HAL
open science

A comparative study of supervised learning techniques for ECG T-wave anomalies detection in a WBS context

Medina Hadjem, Farid Naït-Abdesselam

To cite this version:

Medina Hadjem, Farid Naït-Abdesselam. A comparative study of supervised learning techniques for ECG T-wave anomalies detection in a WBS context. Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS), 2015 International Conference on, Jul 2015, Paris, France. 10.1109/NOTERE.2015.7293505 . hal-01231978

HAL Id: hal-01231978

https://hal.science/hal-01231978v1

Submitted on 21 Nov 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.

A Comparative Study of Supervised Learning Techniques for ECG T-wave Anomalies Detection in a WBS context

Medina Hadjem[a,*] and Farid Naït-Abdesselam[b]
Paris Descartes University, France
[a] medina.hadjem@parisdescartes.fr
[b] naf@parisdescartes.fr

***Abstract*—Today, most of Wireless Body Sensors (WBS) for remote monitoring of cardiovascular disease, rarely include automatic analysis and detection of ECG abnormalities, or are limited to cardiac arrhythmia's. The detection of more complex cardiac anomalies such as Ischemia or myocardial infarction, requires an advanced analysis of ECG wave Known as P, Q, R, S, and T, especially the T-wave, which is often associated with serious cardiac anomalies. The goal of this paper is to study the classification of T-wave abnormalities with consideration to a context of wireless monitoring system. The study approach is based on experimentation and comparison of classification performance and response time of 7 supervised learning models. We performed our experiments on a real ECG data from the EDB medical database from Physionet. Our results show that the decision trees models offer better results with, on average, an Accuracy of 92.54 %, a Sensitivity of 96.06%, a Specificity of 55.41% and an Error Rate 7.41%.**

***Index terms*— ECG, T-wave, WBS, Supervised Learning, Classification**

I. Introduction

Il existe aujourd'hui des capteurs biomédicaux, connus sous le nom de « Capteurs Corporels Sans Fil » ou WBS, pouvant se fixer sur le corps, collecter divers paramètres physiologiques et les transmettre en quasi temps réel à une passerelle telle qu'un smartphone, un PC, une tablette, etc. Cette dernière traite et transmet les informations recueillies à des professionnels de santé distants pour la prise de décisions médicales appropriées.

Diverses applications des capteurs corporels sans fil peuvent être envisagées. L'une des plus prometteuses est la télésurveillance des Maladies CardioVasculaires (MCV) pour la détection automatique d'anomalies. En effet, selon l'Organisation Mondiale de la Santé (OMS) [1], ces maladies représentent la principale cause de décès dans le monde et d'ici 2030, près de 23.3 millions de personnes seront victimes chaque année des MCV. De ce fait, le développement de systèmes de télésurveillance cardiovasculaire est un véritable enjeu de santé publique, la recherche dans ce domaine est très stratégique et est fortement soutenue par les gouvernements.

Plusieurs problématiques doivent être traitées pour la mise en place d'un système de télésurveillance cardiovasculaire (conception, sécurité, fiabilité, éthique etc.). Nous nous intéressons dans ce travail à la détection automatique d'anomalies cardiovasculaires majeures, telles que l'ischémie ou l'infarctus. Aujourd'hui, le diagnostic d'anomalies cardiaques est établi suite à des symptômes qui conduisent le patient à consulter un médecin ou lors d'un examen de routine. Dans les deux cas, le médecin se base sur plusieurs examens cliniques comprenant l'électrocardiogramme (ECG).

L'ECG représente la contraction et relaxation du muscle cardiaque, cette activité électrique est enregistrée par des électrodes placées sur les membres et le thorax et transcrite par un équipement spécial sur un papier millimétré. La technique d'enregistrement ECG la plus largement utilisée est l'ECG à 12 dérivations. Cette technique utilise 10 électrodes à partir desquelles sont dérivés 12 signaux ECG appelés dérivations. On distingue 6 dérivations précordiales (V1, V2, V3, V4, V5, V6) et 6 dérivations frontales (I, II, III, AVR, AVL, aVF). Chaque dérivation ECG représente une vue différente de l'activité cardiaque et consiste en une suite de battements ECG. Un battement est typiquement composé de 5 ondes symbolisées P, Q, R, S et T et sont séparées par des intervalles et segments PR, ST et QT. La Figure 1 illustre un battement ECG typique.

