Improved Detection of Paroxysmal Atrial Fibrillation Using an ECG-based Semisupervised Model

Sara Artal[1], Juan Pablo Martínez[1,2], Antonio Miguel[1], Julia Ramírez[1,2,3]

[1]Aragon Institute of Engineering Research, University of Zaragoza, Zaragoza, Spain
[2]Centro de Investigación Biomédica en Red, Biomateriales, Bioingeniería y Nanomedicina, Spain
[3]William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

Abstract

Paroxysmal atrial fibrillation (PAF) is a common intermittent supraventricular arrhythmia whos identification can be challenging. The 12-lead electrocardiogram (ECG) is a cheap and non-invasive tool, ideal for diagnosing this condition. In recent years, neural networks (NNs) have been used in cardiology for the prediction of cardiac risk, however, their use is still limited due to the scarcity of well-curated labelled data. The objective of this project is to check if pretraining enhances the performance of NNs in a supervised learning task for the diagnosis of PAF, given limited training data with an unbalanced distribution of cases. First, ECG datasets publicly available were downloaded and preprocessed, resulting in a total of 1,602,185 ECG signals. Then, a multilayer convolutional neural network (CNN) was implemented for the diagnosis of PAF. Next, we developed and pretrained, in a self-supervised manner, a NN based on contrastive predictive coding. Both supervised and semisupervised models showed a good performance. By using a simplified CNN together with a self-supervised pretrained NN, accuracy improved from 0.565 to 0.680, the area under the curve rose from 0.692 to 0.714, and specificity at a sensitivity of 0.750 grew from 0.534 to 0.579. Our findings support that the semisupervised learning method improved the performance and simplified the supervised model architecture and training.

1. Introduction

One of the main contributors to cardiovascular death is atrial fibrillation (AF), an arrhythmia that features an irregular and rapid heart rate, which can also occur intermittently, resulting in paroxysmal atrial fibrillation (PAF). There are invasive procedures that allow for the precise identification of PAF, but they cannot be used for large-scale risk stratification, highlighting the need for non-invasive methods with high detection rates [1, 2].

The electrocardiogram (ECG) signal is one of the most widely used medical tests for the early detection of PAF. Recently, neural networks (NNs) have opened up new opportunities for ECG signal analysis, automatically identifying the most important morphological aspects that contribute to AF risk [3, 4]. However, the results of these studies are still limited, potentially because these deep NN models require large annotated training sample sizes to achieve good performance results, which is usually difficult and expensive to obtain. Moreover, these models are sensitive to data imbalance (for example, the prevalence of AF in the general population is estimated to be 2-4%), therefore requiring even larger cohorts to obtain statistical power [5, 6].

In other applications, pretraining a NN to solve simple tasks with unlabelled data has been shown to improve the performance of supervised models when a limited amount of training data is available. In this work, we rely on Contrastive Predictive Coding (CPC) [7], a model that combines prediction of future observations with a probabilistic contrastive loss to learn abstract representations in a self-supervised manner. The main intuition behind CPC is to learn the representations that encode the underlying information shared between the various parts of the signal, while also discarding low-level and local noise information. The result of combining these self-supervised pretrained models with the supervised ones is known as semisupervised learning [8].

The main objective of this project is, thus, to study whether a semisupervised model improves the performance of a supervised model in the diagnosing of PAF, a case in which training data are scarce and the proportion of cases is unbalanced (Figure 1).

2. Materials and Methods

2.1. Materials

The data used in this work come from different databases available in PhysioNet. Table 1 details the main

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.365

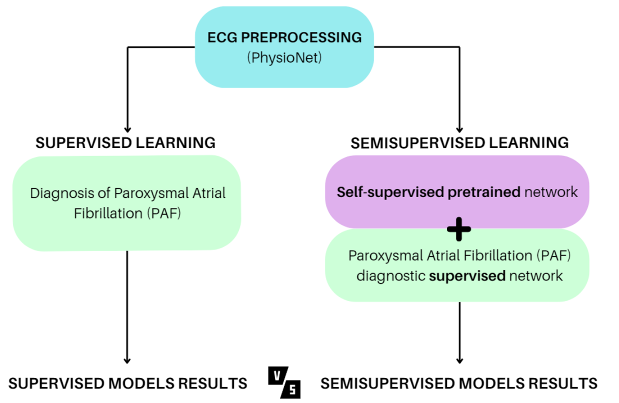

Figure 1. Illustrative diagram for the comparison of supervised and semisupervised models for the detection of PAF.

characteristics of the selected databases.

