# Multi-modal Integrated Approach towards Reducing False Arrhythmia Alarms During Continuous Patient Monitoring: the PhysioNet Challenge 2015

Sardar Ansari[1], Ashwin Belle[1], Kayvan Najarian[1,2]

[1] Department of Emergency Medicine, University of Michigan, Ann Arbor, MI, USA
[2] Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

## Abstract

*This work presents a solution for the Physionet Challenge 2015 regarding false alarm reduction in ICU. False alarms can result in alarm fatigue, i.e. reduced responsiveness of the ICU personnel to the true alarms due to an enormous number of false alarms. As a result, it is necessary to effectively suppress the false alarms while ensuring that the true alarms are not ignored. The challenge data contains five different types of alarms which are treated as independent problems in this paper. A separate subroutine is used for each alarm which is composed of two stages, peak detection and alarm verification. This paper uses a multi-modal peak detection algorithm that uses the information from all the available signals and combines the results from several peak detection algorithms to create a robust peak detection algorithm. The alarm verification stage is alarm dependent, composed of simple decision criteria or a complicated neural network model. The proposed approach achieves an overall score of 74.48 for the real-time event, where only the portions of the signals prior to the alarm are utilized, and 76.57 for the retrospective event, where 30 seconds of the signals after the alarm are used as well.*

## 1. Introduction

Intensive care units (ICU) are arguably the most data rich environment within any hospital system. These units are designed to provide care for patients with severe and life-threatening injuries or illnesses requiring constant monitoring, specialized equipments and highly trained doctors and nurses. To obtain a more comprehensive view of the patho-physiology of each patient in the ICU, they are typically attached to a multitude of monitoring devices such as electrocardiogram (ECG), pulse-oximeter (PPG), arterial blood pressure (ABP) catheter, central venous pressure catheter, ventilators, etc. Each of these devices are sophisticated equipments, many of which have built-in alarms to notify care givers when the measured parameters are observed to be out of what is considered as normal. However, despite the improvements in such physiological monitoring technologies, their alarm mechanisms and management have consistently shown points of failures [1, 2].

Issues with alarm management in ICU settings can be traced back to as early as 1983 when Kerr et. al. published about how the increasing number of alarms are creating confusion in the decision making process for the care givers [3]. Till date, the sheer number of alarms and their unreliability continue to increase and thereby negatively affect the quality of care provided to the ICU patients [4]. One study showed that in a single hospital, nearly 59000 alarms were recorded in a 12 day period of study, majority of which were either false or ignored [5]. In fact a recent study in this field demonstrated that nearly 72% to 99% of all alarms from such physiological monitors are false [6]. Occurrences of such large volumes of monitor alarms per patient as well as the lack of veracity of these alarms have created a copiously documented phenomenon knowns as 'alarm fatigue' [6, 7] where the care-givers attentiveness to alarms starts diminishing leading to the true alarms being ignored.

Alarms in ICU based monitors are fundamentally designed to be highly sensitive in nature. This is in order to avoid missing any occurrence of critical events. However, such sensitivity is almost always achieved while directly compromising the specificity of these alarms. The Pphysionet challenge 2015 explores various ways of reducing the number of false alarms while ensuring that the true alarms are not suppressed. It considers five different alarms: Extreme Bradycardia (BC) where the heart rate (HR) is lower than 40bpm for five consecutive beats, Extreme Tachycardia (TC) where HR is higher than 140bpm for 17 consecutive beats, Ventricular Tachycardia (VC) where five or more ventricular beats are present and HR is higher than 100bpm, Ventricular Flutter/Fibrillation (VF) where fibrillatory, flutter, or oscillatory waveforms are present for a

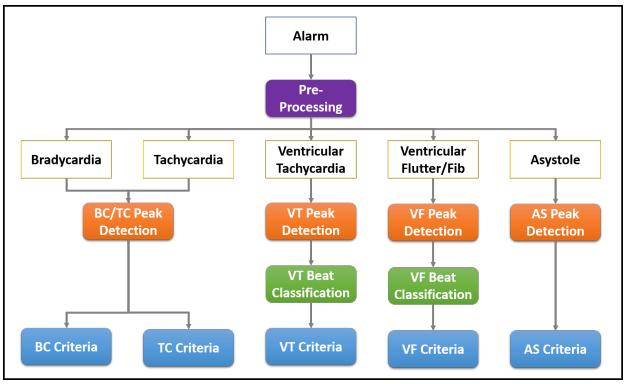



Figure 1: The structure of the proposed method.

duration of at least 4 seconds, and Asystole where there is no QRS for at least 4 seconds.