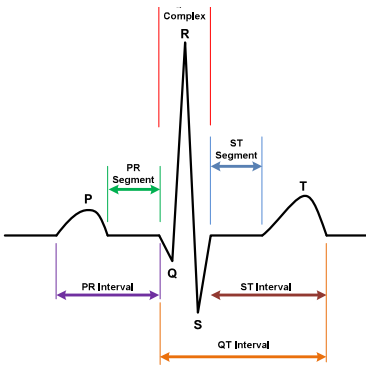

FIGURE 1. Un battement ECG typique avec ondes et segments

L'onde T est une onde majeure de l'ECG et ses variations sont parmi les plus couramment observées. Ces changements peuvent révéler diverses anomalies en fonction de leurs formes, amplitudes et de la présence ou absence d'autres facteurs de risque. Quatre cas d'anomalies de l'onde T sont généralement distingués [2] :

– *Onde T surélevée* : Une augmentation de l'amplitude de l'onde T peut être le signe d'un stade précoce d'un Infarctus du Myocarde (IM), survenant lors des premières minutes suivant l'occlusion de l'artère coronaire et représentant une ischémie sans nécrose. L'intervention à ce stade précoce peut empêcher un infarctus et améliorer le résultat du traitement. D'autres anomalies peuvent être liées à une onde T surélevée telles que l'hyperkaliémie, le Bloc de Branche ou l'Hypertrophie Ventriculaire Gauche.
– *Onde T inversée* : l'inversion de l'onde T peut être une indication d'ischémie ou d'un Syndrome Coronarien Aigü (SCA) résultant de l'occlusion totale ou partielle d'une artère coronaire. En général, l'inversion profonde de l'onde T est souvent un signe d'une lésion cardiaque sous-jacente grave. L'inversion de l'onde T peut également être le signe d'une hyperventilation ou d'une embolie pulmonaire.
– *Onde T aplatie ou biphasique* : Les deux causes principales de l'onde T biphasique sont l'ischémie myocardique et l'hypokaliémie. Les ondes T aplaties sont en général des observations non spécifiques, elles peuvent dans certains cas suggérer une ischémie.

A ce jour, il n'y a pas de paramètres de normalité clairement établis de l'amplitude et de la durée de l'onde T. Habituellement, les spécialistes s'intéressent aux changements d'amplitude ou de forme dans le temps, en comparant l'ECG à un instant t avec des enregistrements précédents. Par conséquent, la pertinence de l'analyse de l'onde T réside dans la possibilité de suivre ses variations dans le temps comme le permet l'utilisation de capteurs corporels. Cette étude doit donc être considérée dans un contexte de télésurveillance continue de l'ECG dans le but d'alerter et non d'établir un diagnostic.

Le but de cette étude est de proposer un modèle de détection des anomalies de l'onde T offrant à la fois de bonnes performances et un temps de traitement adapté à la contrainte temps réel des systèmes de monitoring ECG sans fil. Ce modèle doit être capable de distinguer entre les ondes T normales, surélevées ou inversées (Le cas des ondes T aplaties ou biphasiques ne sera pas traité dans cette étude).

La suite de cette étude est organisée comme suit : La section II étudie brièvement l'état de l'art. La section III décrit l'approche proposée et la section IV discute les résultats obtenus. Enfin, la section V conclut cet article.

II. État de l'art

Le développement au cours de la dernière décennie de la télésanté et des capteurs biomédicaux, a initié plusieurs projets et travaux dans le domaine de la télésurveillance médicale, aussi bien au niveau de la recherche académique qu'en industrie.

Dans cette étude, nous nous intéressons en particulier aux travaux concernant les systèmes de télésurveillance cardiovasculaire basés sur l'ECG. Initialement, la plupart de ces travaux proposaient des systèmes de collecte et de télétransmission de l'ECG sans détection d'anomalies, les prototypes MyHeart [3] et MEDISN [4] en sont un exemple. Plus récemment, ces systèmes évoluent pour intégrer la détection automatique des troubles du rythme cardiaque, comme dans les travaux [5] et [6]. Certains de ces systèmes sont déjà sur le marché tels que CARDIONET [7] et BodyGuardian [8]. D'autres contributions telles que [9] et [10], s'intéressent à la détection d'anomalies plus complexes comme l'infarctus du myocarde ou l'ischémie.

Les méthodes d'analyse proposées dans ces systèmes sont généralement basées sur les caractéristiques temporelles de l'ECG, c'est à dire des ondes P, Q, R, S, T, et des intervalles et segments associés. La plupart analysent l'intervalle RR comme indicateur de la fréquence cardiaque, ou l'élévation du segment ST qui peut être le signe d'un infarctus ou d'une ischémie myocardique. D'autres études analysent également l'onde P pour la détection des arythmies.

