Physionet Challenge 2011: Improving the Quality of Electrocardiography Data Collected Using Real Time QRS-Complex and T-Wave Detection

Thomas Ho Chee Tat, Chen Xiang, Lim Eng Thiam

Institute for Infocomm Research, Singapore, Singapore

Abstract

The Physionet Challenge [1] focused on discerning between usable and unusable electrocardiography (ECG) data tele-medically from mobile embedded devices. Based on our publications [2,3,4], we have designed a method to determine the quality of ECG data and its usability using an adaptation of the Tompkins et al [5] real time QRS detection algorithm. With our modifications to the algorithm to cater to a very short length of data, out method is able to differentiate with accuracy the usability of ECG data in training set A as well as test set B.

1. Introduction

This year's the challenge is set at aiding information gathering tele-medically [1]. Collection of medically over a distance has been an issue every since tele-medicine was established. This was due to the lack of medically trained expertise from the gatherer to discern between usable and unusable medical information. This year's challenge aims at reducing, if not eliminating, all the fallacies that currently plagued usable medical information procurement tele-medically. With this challenge a standard or guidelines can be set up for recognition and following to better expedite the information flow as well as correctness. Hopefully, with such strong conviction, a higher level of health care can be imposed on the world to save lives as well as providing a higher quality of living.

2. Methodology

In this challenge, several approaches in achieving the results were explored. The final results were a combination of adapting some of the approaches as well as eliminating some that were initially hypothetically adopted but eliminated due to circumstances.

2.1. Visual inspection

Our group took part in the effort to grade the collated ECG data for the challenge on the PhysioNetWorks [1]. Our group have done this grading in order to relate to the type of data and the expectations on the grading; as well as to visually see in its raw form how much of and the degree of the artefacts are present.

Our group graded about 2000 sets of data on the PhysioNetworks. During the process we co-related the expectations of the expected grade with what our group felt about the grade; incidentally some of the ECG data that was graded were actually from training set A. Hence, several hypotheses on the expectation were drawn up at this juncture; this was before the reference classifications were released.

Subsequently after the reference classifications were released, another round of visual inspection was conducted. However, this time the inspection was only on the data that has changed from previous its classifications. Many of the existing pre-conditions that were initially deemed unacceptable were acceptable after the release of the reference classifications; prompting many people to question the accuracy of the classifications. Good faith with regards to the accuracy of the reference classifications was adopted on our part when Dr Moody allayed everybody's doubts that the aim of this challenge is to approximate the flaws of the collated ECG data to the best of their efforts and not to replicate them.

2.2. Support vector machine analysis

Besides visual inspection, our group had also sort help from our institutes' data mining team in using support vector machine (SVM) in differentiating the data into the two reference groups. The parameters used in the SVM included Radial Basis Function as the kernel, FFT and the number of detected QRS complexes as the features. The performance on Training Set A that it trained on records 501~1000 and cross-validated with records 1~500. The accuracy achieved was 81%.

Our initial hypothesis was that SVM should be able to separate the data with sufficient high accuracy. However, that was not the case as the SVM somewhat over-fit the data and caused many false negative detections. Nevertheless, we did manage to observe several of the features were useful in our implementation of the algorithm for the challenge.

2.3. The approach

Developing an algorithm that used a SVM is not feasible for deployment onto mobile embedded systems. However, developing an algorithm that mimics the classification engine of a SVM is. Hence, our approach is to adaptation of real time processing of bio-signals on mobile equipment [2, 3, 4]; based on our strengths, followed by the mimicking of the SVM that is developed to suit mobile equipments. The development of the latter followed the guidelines on developing algorithms for mobile embedded devices [6] very closely.

In the approach, the ECG data is read by its channels in reverse order starting from V6. The ECG sample of each channel is read and process sequentially. Figure 1 depicts the overall processes that each sample data goes through in order for the classification of each channel individually in the data set to be determined.

