Politecnico di Milano

School of Industrial and Information Engineering

Department of Electronics, Information and Bioengineering

Atrial Fibrillation detection in PPG signals

Supervisors:

Prof. Luca Mainardi

Prof. Valentina Corino

Author:

Federico Salibra

ID: 884651

Academic year: 2018-2019

Ai miei genitori, a mia sorella e alla mia famiglia tutta, grazie ai quali sono la persona che sono e ai quali posso dire solo un sincero grazie

Ringraziamenti

Vorrei ringraziare sentitamente il Professore Luca Mainardi, relatore di questa tesi, per la disponibilità che mi ha concesso, per la professionalità che ha dimostrato e per il supporto dato in questo momento cruciale della mia carriera universitaria.

Desidero inoltre ringraziare la correlatrice, la Professoressa Valentina Corino, per la sua disponibilità e per tutti i consigli e l'aiuto datomi in questi mesi.

Contents

List of Figures **v**

List of Tables **vii**

Abstract **viii**

Sommario **xvii**

1 Introduction **1**

1.1 Signals 6

1.1.1 Electrocardiogram 7

1.1.2 Intracardiac Electrogram 10

1.1.3 Implantable Loop Recorders 12

1.1.4 Photoplethysmogram 13

1.2 AF detection algorithms 16

2 Materials and methods **18**

2.1 Database 19

2.1.1 Simulated database 19
2.1.2 Clinical database 23
2.2 Peak detection and parameters computation 24
2.2.1 Time domain heart rate parameters 25
2.2.2 Entropy 27
2.2.3 Shape analysis 28
2.2.4 Poincarè plot 29
2.2.5 Turning Point Ratio 31
2.3 Classification 32
2.4 AF Detection algorithms 33
2.4.1 Lee et al. 2012 34
2.4.2 Chong et al. 2015 35
2.4.3 Petrenas et al. 2015 37

3 Results **41**
3.1 Characterization of AF vs. NSR 42
3.2 Classification 49
3.2.1 Simulated data 49
3.2.2 Clinical data 51
3.3 Comparison of AF Algorithms 53

4 Discussion and conclusions **56**
4.1 Discussion and conclusions 57
4.2 Limitation of the study and future developments 58

Bibliografy **60**

List of Figures

1 Colormaps of t-test results xiii

2 Validation and test accuracy for classification algorithm . . xiv

3 Mappe dei colori per risultati dei t-test xxii

4 Accuratezza di validazione e test dell'algoritmo di classificazione xxiv

1.1 Example of different arrythmias 3

1.2 Comparison ECG of a normal heart rhythm and AF 5

1.3 Electrocardiogram waveform 8

1.4 Pulse wave form of PPG 14

1.5 Photoplethysmography 15

2.1 Graphical user interface of PPG simulator 22

2.2 Empatica E4 wristband 23

2.3 Poincarè plot in Chong et al. 2015 [9] 30

2.4 Poincarè plot 31

2.5 Camera placement for the pulsatile acquisition 34
2.6 Flowchart of Chong et al. 2015 36
2.7 Flowchart of Petrenas et al. 2015 37
2.8 Output of each building block of Petrenas et al. 2015 . . . 40

3.1 Shannon Entropy, Sample Entropy, rMSSD and Lorenz Plot SD comparison 43
3.2 Similarity indexes comparison 44
3.3 pNN comparison 45
3.4 *p-value* t-test signals with RR length 50 46
3.5 *p-value* t-test signals with RR length 300 47
3.6 *p-value* t-test signals with SNR=0 48
3.7 *p-value* t-test signals with SNR=30 48
3.8 Validation and test accuracy for classification algorithm . . 50
3.9 TPR and TNR of classification for different lengths of signal: 50, 100 and 300 52
3.10 Validation accuracy and test accuracy of classification for different number of features in clinical data 52
3.11 TPR and TNR of classification for different number of features in clinical data 53

List of Tables

1 Accuracym, TPR and TNR computed for each algorithm, for signals with RR length=300 and different SNR xv
2 Accuracy, TPR and TNR for each algorithm, for real PPG data xv

3 Accuratezza, TPR e TNR per segnali con lunghezza RR=300 e diversi valori di SNR xxiii
4 Accuratezza, TPR e TNR per ogni algoritmo per i dati clinicixxv

3.1 Accuracym, TPR and TNR computed for each algorithm, for signals with RR length=100 and different SNR 54
3.2 Accuracym, TPR and TNR computed for each algorithm, for signals with RR length=300 and different SNR 55
3.3 Accuracy, TPR and TNR for each algorithm, for real signals 55

Abstract

Introduction

Cardiac arrhythmias are the abnormalities or perturbations in the normal beating of heart. A cardiac arrhythmia is characterized by irregular rhythm of heartbeat which could be either too slow (<60 beats/min) or too fast (>100 beats/min) and can happen at any age. The sinus node sends a depolarization wave over the atrium and depolarizing atrioventricular (AV) node propagating over His-Purkinje system and depolarizes ventricle in systematic way.

The symptoms of cardiac arrhythmias are complaints of dizziness, palpitations, fast heart beating, and feeling of weakness [1]. Fatigue and shortness of breath are more often seen in elderly people whereas chest pain and palpitations are particularly common in younger patients [2]. Atrial fibrillation (AF) is often detected in asymptomatic patients, and the arrhythmia may become asymptomatic over time or after treatment [3].