Table 1. Databases used in this project.

Database	Total number of patients	Total number of recordings	ECG leads	Sampling Frequency (Hz)	Duration
ICENTIA 11K	11,000	542,157	1	250	∼70 min
ECG ARRHYTHMIA	45,152	45,152	12	500	10 s
PTB-XL	18,869	21,799	12	500	10 s
GEORGIA	15,742	10,344	12	500	10 s
CHALLENGE 2017	8,528	8,528	1	300	30 to 60 s
CPSC 2018	9,458	6,877	12	500	6 to 60 s
CPSC 2018 EXTRA	9,458	3,453	12	500	6 to 60 s
AFPDB	98	300	2	128	30 or 5 min

2.2. ECG Preprocessing

All ECG signals were first resampled (or decimated) at 500 Hz using a quadrature polyphase filter (filter bank) that reduces edge effects. To remove the baseline noise, an order 4 Butterworth high-pass IIR filter with a cut-off frequency of 0.3 Hz was used.

2.3. Supervised learning model

The objective of the supervised learning task, used to test the assumption initially made, is to diagnose PAF from a segment of ECG in sinus rhythm. Therefore, given an input ECG signal in sinus rhythm, the NN must decide whether or not that subject will have episodes of AF.

A convolutional NN (CNN) was used, which was composed of two parts, the feature extraction and the classification part. As shown in Figure 2, feature extraction is performed using 6 1D convolutional layers of 128 hidden units, with a kernel size of 3, as well as a stride of 2. These layers operate in one dimension of the input signal (ECG lead), implying a global decimation of the signal by a factor 64. After the convolutional layers, we used a bidirectional LSTM layer [9], which enables storing relevant information from the past. Additionally, batch normalization [10] between convolutional layers was used to improve convergence speed, performance, and NN stability. We selected the ReLU as the activation function for the entire NN in order to learn non-linear relationships present in the data. Moreover, to perform the classification, a single fully connected layer was used. This layer will produce a single output that represents the probability of belonging of a given input to the positive class (AF).

The database used in this task is AFPDB (Table 1). The training set consisted of 300 ECG signals and the test set of 200 ECG signals. The model was trained for 100 epochs with 300 samples per epoch, and we used a learning rate with a starting value of 0.003, which was reduced every 20 iterations by 5%. In addition, the data was processed in batches of size 64 and the NN inputs were 10-second ECG fragments, equivalent to 5,000 samples at a sample rate of 500 Hz.

To measure the performance of the model during the test, precision, the ROC (Receiver Operating Characteristic) curve and the AUC (area enclosed under the ROC curve) were obtained. A sensitivity threshold of 0.750 was established to achieve certain quality requirements.

2.4. Self-supervised learning model

The architecture of the developed CPC model, shown in Figure 2, can be divided into two blocks. First, it consists of a non-linear encoder that extracts the characteristics of the input ECG signals. It maps the sequence of raw input samples to a sequence of latent representations (z) [7]. In our case, it is composed of 6 1D convolutional layers, all of them with 1024 hidden units, a kernel size of 3 and a stride of 2. Consequently, the temporal resolution of the latent representations, z, will be potentially lower. Again, batch normalization and the ReLU activation function were used between convolutional layers.

The output of the encoder (z) is the input of the second block. This is an autoregressive (AR) block, which is composed of a GRU layer [11] that stores relevant information for several past time instants by means of a state variable, in order to generate predictions of how the signal will evolve in future samples. At the output of the GRU layer, the context information (c) is obtained, called the contextual latent representation. This will be used to make predictions of future k steps, in our case, a different prediction will be obtained for each one of them (k = 8) [7]. The prediction layer (Pred) consists of two linear layers and a ReLU between them, therefore, we are applying a nonlinear transformation to the context information. Actually, this layer can be seen as a Hinton SimCLR [12].