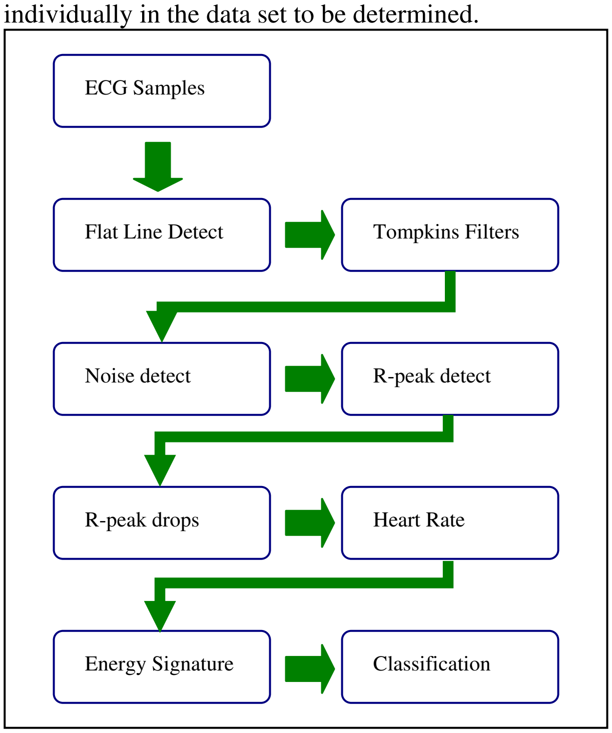

Figure 1. Processing of ECG samples to determine each channel's individual classification

The first detection is for ensuring that the signal quality is not a flat line. In this process, a running window of 0.5 seconds is used for the detection for the entire channel. If there exist on 1 unique value in its first order derivative then the channel is a straight line. Hence, the aim is to ensure as much data collected in the limited amount of time is usable. If not, do not accept the data and take another reading.

The second detection is by passing it through the Tompkins' filters [5] as well as the R-peak detection algorithm. However, since the length of data per channel is insufficient for the algorithm to successfully determine even the first r-peak, the time averaging and r-peak detection module of the algorithm has to be modified. More details on the modification are covered in the next sub-section.

Within Tompkins' r-peak detection algorithm, there is a detection threshold. This threshold now in this modified version of the algorithm beside r-peak detection now has another significant role to play. It was observed that for relative noiseless ECG data, the threshold is kept to a minimal. Whereas for noisy ECG data, the detection threshold is of a significant high values compared to the rest. Therefore, this is another feature that can be used to identify the amount of noise in the data. Likewise due to this expedited process where the detection threshold does not have sufficient duration to stabilize, false detection of r-peaks will occur in noisy ECG data. It was observed that such false detection will lead to r-r intervals being interpreted as greater than 300 beats per minute (BPM). This of course would indicate that this particular r-peak is not a true r-peak and should be eliminate in the subsequent calculations of r-r intervals. Hence, another derived feature to determine the usability of the data would be the drop rate occurrence of the r-peak.

The last 2 classifiers deal with the physical aspects of the data. Machine generated signals or noise attributing from poor skin contact are some examples of the aspects being examined. Supposed that the data is clean of artefacts, the task now is to differentiate between machine generated and human acquired data. The detected r-peaks can be bounded by thresholds that are within the human BPM limits; for our case between 30 to 210 BPM. Falling out the bounds would indicate that the data is not humanly possible and should be consider as a clause for incorrect classification.

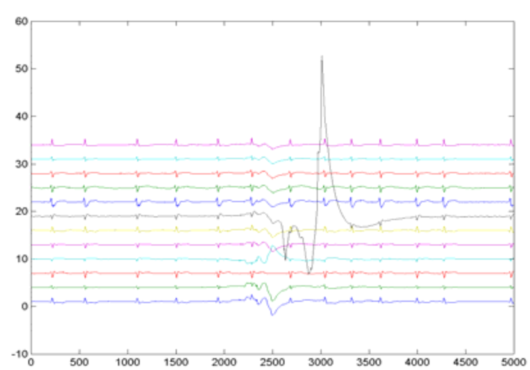

Figure 2. Sudden burst of high strength in amplitude

The final filter for the classification would to detect the artefacts attributing from the amplitude of the ECG data. Sudden bursts of high strength in the amplitude or continual amplitude band meant that were disturbances to the physical contact during data collection.

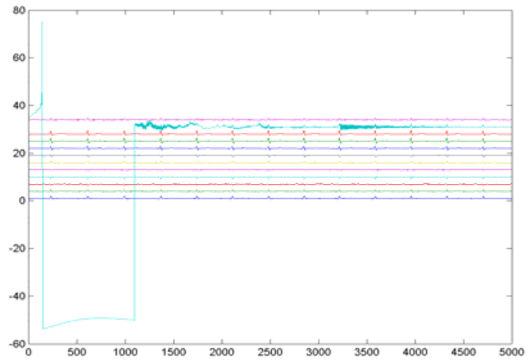

Figure 3. Continual amplitude bands

Our approach will ascertain to a certain degree of how much of such disturbances are acceptable and will classify the ECG data set accordingly.