AF is the most common clinically significant cardiac arrhythmia. In 2014 AF affected approximately 33 million people. It occurs when ectopic foci in the atria suppress or replace the normal sinus mechanism [2]. AF is responsible for a high rate of cardiovascular and cerebrovascular morbidity and mortality, resulting in a high health care cost and public health burden [4]. Common techniques for AF diagnosis are based on electrocardiogram (ECG), and can be either in-hospital or ambulatory monitoring, such as 24/48 hours Holter monitors, or event recorders triggered automatically or by the user when AF symptoms are noticed.

Photoplethysmography is a simple optical technology: it uses one or more light-emitting diodes to illuminate tissue with different wavelengths. The intensity of the non absorbed light at each wavelength is measured by a photo-diode. The light absorption and transmission depends on the traveled light path, optical density of the tissue, volume of blood present in the tissue, and blood composition.

The Pulse Wave Form (PWF) of a photoplethysmogram (PPG) consists of a pulsatile (ac) waveform, which is superimposed on a slowly varying (dc) baseline with lower frequency components. The systolic phase starts with a valley that marks the pulse wave begin (PWB) and ends with the pulse wave systolic peak (PWSP). During the diastolic phase, the small and downward deflection observed (Dicroctic Notch) corresponds to the closure of the aortic or pulmonic valve [5]. Technology of photoplethismography allows a noninvasive measurement of variations of the blood volume

over time. PPG has widespread clinical application, with the technology utilized in commercially available medical devices.

The PPG is accepted as a reliable source for measurements of peripheral arterial oxygen saturation (SpO_2) and pulse rate. Due to its technological and practical advantages, PPG is becoming increasingly popular in wrist-worn devices for pulse rate detection.

The aim of the work is evaluate the potential of a set of PPG-derived measures to discriminate between AF and NSR, and compare the new algorithm with other published algorithm for the AF detection.

Materials and Methods

The starting point of this study is the creation of the database of simulated PPG signals with AF events which is used to compare the performances of three algorithms of AF detection published in literature and one developed at Politecnico di Milano, which is also validated in this thesis. In addition to simulated, real PPG data have been analyzed. The simulated database was created with the PPG simulator proposed by Solosenko [6]. PPG signals affected by AF and without AF have been generated. The database is composed by several signals which differ each other in terms of length, signal to noise ratio (SNR) and percentage of AF, it is possible to set the AF burden value, in order to generate signals completely in AF (AF burden=1) or signals in normal sinus rhythm (NSR) (AF burden=0).

The different length used in this study, in terms of number of RR interval, are 50, 100, 150, 200, 250, 300 beats. The different SNR taken into account are 0, 1, 5, 10, 15, 20, 25 and 30 dB. The simulator is based on a novel phenomenological model for simulating PPG signals in AF conditions. The model uses RR interval series as input for generating a PPG signal.

Real PPG data are recorded from 52 patients admitted to the Ospedale Maggiore Policlinico in Milan, Italy. 23 patients had persistent AF and 29 were healthy subjects (NSR group). All recordings were performed with the subject in a supine position, at rest. The subject was asked to stay as still as possible to reduce motion artifacts. The Empatica E4 wristband was applied on the wrist of the non-dominant arm, with the main part of the device facing upward, in similar way to a regular wrist watch. Two minutes recording was acquired for each subject.

An algorithm for the detection of peaks and valleys is performed on each signal, in this way pulse-to-pulse intervals (PPi) are computed in order to study the signals. For each signal, 24 parameters have been computed and analyzed. Parameters studied are: PPi mean, PPi standard deviation, PPi rMSSD, PPi nrMSSD, ΔPPi standard deviation,Shannon entropy, sample entropy, five indexes of similarity between each pair of PPG pulses, pNN10, pNN20, pNN30, pNN40, pNN50, pNN60, pNN70, pNN80, pNN90, pNN100, Poincarè plot and Lorenz plot. After the computation of the cited parameters, a t-test is performed to determine if there is a significant difference between the two groups (AF and NSR).

An algorithm for the classification is performed in order to classify the signals in the database as AF or NSR. In this study has been used linear SVM for classification. For feature selection, a sequential forward floating search (SFFS) algorithm [7] is used to identify the best subset of features differentiating the two conditions. The quality of each model is assessed through train-validation-test split (55% of the data is used as training set, 25% as validation set, 20% as test set), repeated 100 times (bootstrapping). Training and validation sets are used to choose the best model, the test set to test the model on never-seen data.

The performances of the new algorithm have been compared with three published algorithms: Lee et al. 2012 [8], Chong et al. 2015 [9] and Petrenas et al. 2015 [10].

Results

The parameters obtained are studied for signals in AF and signals in NSR. Boxplots are used to display the distribution of data. For signals with RR length=300 and SNR=0, SampEn, rMSSD, Lorenz plot standard deviation, S1, S2, pNN10, pNN20, pNN30, pNN40, pNN50, pNN60, pNN80, pNN90, pNN100 values are significantly different ($p<0.05$) between NSR and AF. For signals with RR length=300 and SNR=30, all the parameters values are significantly different ($p<0.05$) between NSR and AF.

