Cardiac Abnormalities Detection Through 12-Lead/Reduced-Lead ECG Spectrograms and 2D-CRNN

Jonathan R Torres-Castillo [1], Miguel A Padilla-Castañeda[1]

[1] Institute of Applied Sciences and Technology, Universidad Nacional Autónoma de México, Mexico

Abstract

Heart abnormalities represent around 32 % of the deaths for illnesses in the world. This work presents an automated pattern recognition method for detecting 25 different cardiac arrhythmias and the normal sinus rhythm type. A two-dimensional convolutional recurrent neural network (2D-CRNN) model was employed by using raw-data images and signal spectrograms. With a database of 88,253 ECG signals of 12-leads, a four-step method was created. ***1. Preprocessing****. The data were filtered and downsample or filled in so that they were the same length.* ***2. Representation****. Two sets of images were obtained, one with the time series and the other with spectrograms through Wavelet Synchrosqueezing (WS).* ***3. Feature extraction.*** *A CNN network was chosen to get relevant features of the images; these were flattened in a vector to feed a recurrent neural network.* ***4. Classification.*** *A fully connected layer was used to classify the signals. A 10-fold cross-validation method (CV) was executed. The grid search technique was used to obtain the optimal threshold to improve the Challenge Score (Ch-score).* ***Results:*** *With this model, our UIDT_UNAM team received an unofficial score of 0.34 for all sets of established leads utilizing the hidden test data.*

1. Introduction

Heart abnormalities are the leading cause of death worldwide, with an estimated 32% in total with about 17.9 million people [1]. Cardiac arrhythmias (CAs) are the most common, and, in clinical practice, the gold standard for diagnosing them is using an electrocardiogram (ECG) [2]. ECG is a technique that captures the electrical activity involved in the functioning of the heart and results in a signal composed of cardiac impulses (beats). The analysis of these signals is fundamental to a proper diagnosis. The early detection of CAs and their treatment for sudden cardiac death (SCD) prevention represents a significant opportunity to reduce mortality further [2].

However, diagnosis of CAs usually requires evaluating twelve different points on the human body simultaneously, obtaining a 12-lead ECG; for this, it is necessary to: a) high-cost clinical equipment and b) experienced clinicians with high knowledge in different CAs. The above makes it challenging to obtain an accurate diagnosis in many places, especially in developing countries. In addition, Physicians take quite a bit of time when analyzing a 12-leads ECG since they perform a detailed analysis of the morphology of each beat and then correlate the irregular segments between the different leads to detect a specific CA.

Consequently, this study presents a way to automatically obtain such an analysis to assist the physician in the diagnostic task. It also evaluates the practicality of using 12-lead-ECG versus reduce-lead ECG, allowing new lower-cost clinical equipment design.

Recently, there has been increasing research focused on the automatic detection of CAs using 12-lead-ECGs. The most promising studies are based on Machine Learning (ML) and Deep Learning(DL) methods [3]. Theoretically, most of these algorithms achieve accuracies > 90% in the classification of CAs. However, the promising outcome of those tests is a consequence of using small and homogeneous datasets. Therefore, A. Perez et al. [4] proposed assembling multiple global databases that contain more than 100 classes of CAs and launched the Physionet/CinC 2020 challenge.

In that challenge, they evaluated 41 algorithms in the final stage (test set) developed by different academic and industry groups. One of the results was that no algorithm exceeded 55% of the proposed score, mainly due to the imbalance, variety of classes, and multiple diagnostics by signal.

For 2021, these same organizers posed the same problem with validating different sets of lead ECGs; additionally, they increased the number of data [5]. This suggested a challenge to solve a real-life problem. The algorithm presented in this paper is one of the algorithms submitted to the Physionet/CinC 2021 challenge (PCC 2021).

2. Material and methods

Our approach is initially based on deep learning techniques for diagnosing CAs implemented in [6]. They use a parallel deep neural network in which each signal represents it as an image that contains frequency-time information, and then those images are using as input data in a convolutional neuronal network (2D-CNN).

In this study, the novelty lies in a new deep neuronal network, which not only uses known convolutional networks as VGG-16 and a fully connected network to classify the signals: Instead, we propose a new end-to-end

network with a CNN stage for extraction of features followed by a recurrent neuronal network (RNN) to add temporary features (2D-CRNN).

2.1. Data set

For the official phase of PCC-2021, the organizers have made available 88,253 12-lead ECG recordings containing 132 classes of CAs and normal sinus rhythm (NSR) class, of which 30 types are the most common relevant in clinical practice, and they chose to evaluate the algorithms submitted, a detailed description of these data is presented in [5].

