Reduction of False Alarms in Intensive Care Unit using Multi-feature Fusion Method

Chengyu Liu[1], Lina Zhao[1], Hong Tang[2]

[1] School of Control Science and Engineering, Shandong University, Jinan, China
[2] Department of Biomedical Engineering, Dalian University of Technology, Dalian, China

Abstract

In thist study, we proposed a multi-feature fusion method for accurately classifying the true or false alarms for five life-threatening arrhythmias: asystole, extreme bradycardia (EB), extreme tachycardia (ET), ventricular flutter/fibrillation (VF) and ventricular tachycardia (VT). The proposed method consisted of four steps: 1) signal pre-processing, 2) detection validation and feature calculation, 3) real-time determining and 4) retrospectively determining. Up to four signal channels, that is, two ECGs, one arterial blood pressure (ABP) and/or one photoplethysmogram (PPG) signals were analyzed to obtain the classification features. Multi-features from those signals were merged to reduce the maximum number of false alarms, while avoiding the suppression of true alarms. Two events existed: Event 1 for "real-time" and Event 2 for "retrospectively". The optimal results of true positive ratio (TPR) for the training set were: 100% for asystole, EB, ET and VF types and 94% for VT type. The corresponding results of true negative ratio (TNR) were 93%, 81%, 78%, 85% and 50% respectively, resulting in the corresponding scores of 96.50, 90.70, 88.89, 92.31 and 64.90, as well as with score 80.57 for Event 1 and 79.12 for Event 2. The results of the our open source entries for the Challenge obtained the optimal scores of 88.73 for asystole, 77.78 for EB, 89.92 for ET, 67.74 for VF and 61.04 for VT types, with the final scores 71.68 for Event 1 and 75.91 for Event 2.

1. Introduction

The detailed description for the background to the competition could be found in [1]. This study aimed to develop a multi-feature fusion method to reduce the number of false alarm and to avoid the suppression of true alarms by analysing the simultaneously recorded two channel ECGs, and the possible blood pressure waveform (ABP) and/or photoplethysmogram (PPG) signals.

2. Methods

2.1. Database

Detailed description for the database could be also found in [1]. As a summary, Table 1 details the alarm definitions and the distributions of the five alarm types for the training set used in this challenge, as well as their associated true and false rates.

2.2. Algorithm description

Figure 1 showed the algorithm flow chart. The proposed algorithm for identifying the alarm as true or false one consisted of four steps. Step 1: Signal pre-processing; Step 2: Detection validation and feature calculation; Step 3: Real-time determining and Step 4: Retrospectively determining. Each step consisted of several sub-steps.

In Step 1, the alarm flag (1 denotes true alarm and 0 false) and the time window for baseline feature analysis (T) for each alarm type were firstly initialized as follows:

$$Flag\_alarm = \leq \begin{cases} 0 & if \in \text{ VT type} \\ 1 & \text{other types} \end{cases} \tag{1}$$

T was set as 60 s, that is looking back 60 s in time from the onset of the alarm. Then the invalid values 'NaN' in the ECG signals were corrected using data interpolation. The 5-40 Hz band-pass filter was used for filtering the ECG signals and 5-35 Hz band-pass filter for ABP and PPG signals.

In Step 2, for ECG signals, a threshold-based detection method was used for R-peak location. Different alarm type recordings used different amplitude and time width thresholds. For ABP and PPG signals, pulse foot detection was performed using the wabp.m function (an open source beat detector available at www.physionet.org). Then for each channel, a distance-matrix-based method was used to verify the accuracy for the R-peak or pulse foot detection. This method was summarized as follows: M consecutive R-peak or pulse foot locations with the minimum standard deviation of beat-beat intervals were firstly selected. Then an $M \times M$ distance matrix D was initialized with all 0 elements and was updated using the following rule:

$$D_{i,j} = \begin{cases} 1 & if \ \dfrac{\text{Amp}(i)}{\text{Amp}(j)} > thr_1 \ \ or \ \ \dfrac{\text{Amp}(j)}{\text{Amp}(i)} > thr_1 \\ 0 & else \end{cases}, \tag{2}$$