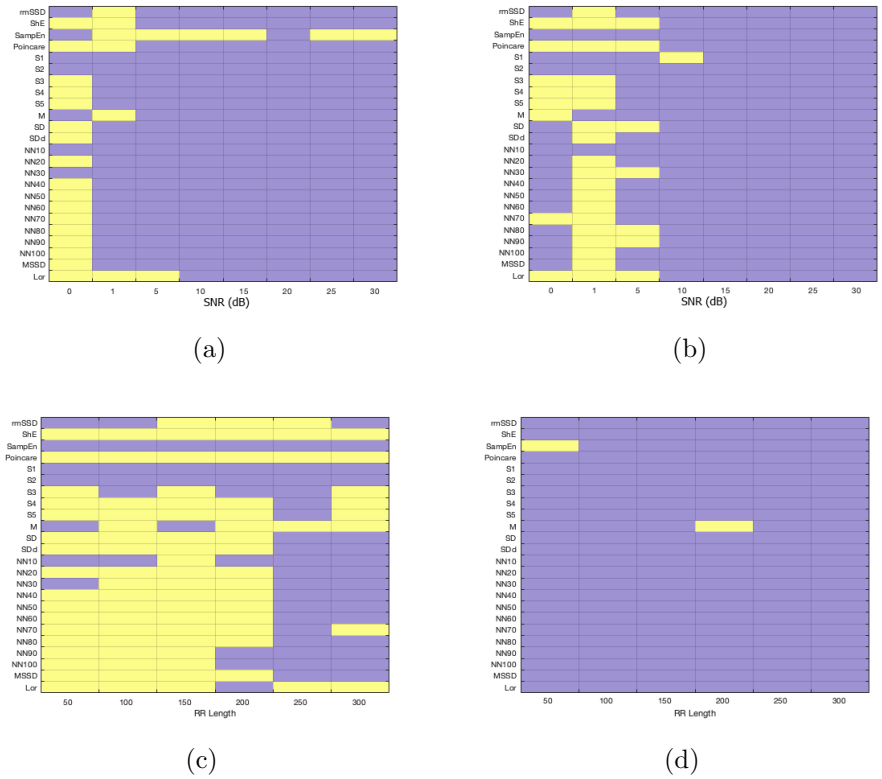
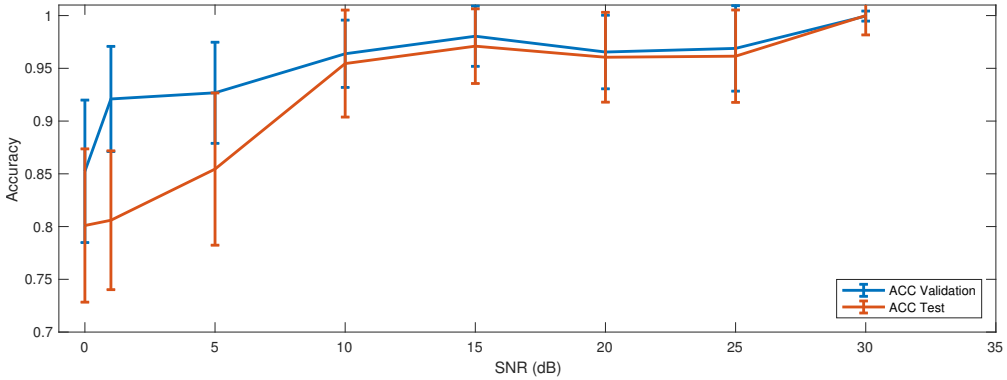


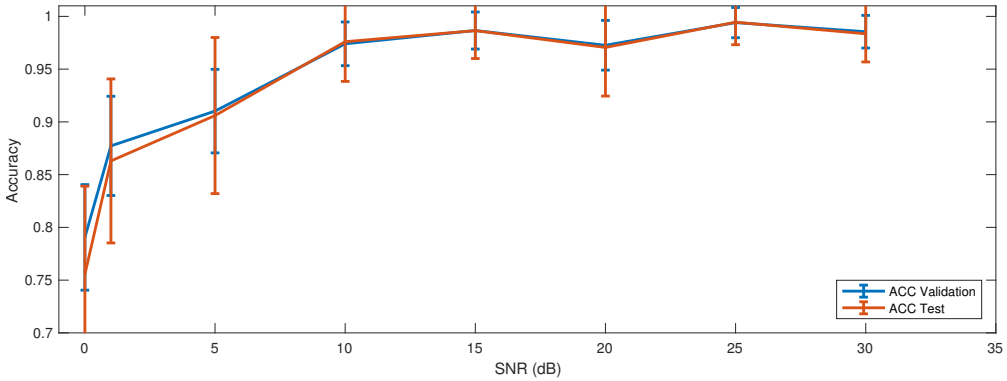
Figure 1: Colormaps of t-test results for signals with RR length=50 (a), RR length=300 (b), SNR=0 (c) and SNR=30 (d).

T-test results for signals with RR length=50 and RR length=300 and for signals with SNR=0 and SNR=30 are shown in Figure 1 where if p-$value$<0.05 the block is purple, otherwise is yellow.

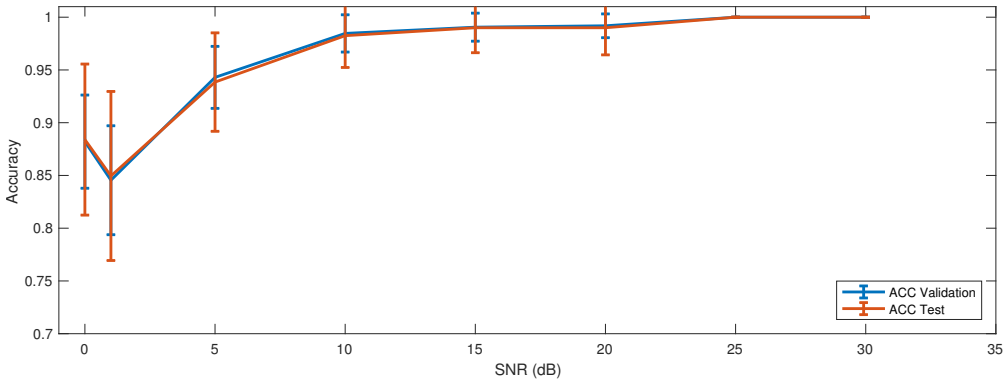
In Figure 2 the validation and test accuracy of the classification algorithm in signals with RR length=50,100 and 300 in function of SNR is shown. From SNR≥10 both values are almost equal for each length. Best level of accuracy are find in SNR=30 for RR length=50 and 100, and in SNR=20 for RR length=100.