The proposed method for alarm verification is introduced in the next section.

## 2. Methods

The verification of each type of alarm can be treated as an independent problem. Hence, five different subroutines are proposed for verifying the different types of alarms, all of which consist of two fundamental steps, peak detection and alarm verification, shown in Figure 1. After introducing the pre-processing step, the paper discusses the general peak detection algorithm that is used in this work. The peak detection algorithm is then customized and fine-tuned for each type of alarm. Some of the subroutines compute peak quality indices as well. These indices are discussed in Section 2.3. The Asystole subroutine uses a different peak detection algorithm which is discussed in Section 2.6. The proposed method starts with a pro-processing step which is outlined in the next section.

### 2.1. Preprocessing

The preprocessing step starts by retrieving the input signals. The signals from two ECG leads are used in this work as well as the ABP and PPG signals. The algorithm first searches for ECG leads II and V and uses any available leads if these two are not available. All the unavailable signals are set to zeros. Only 16 seconds prior to the alarm are used in this work and the rest of the signals are discarded. All the signals are re-sampled to 125Hz. The retrieved signals are then filtered using zero-phased bandpass Butterworth filters between 0.5Hz and 40Hz for the ECG signals and between 0.5 and 10Hz for the ABP and PPG signals.

The data sets for the BC and TC alarms are usually less noisy compared to the VT and VF alarms. Hence, the BC and TC signals are normalized by removing the median of the signal and dividing it by signal's 99th percentile minus the 1st percentile. For the VT and VF signals, the trend is found and removed using a median filter with a span of 1 second. Then, an initial set of peaks are found using Matlab's *findpeaks* function with *MinPeakDistance* equal to 1 second applied to the absolute value of the signals. The signals are then normalized by dividing them by the median of the peak heights.

### 2.2. Peak Detection

The peak detection is performed by polling several different peak detection algorithms for the ECG, ABP and PPG signals. Some of these algorithms are available in the literature and others are devised by the authors. First, the peak detection algorithms for the ECG signals that are used in this work are introduced. The first one is the *sqrs* routine in the Physionet toolbox which implements the algorithm in [8]. The second algorithm is a peak detection code by Sadeghi et al. that is available on Matlab Central [9]. Three more algorithms are devised by the authors. The first uses the absolute value of the derivate of the signal to find the peaks. It first finds an initial set of peaks using the *findpeaks* function in Matlab with *MinPeakDistance* equal to 2 seconds. Then, it finds a threshold equal to the median of the depth of the initial peaks (the average difference between the height of the peak and the neighboring troughs) divided by 3. The final set of peaks is found using *findpeaks* function with *MinPeakDistance* equal to 250ms and *MinPeakProminence* equal to the threshold. This algorithm is referred to as *PF1*. The second algorithm, *PF2*, is similar to *PF1* but is applied to the ECG signal itself instead of its derivative. The last algorithm, referred to as *PF3*, uses the absolute value of the Stockwell Transform of the signal and creates a new signal by adding the values along the frequency dimension. The peaks of this new signal are found using a method similar to the previous algorithms, except that the threshold is computed by dividing the average depths by 5 instead of 3.

Some of the algorithms select points that are not critical, i.e., the derivative is non-zero at the selected points. As a result, a neighborhood around each selected point is searched for the critical point with the highest absolute value to replace the selected point. Moreover, the shallow peaks whose depth (average drop from the peak to the neighboring troughs) is less that one fifth of the 60th percentile of all the peak heights are removed.

Two algorithms have been employed to find the peaks in the ABP signal. The first one uses the *wabp* routine in the Physionet toolbox. The second one is similar to the *PF1* algorithm. The same two algorithms are used to find the peaks in the PPG signal as well. For the ABP signal, the original signal before normalization is used since the *wabp* routine takes advantage of the features of the ABP signal

such as systolic and diastolic blood pressures. Moreover, the PPG signal is scaled between 80 and 120 to resemble an ABP signal before being fed into the *wabp* routine. Similar to the peak detection algorithms for the ECG signals, the peak detection algorithms for the ABP and PPG signals search a neighborhood around the peaks to find the critical points with the highest amplitude and replace the selected points with them. However, the shallow peaks are not removed here since abnormal beats in VT and VF cases can lead to ABP and PPG beats that have a lower amplitude compared to the normal beats.