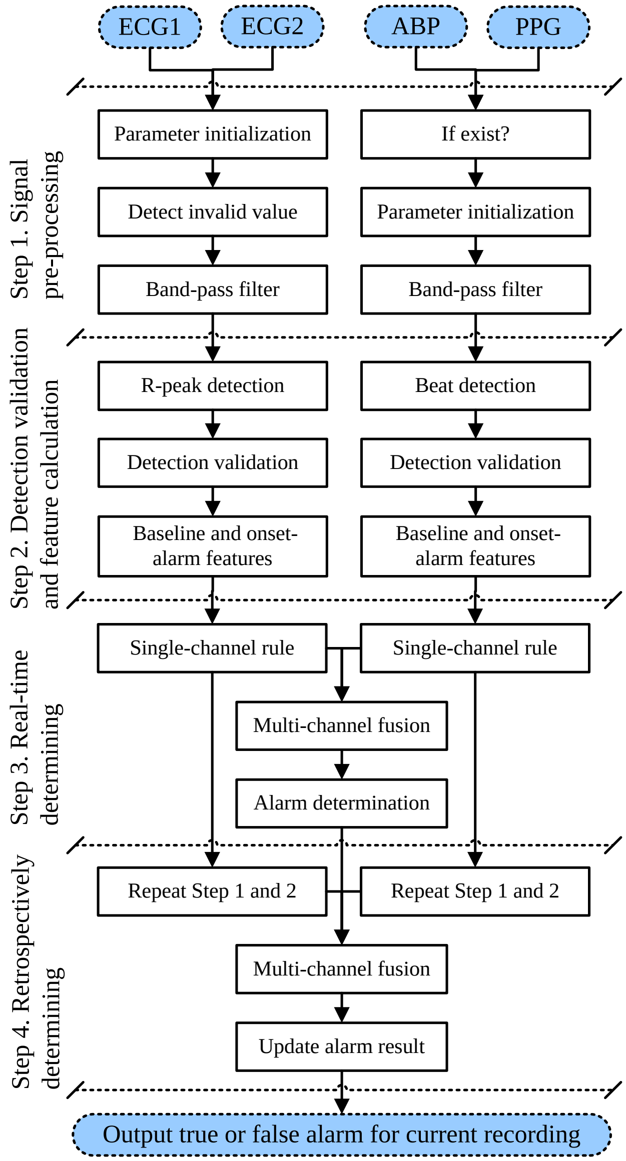

Figure 1. Algorithm flow chart

where Amp(i) means the amplitude of the ith beat and thr_1 is the amplitude threshold. Then the detection accuracy flag was set as:

$$Flag\_DecAcc = \begin{cases} 1 & if \quad \sum_{i=1}^{M}\sum_{j=1}^{M} D_{ij} < thr_2 \\ 0 & else \end{cases} \quad (3)$$

where thr_2 is the percentage threshold and $Flag\_DecAcc = 1$ verified that the current detection results for the M consecutive R-peak or pulse foot locations had high accuracy level and thus the baseline-features for the current channel would be obtained from the analysis of those M consecutive R-peak or pulse foot locations. These baseline features included:

- *HR_base*: baseline heat rate;
- *Template_base*: baseline signal template;
- *MaxAmp_base*: baseline signal maximum amplitude;
- *Range_base*: baseline signal amplitude range.

For the channels with $Flag\_DecAcc = 1$, the obtained *HR_base* values were compared to exclude the potential errors for *Flag_DecAcc* determination. The parameters central HR and threshold of HR ratio were set. If the ratio of *HR_base* between two channels exceeded the threshold, the channel with the *HR_base* far from the central HR value was excluded and the corresponding *Flag_DecAcc* was set as 0 again.

Then if $Flag\_DecAcc = 1$, the onset-alarm features were obtained by analyzing the *T_alarm* window length signal before the sounding of the alarm. Valid locations of R-peak or pulse foot were firstly selected by comparing the ratios between their amplitudes and the obtained *MaxAmp_base* or *Range_base* values. Then the onset-alarm features were obtained by analysing the valid locations:

- *Num_cur*: number of valid locations;
- *HR_cur*: current heat rate during alarm;
- *MaxRR_cur*: maximum of RR interval;
- *MaxHR_cur*: maximum of heart rate;
- *Cor*: Correlation degree with the *Template_base*;
- *Mor*: Morphology change rate compared with the *Template_base*;
- *W_QRS*: QRS complex width for each beat.

In Step 3, the flag for determining each channel as true/false alarm *Flag_Determine* was firstly initialized as 0 for VT and 1 for other four types (1 means judging the alarm as true based on current channel). Then *Flag_Determine* was updated as 0 for asystole, EB, ET and VF types if it meets the following rules (free parameters after optimizing):

For asystole type, $Num\_cur > 0.7 * \dfrac{HR\_base}{60 * T\_alarm}$ && $MaxRR\_cur < 3\text{s}$.

For EB type, $HR\_cur > 50 + 0.06 * HR\_base - 8$ && $HR\_base > 42$.

For ET type, $HR\_cur <= 100$ && $Num\_cur > 5$.

For VF type, $Cor_{mean} > 0.9$ && number of $W\_QRS > 0.1\text{s}$ less than 4 for ECGs, and $Cor_{mean} > 0.9$ && $MaxHR\_cur < 170$ for ABP and PPG.