(a)



(b)



(c)

Figure 2: Validation and test accuracy for classification algorithm in signals with RR length=50 (a), RR length=100 (b) and RR length=300

SNR	Accuracy				TPR				TNR			
	A	P	L	C	A	P	L	C	A	P	L	C
0	0.884	0.5	0.5	0.5	0.82	1	1	1	0.9237	0	0	0
1	0.8495	0.5	0.5	0.5	0.808	1	1	1	0.8730	0	0	0
5	0.9385	0.5	0.5	0.505	0.971	1	1	1	0.9130	0	0	0.01
10	0.9825	0.52	0.505	0.565	0.984	1	0.99	0.99	0.9774	0.04	0.02	0.14
15	0.99	0.75	0.7	0.76	0.983	1	1	1	0.9907	0.5	0.4	0.52
20	0.99	0.86	0.895	0.88	0.99	1	1	0.99	0.9926	0.72	0.79	0.77
25	1	0.88	0.94	0.91	1	1	1	0.98	1	0.76	0.88	0.84
30	1	0.945	0.975	0.945	1	1	1	0.98	1	0.89	0.95	0.91

Table 1: Accuracy, TPR and TNR for signals with RR length=300 and different SNR computed for the classification algorithm (A), Petrenas (P), Lee (L) and Chong (C).

The results for the real PPG data, in which the accuracy is computed for different number of features used in the classification, show that the validation accuracy is maximum with 2 featurs. The TPR is maximum when the features considered are 4.

The comparisons with the different algorithm taken into account is shown in Table 1 for simulated signals with RR length=300, and in Table 2 for the real PPG data.

	Accuracy	**TPR**	**TNR**
Classification A	0.958	0.995	0.913
Petrenas	0.673	1	0.414
Lee	0.577	1	0.241
Chong	0.596	0.957	0.310

Table 2: Accuracy, TPR and TNR for each algorithm, for real PPG data

Discussion and Conclusions

The PPG signal, considering the large quantity of devices which are able to record and thanks to its easy recording method, can give an important help in the screening and detection of AF and other arrythmias. This study aims at evaluating the potential of a set of PPG-derived measures to discriminate between AF and NSR. In signal with high value of SNR the distribution of the parameters is different in AF and NSR, quite the opposite, with low values of SNR the distributions of the parameters are more similar.

The classification algorithm works shows good performances in general, but the best results are obtained with values of SNR$\geq$10 db. Noise, in particular motion artifacts, really affects the detection of AF episodes. Longer signals are better classified, with higher value of accuracy and TPR. The new algorithm works better than the others algorithms even with low values of SNR, discriminating well the two rhythms and reducing the number of false alarms; particularly where classification of the NSR is considered.

Sommario

Introduzione

Le aritmie cardiache sono anomalie o perturbazioni nella normale attivazione o nel battito cardiaco. L'aritmia cardiaca è caratterizzata da un ritmo irregolare del battito cardiaco che potrebbe essere troppo lento (<60 battiti/min) o troppo veloce (>100 battiti/min) e può verificarsi a qualsiasi età. Il nodo senoatriale invia un'onda di depolarizzazione all'atrio e depolarizza il nodo atrioventricolare (AV) che si propaga sul sistema His-Purkinje e depolarizza il ventricolo in modo sistematico.

I sintomi delle aritmie cardiache sono disturbi di vertigini, palpitazioni, battito cardiaco accelerato e sensazione di debolezza [1]. La fatica e la mancanza di respiro sono più frequenti nelle persone anziane, mentre il dolore toracico e le palpitazioni sono particolarmente comuni nei pazienti più giovani [2]. La fibrillazione atriale (FA) viene spesso rilevata nei pazienti asintomatici e l'aritmia può diventare asintomatica nel tempo o dopo il trattamento [3].

La FA è l'aritmia cardiaca clinicamente significativa più comune. Nel 2014 la FA ha colpito circa 33 milioni di persone. Si verifica quando un pattern diffuso e caotico di attività elettrica negli atri sopprime o sostituisce il normale meccanismo del seno [2]. La FA è responsabile di un alto tasso di stato patologico cardiovascolare e cerebrovascolare e mortalità, con conseguenti costi sanitari elevati e per la salute pubblica [4]. Le tecniche comuni per la diagnosi della FA si basano sull'elettrocardiogramma (ECG) e possono essere sia ambulatoriali che domiciliari, come i dispositivi Holter indossati per 24-48 ore o i registratori di eventi attivati automaticamente o dall'utente quando si notano i sintomi della FA.

La fotopletismografia è una semplice tecnologia ottica, che utilizza uno o più diodi emettitori di luce per illuminare tessuti con lunghezze d'onda diverse. L'intensità della luce non assorbita ad ogni lunghezza d'onda è misurata da un foto-diodo. L'assorbimento e la trasmissione della luce dipendono dal percorso della luce percorsa, dalla densità ottica del tessuto, dal volume di sangue presente nel tessuto e dalla composizione del sangue.

La forma d'onda del PPG (PWF) è costituita da una forma d'onda pulsatile (ac), che si sovrappone a una linea di base che varia lentamente (dc) con componenti a frequenza più bassa. La fase sistolica inizia con una valle che segna l'inizio dell'onda del polso (PWB) e termina con il picco sistolico dell'onda del polso (PWSP). Durante la fase diastolica, la deflessione piccola e verso il basso osservata (Dicroctic Notch) corrisponde alla chiusura della valvola aortica o polmonare [5]. La tecnologia della fotopletismogra-

fia consente una misurazione non invasiva delle variazioni del volume del sangue nel tempo. Il segnale PPG ha un'applicazione clinica diffusa, viene utilizzato in dispositivi medici disponibili in commercio.

