An ECG T-wave Anomalies Detection Using a Lightweight Classification Model for Wireless Body Sensors
Medina Hadjem, Farid Naït-Abdesselam

To cite this version:

HAL Id: hal-01191622
https://hal.archives-ouvertes.fr/hal-01191622
Submitted on 2 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
An ECG T-wave Anomalies Detection Using a Lightweight Classification Model for Wireless Body Sensors

Medina Hadjem and Farid Naït-Abdesselam
Paris Descartes University, France
{firstname.lastname}@parisdescartes.fr

Abstract—Various wearable devices are foreseen to be the key components in the future for vital signs monitoring as they offer a non-invasive, remote and real-time medical monitoring means. Among those, Wireless Body Sensors (WBS) for cardiac monitoring are of prominent help to early detect CardioVascular Diseases (CVD) by analyzing 24/24 and 7/7 collected cardiac data. Today, most of these WBS systems for CVD detection, include only limited automatic anomalies detection, particularly regarding ECG anomalies. Severe CVD, such as Myocardial Infarction or Ischemia, needs to achieve an advanced analysis of ECG waves known as P, Q, R, S and T. In particular, the T-wave and its specific changes. In this paper, we focus on T-wave anomalies detection in a context of WBS. Our study suggests an accurate and lightweight T-wave changes detection model which suits well an ECG monitoring system based on WBS architecture. We performed a comparative study of 7 well-known supervised learning classification models, on real ECG data sets from 7 different leads. We compared the results from both perspectives of classification and processing times. Our results show that the C4.5 Decision Tree technique performs better results with 92.54% Accuracy, 96.06% Sensibility, 55.41% Specificity and 7.41% Error Rate.

Index terms—ECG, T-wave, WBS, Supervised Learning.

I. INTRODUCTION

Recent technological advances in wireless technologies, sensors conception and miniaturization are revolutionizing the way how healthcare services are delivered. In fact, existing sensors are able to collect various physiological parameters in real-time and with wireless transmission capabilities, such as Heart Rate, Respiration Rate, Temperature, Oxygen Saturation, Blood Pressure, Electromyogram, Electrocardiogram, etc. These Wireless Body Sensors (WBS) can be attached to the human body and transmit collected data to a gateway device (Smartphone, Tablet, etc.), which processes and transmits gathered information in real-time to remote healthcare professionals so they can make appropriate medical decisions.

Various applications of WBS can be considered. One of the most promising is remote monitoring and detection of Cardiovascular Diseases (CVD). According to the World Health Organization [1], these diseases are the leading causes of death worldwide. In 2008, 17.3 million people died from CVD and almost 23.3 million people will die annually by 2030. This area of research is very challenging and strategic for medical sector and is highly supported by governments.

Today, diagnosis of heart defects is done after physical symptoms that lead patients to consult doctors or during a routine medical examination, where the Electrocardiogram (ECG) is recorded. The ECG represents the contraction and relaxation of cardiac muscle, resulting from depolarization and re-polarization of myocardial cells. These electrical changes are recorded via electrodes placed on the limbs and chest wall and are transcribed on a graph paper.

The most widely used ECG recording technique is the 12-lead ECG. This technique uses 10 recording sites from which are derived 12 signals called leads. Among which, 6 Precordial Leads (V1, V2, V3, V4, V5, V6), 3 Limb leads (I, II, III) and 3 augmented Limb leads (aVR, aVL, aVF). Each lead views the heart from a different angle, and is composed of 5 waves symbolized P, Q, R, S, and T. These waves are separated by intervals and segments as shown on Figure 1 that illustrates a typical one-cycle ECG waveform.

T-wave is a significant ECG wave, it corresponds to the ventricular re-polarization and its changes are one of the most common noted on an ECG. These changes may be normal for some healthy individuals because of age, body configuration and medications or may indicate several heart anomalies based on their abnormal shape or amplitude.

