

An Intelligent Cardiac Health Monitoring and Review System

T. Zhu^{*‡}, M. Osipov^{*‡}, T. Papastylianou^{*‡}, J. Oster^{*}, D. A. Clifton^{*‡}, and G. D. Clifford^{*‡‡}

^{*}Institute of Biomedical Engineering, University of Oxford, UK. Email contact: tingting.zhu@eng.ox.ac.uk,

[†]Depts. of Bioinformatics and Biomedical Engineering, Emory University and Georgia Institute of Technology, USA and

[‡]Centre for Affordable Healthcare Technology, Kellogg College, University of Oxford, UK.

Keywords: Crowdsourcing, Signal-Quality, m-Health, Remote-Monitoring

Abstract

Cardiac monitoring is an important application of mobile-monitoring systems. Such systems typically require experts to label electrocardiogram (ECG) data; however, large inter- and intra-expert variation limits the reliability and accuracy of diagnosis. This paper presents a process of integrating the mobile end of the system with a back-end annotation system for reviewing and scoring the quality of an ECG signal. This process serves as a platform for remote cardiac health monitoring and diagnosis. Furthermore, a “crowdsourcing” methodology was used to provide an adjudication of ECG annotations from a set of “experts”, comprising human trainees and automated algorithms. We show that this can provide diagnoses with equivalent accuracy to that of experts, at a substantially reduced cost.

1 Introduction

Cardiovascular disease, which currently accounts for 30% of deaths worldwide, is predicted to remain as the single leading global burden of injury and death for the next 20 years, particularly affecting lower- and middle-income countries [1]. Smartphones or mobile tablets have become increasingly popular as potential medical diagnostic devices due to their relatively low cost, connectivity, mobility, and processing power. Clinicians now can view or edit patients’ electronic healthcare records (EHRs) on their smartphones or tablets.

The electrocardiogram (ECG) is a standard and powerful tool for assessing cardiovascular health because many heart conditions manifest as abnormalities in the ECG signals. Expert annotation is considered to be the gold standard in ECG diagnosis. However, auditing the quality of these annotations is often lacking, resulting in intra- and inter-expert variation as high as 30% in diagnosis [2, 3]. Currently, there is no standardisation to regulate ECG reporting or method of measuring the level of expertise in medical practitioners [4].

A process of integrating the mobile end and a back-end annotation system for reviewing and scoring the quality of an

ECG signal was developed. It provides a platform for clinicians to gain remote access to patients’ EHRs and allows for health monitoring, diagnosis, and training.

2 System design

2.1 Mobile application

We developed an ECG analysis app that allows for ECG acquisition from sensors connected to the phone via Bluetooth. Subsequent analysis of the ECG can be divided into three steps: 1) beat (QRS) detection, 2) Signal Quality Index (SQI) assessment, and 3) arrhythmia detection. Figure 1 shows a screenshot of the smartphone application. An open-source version of the associated algorithms is available [5].

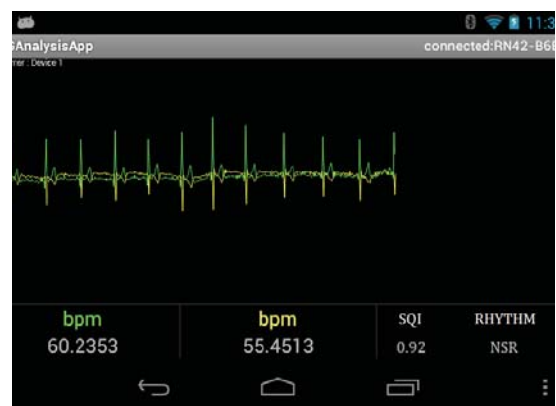


Fig. 1: Screenshot of the smartphone application demonstrating a two channel ECG (i.e.: yellow and green traces) as well as the corresponding heart rate, the signal quality index (SQI), and the rhythm (NSR, or normal sinus rhythm).

1) QRS detection: this is performed using an energy-based approach [6–8] with modifications to deal with saturation and signal loss.

2) SQI: this is computed using previously reported methods [9–11]. A range of individual SQI algorithms were evaluated using an extended version of the dataset from the PhysioNet/Computing in Cardiology Challenge 2011, which comprises 30,000 labelled 10s ECG segments. The best-performing SQI was deemed to be “bSQI”, which involves a

comparison of two QRS detectors. The implemented version includes a support vector machine (SVM) approach for combining individual SQI metrics to form a more robust SQI (see Behar *et al.* [11] for more details).