2.4. Modified QRS detection algorithm

The Tompkins' QRS detection was modified. After the band-pass filters described in [5] are applied, a simplistic first order derivative filter of 10 microseconds is applied. The absolute values are taken after the derivation. Method for peak detection remains the same. However, the line per channel in the dataset is simply too short for the initialization of the algorithm to even make its first R-peak detection. Hence, in order for a r-peak to be detected in this context, the initialization period was shorten. After many observation of running the entire training set as well as the test set, the number of R-peak required for the initialization was reduced from 32 to 4. This was because 4 is short enough for the rest of the remaining data to have meaningful features being extracted as well as it is sufficient for the detection threshold to stabilize for majority of the less noisier data.

That being said noisy channels will have a lot of false detection due to the instability of the detection threshold. Therefore the shortcoming was minimized with the usage of noise threshold detection inherent to Tompkins' algorithm. Next was to clear up any R-peaks that did not make any sense. Since the heart rate can only exist between 30 and 220 BPM, it makes sense to threshold the R-R intervals that would not lead to the heart rate range. On hindsight, the range was shortened to 30 to 210 BPM. This was because during visual inspection, no heart rate was above 210 BPM but there were heart rate as low as 30 BPM.

2.5. Mimicking the SVM

As mentioned earlier, our approaches mimicked the way a SVM classifies data. Our approach is a hybridized version in a sense, being a rule-based classification engine, that we obtained the degrees of separation from a SVM, formulate our own separation pathways and finalized the decision with a weighted scheme per channel. In section 2.2, we have demonstrated the processes that were used to obtain features from the ECG data. In this section we will demonstrate the hybridized rule-based engine that mimicked a SVM.

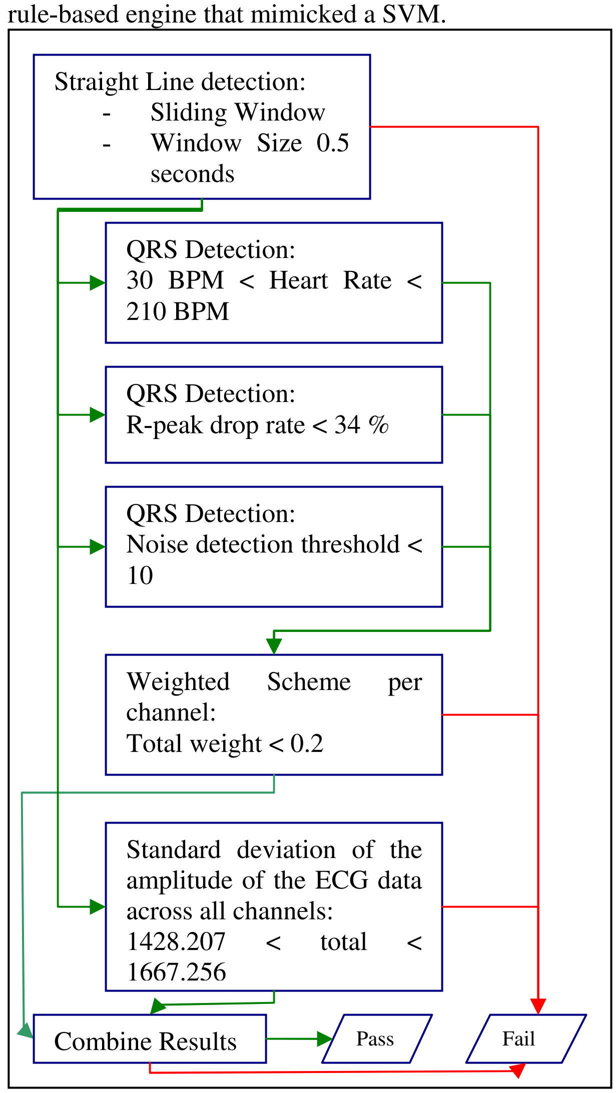

Figure 4. Hybridized rule based classifying engine

Hence, it can be seem that the decision flow to determine the acceptability do not follow a straight forward rule-by-rule based elimination like a traditional decision engine. The current configuration and values of the rule was due to countless arduous testing and re-iteration on the training data set A. The iterations chiefly optimized the variables that are seen in Figure 4. This decision engine is used to determine the final scores.

