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

Sara Artal, Juan Pablo Martínez, Antonio Miguel, Julia Ramírez

Aragon Institute of Engineering Research, University of Zaragoza
Zaragoza, Spain

Background: Paroxysmal atrial fibrillation (PAF) is a common intermittent supraventricular arrhythmia with a challenging identification and detection. The 12-lead electrocardiogram (ECG) is a non-invasive and affordable test ideal for diagnosing this condition. In recent years, neural networks (NNs) have shown promise in the prediction of cardiac risk, including PAF, but their use is still limited due to the scarcity of well-curated labelled data, as well as case-control imbalance. In this project we evaluated the improvement provided by an unsupervisedly pretrained NN in a supervised learning task for the diagnosis of PAF.

Materials and methods: First, ECG datasets publicly available were downloaded and preprocessed, resulting in a total of 1,602,785 ECG signals. Then, a supervised model based on a multilayer convolutional NN (CNN) was implemented for the diagnosis of PAF. Next, we developed and pretrained an unsupervised model based on a predictive coding algorithm, known as contrastive predictive coding. Finally, different semisupervised tasks were implemented, combining the unsupervised model with the supervised one, in order to optimize the detection of PAF.

Results: In the supervised task, using 6 convolutional layers at the feature extraction part, we obtained an accuracy (Acc) of 0.565, an area under the curve (AUC) of 0.692 and a specificity (Sp) of 0.534 for a sensitivity (Se) of 0.750. The semisupervised task, with a simplified CNN (2 convolutional layers) and an unsupervisedly pretrained NN, showed substantial improvements, the Acc improved to 0.625, AUC rose to 0.721, and Sp, at a Se of 0.750, increased to 0.602.

PAF detection measuring Acc, AUC and Sp for the supervised and semisupervised model.

	Number of layers in the feature extraction block	
	6	**2**
Acc	0.565	0.625
AUC	0.692	0.721
Sp	0.534	0.602

Conclusions: Our main finding is that the semisupervised learning model improved the performance and simplified the supervised model architecture and training. This work could be expanded to other scenarios with limited available signals, like predicting ventricular arrhythmias in the general population, helping in the early detection and prevention of these conditions.