

CNN-based Quality Control to Reduce False Positive Rate in AF Screening

Hesam Halvaei¹, Emma Svennberg², Leif Sörnmo¹, Martin Stridh¹

¹ Biomedical Signal Processing Group, Department of Biomedical Engineering, Lund University, Lund, Sweden

² Department of Clinical Sciences, Danderyd University Hospital, Karolinska Institutet, Stockholm, Sweden

Abstract

Early detection of AF is essential and emphasizes the significance of AF screening. However, AF detection in screening ECGs, usually recorded by handheld and portable devices, is limited because of their high susceptibility to noise. In this study, the feasibility of applying a machine learning-based quality control stage, inserted between the QRS detector and AF detector blocks, is investigated with the aim to improve AF detection. A convolutional neural network was trained to classify the detections into either true or false. False detections were excluded and an updated series of QRS complexes was fed to the AF detector. The results show that the convolutional neural network-based quality control reduces the number of false alarms by 26.5% at the cost of 1.5% decrease in sensitivity compared to AF detection without any quality control.

1. Introduction

Atrial fibrillation (AF) is a supraventricular tachyarrhythmia originating from uncoordinated atrial contractions [1]. Detection of AF is usually based on the ECG, and the main characteristics are an *irregular rhythm*, *absence of P-waves*, and *presence of f-waves*. Due to the high risks associated with untreated AF, screening is becoming increasingly common [2–5].

A large number of AF detectors have been proposed over the years, see, e.g., [6–11], that may be divided into rhythm-based AF detectors and rhythm- and morphology-based AF detectors. Rhythm-based AF detectors explore the degree of irregularity in the RR interval series and can be based on entropy measures [9], statistical measures [8], or Poincaré plot features [7]. Rhythm- and morphology-based AF detectors, in addition to rhythm irregularity information, detects atrial activity, i.e., P-waves or f-waves [10, 11].

AF screening is usually performed using handheld and

portable devices. These devices allow for simplified measurements at home but the recorded signals are often of lower quality compared to the conventional ECG recordings [12, 13]. A lower signal quality influences the performance of AF detection algorithms and typically leads to a higher false alarm rate. Therefore, handling of lower signal quality is an important problem to address. In this study, we propose a signal quality control stage and investigate its influence on AF detection performance in short-term screening ECGs. This stage is a CNN that classifies the outputs of the QRS detector described in [14] into either true or false detections. A rhythm-based AF detector was used since atrial activity is difficult to quantify in screening ECGs. The significance of the CNN-based quality control is investigated on the Physionet/CinC 2017 Challenge database [15].

2. Databases

Two databases were used for method development and evaluation: The Physionet/CinC 2017 Challenge database (CinC17DB) [15] and the StrokeStop I database (SSIDB) [3].

SSIDB can be divided into two parts from the expert annotation perspective: signal-level and patient-level annotations. Half of the database is annotated at the signal level, indicating whether AF episodes are present in the ECG. The other half is annotated at the patient-level, indicating whether a patient has AF and does not reveal in which of the ECG recordings AF is present.

2.1. Software-generated categorization

In addition to the database annotations, labels generated by ECG Parser (<https://cardiolund.com>), a CE-marked ECG analysis software, were used. The software divides the database into the following categories: poor quality, sinus rhythm, bigeminy, trigeminy, SVES, VES, fast/slow rhythm, pause/AV block, and irregular signal, see Fig. 1.