3) Arrhythmia: atrial fibrillation (AF) is a heart condition characterised by an irregular heart rate due to the chaotic firing of the atrio-ventricular node of the heart, and is known to be the most common sustained cardiac arrhythmia [12]. It is estimated that 25% of people over the age of 40 years will develop AF during their lifetime [13]. AF was identified to be an independent risk factor for stroke resulting in four-fold excess risk [14]. This condition is commonly only detected after a patient exhibits a serious complication related to AF, such as stroke or heart failure. The recent advent of mobile technology provides a screening and follow-up pathway for AF management. The impact of such technology could be life-saving and might reduce associated healthcare costs significantly. The AF-detection algorithm used in our work is based on an SVM-approach to fusing multiple features [15].

2.2 Server application

The server infrastructure is based on a client-server model [16], which uses OpenMRS as an open-source medical record system at the back-end, SanaMobile as the intermediate server system, and which has a front-end Android client. The server-side software is able to capture data uploaded from the Android client (SanaApp), and a set of APIs allow data to be entered and viewed in both CrowdLabel (an annotation interface with faculty to merge independent annotations) and OpenMRS. In addition to the signal analysis described above, we integrated a novel algorithm into our server system, which uses a probabilistic methodology for combining multiple (potentially disagreeing) annotations of medical data. In previous studies, we described this method in its application to improving the estimation of ECG intervals and R-peak locations using multiple expert / non-expert annotators [17, 18].

2.3 System use

The system architecture and user interaction are shown in Figure 2. During a visit to a patient, a healthcare worker may acquire ECGs from the patient and upload them using SanaApp for remote evaluation.

Once a new ECG recording is submitted, annotators (such as trainees) may be notified via e-mail. Annotators may then log in to CrowdLabel, and perform ECG annotation and analysis using CrowdLabel tools [19]. There may be multiple annotators analysing the same ECG recording; their resulting annotations are combined using a voting algorithm [20] when a new annotation is added to the system.

A clinician may review patients' records periodically using OpenMRS with the "Sana Queue" plugin (see Figure 3). In the Queue, all pending cases are displayed. After choosing a case, the clinician can review the ECG recording from the

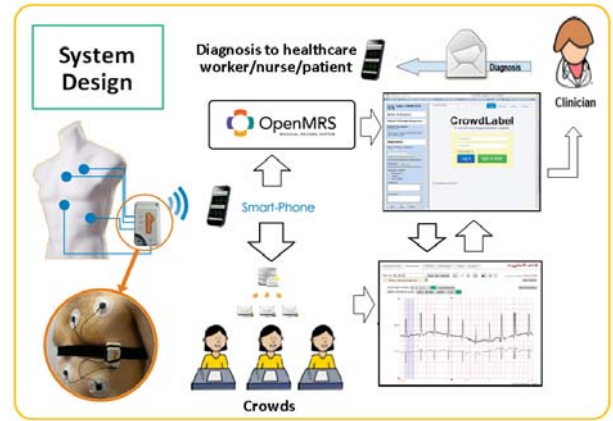


Fig. 2: System architecture and user interaction.

CrowdLabel interface, and may access the results of the voting algorithm used to combine annotations.

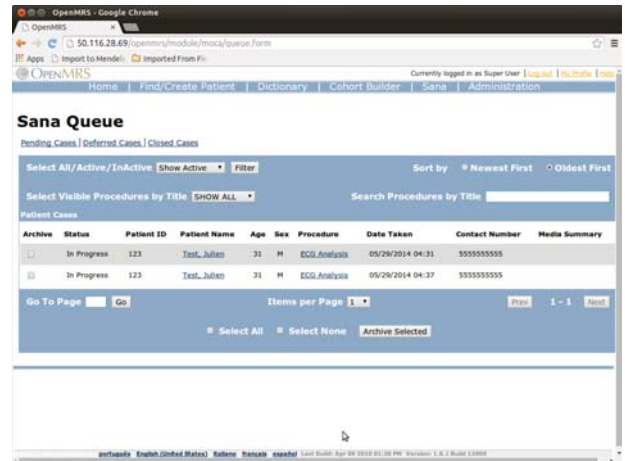


Fig. 3: The OpenMRS Sana Queue interface.

The clinician may also perform ECG analysis using the voting algorithm in real-time by interacting with the *encounter user interface (UI)* (see Figure 4). After reviewing the uploaded case, the clinician may submit diagnoses and recommendations and send these to the healthcare worker via the OpenMRS interface.

3 System validation

3.1 ECG app

The QRS detector was evaluated on the MIT-BIH arrhythmia database, including 48 recordings of 30min ECG signals [21]. It exhibited a sensitivity (Se) of 99.0% and a specificity (Sp) of 98.7%, and showed robust performance on high noise ambulatory data, including recovery from saturation and large movement artifacts. The SQI was shown to have an accuracy (Ac) of 96.9%, $Se = 98.3%$ and $Sp = 95.3%$ (see Behar *et al.* [11] for more details). Accuracy of up to 99.3% was achieved when



Fig. 4: The OpenMRS Sana *encounter* UI. The voting algorithm can be triggered by pressing the buttons highlighted in red.