Parmi ces travaux, peu abordent l'analyse de l'onde T, les travaux les plus significatifs sur cette onde, concernent la détection de l'alternance de l'onde T ou T-wave Alternance (TWA), qui consiste en une variation de l'amplitude et/ou de la forme de l'onde T à raison d'un battement sur deux. Ceci est lié à une instabilité cardiaque et peut indiquer un risque de mort subite. Plusieurs contributions sur la détection de TWA ont été proposées, celles-ci sont basées sur différentes techniques de traitement du signal, d'exploration de données ou d'apprentissage automatique. De bons résultats ont été obtenus en utilisant l'Analyse en Composante Principale (ACP) [11], la méthode spectrale [12] et la Transformée de Fourier Rapide (TFR) [13]. Plusieurs autres techniques de détection de l'alternance de l'onde T sont comparées et discutées dans [14]. L'analyse de l'onde T a également été utilisée dans la détection de l'ischémie myocardique, comme paramètre supplémentaire au segment ST [15]. D'autres contributions proposent une détection générale des épisodes de changement de l'onde T comme dans [16]. A notre connaissance, il n'y a que très peu de travaux sur les systèmes de télésurveillance ECG intégrant l'analyse de l'onde T, [17] est une des rares contribution qui aborde la conception d'un système de diagnostic d'anomalies ECG, basé sur une plateforme de capteurs corporels sans fil et intégrant l'analyse de l'onde T.

Dans cette étude, nous expérimentons 7 modèles d'apprentissage supervisé connus afin de proposer un modèle qui combine de bonnes performances de classification de l'onde T et un temps de réponse convenable à une implémentation dans un système de télésurveillance temps réel.

III. Approche proposée

L'approche proposée dans cette étude est axée sur 3 étapes principales : prétraitement et extraction des paramètres ECG, expérimentation de plusieurs modèles de classification et proposition d'intégration du modèle sélectionné dans un système de télésurveillance ECG.

A. Prétraitement et extraction des paramètres ECG

Le prétraitement est un pré-requis nécessaire à l'élimination des divers bruits qui contaminent l'ECG et qui peuvent fausser les résultats de l'analyse. Ces bruits peuvent couvrir une large plage de fréquence de [0.01Hz-150Hz] et par conséquent

interférent avec les fréquences utiles de l'ECG. Ci-dessous, nous décrivons brièvement les bruits qui peuvent contaminer l'ECG et les techniques utilisées pour les atténuer :

- La respiration et le mouvement des patients sont responsables de déplacements de la ligne isoélectrique de l'ECG. Leur fréquence se situe généralement entre 0.15Hz et 0.5Hz (en dessous des fréquences utiles de l'ECG). Afin d'éliminer ce bruit, nous utilisons un filtre Passe-haut à réponse impulsionnelle finie, avec une fréquence de coupure de 0.5Hz.
- Le bruit induit par le mouvement des électrodes (entre 1Hz à 10Hz) et par l'activité électrique musculaire (entre 25Hz à 100Hz), sont plus complexes à éliminer car ils se confondent avec les fréquences utiles de l'ECG, ce qui empêche l'utilisation de filtres qui sont souvent remplacés par la technique de la Transformée en Ondelettes Discrète ou Discrete Wavelet Transform (DWT). Dans cette étude, nous utilisons la Transformée en Ondelettes Non Décimée ou Undecimated Wavelet Transform (UWT), une variante de la DWT qui présente un meilleur compromis entre filtrage de l'ECG et maintien de son information utile. L'utilisation de cette technique est décrite dans [18].
- Le bruit relatif aux circuits électriques et à la ligne d'alimentation oscille autour de 50Hz ou de 60Hz, il est possible de l'éliminer par filtre passe-bas et il est généralement filtré par le matériel d'acquisition de l'ECG, c'est à dire les capteurs.

Suite au prétraitement, une analyse doit être effectuée pour estimer les positions temporelles des différentes ondes ECG (ondes du complexe QRS, l'onde P et l'onde T). Dans cette étape, l'ECG pré-traité (filtré) est utilisé uniquement pour détecter les positions des complexes QRS, les autres paramètres temporels seront extraits en utilisant le signal d'origine, pour éviter l'impact d'éventuels distorsions introduites par le filtrage. La technique utilisée pour l'estimation des paramètres temporels est inspirée par la méthode Pan & Tompkins, connue pour la détection des pics QRS [19]. Le tableau I synthétise les paramètres temporels extraits pour chaque battement ECG.

TABLE I
PARAMÈTRES TEMPORELS EXTRAITS DE CHAQUE BATTEMENT ECG

Paramètre	Description
Ponset	Instant début de l'onde P en *sec*
Poffset	Instant fin de l'onde P en *sec*
QRSonset	Instant début du complexe QRS en *sec*
QRSoffset	Instant fin du complexe QRS en *sec*
Tonset	Instant début de l'onde T en *sec*
Toffset	Instant fin de l'onde T en *sec*

A partir des ces paramètres temporels, nous calculons les paramètres ECG finaux sur lesquels se basera la classification. Seuls les paramètres les plus pertinents liés à des anomalies de l'onde T seront retenus, à la fois pour réduire la taille des données à traiter, et augmenter la précision du modèle. Après avoir recueilli plusieurs informations médicales sur les anomalies de l'onde T [2], nous avons retenu les paramètres décrits dans le tableau II. Ils sont calculés sur la base des paramètres temporels extraits précédemment.