Il PPG è accettato come una fonte affidabile per le misurazioni della saturazione di ossigeno arteriosa periferica e della frequenza del polso. Grazie ai suoi vantaggi tecnologici e pratici, il segnale PPG sta diventando sempre più popolare nei dispositivi da polso per il rilevamento della frequenza cardiaca.

Lo scopo di questo lavoro è quello di valutare il potenziale di un gruppo di misure derivate dal PPG per discriminare FA e NSR, e di fare un confronto tra il nuovo algoritmo e altri algoritmi pubblicati per la detezione della FA.

Materiali e metodi

Il punto di partenza dello studio è la creazione del database di segnali PPG con FA simulati, utilizzato per confrontare le performances di tre algoritmi per la detezione della FA e l'algoritmo di classificazione sviluppato al Politecnico di Milano. Sono stati analizzati sia dati simulati sia dati reali. Il database simulato è stato creato con il simulatore PPG proposto da Solosenko [6]. Sono stati generati segnali PPG affetti da FA e senza FA. Il database è composto da diversi segnali che variano in termini di lunghezza, rapporto segnale rumore (SNR) e percentuale di FA, è possi-

bile modificare il parametro *AF burden* in modo tale da generare segnali completamente in FA (AF burden=1) o segnali in NSR (AF burden=0). Le diverse lunghezze utilizzate in questo studio, in termini di numero di intervalli RR, sono 50, 100, 150, 200, 250, 300 battiti. I diversi SNR presi in considerazione sono 0, 1, 5, 10, 15, 20, 25 e 30 dB. Il simulatore si basa su un nuovo modello fenomenologico per simulare segnali PPG in condizioni di FA. Il modello utilizza la serie di intervalli RR come input per la generazione del segnale PPG.

Per i dati clinici, i segnali PPG sono stati registrati da 52 pazienti ricoverati all'Ospedale Maggiore Policlinico di Milano, Italia. 23 pazienti presentavano FA persistente e 29 erano soggetti sani (gruppo NSR). Tutte le registrazioni sono state eseguite con il soggetto in posizione supina, a riposo. È stato chiesto al soggetto di rimanere il più fermo possibile per ridurre gli artefatti da movimento. Il cinturino Empathic E4 è stato applicato sul polso del braccio non dominante, con la parte principale del dispositivo rivolta verso l'alto, in modo simile a un normale orologio da polso. È stata acquisita una registrazione di due minuti per ciascun soggetto.

Un algoritmo per il rilevamento di picchi e valli è eseguito su ciascun segnale, vengono calcolati gli intervalli da battito a battito (PPi) per studiare i segnali. Per ogni segnale, sono stati calcolati e analizzati 24 parametri. I parametri studiati sono: media del PPi, deviazione standard delPPi, PPi rMSSD, PPi nrMSSD, deviazione standard di *Delta*PPi, entropia di Shannon, sample entropy, cinque indici di similarità tra ogni coppia di bat-

titi, pNN10, pNN20, pNN30, pNN40, pNN50 , pNN60, pNN70, pNN80, pNN90, pNN100, Poincarè plot e Lorenz plot. Dopo il calcolo dei parametri citati, viene eseguito un t-test per determinare se esiste una differenza significativa tra i due gruppi (FA e NSR).

Un algoritmo per la classificazione viene utilizzato al fine di classificare i segnali nel database come FA o NSR. In questo studio sono state utilizzate SVM lineari per la classificazione. Per la selezione dei parametri, un algoritmo di ricerca sequenziale in avanti (SFFS) [7] viene utilizzato per identificare il miglior sottoinsieme di parametri che differenziano le due condizioni. La qualità di ciascun modello viene valutata mediante suddivisione addestramento-validazione-test (il 55% dei dati viene utilizzato come set di addestramento, il 25% come set di convalida, il 20% come set di test), ripetuto 100 volte (bootstrap). I set di training e validazione vengono utilizzati per scegliere il modello migliore.

Le performance del nuovo algoritmo di classificazione vengono confrontate con tre diversi algoritmi pubblicati: Lee et al. 2012 [8], Chong et al. 2015 [9] e Petrenas et al. 2015 [10].

Risultati

I parametri ottenuti sono studiati per segnali affetti da FA e segnali in ritmo sinusale (NSR). I boxplot vengono utilizzati per visualizzare la distribuzione dei dati. Per segnali con lunghezza RR=300 e SNR=0, Sam-

pEn, rMSSD, deviazione standard del Lorenz plot, S1, S2, pNN10, pNN20, pNN30, pNN40, pNN50, pNN60 , pNN80, pNN90, pNN100 sono significativamente diversi ($p<0.05$) tra NSR e FA. Per segnali con lunghezza RR=300 e SNR=30, tutti i valori dei parametri sono significativamente diversi ($p<0.05$) tra NSR e FA.

I risultati del t-test per segnali con lunghezza RR=50 e lunghezza RR=300 e per segnali con SNR=0 e SNR=30 sono mostrati nella Figura 3 dove se il *p-value*<0.05 il blocco è viola, altrimenti è giallo.

