameters used in the random search for the model selection.

	Hyperparameters
filter	$[5, 10, 15]$
kernel	$[5, 10, 15, 20, 30, 40]$
weight decay	$[5 \cdot 10^{-6}, 1 \cdot 10^{-5}, 1 \cdot 10^{-3}, 5 \cdot 10^{-3}]$
dropout	$[0, 0.1, 0.2, 0.3, 0.4, 0.5]$
learning rate	$[0.01, 0.005, 0.004, 0.003]$

validation accuracy is evaluated for each outer fold by averaging the accuracy achieved in the inner folds. Then, for each outer fold, the test results are evaluated considering the hyperparameters that obtained the best validation accuracy in the inner folds. The test results are computing averaging the results obtained in the outer folds of the DCV. Finally, the best configuration is considered to apply the proposed approach presented in Sect. 5.

5 Interpretation of ECGs Analysis in BrS Diagnosis

In this section, we describe the proposed approach used for the interpretation of time-series analysis in the case of BrS classification. Initially, we employed the Pan-Tompkins algorithm [11] to accurately identify R-peaks in each ECG recording proceeding with a beat alignment process. To ensure a higher precision of quantitative results we manually refined borderline cases. To capture the relevant electrical activity, we utilized a fixed time window of 1 s, comprising 250 timesteps before and after each R-peak. The signal was then processed by averaging each sample from the dataset within respective classes computing their gradient class activation mappings (Grad-CAM [12]) adapted for the temporal case as detailed in this section.

Gradient class activation mappings allowed us to point out highly contributing regions from an input ECG starting from a prediction, while identifying the features negatively or positively impact the outcome. The original idea is preserved but reformulated for a 1D input scenario; using the gradients of the final convolutional layer's output with respect to the class score to generate a region heatmap.

Technically, we can define $L(c)$ as the final convolutional layer in the CNN and $A_k^L(p)$ as the output for the k-th feature map in the layer at position p. The global average pooling of the gradient of the output class score y^c with respect to the feature maps A_k^L is computed as $\alpha_k^c = \frac{1}{Z} \sum_p \frac{\partial y^c}{\partial A_k^L(p)}$, where Z

is the number of spatial positions in the feature map. Afterwards, we compute a weighted combination of the activation maps A_k^L using the coefficients α_k^c as weights $L^c(p) = \sum_k \alpha_k^c A_k^L(p)$. Differently from the original formulation, our importance coefficients $L^c(p) \in \mathbb{R}$ do not rule out negative values, since we avoided the usage of the outer ReLU activation. This allowed for a more comprehensive interpretation of the binary prediction enclosing positive and negative contributions.

6 Experimental Results

In this section, we show the classification results of model interpretation. Table 3 shows the test results obtained by the CNN model using $N_b = 3, 4, 5, 6, 7$ convolutional blocks, weight decay, dropout and leads V1 and V2. The results highlight that using V2 lead obtains better results than V1 lead. In particular, the use of 6 convolutional blocks with the use of V2 achieves the best test result with accuracy, sensitivity and specificity of 90.53%, 87.73% and 92.34%, respectively. Considering the configuration with 6 convolutional blocks, we evaluated the model with the use of both V1 and V2 leads achieving accuracy, sensitivity and specificity of 89.87%, 89.43% and 90.20%, respectively. In this case, the configuration achieved a slightly low accuracy of 89.87% w.r.t. the best configuration that achieved 90.53.%. However, the use of 6 convolutional and both V1 and V2 leads allowed us to obtain a better sensitivity of 89.43% w.r.t. best configuration that obtained 87.73%. The results obtained in the BrS classification task are generally competitive with the recent literature approach [10] in which the model obtains a sensitivity of 88.4% and a specificity of 89.1% (although evaluated with a different dataset). Moreover, the results obtained in this paper significantly outperforms the results achieved in preliminary works [15] [4] (up to a maximum of 80% accuracy).

After the model selection, the architecture mentioned above (6 convolutional blocks) have been employed for the interpretability assessment focusing on lead V2. Following the principles detailed in Sect. 5, we computed the importance coefficients $L^c(p)$ averaged over each test fold from the double-CV. A bar plot visualization of the coefficients is showcased in the bottom row in Fig. 3 (red bars stands for positive class contribution, while blue bars for negative class contribution). For instance, in the left column (BrS patterns) each coefficient greater than zero supports the actual prediction, thus being colored in red. Mutually, negative coefficients refer to timesteps disowning the ground truth label, thus favoring a control-like prediction instead. Additionally, we provide the outcome of a manual ECG wave annotation phase performed by clinicians to identify the different ECG segments and better comprehend the correlation between interpretability outcomes and clinical practice (top row). The middle row depicts an

Table 3. Test accuracy (Acc), Sensitivity, Specificity and the standard deviation inside the brackets obtained by the CNN model with the use of $N_b = 3, 4, 5, 6, 7$ convolutional blocks, weight decay (wd), dropout and leads V1 and V2.