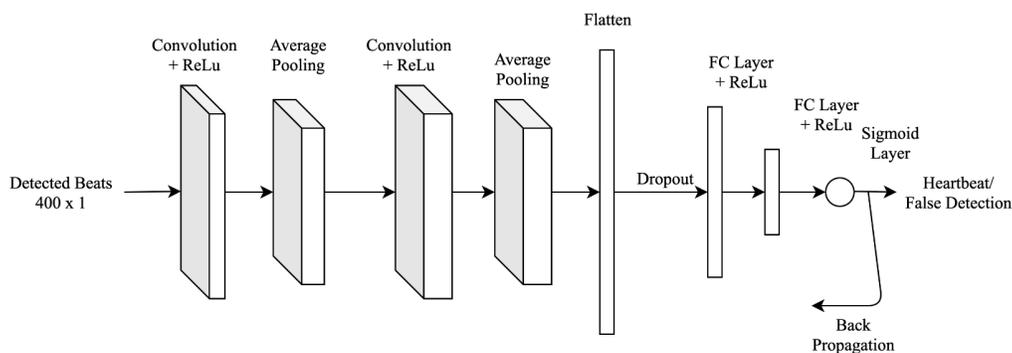


Figure 2. Structure of the proposed CNN for quality control

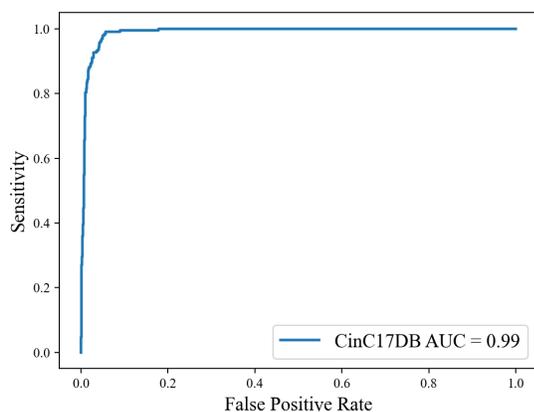


Figure 3. The ROC for CinC17DB detections test set.

recordings better. The ROC curve for the CinC17DB detection test set is displayed in Fig. 3. It shows that an AUC of 0.99 was obtained for CinC17DB.

The trained CNN was applied to the group of irregular signals in the CinC17DB to investigate the effect of quality control on AF detection performance. The AF detection sensitivity and FPR for different AF detection thresholds are displayed in Fig. 4, with and without using the quality control stage. A high sensitivity achieved without quality control came at the expense of a high false positive rate, see Fig. 4.

By inserting the quality control stage, the false positive rate was reduced by 17%-26.5%, while the sensitivity decreased marginally. For the highlighted threshold at 0.39 in Fig. 4, the CNN-based quality control stage reduced the false positive rate by 26.5% compared to AF detection alone. This improvement was achieved at the expense of 1.5% reduction in sensitivity.

5. Discussion

In this study, we applied a CNN to identify and exclude false QRS detections produced by the QRS detector, which deteriorate the performance of the rhythm-based AF detector. The quality control stage could have been merged with the QRS detector, the AF detector, or the whole pipeline could have been integrated into one ML-based box. However, merging would disrupt the traceability of decisions, while using the proposed pipeline, the decisions can be traced backwards. Besides, well-established methodology for QRS and AF detection does not have to be reinvented. The resulting pipeline provided a significant reduction in the false positive rate at the expense of small reduction in sensitivity.

Different QRS detectors and AF detectors could have been used. However, false detections by the QRS detector in noisy signals and numerous false AF alarms in screening databases are well-documented. The solution proposed in this work may be combined with any other modules in different settings and the results in this paper represents an example of performance improvement that can be achieved.

Acknowledgments

This research was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.766082 (MY-ATRIA project).

References

- [1] Platonov PG, Corino VD. A clinical perspective on atrial fibrillation. In *Atrial Fibrillation from an Engineering Perspective*, (L. Sörnmo, ed.). Springer, 2018; 1–24.
- [2] Engdahl J, Andersson L, Mirskaya M, Rosenqvist M. Stepwise screening of atrial fibrillation in a 75-year-old population: implications for stroke prevention. *Circulation* 2013; 127(8):930–937.