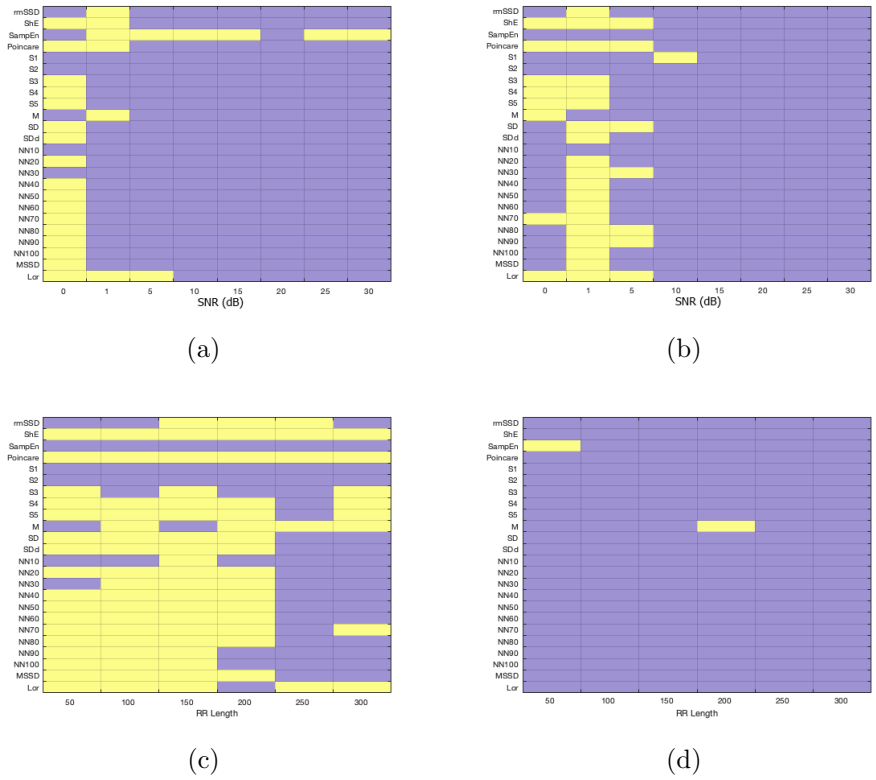


Figura 3: Mappe dei colori per risultati dei t-test per segnali con lunghezza RR=50 (a), 300 (b), e con SNR=0 (c) and 30 (d).

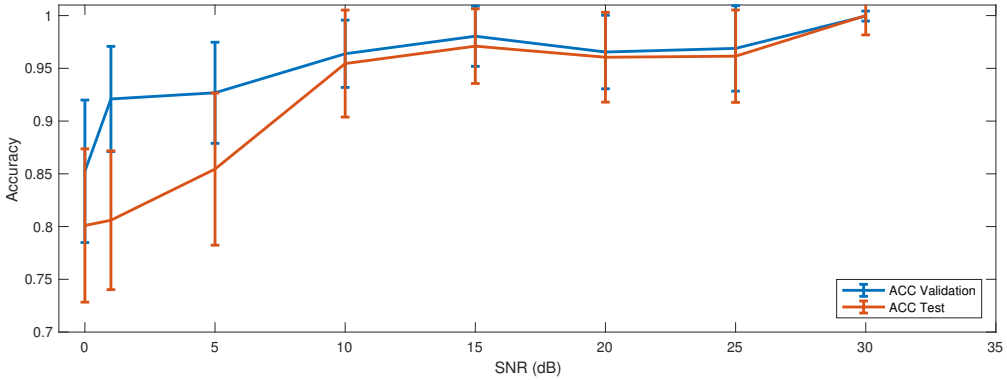
SNR	Accuratezza				TPR				TNR			
	A	P	L	C	A	P	L	C	A	P	L	C
0	0.884	0.5	0.5	0.5	0.82	1	1	1	0.9237	0	0	0
1	0.8495	0.5	0.5	0.5	0.808	1	1	1	0.8730	0	0	0
5	0.9385	0.5	0.5	0.505	0.971	1	1	1	0.9130	0	0	0.01
10	0.9825	0.52	0.505	0.565	0.984	1	0.99	0.99	0.9774	0.04	0.02	0.14
15	0.99	0.75	0.7	0.76	0.983	1	1	1	0.9907	0.5	0.4	0.52
20	0.99	0.86	0.895	0.88	0.99	1	1	0.99	0.9926	0.72	0.79	0.77
25	1	0.88	0.94	0.91	1	1	1	0.98	1	0.76	0.88	0.84
30	1	0.945	0.975	0.945	1	1	1	0.98	1	0.89	0.95	0.91

Tabella 3: Accuratezza, TPR e TNR per segnali con lunghezza RR=300 e diversi valori di SNR calcolate per l'algoritmo di classificazione (A), Petrenas (P), Lee(L) e Chong(C).

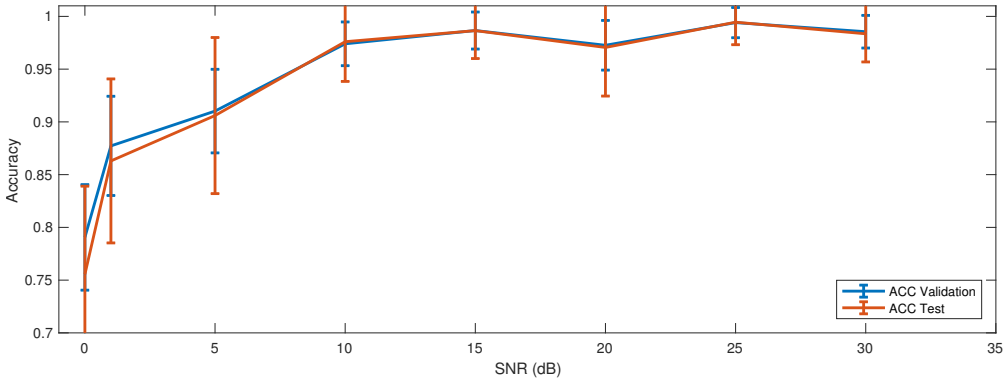
Nella figura 4 è mostrata l'accuratezza di validazione e test dell'algoritmo di classificazione nei segnali con lunghezza RR=50, 100 e 300 in funzione del SNR. Per SNR$\geq$10 entrambi i valori sono quasi uguali per ogni lunghezza. Il miglior valore di accuratezza è in SNR=30 per lunghezza RR=50 e 100 e in SNR=20 per lunghezza RR=100.