In addition, *Flag_Determine* was updated as 1 for VT type if it meets the following rule: for ECG singals, $Cor < 0.97$ last for at least 3 s && the ECG amplitude change last for as least 3 s, or $W\_QRS > 0.12\text{s}$ more than 4 beats; for ABP and PPG, $Cor_{mean} > 0.85$ && $MaxHR\_cur > 120$.

After the single-channel determination, the alarm flag for the current recording was updated using the multi-channel information fusion.

For asystole, EB, ET and VF types, *Flag_alarm* was updated as 0 if any channel k meets the following rule:

Table 1. Definitions and distributions of the 5 alarm types for the training set used in this challenge. Average true alarm rate = 39.2%.

Alarm type	Definition	Training set (N=750)			
		False (%*)	True (%#)	Total (%&)	FAR
Asystole	No QRS for at least 4 seconds	100 (21.9%)	22 (7.5%)	122 (16.3%)	82.0%
EB	Heart rate lower than 40 bpm for 5 consecutive beats	43 (9.4%)	46 (15.7%)	89 (11.9%)	48.3%
ET	Heart rate higher than 140 bpm for 17 consecutive beats	9 (2.0%)	131 (44.6%)	140 (18.7%)	6.4%
VT	5 or more ventricular beats with heart rate higher than 100 bpm	252 (55.3%)	89 (30.3%)	341 (45.5%)	73.9%
VF	Fibrillatory, flutter, or oscillatory waveform for at least 4 seconds	52 (11.4%)	6 (2.0%)	58 (7.7%)	90.0%
All		456 (100%)	294 (100%)	750 (100%)	60.8%

Note: '*' means the percentage of all alarms that are false; '#' means the percentage of all alarms that are true; '&' means the percentage of all alarms; FAR: false alarm rate.

Table 2. Setting for the fixed and free parameters.

Parameter setting	Asystole	EB	ET	VF	VT
Fixed parameters					
Time window for baseline feature analysis (s)	60	60	60	60	60
Time window for post alarm segment analysis (s)	30	30	30	30	30
Band-pass filter for ECG signals (Hz)	5-40	5-40	5-40	5-40	5-40
Band-pass filter for ABP and PPG signals (Hz)	5-35	5-35	5-35	5-35	5-35
Dimension M of the distance matrix	10	10	10	10	10
Central heart rate (beats/min)	75	75	75	100	100
Threshold of HR ratio	1.5	1.5	1.5	1.5	1.5
Free parameters (after optimization)					
Time window for alarm segment analysis (s)	8	8	8	8	14
R-peak detection – amplitude threshold (ratio of Max_{amp})	0.75	0.75	0.5	0.5	0.3
R-peak detection – time width threshold (ratio of Mean_{RR})	0.7	0.7	0.4	0.5	0.25
The optimal parameter combination for each alarm type	Please see the detailed values in Step 3 and 4				

Table 3. Optimal results of the proposed multi-feature fusion method for false alarm suppression.

Alarm type	Training set (N=750)							Test set (N=500)		
	Number of TP	Number of FN	Number of FP	Number of TN	TPR (%)	TNR (%)	Score	TPR (%)	TNR (%)	Score
Asystole	22	0	7	93	100	93	96.50	89	93	88.73
EB	46	0	8	35	100	81	90.70	90	91	77.78
ET	131	0	2	7	100	78	88.89	98	60	89.92
VF	6	0	8	44	100	85	92.31	89	69	67.74
VT	84	5	126	126	94	50	64.90	79	69	61.04
Event 1	151	2	74	148	99	67	**80.57**	89	78	**71.68**
Event 2	138	3	77	157	98	67	**79.12**	93	78	**75.91**

$Flag\_DecAcc_k = 1$ && $Flag\_Determine_k = 0$, where k denotes the channel of ECG1, ECG2 or the possible channel ABP and PPG.

For VT type, *Flag_alarm* was updated as 1 if any ECG channel meets: $Flag\_DecAcc = 1$ && $Flag\_Determine_k = 1$ && Min(Cor)<0.8 for 5 s window && Max(Cor)>0.9 for 5 s window && Max(Mor)>0.65 for 5 s window, or any ABP/PPG channel meets: $Flag\_DecAcc = 1$ && $Flag\_Determine_k = 1$ && (either of the two ECG channels $Flag\_DecAcc = 1$ && $Flag\_Determine_k = 1$)

In Step 4, the 30 s signals from the retrospective recordings were performed the analysis of the Step 1 and 2. Herein, it did not need to obtain the baseline features but calculated the post-alarm features. Then the multi-channel information was used to update the final alarm determination results.

The setting for the fixed and free parameters was shown in Table 2.