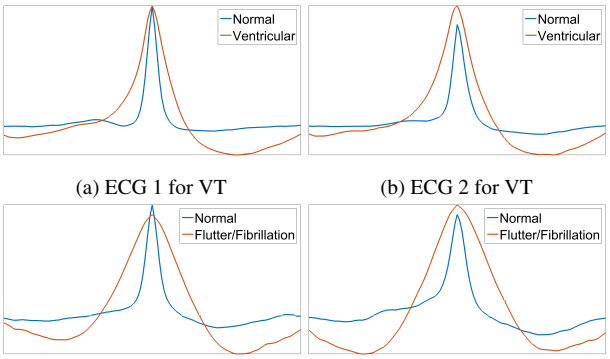
The polling is done by creating an indicator signal using the peaks that are detected by each algorithm. For each set of detected peaks, the indicator signal is non-zero in a 100ms neighborhood around each peak and is zero everywhere else. The height of the non-zero sections are set to 1 for ECG signals and 1.5 for the ABP and PPG signals in the BC/TC Peak Detection subroutine. On the other hand, the heights of the non-zero sections are determined by the quality of the peaks, which are discussed in the next section, for the VT and VF Peak Detection subroutines.

All the indicator signals for the ECG signals are added together to create a cumulative ECG indicator signal. The cumulative ABP and PPG indicator signals are created similarly. The cross-correlations between the cumulative indicator signals are used to align the ABP and PPG signals with the ECG signals. The aligned indicator signals are then added to create a global indicator signal that is composed of all the peaks that are detected by a plurality of peak detection algorithms from all the available signals. The global indicator signal is then smoothened and thresholded to detect the heart beats. Thresholds of 1.75, 1.5 and 1.2 are used for the global indicator signals for the BC/TC, VT and VF Peak Detection routines. The BC/TC Peak Detection subroutine only uses the *sqrs*, *Sadeghi* and *PF1* for ECG peak detection while the VT Peak Detection subroutine uses *sqrs*, *PF1* and *PF3*, and the VF Peak Detection subroutine uses *PF1*, *PF2* and *PF3*.

## 2.3. Assessing Peak Quality

The subroutines for the VT and VF alarms use quality indices for the detected peaks to create the indicator signals. The quality of the ECG peaks are computed by creating a window around each peak with the peak in the center. The size of the window is determined by $f_s/f_{max}$ where $f_s$ is the sampling rate and $f_{max}$ is the maximum HR specified for each type of alarm. For each peak, the absolute value of the Pearson correlation between its window and the two neighboring peak windows are computed and the peak quality is set to be the maximum correlation.

The quality of the ABP peaks is assessed using the code that was provided by the Challenge. The algorithm first extracts 12 features from the ABP signal including the pressure values and the duration of systole and diastole, pulse pressure, etc. Then, it checks if each feature is within the normal range for that feature. Finally, it computes a quality index based on the number of features that are outside their normal range.




(a) ECG 1 for VT (b) ECG 2 for VT

(c) ECG 1 for VF (d) ECG 2 for VF

Figure 2: The typical patterns for the normal and abnormal beats.

The quality of the PPG peaks are determined using a code that was provided in the sample entry by the Challenge. It first computes a beat template using all the beats in the signal. The correlation coefficient between a window around each peak and the template is considered as the peak quality.

## 2.4. Bradycardia and Tachycardia

After finding the peaks, the decision on whether the alarm is true or not is made based on the timing of the peaks. A relaxed criteria is used compared to the definition of the alarms in order to reduce the false negatives. For both alarms, the minimum and maximum HR are assumed to be 30bpm and 240bpm, respectively. For the BC alarm, if the HR is below 50bpm for at least three beats, the alarm is declared to be true, and vice versa. A HR of higher than 130bpm for more than 13 beats is considered as true TC alarm; otherwise, the alarm is designated as false.

## 2.5. Ventricular Tachycardia and Flutter/Fibrillation

The verification of the VT and VF alarms consist of two steps, phase wrapping and machine learning. For phase wrapping, the ECG signals are normalized in a similar manner as the peak detection step for the VT and VF alarms. Next, the mid points between the peaks are found and each period is defined to start from the mid point before that peak and end at the mid point after that peak. If any of the two sides of the period are longer than 1 second, that side is trimmed down to 1 second. Then, the shorter

side of the period is zero-padded to position the peak in the middle of the period. The periods for which the amplitude of the peak is smaller than the median of the period are inverted to create all-positive peaks. Finally, the periods are interpolated into 125 samples. The results are equally sized, normalized, all positive periods for each peak with the peak in the center. Figure 2 shows the typical shapes for the normal and abnormal classes. The periods are then used as instances for machine learning to classify the beats.