For the self-supervised task, all databases in Table 1 were available except AFPDB, as it was used for the control supervised model. In total, we had 1,602,185 signals, of which 1,600,000 signals were allocated to the training set, and the remaining signals, 2,185, to the test set. As an initial approach to train the self-supervised model, we use a training set that consists of 637,448 ECG signals (lead I). Moreover, the model was trained for 3,000 epochs with 50,000 samples per epoch, as well as with a learning rate with a starting value of 0.001, which was reduced every 40 iterations by 5%. The data was processed in batches of size 256 and the NN inputs were 3-second ECG fragments, equivalent to 1,500 samples at a sample rate of 500 Hz.

2.5. Semisupervised learning models

The aim of the semisupervised learning model is to assess the added value of the pretrained self-supervised learning model in the PAF diagnosis task. For this purpose, the internal representations of the self-supervised model, z (latent representation) and c (contextual latent representation), were used as the input for the training of the supervised learning model. Thus, the number of channels of the supervised model in the first convolutional layer was modified to match the dimension of the output representation of the self-supervised model (z or c). Figure 2 presents a diagram detailing the architecture and dimensions of the implemented semisupervised learning model for the diagnosis of PAF.

Besides, multiple experiments were performed in which the number of convolutional layers in the feature extraction part of the supervised model was modified. Starting from the original supervised model, which had 6 convolutional layers, new supervised models were implemented in which the number of convolutional layers ranged from 6 layers to none. In all cases, after the convolutional layers, the bidirectional LSTM layer was mantained, except in the last case, where the linear classification layer was simply left.

3. Results

In the supervised PAF diagnostic task, we obtained an accuracy (considering a 0.500 threshold) of 0.565, an area under the curve (AUC) of 0.692 and a specificity of 0.534 for a sensitivity of 0.75 (Table 2, metrics obtained using the raw samples as inputs for the supervised model with 6 convolutional layers). Therefore, the proposed binary classification model has a certain capacity for discrimination, without being highly precise.

On the other hand, the results obtained for the different experiments made for the semisupervised task are also shown in Table 2. Of all these combinations, the one that achieved the best results has been the union of the original supervised model and the self-supervised model, by passing the variable c (contextual latent representation) as input to the supervised model, while reducing the number

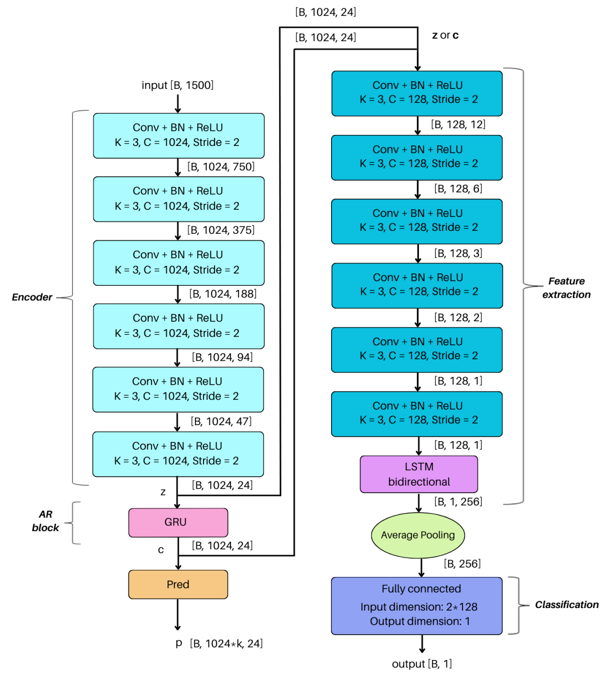

Figure 2. Architecture of the semisupervised learning model. On the left, the self-supervised model architecture (encoder and AR block) and, on the right, the supervised model architecture with 6 1D convolutional layers at the feature extraction part.

of convolutional layers in the supervised feature extraction part from 6 to 2. In this way, the results improve considerably with respect to those obtained in the supervised control task. We now get an accuracy, for a 0.500 threshold, of 0.680, an AUC of 0.714 and a specificity of 0.579 for a sensitivity of 0.750.

4. Discussion and Conclusions

In this study, we analyzed the performance improvement produced by using a pretrained predictive coding model in a supervised PAF diagnostic task. Our main finding is that the self-supervised model is capable of providing notable improvements to the original supervised learning task in a setting where available data is limited and unbalanced.