I risultati per i dati clinici, in cui viene calcolata l'accuratezza per diversi numeri di parametri utilizzati nella classificazione, mostrano che l'accuratezza della validazione è massima con 2 parametri. Il TPR è massimo quando i parametri considerati sono 4.

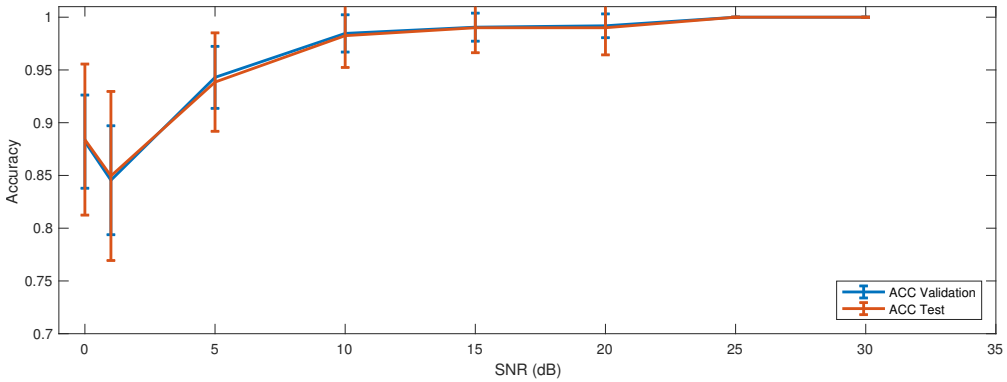
I confronti con i diversi algoritmi presi in considerazione sono mostrati nella Tabella 3 per segnali simulati con lunghezza RR=300 e nella Tabella 4 per i dati clinici.



(a)



(b)



(c)

Figura 4: Accuratezza di validazione e test dell'algoritmo di classificazione per segnali con lunghezza RR=50 (a), 100 (b) e 300 (c)

	Accuratezza	TPR	TNR
Classification A	0.958	0.995	0.913
Petrenas	0.673	1	0.414
Lee	0.577	1	0.241
Chong	0.596	0.957	0.310

Tabella 4: Accuratezza, TPR e TNR per ogni algoritmo per i dati clinici

Discussione e conclusioni

Considerando la grande quantità di dispositivi che sono in grado di registrare e grazie al suo semplice metodo di registrazione, il segnale PPG può fornire un aiuto importante nello screening e nel rilevamento di AF e altre aritmie. Questo studio mira a valutare il potenziale di una serie di misure derivate dal PPG per discriminare tra AF e NSR. I risultati dimostrano che in segnali con alto valore di SNR la distribuzione dei parametri è differente in AF e NSR, al contrario, con valori bassi di SNR le distribuzioni dei parametri sono più simili.

L'algoritmo di classificazione mostra buone performaces, i migliori risultati si ottengono con valori di SNR$\geq$10 db. Il rumore, in particolare gli artefatti causati dai movimenti, influenza molto il rilevamento di episodi AF. I segnali più lunghi sono meglio classificati, con un valore più alto di precisione e TPR. Il nuovo algoritmo di classificazione funziona meglio degli altri algoritmi considerati anche con valori bassi di SNR, discriminando bene i due ritmi e riducendo il numero di falsi allarmi; questi risultati sono ottenuti in particolare quando la classificazione del NSR viene considerata.

Chapter 1

Introduction

This chapter shows a global view on atrial fibrillation, its epidemiology and the methods to diagnose it. Signals that are involved in diagnosis are described, in particular photopletysmogram.

Atrial Fibrillation

Cardiac arrhythmias are the abnormalities or perturbations in the normal activation or beating of heart myocardium. A cardiac arrhythmia is characterized by irregular rhythm of heartbeat which could be either too slow (<60 beats/min) or too fast (>100 beats/min) and can happen at any age. The sinus node sends a depolarization wave over the atrium and depolarizing atrioventricular (AV) node propagating over His-Purkinje system and depolarizes ventricle in systematic way.

There are many types of cardiac arrhythmias or abnormal heart beating. The normal rhythm of heart is called as sinus rhythm (NSR) which can be disturbed through failure of automaticity as a sick sinus syndrome (SSS) or as an inappropriate sinus tachycardia. The seriousness of cardiac arrhythmias depends on the presence or absence of structural heart disease [1]. The symptoms of cardiac arrhythmias are complaints of dizziness, palpitations, fast heart beating, and feeling of weakness [1]. Fatigue and shortness of breath are more often seen in elderly persons whereas chest pain and palpitations are particularly common in younger patients [2].