Four cases of T-wave changes are generally identified [2]:

![Fig. 1. An ECG heart beat with waves and segments](image-url)
– *Elevated*: An increase in T-wave amplitude may indicate an early stage of myocardial infarction (MI), occurring in the first few minutes after occlusion of a coronary artery. Intervention at this early stage can prevent infarction and improve the outcome of treatment. In practice, most patients consult after chest pain which is generally the stage of ST segment elevation, this is why ST elevation is often considered as the most sensitive marker of MI. Other anomalies can be related to a prominent T-wave such as hyperkalemia (Excess of potassium in the blood), Bundle Branch Block, Left Ventricular Hypertrophy.

– *Inverted*: T-wave inversion may indicate Ischemia or Acute Coronary Syndrome (ACS) resulting from complete or partial thrombotic occlusion of a coronary artery. In general, Deep T-wave inversion in some leads usually relates to a serious underlying cardiac lesion. T-wave inversion may also be indicative of hyperventilation or pulmonary embolism.

– *Flattened or Biphasic*: The two main causes of biphasic T-waves are Myocardial Ischemia and Hypokalaemia (Potassium deficiency in the blood). Flattened T-waves are a non-specific finding, they may indicate Ischemia in some cases. There is no clearly established normality parameters of the T-wave amplitude and duration. Usually, specialists are interested in changes of amplitude over time, by comparing ECG with previous recordings. The observation of these changes, only at a given time, can’t lead to an automatic diagnosis and other medical examinations are often necessary. Therefore, the relevance of T-wave analysis resides in the possibility of time tracking of its parameters, as it is the case using WBS. Therefore, this study should be considered in a context of continuous ECG monitoring system for a purpose of medical alerting and not diagnosis.

The aim of this study is to find an efficient and lightweight model for T-wave changes detection. This model should distinguish between normal, elevated and inverted T-waves (Flattened or biphasic T-wave will not be studied in this paper), and offers the best classification performance with a processing time suitable to real-time constraint of WBS systems.

The rest of this paper is organized as follows: Section II surveys related work. Section III presents the proposed approach and Section IV discusses experimental results. Finally, Section V concludes this paper.

II. RELATED WORKS

With the development of e-Health in recent years, wireless body sensors monitoring systems have been the subject of various research in both public and industry sector. Regarding ECG monitoring, these systems are able to capture patient’s ECG wirelessly and send it in real-time to remote applications, so that it can be analyzed by specialists. Many prototypes and frameworks were proposed, for example, MyHeart [3] and MEDISN [4]. More recently, these systems are evolving to incorporate cardiac anomalies detection like arrhythmia, as in [5] and [6]. Some systems are already on the market such as CardioNet [7] and BodyGuardian [8]. Other contributions propose to detect more complex anomalies as Myocardial Infarction like in [9] and [10].

The detection methods proposed in such ECG monitoring systems are generally based on temporal characteristics of ECG waves (P, Q, R, S, T). Most of them analyze the RR Interval as an indicator of heart rate, or ST segment elevation or depression, which is generally a sign of Myocardial Infarction or Ischemia. Other studies analyze the P-wave and PR Interval anomalies to detect Arrhythmia. Among these works, little focus on T-wave analysis changes, which are yet very used by doctors in the diagnosis process of various cardiac anomalies as mentioned in section I.

The most significant works on T-wave analysis, focus on the T-wave alternans (TWA) detection, which is a periodic beat to beat variation in T-wave amplitude or shape. This is related to cardiac instability and may indicate a risk of sudden cardiac death. Several contributions on TWA detection were proposed, especially under the 2008 Physionet/Computing In Cardiology Challenge [11]. The proposed solutions are based on different techniques of signal processing, data mining or machine learning. Good scores were obtained with Principal Component Analysis (PCA) [12], Spectral Method [13] and Fast Fourier Transform (FFT) [14]. Many other TWA detection techniques are compared and discussed in [15]. T-wave analysis was also used for Myocardial Ischemia detection, as an additional parameter to ST segment, like in [16]. Other contributions propose a general detection of T-wave changes episodes like in [17]. According to our research, there is very little ECG monitoring systems including T-wave analysis. A proposal was made in [18] to develop a real-time ECG anomalies diagnosis, including T-wave analysis on a sensor node platform.