Method	Se (%)	PPV (%)	Acc (%)
Colloca et al. [15]	99.0	96.1	97.8
Sarkar et al [22]	94.4	97.4	96.4

Table 1: Atrial Fibrillation classification results on the MIT-BIH AF database.

combining multiple SQIs using an SVM approach [11], and was also shown to be robust with pathological ECGs. The AF detector was evaluated on the MIT-BIH AF database [21], which consists of 25 long term recordings of 10 hours each. The AF algorithm achieved performance which exceeded all other published results; the results for the AF detector are presented in Table 1 and compared to the next best reported algorithm in the literature (Sarkar *et al.* [22]).

3.2 The voting system: description and evaluation

1) Adult QT database:

The data were drawn from the 2006 PhysioNet/Computing in Cardiology (PCinC) Challenge [23]. Each participant was required to submit a Q onset with accompanying T offset for each of the 549 recordings in the Physikalisch-Technische Bundesanstalt Diagnostic ECG Database (PTBDB) [24]. The records were obtained from 290 subjects (209 men with mean age of 55.5 and 81 women with mean age of 61.6). About 20% of the subjects were healthy controls. The PTBDB contained records with a variety of ECG morphologies having QT intervals ranging from 256ms to 529ms.

Our proposed voting algorithm, the Probabilistic Label Aggregator (PLA) with contextual features, was applied to manual and automatic annotations [20] to infer a single ground truth for each annotation. The root-mean-squared errors (RMSEs) were calculated using the reference provided by the Challenge and compared with the best-performing scores that were pub-

lished in the 2006 PCinC Challenge. In addition, the PLA was compared to the median and mean voting strategies. The results are detailed in [20]. In summary, the PLA RMSE was 6.04ms when considering all human annotators. It outperformed the mean voting strategy for 20 annotators and the median voting approach after for nine annotators or less (median RMSE = 4.71ms for 20 annotators). Using 15 out of 20 annotators (RMSE = 6.62ms) the PLA achieved a similar error as the best human score (RMSE = 6.65ms) provided in the Challenge. In automatic entries, the PLA consistently outperformed the median and mean voting approaches (14.44ms and 17.67ms respectively) and achieved the minimum RMSE of 13.97ms. This was also lower (i.e. better) than the best performing automatic entrant (RMSE = 16.36ms).

2) Foetal QT database:

A total of 501, 30s segments were extracted from 15 healthy foetal ECG recordings from a private database. Twenty-three volunteers participated in the study and provided a total of 7,307 foetal QT (FQT) annotations, which were aggregated using the PLA [19]. The participants were ranked based on the RMSE when compared to the aggregated annotation generated. The variance of each annotator was further estimated and compared. The annotator with a lower variance indicates high consistency and hence higher precision. The results of standard deviation (σ) of the change of FQT annotations (ΔFQT) of the three annotators (best, medium, and worst performance rated by the PLA) were compared in different signal quality scenarios: When the signal quality of a segment was very good, the σ of the ΔFQT s was 13.35 ms, 19.24 ms, and 18.39 ms for best, medium, and worst annotator respectively; When the signal quality of a segment was poor, the σ of ΔFQT was much larger (35.52 ms, 62.96 ms, and 75.65 ms). Nevertheless, the PLA selected best annotator has always proved to have the least variance across different quality segments.

4 Conclusion

This paper has demonstrated an intelligent cardiac health monitoring and review system. We are currently collaborating with partners in the telecommunications industry to tailor the system to the needs of individual m-Health users. A pilot-study will be conducted with the aim of collaborating with hospitals to assemble a large dataset of ECGs and associated annotations; the latter will then be used to further test and validate the proposed system.

Acknowledgements

The Centre for Affordable Healthcare Technology is supported by ARM and the John Fell Fund. Funding for this work was also provided by China Mobile Research and the EPSRC. TZ, MO, and TP acknowledge the support of the RCUK Digital Economy Programme grant number EP/G036861/1. JO is supported by Wellcome Trust Centre Grant No. 098461/Z/12/Z (Sleep, Circadian Rhythms & Neuroscience Institute). DAC is

supported by the Royal Academy of Engineering and Balliol College, Oxford.