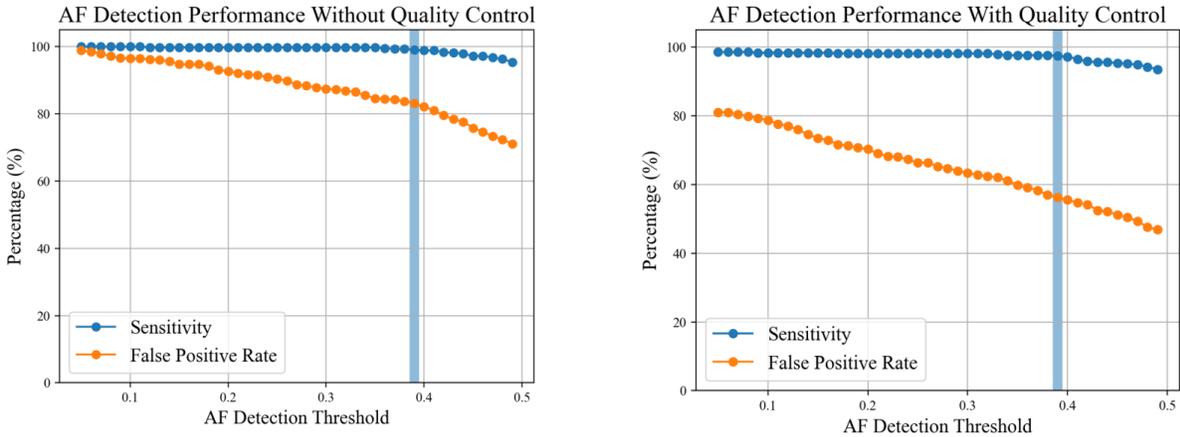


Figure 4. AF detection performance on CinC17DB irregular signals

[3] Svennberg E, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Mass screening for untreated atrial fibrillation: the Strokestop study. *Circulation* 2015; 131(25):2176–2184.

[4] Proietti M, Mairesse GH, Goethals P, Scavee C, Vijgen J, Blankoff I, Vandekerckhove Y, Lip GY. A population screening programme for atrial fibrillation: a report from the Belgian heart rhythm week screening programme. *Eurpace* 2016;18(12):1779–1786.

[5] Chan Ny, Choy Cc. Screening for atrial fibrillation in 13 122 Hong Kong citizens with smartphone electrocardiogram. *Heart* 2017;103(1):24–31.

[6] Dash S, Chon K, Lu S, Raeder E. Automatic real time detection of atrial fibrillation. *Ann. Biomed. Eng.* 2009; 37(9):1701–1709.

[7] Lian J, Wang L, Muessig D. A simple method to detect atrial fibrillation using RR intervals. *Am. J. Cardiol.* 2011; 107(10):1494–1497.

[8] Huang C, Ye S, Chen H, Li D, He F, Tu Y. A novel method for detection of the transition between atrial fibrillation and sinus rhythm. *IEEE Trans. Biomed. Eng.* 2010;58(4):1113–1119.

[9] Petrėnas A, Marozas V, Sörnmo L. Low-complexity detection of atrial fibrillation in continuous long-term monitoring. *Comput. Biol. Med.* 2015;65:184–191.

[10] Ródenas J, García M, Alcaraz R, Rieta JJ. Wavelet entropy automatically detects episodes of atrial fibrillation from single-lead electrocardiograms. *Entropy* 2015;17(9):6179–6199.

[11] Ladavich S, Ghoraani B. Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity. *Biomed. Signal Proc. Control* 2015;18:274–281.

[12] Svennberg E, Stridh M, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Safe automatic one-lead electrocardiogram analysis in screening for atrial fibrillation. *Eurpace* 2017;19(9):1449–1453.

[13] Freedman B, Camm J, Calkins H, Healey JS, Rosenqvist M, Wang J, Albert CM, Anderson CS, Antoniou S, Ben-

jamin EJ, et al. Screening for atrial fibrillation: a report of the AF-screen international collaboration. *Circulation* 2017;135(19):1851–1867.

[14] Lagerholm M, Peterson C, Braccini G, Edenbrandt L, Sörnmo L. Clustering ECG complexes using Hermite functions and self-organizing maps. *IEEE Trans. Biomed. Eng.* 2000;47(7):838–848.

[15] Clifford GD, Liu C, Moody B, Li-wei HL, Silva I, Li Q, Johnson A, Mark RG. AF classification from a short single lead ECG recording: The Physionet/Computing in Cardiology Challenge 2017. In *2017 Comput. Cardiol. (CinC)*, 2017; 1–4.

[16] Sörnmo L, Petrėnas A, Marozas V. Detection of atrial fibrillation. In *Atrial Fibrillation from an Engineering Perspective*, (L. Sörnmo, ed.). Springer, 2018; 73–135.

Address for correspondence:

Hesam Halvaei
Lund University, Department of Biomedical Engineering, Box 118, 22100, Lund, Sweden
hesam.halvaei@bme.lth.se