TABLE II
PARAMÈTRES ECG SÉLECTIONNÉS POUR LA CLASSIFICATION DE L'ONDE T

Paramètre	Description
Tamp	Amplitude en mV de l'onde T
Tdur	Durée en *sec* de l'onde T entre Tonset et Toffset
STamp	Amplitude en mV du segment ST entre QRSoffset et Tonset
QTdur	Durée en *sec* de l'intervalle QT entre QRSonset et Toffset

B. Sélection d'un modèle de classification de l'onde T

L'amplitude et la forme de l'onde T peut varier d'un individu à l'autre et il n'existe pas de paramètres de normalité établis pour classer les ondes T, c'est pourquoi, une approche de classification par apprentissage supervisé semble pertinente. Elle permet de construire un modèle prédictif sur des données observées et préalablement annotées par un spécialiste, afin de pouvoir établir des prédictions sur de nouvelles données inconnues. Plus formellement, un problème de classification consiste en un ensemble d'objets X appelé l'espace d'entrée et un ensemble fini C de classes appelé l'espace de sortie. Nous considérons un échantillon $E = \{(x_1, c_1), ..., (x_n, c_n)\}$ avec une probabilité de distribution inconnue D sur $X \times C$. L'objectif de la classification est de trouver la fonction $h \in C^X$ qui définit la relation entre les objets et les classes, en se basant uniquement sur l'échantillon E, généré indépendamment et au hasard selon D. Comme nous sommes à la recherche d'un modèle approprié pour un système à contrainte temps réel, nous nous intéressons également au temps de réponse du système. Il existe plusieurs modèles d'apprentissage supervisé, dans cette étude, nous comparons 7 modèles connus : 3 modèles linéaires (Naïve Bayes Classifier (NBC), Support Vector Machine (SVM) et Logistic Regression (LR)), 2 modèles non-linéaires (K-Nearest Neighbors (K-NN) et AdaBoost) et 2 modèles d'arbres de décision (Decision Trees (C4.5) et Random Forest (RF)). Un aperçu de ces modèles peut être trouvé dans [20]. Les résultats obtenus sont détaillés au niveau de la section IV.

C. Intégration du modèle dans un système de télésurveillance

Nous considérons un système de télésurveillance ECG basé sur des capteurs corporels sans fil. Le modèle sélectionné peut être intégré en tant qu'une application mobile fonctionnant sur un smartphone. Cette application reçoit l'ECG en quasi temps réel depuis les capteurs, pré-traite et extrait les paramètres liés à l'onde T (*Tamp, Tdur, STamp, QTdur*) et utilise ces paramètres pour classer l'onde T associée en utilisant le modèle sélectionné. En fonction du résultat de classification, l'application mobile remonte une alarme à l'assistance médicale. Le modèle est d'abord construit grâce à des données ECG réelles pré-annotées par des spécialistes, puis sera mis à jour périodiquement avec les données collectées du patient pour la personnalisation. Afin de limiter le nombre d'alarmes, nous définissons un seuil Min qui représente le nombre minimum d'ondes T classées comme anormales durant une fenêtres de temps w (Min et w sont des paramètres de l'application mobile). L'algorithme 1 résume les principales étapes réalisées par l'application mobile cible.

Algorithm 1 Classification de l'onde T de l'ECG

Dériver le modèle de classification sur un ensemble d'apprentissage TS ;
Positionner la taille de la fenêtre w ; $i = 0$;
while battement ECG reçu k **do**
  filtrer l'ECG k (filtre passe-haut de 0.5Hz et UWT) ;
  Extraire les Pics, Onsets et Offsets des ondes P,QRS et T ;
  Calculer les paramètres *Tamp*, *Tdur*, *STamp*, *QTdur* ;
  Appliquer le modèle sur ces paramètres afin de classer l'onde T de l'ECG en normale, surélevée ou inversée ;
  if $i = w$ **then**
    if (Nombre($Class \neq Normal$) > Min) **then**
      Remonter une alarme ;
    end if
    $i = 0$;
    mettre à jour TS avec les nouvelles données ECG ;
  else
    $i = i + 1$;
  end if
end while

IV. Résultats

Afin d'expérimenter l'approche proposée, nous utilisons la base de données européenne (EDB) de Physionet [21]. Il s'agit d'une collection d'enregistrements ECG de divers patients, spécifiquement destinée à être utilisée pour l'évaluation d'algorithmes de détection de changements du segment ST et de l'onde T. Elle contient une sélection représentative des anomalies de l'onde T comprenant 401 épisodes de changement de l'onde T. La durée de ces épisodes varie de 30 secondes à plusieurs minutes. Chaque enregistrement a une durée de 2 heures (7200 battements) et contient deux dérivations (différentes pour chaque enregistrement). Deux cardiologues ont annoté chaque enregistrement et ont associé à chaque battement un code anomalie : (N) si le battement est normal et (T) si un changement de l'onde T est observé. La direction du changement est également indiquée, (T+) pour une élévation et (T-) pour une inversion.