Architecture	Test Accuracy (%)	Sensitivity	Specificity
3 blocks V1	83.65(3.90)	80.50(0.078)	85.84(0.067)
4 blocks V1	85.28(7.40)	87.00(0.015)	84.22(0.120)
5 blocks V1	86.91(0.049)	88.67(0.058)	85.78(0.063)
6 blocks V1	85.62(0.020)	85.36(0.001)	85.81(0.032)
7 blocks V1	83.00(0.033)	76.43(0.069)	87.50(0.064)
3 blocks V2	83.65(3.90)	80.50(0.078)	85.84(0.067)
4 blocks V2	87.28(0.060)	82.77(0.094)	90.24(0.103)
5 blocks V2	87.31(0.007)	86.99(0.001)	82.52(0.010)
6 blocks V2	**90.53(0.028)**	**87.73(0.083)**	**92.34(0.011)**
7 blocks V2	87.59(0.012)	83.73(0.057)	90.17(0.037)
6 blocks V1, V2	89.87(0.028)	89.43(0.078)	90.20(0.032)

average ECG beat for positive and negative classes and their standard deviation (translucent region). Interestingly, we can infer how the saddleback pattern typical of Type 1 BrS ECG is associated with the syndrome in the left pane, which can be distinguished from the average morphology of healthy subjects (V2 lead, right column). In Fig. 3, the grey dashed lines represent the focal areas for a correct model prediction. In both positive and negative cases, there is a noticeable concentration of coefficients throughout the ST segment (from S wave offset to T onset), a portion heedful to healthcare practitioners to identify patterns to establish a BrS diagnosis. Furthermore, the model leverages the waves S and T in the positive case, as shown by the two distinct points defining average beat contribution. Conversely, for the averaged morphology of patients without the disease the network examines mainly the QRS segment. Although, it should be acknowledged that the multitude of overlapping opposite sign coefficients (visible over time regions in bottom row plots) results from the aggregation of various ECG recordings from diverse patients in the test set. Thus, the presence of noise is attributable to the process of joining singular variations in signal morphology across various patients, which is consistent with the complexity of the task.

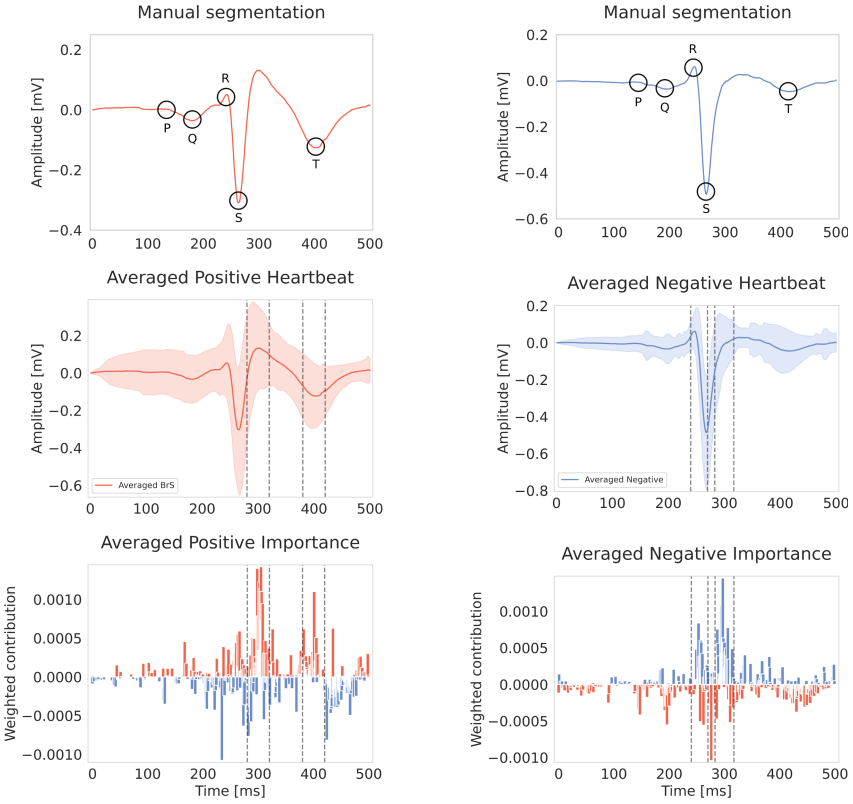

Fig. 3. (Top row) Outcome of a manual wave annotation process performed by clinicians. (Middle row) Averaged amplitude plot from ECGs including positive (left) and negative (right) class. (Bottom row) Weighted importance coefficients averaged over test folds. The areas highlighted by the grey dashed lines represent the signal ranges where the model is focused on making the prediction. (Color figure online)

7 Conclusions

In this paper, we proposed a predictive approach supporting BrS diagnosis and a post hoc interpretability assessment leveraging CNNs. The dataset has been collected from several clinical centers within the BrAID project [1]. It comprises of 306 ECGs (123 of positive class and 183 of negative class) obtained from scanned paper and digital format. An extensive pre-processing pipeline was performed to produce the dataset through a multitude of technical hurdles (such as different clinical centers and various ECG formats) to address having a temporal input suitable for a machine learning model.