In this study, we experiment 7 supervised learning methods (Naïve Bayes, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, AdaBoost, Decision Trees and Random Forest). The goal is to suggest a model that is both accurate to detect T-wave changes (Elevation and Inversion) and lightweight to suit a context of WBS implementation.

III. PROPOSED APPROACH

The proposed approach consists of three main steps. ECG Pre-processing and features extraction, test of 7 classification methods to select the most suitable, and finally propose an integration algorithm of the selected model in a WBS ECG monitoring system.

A. ECG Pre-processing and Features Extraction

Pre-processing is a required step for any ECG analysis in order to remove various noises. Theses noises can seriously affect the quality of collected ECG and distort the analysis results. They interfere with the ECG components in the frequency domain, specifically in the range [0.01Hz-150Hz]. The most significant are:

- Power line interference around 50Hz or 60Hz.
- Baseline wandering with a frequency bellow 0.5Hz.
- Electrodes motion artifacts ranging from 1Hz to 10Hz.
Electromyographic (EMG) noise from 25Hz to 100Hz.

Power line noise comes from electronic circuits of ECG recorder. It is a narrow-band signal centered at a frequency of 50Hz or 60Hz with a bandwidth of less than 1Hz, it has a form of a sinusoid with a significant amplitude compared to the magnitude of the digitized signal. Generally, this noise is filtered by ECG signal acquisition hardware.

Baseline wandering is usually due to patient movements and respiration. It ranges between 0.15Hz and 0.5Hz and causes changes in ECG isoelectric line. It can be removed, without loss in original signal quality, by a highpass digital filter or by a standard Wavelet Transform (WT). In this study, we use a 0.5Hz FIR highpass filter (Finite Impulse Response).

Electrodes motion artifacts and EMG noise are more difficult to remove because they may be a complex stochastic processes within a wideband, and traditional digital filters can’t remove them. Instead, Discrete Wavelet Transform technique (DWT) is widely used. In this study, we use Undecimated Wavelet Transform (UWT), a variant of DWT with better balance between smoothness and accuracy and ensures no loss of signal sharpest features. UWT for ECG denoising is described in [19].

After pre-processing, an ECG segmentation is performed to estimate time positions of the main ECG waves (QRS complex, P-wave and T-wave). These time positions will serve to estimate time positions of the main ECG waves (QRS complex, P-wave and T-wave). These time positions will be used in the classification stage. To avoid the effect of any changes introduced by the pre-processing, the preprocessed ECG signal is only used to detect QRS complexes positions, other ECG waves positions are estimated using the original signal. The method used for segmentation is the ECG Feature Extractor technique developed by Labview [20], which is inspired by the well known Pan & Tompkins technique for QRS peaks detection [21]. Table I synthesizes the temporal parameters extracted for each ECG beat.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ponset</td>
<td>Start time of the ECG P-wave</td>
</tr>
<tr>
<td>Poffset</td>
<td>End time of the ECG P-wave</td>
</tr>
<tr>
<td>QRSonset</td>
<td>Start time of the ECG QRS complex</td>
</tr>
<tr>
<td>QRSoftset</td>
<td>Start time of the ECG QRS complex</td>
</tr>
<tr>
<td>Tonset</td>
<td>End time of the ECG T-wave</td>
</tr>
<tr>
<td>Toffset</td>
<td>End time of the ECG T-wave</td>
</tr>
</tbody>
</table>

Based on previous temporal parameters, we calculate the ECG features used in classification stage. In order to both reduce the data size to process and increase the classifier accuracy, only the most relevant features related to T-wave abnormalities are selected. After collecting various medical information about T-wave abnormalities, mainly from [2], we selected the ECG features described in table II.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamp</td>
<td>T-wave amplitude in mV calculated by peak research between Tonset and Toffset</td>
</tr>
<tr>
<td>Tdur</td>
<td>T-wave duration in seconds between Tonset and Toffset</td>
</tr>
<tr>
<td>STamp</td>
<td>ST segment amplitude in mV approximated by averaging the extrema between QRSoftset and Tonset</td>
</tr>
<tr>
<td>QTdur</td>
<td>QT interval duration in seconds between QRSonset and Toffset</td>
</tr>
</tbody>
</table>