Pour chaque enregistrement de la base de données EDB, le prétraitement et l'extraction des paramètres ECG sont effectués pour chaque dérivation par un script que nous avons développé en utilisant l'outil LabVIEW [22]. Le résultat de cette étape est une série de paramètres temporels ECG caractérisant chaque battement de chaque dérivation. A partir de ces paramètres, nous calculons pour chaque battement les paramètres pertinents de l'onde T *Tamp*, *Tdur*, *STamp*, *QTdur*.

La Figure 2 montre des exemples de signaux ECG avec une onde T normale, élevée et inversée, extraites de l'enregistrement $e0104$ de la base EDB.

La Figure 3 montre les signaux précédents après filtrage (filtre numérique FIR passe-haut à 0.5Hz et la technique UWT). Nous pouvons constater que l'ondulation de la ligne de base a été éliminée et rétablie au niveau isoélectrique, tandis que les principales caractéristiques des signaux ECG d'origine ont

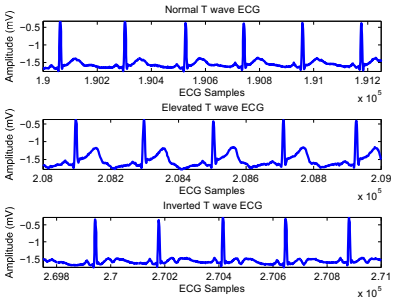

FIGURE 2. Des ECGs de la base EDB avec diverses ondes T

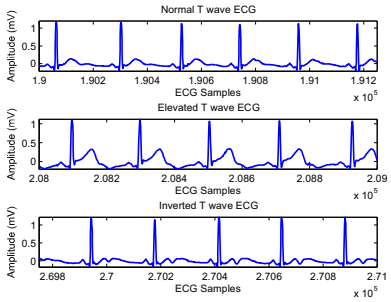

FIGURE 3. ECG précédents après pré-traitement

été conservées. Nous constatons aussi que le signal est moins bruyant, en raison de l'élimination des bruits à large bande (EMG, mouvements), et ce sans distorsion du signal.

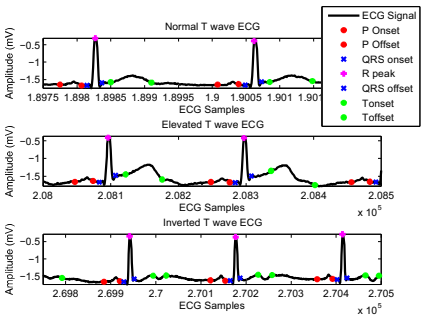

FIGURE 4. ECGs précédents après extraction des paramètres temporels

La Figure 4 montre les paramètres temporels extraits *Ponset*, *Poffet*, *pic R*, *QRSonset*, *QRSoffet*, *Tonset* et *Toffet*.

La Figure 5 montre la variation de l'amplitude et de la durée de l'onde T au cours d'un épisode d'élévation, d'inversion et de normalité. Comme nous pouvons le voir, les variations des paramètres spécifiques à l'onde T (*Tamp* et *Tdur*) sont significativement représentatives de la survenance ou non d'un épisode d'élévation et d'inversion. Ceci est moins évident dans le cas des deux autres paramètres (*STamp* et *QTdur*), puisque ces paramètres sont moins spécifiques à l'onde T.

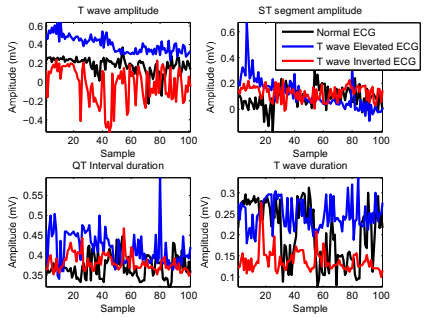

FIGURE 5. Variation des paramètres de l'onde T extraits pour chaque classe

Afin de construire le corpus de données pour la classification, on associe à chaque battement ECG, en plus des quatre paramètres de l'onde T cités précédemment, le code anomalie associé au battement annoté par les spécialistes. Ces codes d'anomalies représentent les classes : onde T normale (N), onde T élevée (T+), onde T inversée (T-). Afin d'obtenir des résultats plus pertinents, nous regroupons les enregistrements ECG par dérivation, ceci nous permettra d'étudier les différences entre les dérivations et d'avoir pou chaque corpus des battements ECG issus de patients différents. Nous avons construit 7 corpus de données, correspondant aux 7 dérivations différentes de la base EDB (V1, V2, V3, V4, V5, I, III). Chaque corpus est composé des données de plusieurs enregistrements. La taille de chaque corpus varie de 45.000 à 359.526 battements au total. La classification se fait sur chaque enregistrement séparément, puis les résultats définitifs, par dérivation, sont obtenues en moyennant les résultats de tous les enregistrements d'une même dérivation. Le tableau III synthétise le corpus obtenu.