For the analysis and training task, we only take the signals with valid CAs, and for the signals with multiple CAs diagnostics, we make a copy for each class. Likewise, we replaced four classes that presented similarity according to [5], obtaining 25 classes of CAs and NSR class at the end.

Figure 1 shows the histogram of the data per class of CAs after this procedure. There we have a total of 111,425 12-lead ECG records. This data set was stratifically divided into training/validation and testing sets, using 85% and 15% of the data, respectively.

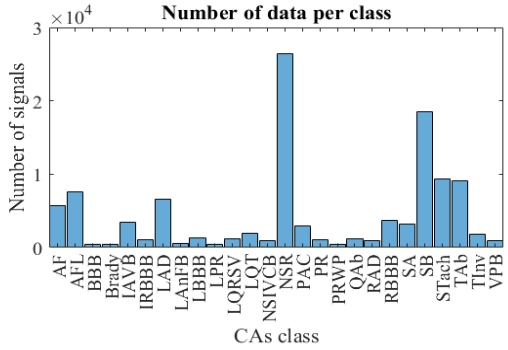

Figure 1. Histogram of the data per CA class in the database.

2.2. Preprocessing signals

In order to obtain a homogeneous data set, each signal is preprocessed, matching the sampling frequency to 200 Hz and using a Butterworth bandpass filter between 1 - 100 Hz.

Then, we took the first 10 seconds of each sign for each lead, and if the signal was shorter over time, we performed an upsampling using zero padding.

This interval size is the average recording time by signal throughout the database; this allows us to obtain the relevant information of each signal.

2.3. Two-dimensional representation

At this stage, we obtained the two-dimensional representations of each signal for each lead. We initially form the first set of images by transforming each signal into a 128 x 128 pixel image containing the time series, as shown in figures 2a), 2b), and 2c).

The second set is formed using the Wavelet Synchrosqueezing Transform (WSST) in each lead; with this method, we obtain information about the behavior of each signal in the time-frequency-amplitude space.

2.3.1. Wavelet Synchrosqueezing Transform

ECG signals are non-stationary signals in which changes in behavior occur at an instant of time that are unpredictable given previous data. For this reason, it is essential to capture this changing information as accurately as possible.

The WSST is a suitable method for analyzing this type of signal with which we obtained a time-frequency representation (spectrogram) of the signal. This graph contains the instantaneous frequencies over a short period: This is achieved by decomposing the signal into different sub-bands orthogonal to each other [7].

The first step is to obtain the continuous wavelet transform (CWT). [8].

$$W_x(a,b) = a^{-\frac{1}{2}} \int x(t)\bar{\psi}\left(\frac{t-b}{a}\right)dt \qquad (1)$$

where $x(t)$ is the preprocessed ECG signal, ψ is the base wavelet function, a is a scale parameter, and b is a translation parameter.

The second step is to obtain an analytical signal, extracting the instantaneous phases by deriving W_x respect to the translation b.

$$\omega_{a,b} = -i\left(W_{a,b}\right)^{-1}\frac{\partial}{\partial b}W_{a,b} \qquad (2)$$

Dividing ω by $i2\pi$ we obtain the instantaneous frequencies (IF) of the signal, with which we obtain concentrated high-resolution time-frequency patterns.

Finally, we reassign (synchrosqueezing) the ω values to center the energy of the spectrogram towards the IF curves obtaining a sharper spectrogram. An example of these plots is shown in figures 2d), 2e), and 2f).

As a result, we obtained a set of 24 images (12 raw data, 12 spectrograms) for each signal in the case of the 12-lead ECG from the arrhythmia datasets used here.

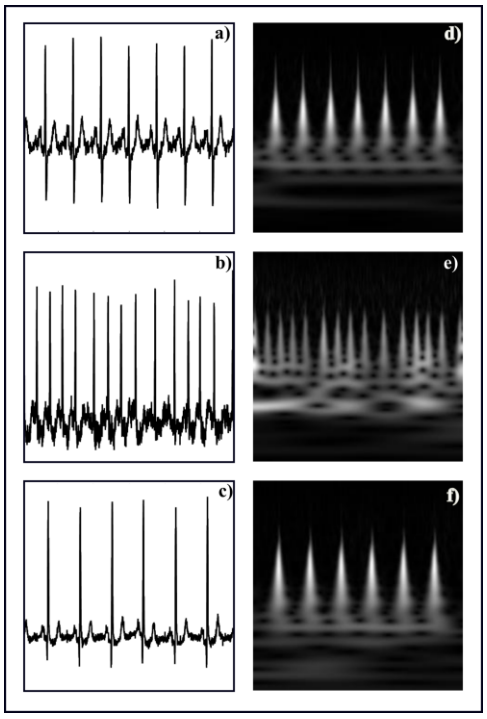

Figure 2. 2D representation of lead No. 2 of different types: a, d) normal sinus rhythm (NSR), b, d) St depression (STD), and c, f) Premature atrial contraction (PAC).