The implementation of this technique entails a methodological improvement. The quality of our results has increased, achieving greater accuracy and consistency through the use of a simpler NN architecture with fewer convolutional layers at the feature extraction part. In the original supervised task we used 6 convolutional layers, now, in semisupervised tasks, even with a single convolutional layer in the supervised model we are able to achieve acceptable results. Moreover, only a small portion of the available signals (637,448 ECGs) was used to train the self-supervised model, so results are expected to further

Table 2. PAF task detection experiment in the AFPDB database measuring accuracy (Acc), area under the curve (AUC) and specificity (Sp) for different supervised and semisupervised model architectures.

Metric	Input features	Number of convolutional layers in the supervised model						
		6	5	4	3	2	1	0
Acc	raw samples	0.565	0.565	0.525	0.490	0.575	0.575	0.420
	self-sup. z	0.605	0.590	0.555	0.615	0.585	0.570	0.490
	self-sup. c	0.485	0.630	0.590	0.580	**0.680**	0.670	0.520
AUC	raw samples	0.692	0.583	0.634	0.516	0.637	0.567	0.499
	self-sup. z	0.634	0.605	0.589	0.633	0.642	0.653	0.547
	self-sup. c	0.604	0.687	0.629	0.608	**0.714**	0.661	0.610
Sp	raw samples	0.534	0.364	0.523	0.318	0.398	0.364	0.307
	self-sup. z	0.477	0.455	0.329	0.477	0.420	0.398	0.341
	self-sup. c	0.523	0.568	0.421	0.455	**0.579**	0.546	0.432

improve with the full database (over 1.5M ECGs). Additionally, the simplification of the supervised architecture comes with a reduction in the number of operations and parameters to be calculated, decreasing memory and time requirements during the training process of the models, as well as the risk of overfitting, achieving more generalizable models. However, the cost savings are not applicable to the testing phase, as both the self-supervised and the supervised models must be evaluated at this stage.

Currently, one of the difficulties of the application of NNs for PAF detection using the ECG is to have sufficiently large volumes of labelled data. The proposed method, in addition to improving the results compared to a purely supervised model, opens the door to the use of NNs in problems where the number of available signals is reduced, as is the case with low-prevalence cardiovascular diseases. Our findings show the enormous potential of NNs to significantly improve the prediction of cardiac risk.

Acknowledgments

This work was supported by projects PID2021-128972OA-100 and CNS2023-143599 funded by Spanish Ministry of Science and Innovation (MCIN) and FEDER, and to the Biomedical Signal Interpretation and Computational Simulation (BSICoS) Group T39-23R 2014-2020. SA acknowledges funding from a scholarship for undergraduate students in the framework of their bachelor degree thesis granted by the Aragon Institute of Engineering Research. JR acknowledges funding from fellowship RYC2021-031413-I from MCIN.

References

[1] Welton N, McAleenan A, Thom H, Davies P, Hollingworth W, Higgins J, et al. Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis. Health Technology Assessment 2017;21(29).

[2] Faust O, Ciaccio E, Acharya U. A review of atrial fibrillation detection methods as a service. International Journal of Environmental Research and Public Health 2020;17(9).

[3] Attia Z, Noseworthy P, López F, Asirvatham S, Deshmukh A, Gersh B, et al. An artificial intelligence-enabled ecg algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. The Lancet 2019;394(10201):861–867.

[4] Kashou A, May A, Noseworthy P. Artificial intelligence-enabled ecg: a modern lens on an old technology. Current cardiology reports 2020;22:1–8.

[5] Liu H, Zhao Z, She Q. Self-supervised ecg pre-training. Biomedical Signal Processing and Control 2021; 70:103010.

[6] Buda M, Maki A, Mazurowski M. A systematic study of the class imbalance problem in convolutional neural networks. CoRR 2017;abs/1710.05381.

[7] van den Oord A, Li Y, Vinyals O. Representation learning with contrastive predictive coding. arXiv preprint arXiv180703748 2018;.

[8] Farouk M, Schwenker F. Semi-supervised Learning. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013; 215–239.

[9] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation 1997;.

[10] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv150203167 2015;.

[11] Cho K, van Merriënboer B, Bahdanau D, Bengio Y. On the properties of neural machine translation: Encoder–decoder approaches. arXiv preprint arXiv14091259 2014;.

[12] Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. Proceedings of Machine Learning Research 2020; 119:1597–1607.

Address for correspondence:

Julia Ramírez
c/Mariano Esquillor s/n, Lab 6.1.07, 50018, Zaragoza, Spain
Julia.Ramirez@unizar.es