References

- [1] World Health Organisation, “Cardiovascular diseases,” <http://www.who.int/mediacentre/factsheets/fs317/en/index.html>, September 2014.
- [2] R. Bond, T. Zhu, D. Finlay *et al.*, “Assessing Computerised Eye Tracking Technology for Gaining Insight into Expert Interpretation of the 12-lead Electrocardiogram: An Objective Quantitative Approach,” *J Electrocardiol*, 2014.
- [3] S. K. Warfield, K. H. Zou, and W. M. Wells, “Validation of image segmentation by estimating rater bias and variance,” *Philos Trans A Math Phys Eng Sci*, vol. 366, pp. 2361–2375, 2008.
- [4] S. M. Salerno, P. C. Alguire, and H. S. Waxman, “Competency in interpretation of 12-lead electrocardiograms: a summary and appraisal of published evidence,” *Ann Intern Med*, vol. 138, no. 9, pp. 751–760, 2003.
- [5] J. Oster, J. Behar, R. Colloca, Q. Li, Q. Li, and G. D. Clifford, “Open source Java-based ECG analysis software and Android app for atrial fibrillation screening,” in *Comput Cardiol*, vol. 40. IEEE, 2013, pp. 731–734.
- [6] G. D. Clifford, “Signal processing methods for heart rate variability,” Ph.D. dissertation, Department of Engineering Science, University of Oxford, 2002.
- [7] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” *IEEE Trans Biomed Eng*, vol. 32, no. 3, pp. 230–236, 1985.
- [8] M.-E. Nygård and L. Sörnmo, “Delineation of the QRS complex using the envelope of the ECG,” *Med Biol Eng Comput*, vol. 21, no. 5, pp. 538–547, 1983.
- [9] Q. Li, R. G. Mark, and G. D. Clifford, “Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter,” *Physiol Meas*, vol. 29, no. 1, pp. 15–32, 2008.
- [10] G. D. Clifford, J. Behar, Q. Li, and I. Rezek, “Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms,” *Physiol Meas*, vol. 33, no. 9, p. 1419, 2012.
- [11] J. Behar, J. Oster, Q. Li, and G. Clifford, “ECG signal quality during arrhythmia and its application to false alarm reduction,” *IEEE Trans Biomed Eng*, vol. 60, no. 6, pp. 1660–6, 2013.
- [12] D. G. Wyse, A. L. Waldo, J. P. DiMarco *et al.*, “A comparison of rate control and rhythm control in patients with atrial fibrillation,” *N Engl J Med*, vol. 347, no. 23, pp. 1825–1833, 2002.
- [13] D. M. Lloyd-Jones, T. J. Wang, E. P. Leip *et al.*, “Lifetime risk for development of atrial fibrillation the Framingham heart study,” *Circulation*, vol. 110, no. 9, pp. 1042–1046, 2004.
- [14] P. A. Wolf, R. D. Abbott, and W. B. Kannel, “Atrial fibrillation as an independent risk factor for stroke: the Framingham study,” *Stroke*, vol. 22, no. 8, pp. 983–988, 1991.
- [15] R. Colloca, A. E. Johnson, L. Mainardi, and G. D. Clifford, “A Support Vector Machine approach for reliable detection of atrial fibrillation events,” in *Comput Cardiol*, vol. 40. IEEE, 2013, pp. 1047–1050.
- [16] G. D. Clifford, R. Hall-Clifford, J. A. Blaya, and H. S. Fraser, “Medical Information Systems: A Foundation for Healthcare Technologies in Developing Countries,” *Biomed Eng Online*, vol. 7, no. 18, 2008.
- [17] T. Zhu, A. E. Johnson, J. Behar, and G. D. Clifford, “Bayesian voting of multiple annotators for improved QT interval estimation,” in *Comput Cardiol*, vol. 40. IEEE, 2013, pp. 659–662.
- [18] I. Silva, J. Behar, R. Sameni, T. Zhu, J. Oster, G. Clifford, and G. Moody, “Noninvasive fetal ECG: The PhysioNet/Computing in Cardiology Challenge 2013,” in *Comput Cardiol*, vol. 40. IEEE, 2013, pp. 149–152.
- [19] T. Zhu, J. Behar, T. Papastilianou, and G. D. Clifford, “Crowdlabel: A Crowd-sourcing Platform for Electrophysiology,” in *Comput Cardiol*, vol. 41. IEEE, 2014.
- [20] T. Zhu, A. E. Johnson, J. Behar, and G. D. Clifford, “Crowd-Sourced Annotation of ECG Signals Using Contextual Information,” *Ann Biomed Eng*, vol. 42, no. 4, pp. 871–884, 2014.
- [21] A. L. Goldberger, L. A. Amaral, L. Glass *et al.*, “Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [22] S. Sarkar, D. Ritscher, and R. Mehra, “A detector for a chronic implantable atrial tachyarrhythmia monitor,” *IEEE Trans Biomed Eng*, vol. 55, no. 3, pp. 1219–1224, 2008.
- [23] G. Moody, H. Koch, and U. Steinhoff, “The PhysioNet/Computers in Cardiology Challenge 2006: QT interval measurement,” in *Comput Cardiol*, vol. 33, 2006, pp. 313–316.
- [24] S. Boussejot R Kreiseler D, “Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das internet,” *Biomedizinische Technik*, vol. 23, pp. 215–220, 1995.