B. T-wave Classifier selection

As there is no established parameters of normality to easily classify T-wave changes, supervised learning approaches seem to be suitable. These machine learning techniques allow to build a predictive model on observed data in order to make prediction on new unknown data. More formally, a supervised learning classification problem, consists in a set X of objects called the input space and a finite set C of classes called the output space. We consider a sample $E = \{(x_1, c_1), ..., (x_n, c_n)\}$ with an unknown probability distribution $D$ over $X \times C$. The goal of the classification is to find the function $h \in C^X$ that defines the relation between objects and classes, based only on the sample $E$ and drawn independently at random according to $D$. As we are looking for a model suitable for an implementation in a WBS context, we focus, in addition to accuracy, on processing time of classification. To find the classifier that best suits this objective, we compare 7 ones : 3 linear classifiers (NBC, SVM and Logistic Regression) generally characterized by a lower complexity, 2 non-linear classifiers (k-NN and AdaBoost) and 2 Decision Trees classifiers (C4.5 and RF) that generally have good performances. An overview of these classifiers can be found in [22]. The results are detailed in section IV.

C. Integration of the model in a WBS ECG monitoring system

We consider a WBS ECG monitoring system composed of wireless ECG sensors transmitting data to a smartphone in real-time. The selected classification model can be implemented as a mobile application running on the smartphone. This application receives ECG wirelessly from sensors, pre-processes and extracts ECG features related to T-wave ($Tamp, Tdur, STamp, QTdur$), and finally classifies these features using the selected model. The model is first built based on annotated real ECG data, and updated in real-time with the classification results of the received ECG. The application raises an alarm to medical assistance only if a minimum number of ECG beats are classified as abnormal during a windows time $\omega$ (to define in the application). Algorithm 1 summarizes all steps achieved by the mobile application of the target WBS system.

IV. Experimental Results

In order to experiment our proposed approach, we used the European ST-T ECG Database (EDB) from Physionet [23]. It’s a collection of real patients ECG records, specifically intended for evaluation of ST and T-wave changes algorithms. This database consists of 90 annotated ECG recordings from 79 men and women of different ages. It contains a representative selection of ECG T-wave abnormalities, including 367 ST segment and 401 T-wave change episodes. Each episode ranges from 30 seconds to several minutes. Each record is 2 hours duration (~7200 ECG beats) and contains 2 leads sampled at

---

TABLE I

Temporal parameters extracted from each ECG beat

TABLE II

Selected ECG features for T-wave classification
obtained by extracting QRS complex peaks, and then deduct P previous ECGs. As described in Section III, these values were ST amplitude we calculate, for each beat, the relevant T-wave features characterizing each beat of each lead. From these parameters, the result of this stage is a series of ECG temporal parameters.

For our experiments, we use the Filtering VI (0.5Hz FIR high-pass filter), the Wavelet Denoise VI (UWT with Daubechies6 wavelet) and ECG Feature Extractor VI. The platform that offers many modules named Virtual Instruments (VI). For each record, we associated to all beats in the episode. An anomaly code is associated to each beat: (N) if the beat is normal and (T) if a T-wave change is observed. The direction of the change, (T+) for an elevation and (T-) for an inversion, is indicated at specific beats that delimit a current episode (Onset, Extrema and End beats), and this direction is associated to all beats in the episode.

For each record in the EDB database, the pre-processing and features extraction is performed for each lead by a script we developed using Labview [24]. It’s a graphical programming platform that offers many modules named Virtual Instruments (VI). For our experiments, we use the Filtering VI (0.5Hz FIR high-pass filter), the Wavelet Denoise VI (UWT with the Daubechies6 wavelet) and ECG Feature Extractor VI. The result of this stage is a series of ECG temporal parameters characterizing each beat of each lead. From these parameters, we calculate, for each beat, the relevant T-wave features Tamp, Tdur, STamp, QTdur.