With these representations, we can capture information in multiple spaces that are not perceptible in standard time series analysis. However, it is not so subtle to evaluate these images manually, so we show below an automatic method that extracts features from them to classify between different types of CA.

2.4. Deep learning model

CNNs are deep learning (DL) methods focused on analyzing information formed by a series of convolutional operations in which filters of different sizes are applied, which allows obtaining features in an automated way. The weights of these filters are adapted in the training stage to get the best accuracy in the classification task by employing the backpropagation algorithm on a large amount of data.

2.4.1. Model description

The model is composed in principle by two CNNs in parallel, in which the input data are the set of images obtained in section 2.3.

Each CNN functions as a feature extractor and has a Resnet18 based structure, where residual blocks help to keep the gradient from vanishing at minimal values.

After obtaining the features in parallel, we used a flatten layer to join them as a time series.

To improve the understanding of these sequences, we added a Bidirectional Long short-term memory (BiLSTM) network with two-layer.

At the end of this array, a fully connected network with a Softmax function takes the output of the BiLSTM network and predicts the best label for the input data. The model and parameters of this network are shown in figure 3.

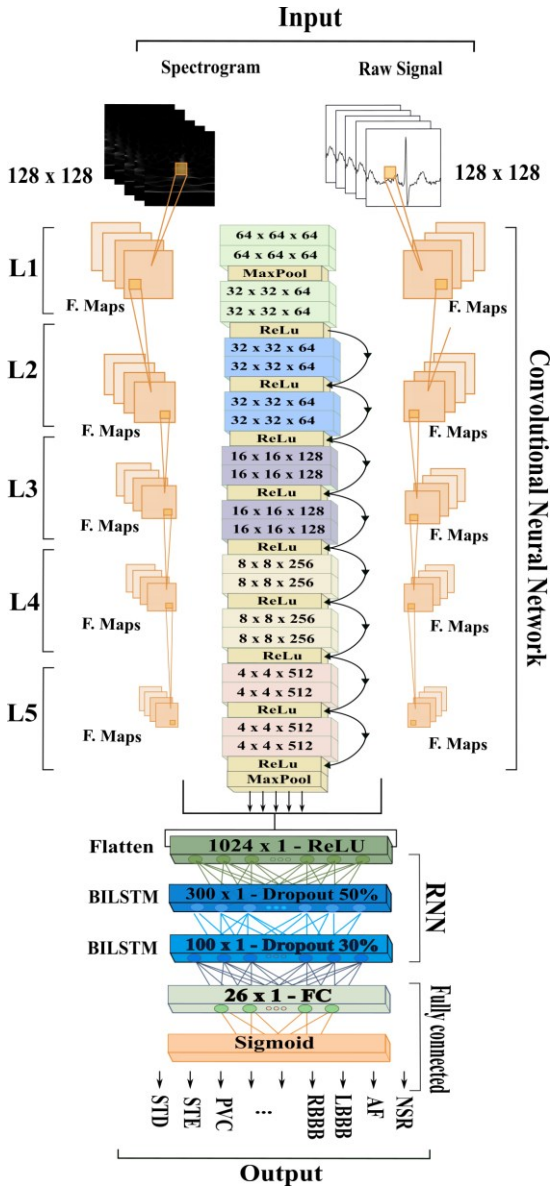

Figure 3. 2D-CRNN model architecture.

3. Results and Discussion

We used the 10-fold cross-validation (CV) method to

train and evaluate the model. Different hyperparameters were tested during the training phase to find the best model, the main parameters we took were the learning rate and the batch size, and the best performance was obtained with the values of 0.001 and 256, respectively. We also performed a grid search to find the optimal threshold (Th) value to indicate multiple CAs in the same signal; the best Th value found was equal to 0.07.

The number of epochs was set at 50 with an early stop criterion set at three validations if their accuracy did not improve. The optimal batch size value for the training process was 256.

In the test process, the mean F1-score and Ch Score obtained in the CV process and the official results when submitting this model to the validation hidden data set can be seen in Table 1.