TABLE III
COMPOSITION DU CORPUS

Dérivation ECG	Nombre d'enregistrements	Nombre d'ECGs	Nombre d'ECGs «N»	Nombre d'ECGs «T+»	Nombre d'ECGs «T-»
V1	11	71165	64111	2642	4413
V2	10	78762	55006	11969	1788
V3	7	46609	41470	3269	1872
V4	34	223843	200703	10514	12627
V5	51	355428	335088	10654	9693
I	19	129914	126112	1182	2621
III	46	309388	286613	16443	6338

Pour réaliser nos expérimentations, nous utilisons Weka [23] avec l'option de classification « 10-fold cross validation », et pour comparer la performance des 7 modèles sélectionnés, nous utilisons les paramètres ROC décrits dans le tableau IV. TP, FP, TN, FN sont respectivement, le nombre de vrais positifs, faux positifs, vrais négatifs et faux négatifs.

Le tableau V montre les résultats obtenus ,par dérivation, pour chaque modèle. En plus des paramètres ROC, nous calculons le temps T passé pour le classement d'un segment de 10 minutes d'ECG (environ 600 battements ECG).

TABLE IV
PARAMÈTRES ROC UTILISÉS POUR LA COMPARAISON DES 7 MODÈLES DE CLASSIFICATION

Paramètre	Description	Formule
Exactitude(Acc)	Taux de classification correcte	$\frac{TP+TN}{TP+TN+FP+FN}$
Sensibilité(Se)	Taux de vrais positifs	$\frac{TP}{TP+FN}$
Spécificité(Sp)	Taux de vrais négatifs	$\frac{TN}{FP+TN}$
Taux d'erreur (Err)	Taux de classification incorrecte	$\frac{FP+FN}{TP+TN+FP+FN}$

TABLE V
RÉSULTATS DE LA CLASSIFICATION

Dérivation	Méthode	Acc(%)	Se(%)	Sp(%)	Err(%)	T(ms)
V1	BNC	89.13%	93.47%	57.30%	10.87%	0.013
V1	LR	90.99%	97.39%	41.38%	9.01%	0.003
V1	K-NN	91.17%	94.76%	58.65%	8.83%	0.507
V1	AdaBoost	90.20%	97.69%	36.38%	9.80%	0.002
V1	C4.5	92.65%	96.84%	53.37%	7.35%	0.001
V1	RF	92.52%	96.78%	55.57%	7.48%	0.005
V1	SVM	89.98%	99.07%	23.26%	10.02%	0.290
V2	BNC	80.43%	78.42%	64.36%	19.57%	0.013
V2	LR	83.47%	85.32%	62.08%	16.53%	0.004
V2	K-NN	83.24%	85.13%	71.89%	16.76%	0.508
V2	AdaBoost	83.67%	86.21%	63.99%	16.33%	0.002
V2	C4.5	86.67%	90.04%	72.41%	13.33%	0.001
V2	RF	86.20%	90.10%	71.00%	13.80%	0.008
V2	SVM	83.57%	86.61%	54.98%	16.43%	0.523
V3	BNC	86.58%	88.72%	62.68%	13.42%	0.013
V3	LR	91.86%	98.10%	44.17%	8.14%	0.003
V3	K-NN	90.75%	94.65%	60.81%	9.25%	0.580
V3	AdaBoost	90.44%	98.84%	34.57%	9.56%	0.002
V3	C4.5	92.97%	97.44%	56.04%	7.03%	0.001
V3	RF	92.79%	97.36%	56.21%	7.21%	0.006
V3	SVM	90.45%	99.56%	26.95%	9.55%	0.314
V4	BNC	90.30%	94.07%	58.77%	9.70%	0.012
V4	LR	92.84%	96.67%	51.71%	7.16%	0.004
V4	K-NN	92.96%	95.08%	66.29%	7.04%	0.537
V4	AdaBoost	91.50%	95.86%	45.89%	8.50%	0.002
V4	C4.5	94.12%	96.53%	64.11%	5.88%	0.001
V4	RF	94.24%	96.94%	64.20%	5.76%	0.005
V4	SVM	92.06%	97.12%	38.01%	7.94%	0.279
V5	BNC	90.00%	93.67%	45.84%	10.00%	0.014
V5	LR	92.84%	97.74%	36.02%	7.16%	0.004
V5	K-NN	92.50%	95.23%	55.56%	7.50%	0.580
V5	AdaBoost	92.98%	96.99%	40.20%	7.02%	0.002
V5	C4.5	93.96%	97.36%	49.58%	6.04%	0.001
V5	RF	93.87%	97.31%	51.55%	6.13%	0.005
V5	SVM	92.54%	98.40%	27.02%	7.46%	0.265
I	BNC	91.46%	95.15%	35.61%	8.54%	0.014
I	LR	93.29%	98.76%	20.19%	6.71%	0.003
I	K-NN	92.47%	95.66%	48.34%	7.53%	0.545
I	AdaBoost	93.11%	97.83%	25.60%	6.89%	0.002
I	C4.5	93.78%	98.08%	33.73%	6.22%	0.001
I	RF	93.83%	98.00%	38.20%	6.17%	0.006
I	SVM	93.31%	99.63%	12.70%	6.69%	0.229
III	BNC	88.63%	93.24%	50.26%	11.37%	0.011
III	LR	92.16%	97.44%	34.16%	7.84%	0.003
III	K-NN	92.42%	94.59%	60.48%	7.58%	0.476
III	AdaBoost	92.57%	96.13%	44.19%	7.43%	0.002
III	C4.5	93.94%	96.19%	58.29%	6.06%	0.001
III	RF	93.89%	96.54%	57.67%	6.11%	0.004
III	SVM	92.77%	97.62%	29.51%	7.23%	0.232

