Deep Transfer Learning for Detection of Atrial Fibrillation using Holter ECG Color Maps

Todor Stoyanov [1], Vessela Krasteva [1], Stefan Naydenov [2], Ramun Schmid [3], Irena Jekova [1]*

[1]Institute of Biophysics and Biomedical Engineering, Sofia, Bulgaria
[2]Medical University of Sofia, Sofia, Bulgaria
[3]Schiller AG, Baar, Switzerland

Long-term Holter-ECG monitoring is recommended for patients with suspected arrhythmias, as they may be transient and missed on routine resting 12-lead ECG. Often, identification of short-term arrhythmic events is challenging in 24-72-hour recordings. Our recent clinical case series study on supraventricular arrhythmias revealed that cardiologists can rapidly analyze Holter-ECG recordings manually and accurately diagnose by observing a compressed rhythm representation via ECHOView color maps (medilog DARWIN2, Schiller AG, Switzerland). This study aims to show that pre-trained image recognition neural networks can be effectively tuned to use similar color maps to automatically classify atrial fibrillation (AF).

Data were extracted from a recently published large ECG-Holter monitoring database with paroxysmal AF (IRIDIA-AF), including 167 records of 2-lead ECG sampled at 200 Hz, with duration of 1609-hours (AF) and 6690-hours (total). From each lead, we generated images of stacked color bars, transforming the ECG waveform amplitude of sequential heartbeats into a color code. The height (300-pixels) is defined by the heartbeat time-resolution (1500ms@200Hz); the width equals the number of heartbeats in the analysis episode (30-seconds), varying with heart rate. Rescaled images were used for network-based deep transfer learning of VGG16 model, where the 5-bottom convolutional layers were pre-trained in the source domain, while the 3-top dense layers were significantly reduced (128-128-1 neurons) and retrained with IRIDIA-AF for a binary classification (AF vs. non-AF).

On 13,536 training and 6,624 validation images with equal AF/non-AF proportions, the VGG16-model presented true positive rate TPR (99.97% training, 97.16% validation), true negative rate TNR (100%, 97.83%), F1-score (99.99%, 97.49%), AUC (0.999, 0.994). Testing with full-length recordings of 10 independent patients (15,997 non-AF, 12,019 AF) presented TPR=98.76%, TPR=98.49%, F1-score=98.38%.

Colormap images consisting stacked heartbeat amplitudes are an interpretable visual modality for pre-trained image classification networks that can be effectively used to automatically identify short AF episodes in Holter-ECG recordings.

Acknowledgement: This work was supported by the Bulgarian National Science Fund, grant number КП-06-Н42/3.