Figure 2 shows the decision mechanism for the true/false alarm identification used in this study.

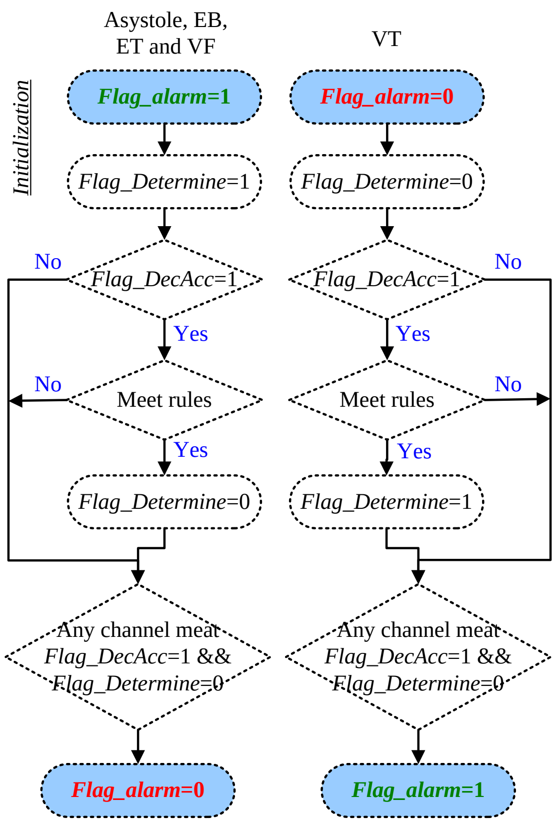

Figure 2. Decision mechanism

2.3. Algorithm scoring

Please refer to [1] for the scoring method.

3. Results

Table 3 shows the optimal results of our proposed method on both training (N=750) and testing (N=500) sets. The optimal results of true positive ratio (TPR) for the training set were: 100% for asystole, EB, ET and VF types and 94% for VT type. The corresponding results of true negative ratio (TNR) were 93%, 81%, 78%, 85% and 50% respectively, resulting in the corresponding scores of 96.50, 90.70, 88.89, 92.31 and 64.90, as well as with score 80.57 for Event 1 and 79.12 for Event 2. It can be seen that good alarm suppression performances achieved for asystole, EB, ET and VF types (all scores>85) but moderate performance achieved for VT type. Besides, this method caused 5 suppressions for true alarms for VT type while no suppressions for true alarms for other four types. This is generally because the noises on the ABP and PPG is coincident with the VT-like noises on the ECGs, which is often the case, inducing that it is hard to suppress such alarms. The open source entries of our method for the Challenge results obtained the optimal scores of 88.73 for asystole, 77.78 for EB, 89.92 for ET, 67.74 for VF and 61.04 for VT types, with the final scores 71.68 for Event 1 and 75.91 for Event 2.

Table 4 shows the referenced final scores from the non-competition entries. The results are all test scores on the hidden data. The voting algorithm took the top 13 best independent performers' final submissions (judged by the training scores) and voted the submissions together in an unweighted and trivial manner. Our method performed much better than example algorithms but had a great gap from the results from the Voting algorithm.

In addition, our open source algorithm was very efficient and it did not need any pre-learning state, indicated by the relative small running times:

- Average running time (training): 10.2% of quota;
- Maximum running time (training): 13.1% of quota;
- Average running time (test): 10.3% of quota;
- Maximum running time (test): 12.8% of quota.

Table 4. Results from the non-competition entries.

Method	Event 1			Event 2		
	TPR	TNR	Score	TPR	TNR	Score
EA 1	76%	44%	41.41	73%	46%	40.83
EA 2	86%	38%	45.07	84%	38%	44.37
EA 3	64%	76%	45.59	61%	77%	47.35
VA	94%	90%	84.26	94%	94%	87.04

Note: EA: Example Algorithm; VA: Voting Algorithm with N=13.

4. Conclusions

We have proposed a multi-feature fusion method for accurately classifying the true or false alarms for five life-threatening arrhythmias. This method achieved good performances for asystole, EB, ET and VF types and moderate performance for VT type. Further development by incorporating the artificial intelligence methods will facilitate to improve the performance of this method.

References

[1] Clifford G D, Silva I, Moody B, Li Q, Kella D, Shahin A, Kooistra T, Perry D, Mark R G. The PhysioNet/Computing in Cardiology Challenge 2015: Reducing False Arrhythmia Alarms in the ICU. Computing in Cardiology. 2015; 42:

Address for correspondence.

Chengyu Liu
School of Control Science and Engineering
Shandong University
17923 Jingshi Road, Jinan, China 250061
Tel: +86-159-53148364
E-mail: bestlcy@sdu.edu.cn.