Algorithm 1 Mobile application algorithm for a WBS system

Build the classification model on an initial training set TS;
Set the windows size \( w \); \( i = 0 \);
while Captured ECG beat \( k \) do
   Denoise ECG \( k \) with 0.5Hz highpass filter and UWT;
   Extract Peaks, Onsets and Offsets of P, QRS and T;
   Calculate the classification features Tamp, Tdur, STamp, QTdur;
   Apply the built Model on these features to classify the beat as Normal, Elevated or Inverted T-wave;
   if \( i = w \) then
      if (Number(Class ≠ Normal) > Min) then
         Raise an alarm;
      end if
      \( i = 0 \);
   else
      Update the model with the new classification results;
      \( i = i + 1 \);
   end if
end while

250 per second. The leads are different for each record so there is 7 distinct leads in the entire database (I, III, V1, V2, V3, V4, V5). Two cardiologists annotated each record beat-by-beat for changes in ST segment, T-wave morphology, rhythm, and signal quality. An anomaly code is associated to each beat: (N) if the beat is normal and (T) if a T-wave change is observed. The direction of the change, (T+) for an elevation and (T-) for an inversion, is indicated at specific beats that delimit a current episode (Onset, Extrema and End beats), and this direction is associated to all beats in the episode.

For each record in the EDB database, the pre-processing and features extraction is performed for each lead by a script we developed using Labview [24]. It’s a graphical programming platform that offers many modules named Virtual Instruments (VI). For our experiments, we use the Filtering VI (0.5Hz FIR high-pass filter), the Wavelet Denoise VI (UWT with the Daubechies6 wavelet) and ECG Feature Extractor VI. The result of this stage is a series of ECG temporal parameters characterizing each beat of each lead. From these parameters, we calculate, for each beat, the relevant T-wave features Tamp, Tdur, QTdur.

Figure 2 shows examples of normal, elevated and inverted T-wave ECG signals, extracted from the EDB record e0104.

Figure 3 shows the previous signals processed by a 0.5Hz FIR highpass digital filter and UWT. We can see that the baseline wandering was removed since the baseline was restored to Iso-level, while the main characteristics of the original ECG signals were kept. We can see also that the signal is less noisy, due to removal of wideband noises without its distortion.

Figure 4 shows the extracted temporal parameters P onset, P offset, R peak, QRS onset, QRS offset, T onset and T offset of the previous ECGs. As described in Section III, these values were obtained by extracting QRS complex peaks, and then deduct P and T peaks and their corresponding onsets and offsets.

Figure 5 shows the variation of calculated T-wave amplitude...
and duration parameters, during normal, elevation and inversion episodes. As we can see, the variation of features specific to T-wave (Tamp and Tdur) are significantly representative of the occurrence or not of elevation and inversion episodes. This is less significant in the case of the two other parameters (STamp and QTdur), since these parameters are less specific to T-wave but stay very related to it.

In order to build the data corpus for the classification, we associate to each beat, in addition to the four T-wave parameters, the annotation code associated to the beat (We implemented a script to read and extracted annotations from EDB). These annotations represent the classes of data (Normal (N), Elevated (T+), Inverted (T-)). To obtain more relevant results, we grouped the records by lead, so we can analyze the performance by lead on a greater number of beats. We get 7 sets of leads (V1, V2, V3, V4, V5, I, III). Each composed between a total of 45,000 to 359,526 beats. Classification is performed on each record separately, then the final results, per lead, are obtained by averaging the results of all records of a given lead. Table III summarizes the obtained corpus.

To compare the performance results of the seven classifiers selected, we use the ROC parameters described in the table IV. TP, FP, TN, FN are respectively, the number of True Positives, False Positives, True Negatives and False Negatives.

Table V shows the results obtained by lead for each classifier. In addition to ROC parameters, we give the time T elapsed for classification of 10 minutes ECG (about 600 beats).