Initially, we performed an experimental assessment employing a double cross-validation approach to evaluate the model in BrS type-1 classification task. Within model selection, we considered several model configurations by varying

the number of convolutional blocks and the input leads (V1 and V2). Specifically, a configuration with 6 convolutional blocks using the lead V2 obtained the best test result with an accuracy, sensitivity, and specificity of 90.53%, 87.73%, and 92.34%, respectively. Accordingly, we used V2 to analyze the model in the BrS classification from the ECG time series.

Finally, we computed the important areas of the ECG time series which allow the model to make the prediction. First, we computed two plots, one averaged from the values of positive beats and the other from negative beats. Analyzing the plots of the averaged beats we can note the typical saddleback pattern in Type 1 BrS. Moreover, by interpreting the predictions on lead V2, we found a higher correlation in the ST segment amongst importance coefficients and positive output class, confirming the suggestion from healthcare clinicians. Overall, the interpretability results are consistent with the patterns analyzed by clinicians to perform diagnosis of BrS. We were able to display through the application of our framework the impact on the decision-making process of the predictions for positive and negative test samples. The approach offers a way to improve the transparency and accountability of the classifier, potentially allowing clinicians to revise the output class based on the interpretability outcome.

Our research highlighted that the proposed predictive model based on convolutional neural networks is comparable with the results provided by state-of-the-art approaches for BrS classification. The results of the proposed method based on the interpretation of the relevant input areas in the BrS classification show that this architecture could potentially support clinicians in the early diagnosis of a rare ECG pattern. Through the insights provided on the focal signal regions, clinicians can better evaluate its suitability for making informed decisions in clinical practice.

In conclusion, we believe the proposed approach can contribute to and stimulate the development of novel approaches in future works to interpret model predictions on time-series applications.

Acknowledgements. The project BrAID was funded within the "Regione Toscana Bando Ricerca Salute 2018" (Decreto n. 15397, 26/09/2018). Moreover we would like to acknowledge all of the clinical centers part of BrAID who has collected the data used in this paper. In particular: Clinical Physiology Institute (IFC) of the Italian National Research Council (CNR), Fondazione Toscana Gabriele Monasterio per la Ricerca Medica e di Sanità Pubblica (FTGM), Azienda Sanitaria Toscana Sud Est - U.O.C Cardiologia ASL 8 Arezzo (AUTSE), Azienda Sanitaria Toscana Nord Ovest - U.O.C. Cardiologia, Ospedale Versilia (AUTNO), Azienda Ospedaliera Universitaria Careggi - SOD Aritmologia - Firenze (AOUC), Azienda Ospedaliero Universitaria Pisana - Cardiologia 2 Cisanello (AOUP).

References

1. Braid. https://www.ifc.cnr.it/index.php/en/news/685-kick-off-del-progetto-braid. Accessed 03 Apr 2023
2. Ahsan, M.M., Siddique, Z.: Machine learning-based heart disease diagnosis: a systematic literature review. Artif. Intell. Med. 102289 (2022)

3. Brugada, P., Brugada, J.: Right bundle branch block, persistent ST segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome: a multicenter report. J. Am. Coll. Cardiol. **20**(6), 1391–1396 (1992)
4. Dimitri, G.M., Gallicchio, C., Micheli, A., Morales, M.A., Ungaro, E., Vozzi, F.: A preliminary evaluation of Echo State Networks for Brugada syndrome classification. In: 2021 IEEE SSCI, pp. 01–08. IEEE (2021)
5. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth AISTATS, pp. 249–256. JMLR Workshop and Conference Proceedings (2010)
6. Hannun, A.Y., et al.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. **25**(1), 65–69 (2019)
7. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: ICML, pp. 448–456. PMLR (2015)
8. LeCun, Y., et al.: Backpropagation applied to handwritten zip code recognition. Neural Comput. **1**(4), 541–551 (1989)
9. Liao, S., et al.: Use of wearable technology and deep learning to improve the diagnosis of Brugada syndrome. Clin. Electrophysiol. **8**(8), 1010–1020 (2022)
10. Liu, C.M., et al.: A deep learning-enabled electrocardiogram model for the identification of a rare inherited arrhythmia: Brugada syndrome. Can. J. Cardiol. **38**(2), 152–159 (2022)
11. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **BME-32**(3), 230–236 (1985)
12. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE ICCV, pp. 618–626 (2017)
13. Sheikh, A.S., Ranjan, K.: Brugada syndrome: a review of the literature. Clin. Med. **14**(5), 482 (2014)
14. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**(1), 1929–1958 (2014)
15. Vozzi, F., et al.: Artificial intelligence algorithms for the recognition of Brugada type 1 pattern on standard 12-leads ECG. Europace **24**(Suppl._1), euac053–558 (2022)
16. Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., Lang, K.J.: Phoneme recognition using time-delay neural networks. IEEE Trans. Acoust. Speech Signal Process. **37**(3), 328–339 (1989)