2.6. Identifying misplaced electrodes

As said previously, we did implement the algorithm to determine misplaces electrodes, these were mainly for RA and LA swapped electrodes. First, all the r-peaks in the channel V6 were identified. This was because overall on the whole the r-peak detection on channel V6 had the least error. Since all 12 channel data were collected concurrently, therefore all their r-peak positions should be at the same point. Hence, we can use V6 as reference points for r-peaks.

To identify RA-LA swapped, we compared the deflection of the r-peaks of ECG-I with reference to the r-peak positions on V6. Taking a window on the r-peak, the baseline was calculated and we calculate the difference of the sample value of the r-peak to the baseline. The gradient before and after the r-peak for 5 samples are also calculated. Hence, with these 3 values, the swapped leads can be identified

3. Results

There were a total of 5 submissions. The first submission was the most preliminary of all the submissions. However, the submission catered to a lot initial conditions such as misplaced electrodes, dropped electrodes, loosely attached electrodes, etc. The remaining 4 submissions were only after the reference classifications were released. In the reference classifications, many conditions that previously failed in the initial classifications were relaxed and thus now allowed to pass. However, it also managed to obfuscate the cutoff line defining the datasets' classifications.

Table 1. Scores for all 5 submissions.

Submission	Score (%)
1	80.8
2	84.2
3	92.0
4	89.8
5	91.2

3.1. Preliminary results

The initial score of 75.1% was achieved in the first submission. In this submission, many degrees of separation were considered. These included identifying misplaced electrodes, identifying noise and arfefacts in the ECG channels. However, the boundaries that affected the degrees of separation were not tightened and hence there was high number of false positive detections during this submission. It was only after the reference classifications were released that the direction to tighten boundaries could be derived. Tightening the boundaries before the release of the reference classifications might have lead to pursuing in a wrong direction and hence a waste of effort. After the reference classifications were released, our score increased to 80.8%.

3.2. Final results

In all our submissions the best of our results was 92%. In the last 2 submissions, the results were lower than the 3rd submission due to the optimization of the parameters. This was the case that we have already reached the plateau of our optimization. Due to the fact than some data were actually data that are similar from an engineer's perspective.

4. Conclusion

This was a very good challenge for us as the main are of our research dealt with ECG data collection and evaluation tele-medically. In the future, we would like to incorporate what we have done for this challenge into our main research development of cardiac applications for telemedicine. Even though we accomplished a score of 92% in the challenge, we will continue down this path until we are able to achieve better performance.

Acknowledgements

We acknowledge the expertise and help on SVM rendered by Dr Feng Mengling of the Data Mining Department and Mr Loy Liangyu of the Embedded systems Department from our institute.

References

[1] Moody GB. Improving the quality of ECGs collected using mobile phones: The 12th Annual Physionet/Computing in Cardiology Challenge. Computing in Cardiology Challenge 2011;38

[2] Chen X, Ho CT, Lim ET, Kyaw TZ. Cell Phone Based Online ECG Processing for Ambulatory and Continuous Detection. Computers in Cardiology 2007; 34:653 - 6.

[3] Lim E.T, Chen X, Tin Z.K, Meena S,"Smart Phone-based Automatic QT Interval Measurement" 34th Computer in Cardiology, Sept 30-October 3, 2007. Durham, North Carolina, USA

[4] Ho C.T.T, Chen X, "An Augmentative and Portable QTc-Observer(QTO-Q2) to Facilitate More Purposeful Outpatient Monitoring" 37th Computing in Cardiology, Sept 26-29, 2010, Belfast, Northern Ireland, United Kingdom

[5] Pan J, Tompkins W.J, "A real-time QRS detection algorithm." IEEE Trans. Biomed. Eng., vol, BME-32, no.3, pp. 230-236, 1985.

[6] Lynn P. A, "Online digital filter for biological signals: Some fast designs for a small computer," Med. Biol. Eng. Comput., vol. 15, pp. 534-540, 1977.

Address for correspondence.

Thomas Chee Tat Ho
1 Fusionopolis Way, #21-01 Connexis. Singapore 138632
E-mail: ctho@i2r.a-star.edu.sg