The results show that Accuracy and Sensitivity parameters are greater than 90%. This indicates a good anomalies detection rate for most of classifiers and leads (performance are slightly lower for V2 lead). Error Rate varies by method and lead, but is generally below 10%. Specificity varies around 50%, this rate is caused by False Positives FP, which are not critical errors but may cause false alarms. To reduce this problem, we propose to raise an alarm, only if the number of T-wave changes detected is greater than a minimum Min in a time windows τ. Regarding the performance of each classification method, it appears that C4.5 achieves the best results with the lowest processing time T. NBC is the worst performer, with a processing time 11 to 14 times higher than C4.5 time, for all ECG leads. This may suggest that all studied leads are equivalent for the T-wave analysis. The performance of RF is close to C4.5, but with a longer processing time, 4 to 6 times. Other methods perform less than C4.5 and RF and have equivalent performance, but with a large disparity in processing time. LR and AdaBoost are 2 to 4 times slower than C4.5, while K-NN is about 500 times slower for a wireless ECG monitoring system implementation. We first present the ECG pre-processing stage done in order to remove various noises, using digital highpass filter and Undecimated Wavelet Transform (UWT). Then, we achieve an ECG temporal features extraction to calculate 4 ECG features related to T-wave (Tamp, STamp, Tdur and QTdur). These features are used to test 7 supervised learning classifiers (NBC, SVM, LR, K-NN, AdaBoost, C4.5 and RF), to select the best one both in terms of classification performance and processing time. We test our proposed approach using the EDB medical database composed of 90 real ECG records, and we use the ROC parameters to compare the models. Experimental results show that C4.5 performs the best results both in classification and

![Fig. 5. Variation of extracted T-wave features for each class](image)

### TABLE III
**COMPOSITION OF THE CORPUS**

<table>
<thead>
<tr>
<th>ECG Lead</th>
<th>Total records</th>
<th>Total beats</th>
<th>Total “N” beats</th>
<th>Total “T+” beats</th>
<th>Total “T-” beats</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>11</td>
<td>71165</td>
<td>64111</td>
<td>2562</td>
<td>44137</td>
</tr>
<tr>
<td>V2</td>
<td>10</td>
<td>78702</td>
<td>53906</td>
<td>11909</td>
<td>1788</td>
</tr>
<tr>
<td>V3</td>
<td>7</td>
<td>46009</td>
<td>41479</td>
<td>3209</td>
<td>1872</td>
</tr>
<tr>
<td>V4</td>
<td>34</td>
<td>223843</td>
<td>200703</td>
<td>10514</td>
<td>12627</td>
</tr>
<tr>
<td>V5</td>
<td>51</td>
<td>355228</td>
<td>335088</td>
<td>10654</td>
<td>9693</td>
</tr>
<tr>
<td>I</td>
<td>19</td>
<td>129914</td>
<td>126112</td>
<td>1182</td>
<td>2621</td>
</tr>
<tr>
<td>III</td>
<td>46</td>
<td>309388</td>
<td>200613</td>
<td>16443</td>
<td>6338</td>
</tr>
</tbody>
</table>

To perform the classification experiments, we choose Weka [25], a well-known machine learning tool. It is a large collection of machine learning algorithms for classification, clustering, association rules, regression and visualization.

### V. Conclusion

In this paper, we study the T-wave changes detection problem for a wireless ECG monitoring system implementation. We first present the ECG pre-processing stage done in order to remove various noises, using digital highpass filter and Undecimated Wavelet Transform (UWT). Then, we achieve an ECG temporal parameters extraction to calculate 4 ECG features related to T-wave (Tamp, STamp, Tdur and QTdur). These features are used to test 7 supervised learning classifiers (NBC, SVM, LR, K-NN, AdaBoost, C4.5 and RF), to select the best one both in terms of classification performance and processing time. We test our proposed approach using the EDB medical database composed of 90 real ECG records, and we use the ROC parameters to compare the models. Experimental results show that C4.5 performs the best results both in classification and
time processing, with an average of 92.54% Accuracy, 96.06% Sensibility, 55.41% Specificity and 7% Error rate. To validate these results in a real user environment, we are currently working on an end-to-end implementation of C4.5 based model using market wireless ECG sensors.

### REFERENCES