Les résultats montrent que l'exactitude et la sensibilité sont supérieures à 90%, ce qui indique un bon taux de détection des anomalies par la plupart des modèles et pour toutes les déri-

vations. Toutefois, les performances sont légèrement inférieur pour la dérivation V2. Le taux d'erreur varie selon le modèle et la dérivation, mais il est généralement inférieure à 10%. La spécificité varie autour de 50%, ce taux est dû aux faux positifs FP, qui sont des erreurs non critiques, mais qui causent des fausses alarmes. Pour réduire cette problématique, nous avons introduit deux paramètres qui permettent de ne déclencher une alarme que si le nombre d'anomalies détectées de l'onde T, est supérieur à un minimum Min durant une fenêtre temporelle w. En ce qui concerne la performance de chaque modèle de classification, il apparaît que la méthode C4.5 basé sur les arbres de décision offre les meilleurs résultats, aussi bien en termes de classification que de temps de traitement, tandis que la méthode NBC présente les moins bonnes performances, avec un temps de traitement de 11 à 14 fois plus élevé que celui de C4.5. Nous constatons ces conclusions pour toutes les dérivations de l'ECG, ceci peut suggérer que les dérivations étudiées sont équivalentes pour l'analyse de l'onde T. La performance de la méthode RF est proche de C4.5 mais avec un temps de traitement plus long, 4 à 6 fois plus. les autres modèles présentent de moins bonnes performances que C4.5 et RF et ont des performances assez équivalentes, mais avec une grande disparité en terme de temps de traitement. La méthode LR et AdaBoost sont de 2 à 4 fois plus lent que C4.5, tandis que K-NN est environ 500 fois plus lent et SVM près de 300 fois plus lent.

Il ressort de ces premiers résultats que C4.5 est un bon choix pour une mise en œuvre d'un système de détection des changements de l'onde T de ECG, dans un environnement temps réel basé sur des capteurs corporels.

V. Conclusion

Dans cette étude, nous nous sommes intéressés à la problématique de détection des anomalies de d'onde T de l'ECG dans un contexte de télésurveillance sans fil. nous avons étudié la détection de deux anomalies de l'onde T (surélevée ou inversée). L'approche proposée combine les techniques de traitement du signal pour les pré-requis de filtrage et l'extraction de paramètres temporels avec d'apprentissage supervisé pour la détection d'anomalies. Nous avons expérimenté et comparé 7 modèles d'apprentissage supervisé (NBC, SVM, LR, K-NN, AdaBoost, C4.5 et RF), à la fois en termes de performance que de temps de traitement, en utilisant des ECG réels issus de la base de données médicale EDB de Physionet. Les résultats expérimentaux montrent que le modèle C4.5 offre de meilleurs résultats avec une moyenne de 92.54% d'exactitude, 96.06% de sensibilité, 55.41% de spécificité et 7% de taux d'erreur. Ce modèle peut être intégré en tant qu'application mobile d'un système de télésurveillance ECG. Afin de valider ces résultats dans un contexte utilisateur, nous travaillons actuellement sur l'implémentation d'un prototype bout en bout avec des capteurs corporels sans fil.

Références

[1] "World health organisation," 2015. [Online]. Available : http://www.who.int/mediacentre/factsheets/fs317/en/

[2] F. Morris, W. J. Brady, and J. Camm, *ABC of clinical electrocardiography*. John Wiley & Sons, 2009, vol. 93.

