# Application of Deep Learning for Quality Assessment of Atrial Fibrillation ECG Recordings

Álvaro Huerta, Arturo Martínez-Rodrigo, Miguel A Arias, Philip Langley, José J Rieta, Raúl Alcaraz

Research Group in Electronic, Biomedical and Telecommunication Engineering, University of Castilla-La Mancha, Cuenca, Spain

**Background and Aim**. Noise of any origin is the unavoidable key issue in biomedical signal interpretation, leading to significant reductions in diagnostic and decision capabilities of automatized ECG-based systems. Aimed at overcoming this problem, a variety of algorithms has been proposed to automatically discern between sufficiently clean and poor-quality ECG intervals. Although accuracies higher than 90% have been reported on sinus rhythm (SR) recordings, performance is drastically decreased when validated on atrial fibrillation (AF) recordings. This work introduces a novel algorithm to reliably identify poor-quality ECG segments within the challenging environment of recordings alternating SR and AF episodes.

**Methods**. This research is based on the high learning capability of the convolutional neural network AlexNet. It was trained with 2D images obtained when turning 5 s-length ECG segments into scalograms through a continuous Wavelet transform. For its validation, the training set proposed for the PhysioNet/CINC Challenge 2017 was used. The 8528 recordings available were firstly segmented into 5 s-length intervals and then grouped into two categories. Because the dataset contained four kinds of rhythms, i.e., SR, AF, other rhythms (OR) and noisy signals, segments from SR, AF and OR constituted the high-quality group, and noisy intervals the poor-quality one. Both groups then consisted of 47349 and 1168 segments, respectively.

**Results**. Given the high imbalance between both categories, the method was exposed to 40 learning–testing cycles. In each iteration all poor-quality ECG intervals were maintained and 1150 samples were randomly selected from the high-quality group. In average for all cycles, values of accuracy, sensitivity, and specificity were 91.3$\pm$2.4%, 90.4$\pm$2.7%, and 93.2$\pm$1.3%, respectively. Moreover, the mean rate of AF intervals correctly classified was 92.4$\pm$3.3%, thus improving by more than 20% performances of most previous ECG quality assessment algorithms dealing with AF signals.

**Significance**. The proposed deep learning method facilitates automatized diagnosis with higher accuracy in long-term monitoring of patients with intermittent AF, since potential confounding bias of poor-quality ECG intervals could be detected and excluded from later interpretations.