Table 1. The metrics obtained using different combinations of leads and the 2D-CRNN model in the training database (F1-Score and Ch-Score-CV) and in the hidden database.

Leads	F1-Score	Ch-Score (CV)	Ch-Score Hidden Validation Data	Un-official Ch-Score Hidden Test Data
12	0.482±0.102	0.566±0.095	0.518	0.33
6	0.465±0.158	0.548±0.183	0.494	0.33
4	0.472±0.098	0.539±0.112	0.519	0.35
3	0.452±0.111	0.530±0.124	0.524	0.35
2	0.455±0.101	0.529±0.065	0.512	0.35

We compared these results against the CNN network without the BiLSTM network and obtained a Ch score of 0.395 +/- 0.154. Therefore, we find it convenient to use the BiLSTM network.

We attribute the classification errors mainly to the morphological similarity between the different cardiac anomalies present in the data set and the class size difference. Initially, we used the sample replication method to address the problem of unbalanced classes, but this caused overtraining and affected the official results. So it is necessary to address this problem with other data augmentation techniques. We also hope to improve the loss function to be in line with Ch-Score metrics. In addition, we intend to extract hand-crafted features that we have implemented in other biosignal analyses to join them to the CNN network features and then bring them into the RNN network [9].

Making these adjustments to the classification methodology is proposed as future work presented here.

5. Conclusion

This paper proposes a classification model of ACs where the backbone uses two-dimensional representations of each signal. As a result, our architecture achieved an average Ch-score of 0.34. The 2D-CRNN model presents an average performance respect to the results of other models presented in the challenge, assuming that only network parameters were optimized but still needs to improve the loss function and address the problem of unbalanced classes.

Acknowledgments

This project was support by DGAPA-PAPIIT TA100920.

References

[1] World Health Organization, “Cardiovascular diseases,” 2021. www.who.int/health-topics/cardiovascular-diseases (accessed Aug. 01, 2021).

[2] N. T. Srinivasan and R. J. Schilling, “Sudden cardiac death and arrhythmias,” *Arrhythmia Electrophysiol. Rev.*, vol. 7, no. 2, pp. 111–117, 2018, doi: 10.15420/aer.2018.15.2.

[3] Z. Ebrahimi, M. Loni, M. Daneshtalab, and A. Gharehbaghi, “A review on deep learning methods for ECG arrhythmia classification,” *Expert Syst. with Appl. X*, vol. 7, p. 100033, 2020, doi: 10.1016/j.eswax.2020.100033.

[4] E. A. Perez Alday *et al.*, “Classification of 12-lead ECGs: The PhysioNet/Computing in Cardiology Challenge 2020,” *Physiol. Meas.*, vol. 41, no. 12, p. 124003, 2020, doi: 10.1088/1361-6579/abc960.

[5] M. A. Reyna *et al.*, “Will two do? Varying dimensions in electrocardiography: The PhysioNet - Computing in Cardiology Challenge 2021,” *Comput. Cardiol. 2021*, no. 48, p. 1:4, 2021, [Online]. Available: https://doi.org/10.13026/jz9p-0m02.

[6] J. R. Torres, K. De Los Rios, and M. A. Padilla, “Cardiac Arrhythmias Identification by Parallel CNNs and ECG Time-Frequency Representation,” *Comput. Cardiol. (2020).*, vol. 2020-Septe, pp. 1–4, 2020, doi: 10.22489/CinC.2020.456.

[7] I. Daubechies, J. Lu, and H. T. Wu, “Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool,” *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 243–261, 2011, doi: 10.1016/j.acha.2010.08.002.

[8] G. Thakur and H. T. Wu, “Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples,” *SIAM J. Math. Anal.*, vol. 43, no. 5, pp. 2078–2095, 2011, doi: 10.1137/100798818.

[9] J. R. Torres-Castillo, C. O. López-López, and M. A. Padilla-Castañeda, “Neuromuscular disorders detection through time-frequency analysis and classification of multi-muscular EMG signals using Hilbert-Huang transform,” *Biomed. Signal Process. Control*, vol. 71, p. 103037, Jan. 2022, doi: 10.1016/j.bspc.2021.103037.

Address for correspondence:

Jonathan R Torres-Castillo & Miguel A Padilla-Castañeda
Instituto de Ciencias Aplicadas y Tecnología (ICAT), UNAM, Cto. Exterior, Cd. Universitaria, Mexico City, 04510, Mexico
jonathanrtc@comunidad.unam.mx
miguel.padilla@icat.unam.mx