[3] J. Luprano, J. Solà, S. Dasen, J.-M. Koller, and O. Chételat, "Combination of body sensor networks and on-body signal processing algorithms : the practical case of myheart project," in *International Workshop on Wearable and Implantable Body Sensor Networks*. IEEE, 2006, pp. 4–pp.

[4] J. Ko, J. H. Lim, Y. Chen, R. Musvaloiu-E, A. Terzis, G. M. Masson, T. Gao, W. Destler, L. Selavo, and R. P. Dutton, "Medisn : Medical emergency detection in sensor networks," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 10, no. 1, p. 11, 2010.

[5] H. Zhou, K. M. Hou, J. Ponsonnaille, L. Gineste, and C. De Vaulx, "A real-time continuous cardiac arrhythmias detection system : Recad," in *27th Annual International Conference of the Engineering in Medicine and Biology Society*. IEEE, 2006, pp. 875–881.

[6] D. D. Patil, D. Patil, S. Pandharpatte, R. Dhekane, T. Mohol, and V. Wadhai, "Intelligent arrhythmia diagnostics system," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, no. 6, 2012.

[7] "Mcot cardionet," 2015. [Online]. Available : http://www.cardionet.com

[8] "Bodyguardian remote monitoring system," 2015. [Online]. Available : http://www.preventice.com/products/bodyguardian/

[9] Z. Sankari and H. Adeli, "Heartsaver : A mobile cardiac monitoring system for auto-detection of atrial fibrillation, myocardial infarction, and atrio-ventricular block," *Computers in biology and medicine*, vol. 41, no. 4, pp. 211–220, 2011.

[10] A. Huang, C. Chen, K. Bian, X. Duan, M. Chen, H. Gao, C. Meng, Q. Zheng, Y. Zhang, B. Jiao *et al.*, "We-care : an intelligent mobile telecardiology system to enable mhealth applications," 2014.

[11] G. Bortolan and I. Christov, "Principal component analysis for detection and assessment of t-wave alternans," in *Computers in Cardiology*. IEEE, 2008, pp. 521–524.

[12] A. Khaustov, S. Nemati, and G. Clifford, "An open-source standard t-wave alternans detector for benchmarking," in *Computers in Cardiology*. IEEE, 2008, pp. 509–512.

[13] D. Zheng, S. Stevens, P. Langley, K. Wang, A. Haigh, S. King, and A. Murray, "T-wave alternans : A comparison of different measurement techniques," in *Computers in Cardiology*. IEEE, 2008, pp. 597–600.

[14] J. P. Martínez and S. Olmos, "Methodological principles of t wave alternans analysis : a unified framework," *Biomedical Engineering, IEEE Transactions on*, vol. 52, no. 4, pp. 599–613, 2005.

[15] A. R. M. Dehnavi, I. Farahabadi, H. Rabbani, A. Farahabadi, M. P. Mahjoob, and N. R. Dehnavi, "Detection and classification of cardiac ischemia using vectorcardiogram signal via neural network," *Journal of research in medical sciences : the official journal of Isfahan University of Medical Sciences*, vol. 16, no. 2, p. 136, 2011.

[16] C. Papaloukas, D. I. Fotiadis, A. Likas, C. S. Stroumbis, and L. K. Michalis, "Use of a novel rule-based expert system in the detection of changes in the st segment and the t wave in long duration ecgs," *Journal of electrocardiology*, vol. 35, no. 1, pp. 27–34, 2002.

[17] F. J. Rincón, L. Gutiérrez, M. Jiménez, V. Diaz, N. Khaled, D. Atienza, M. Sánchez-Elez, J. Recas, and G. De Micheli, "Implementation of an automated ecg-based diagnosis for a wireless body sensor platform," in *Proceedings of BIODEVICES*, vol. 1, no. EPFL-CONF-130660. Springer, 2009, pp. 88–96.

[18] V. P. Raj and T. Venkateswarlu, "Ecg signal denoising using undecimated wavelet transform," in *3rd IEEE International Conference on Electronics Computer Technology (ICECT)*, vol. 3, 2011, pp. 94–98.

[19] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *Biomedical Engineering, IEEE Transactions on*, no. 3, pp. 230–236, 1985.

[20] P. Cunningham, M. Cord, and S. J. Delany, "Supervised learning," in *Machine Learning Techniques for Multimedia*. Springer, 2008, pp. 21–49.

[21] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg, and C. Marchesi, "The european st-t database : standard for evaluating systems for the analysis of st-t changes in ambulatory electrocardiography," *Eur Heart J*, vol. 13, no. 9, pp. 1164–72, 1992.

[22] "Labview for ecg signal processing," 2012. [Online]. Available : http://www.ni.com/white-paper/6349/en/

[23] I. H. Witten, E. Frank, L. E. Trigg, M. A. Hall, G. Holmes, and S. J. Cunningham, "Weka : Practical machine learning tools and techniques with java implementations," 1999.