We have used a 2-layer neural network (NN) with the raw periods as input and the type of peak as output. For the VT alarms, two classes are defined, the normal class and the ventricular class. Similarly, the VF alarms consist of normal and fibrillatory/flutter classes. The data has been manually annotated by the authors and the annotations have been provided as class labels. The 250 samples for the two ECG signals in each row are down-sampled by a factor of 2. Hence, the input layer for the NN contains 125 nodes and a bias node. The size of the hidden layer and the type of the neurons are chosen by nested cross-validation (CV), with 5 validation folds and 5 testing folds. This led to 150 tangent sigmoid neurons for the VT model and 50 tangent sigmoid neurons for the VF model in the hidden layer, plus the bias term. The VT and VF samples are unbalanced with more instances in the normal class. Hence, the abnormal beats are replicated five times to reduce the number of incorrectly classified abnormal beats.

After classification, the detected peaks are used to verify the alarms. The decision criteria are relaxed to avoid false negatives. A VT alarm is declared as true if there are at least 3 consecutive ventricular beats and the HR is above 80bpm. For the VF alarms, the alarm is verified if there is fibrillatory/flutter pattern for at least 2 seconds.

## 2.6. Asystole

In the presented method, in order to detect Asystole, the PPG and ABP signals are utilized. Each window of the signals is processed using the *wabp* routine to find the peaks. If the heart rate is found to be either greater than 140 or lesser than 20 then it is assumed that the window is noisy and thus not used for further computation. If a distance of 4 seconds or greater is detected between any two consecutive peaks in either of the signals, then the alarm is flagged as true, and vice versa.

# 3. Results

The subroutines that are proposed in the previous sections are applied to the Physionet training dataset (publicly unavailable) and the results are shown in Table 1. The best results are obtained for the AS, TC and VF alarms. However, the largest section of the dataset belongs to the VT alarms with a TNR of 78%. The relatively high number of FN, which weight five times higher, lowers the overall score. Two overall scores are measured, one for the real-time event where only the signal prior to the alarm is used for verification. The retrospective score is measured using the signals prior to the alarm as well as 30 seconds after.

Table 1: The alarm verification results.

| Alarm Type | TPR | TNR | Score |
|---|---|---|---|
| Asystole | 94 | 82 | 81.76 |
| Bradycardia | 77 | 86 | 60.15 |
| Tachycardia | 98 | 60 | 89.92 |
| Ventricular Flutter Fib | 100 | 90 | 91.38 |
| Ventricular Tachycardia | 78 | 85 | 70.69 |
| Real-time | 89 | 84 | 74.48 |
| Retrospective | 89 | 87 | 76.57 |

# References


[1] Görges M, Markewitz BA, Westenskow DR. Improving alarm performance in the medical intensive care unit using delays and clinical context. Anesthesia Analgesia 2009; 108(5):1546–1552.

[2] Siebig S, Kuhls S, Imhoff M, Gather U, Schölmerich J, Wrede CE. Intensive care unit alarmshow many do we need? Critical care medicine 2010;38(2):451–456.

[3] Kerr J, Hayes B. An alarming situation in the intensive therapy unit. Intensive care medicine 1983;9(3):103–104.

[4] Deb S, Claudio D. Alarm fatigue and its influence on staff performance. IIE Transactions on Healthcare Systems Engineering 2015;5(3):183–196.

[5] for the Advancement of Medical Instrumentation A, et al. A siren call for action: priority issues from the medical device alarms summit. Arlington VA Association for the Advancement of Medical Instrumentation 2011;.

[6] Sendelbach S, Funk M. Alarm fatigue: a patient safety concern. AACN advanced critical care 2013;24(4):378–386.

[7] Cvach M. Monitor alarm fatigue: An integrative review. Biomedical Instrumentation Technology 2012;46(4):268–277.

[8] Engelse W, Zeelenberg C. A single scan algorithm for qrs-detection and feature extraction. Computers in cardiology 1979;6(1979):37–42.

[9] Sedghamiz H. An online algorithm for r,s and t wave detection. http://www.mathworks.com/matlabcentral/fileexchange/45404-ecg-q-r-s-wave-online-detector, 2013.



Address for correspondence:

Sardar Ansari
2800 Plymouth Road, Bldg. 10-A109
Ann Arbor, MI 48109
sardara@med.umich.edu