The pathogenesis of cardiac arrhythmias has three basic mechanisms: the enhanced or suppressed automaticity, triggered activity, or re-entry. The automaticity is a natural property of all myocytes. The various factors which may suppress or enhance automaticity are heart ischemia, scarring, electrolyte disturbance, heart medications, old age, and other factors. The

Figure 1.1: Example of electrocardiograms of different arrythmias. From top to bottm: sinus rhythm, nodal rhythm, heart block, atrial flutter, atrial fibrillation and ventricular fibrillation.

suppression of automaticity of the sinoatrial (SA) node can result in sinus node dysfunction and in SSS which is a common indication for permanent pacemaker implantation. The enhanced automaticity can result in multiple arrhythmias for both atrial and ventricular arrhythmias. The triggered activity which usually occurs after earlier and delayed depolarizations initiates spontaneous multiple depolarizations, and results in precipitating ventricular arrhythmias. The re-entry is a common mechanism of arrhythmogenesis which includes bidirectional conduction and unidirectional block. The micro level re-entry occurs with VT from conduction around the scar of myocardial infarction (MI), and macro level re-entry occurs via conduction through Wolff–Parkinson–White syndrome concealed accessory pathways [1].

Atrial fibrillation (AF) is the most common clinically significant cardiac arrhythmia. It occurs when a diffuse and chaotic pattern of electrical activity in the atria suppresses or replaces the normal sinus mechanism [2]. AF is responsible for a high rate of cardiovascular and cerebrovascular morbidity and mortality, resulting in a high health care cost and public health burden [4], this condition accounts for 1% of the National Health Service budget in the United Kingdom and $16–26 billion of annual US expenses. In Denmark the cost attributable to AF in the study [11] was estimated to be €73–98 million for year and in Italy the estimated mean costs per patient per year is €613-1213 [12]. Atrial fibrillation is often detected in asymptomatic patients, and the arrhythmia may become asymptomatic over time or after treatment [3].

In 2014 AF affected approximately 33 million people. Australia, Europe, and the USA had highest reported prevalence of AF (1% in the adult population), but the prevalence of AF in low-income and middle-income countries was probably underestimated [13].The worldwide age-adjusted prevalence of AF, as estimated in the 2010 Global Burden of Disease Study[14], is 596 per 100,000 men and 373 per 100,000 women. Approximately 3–5 million individuals in the USA have AF. The frequency of AF in the general population is progressively increasing as a result of greater life expectancy, increased prevalence of risk factors for AF, and improved survival after myocardial infarction [4]. With population ageing, AF is expected to affect more than 8 million people in the USA by 2050, and in Europe, AF is projected to increase from the current estimated prevalence of 8.8 million to approximately 18 million in 2060 [13]. The reasons of this

Figure 1.2: Comparison ECG of a normal heart rhythm and AF; (a) ECG of a normal heart rhythm. (b) ECG of a AF episode, it is possible to visualize the hallmarks of AF, that are: irregular rhythm, absence of distinct P-waves, presence of f-waves and rapid ventricular rhythm

increase are attributed to the increasing life expectancy worldwide and its now well established the fact that the elder the patient, the greater the risk to develop AF [15].

It is believed that the natural history of AF is progressive, initially being nonsustained and induced by trigger activity (paroxysmal AF). However, these bouts of AF induce electric alteration within the atrial myocardium (electrical remodelling), which may also promote or accelerate myocardial apoptosis and fibrosis (anatomical remodelling). Eventually, the alteration of the atrial myocardial substrate contributes to maintain AF (persistent AF) by means of complex self-sustained electrical activity [16], therefore it is important an early detection of the pathology.

Common techniques for AF diagnosis are based on electrocardiogram (ECG), and can be either in-hospital or ambulatory monitoring, such as 24/48 hours Holter monitors, or event recorders triggered automatically or by the user when AF symptoms are noticed.

From an electrocardiographic point of view, AF is characterized by the absence of P waves. The P waves are replaced by absolutely uneven small amplitude deflections (f-waves); these are followed by QRS complexes that follow each other at different intervals [17]. However, these methods suffer from several drawbacks. For paroxysmal AF, the arrhythmia episodes are irregular and might not occur during the measurement period (24/48 hour Holter monitoring only detect AF in 30% to 60% of the cases [18]). If worn for longer periods, the electrodes can easily become uncomfortable and possibly cause skin irritations. Monitoring tool obtrusiveness and high cost can also lead to low patient acceptance rate. Since AF patients experience an irregular flow in the blood vessels photoplethysmography can be used as an alternative method for ECG when estimating heart rate (HR) and heart rate variability (HRV) [19].

1.1 Signals

The signals that may be used for AF detection are

- Electrocardiogram (ECG) [20] [21] [22] [18] [23];
- Intracardiac Electrogram (IEGM) [24];
- Implantable Loop Recorders (ILR) [25];
- Photoplethysmogram (PPG) [10], [26], [8], [9].

1.1.1 Electrocardiogram

The surface ECG characteristics are a direct reflection of pathophysiologic events in the atrium and can be used in studying AF. By analyzing the ECG through a variety of signal processing techniques it is possible to find clinically useful information that can be used to better understand and treat AF. Some of the advantages of using the ECG are the ability to record data for a long time period, the minimal risks involved compared with invasive electrophysiologic study, and the ECG's reflection of the global activity in the atria and ventricles during AF [27]. The ECG it is a noninvasive study which measures the electrical currents or impulses that the heart generates during a cardiac cycle. At every beat, the heart is depolarized to trigger its contraction. This electrical activity is transmitted throughout the body and can be picked up on the skin with electrodes. This is the principle behind the ECG. An ECG machine records this activity via electrodes on the skin and displays it graphically. The ECG machine processes the signals picked up from the skin and produces a graphical representation of the electrical activity of the patient's heart. The basic pattern of the ECG is logical:

- electrical activity towards a lead causes an upward deflection;
- electrical activity away from a lead causes a downward deflection;
- depolarization and repolarization deflections occur in opposite directions-

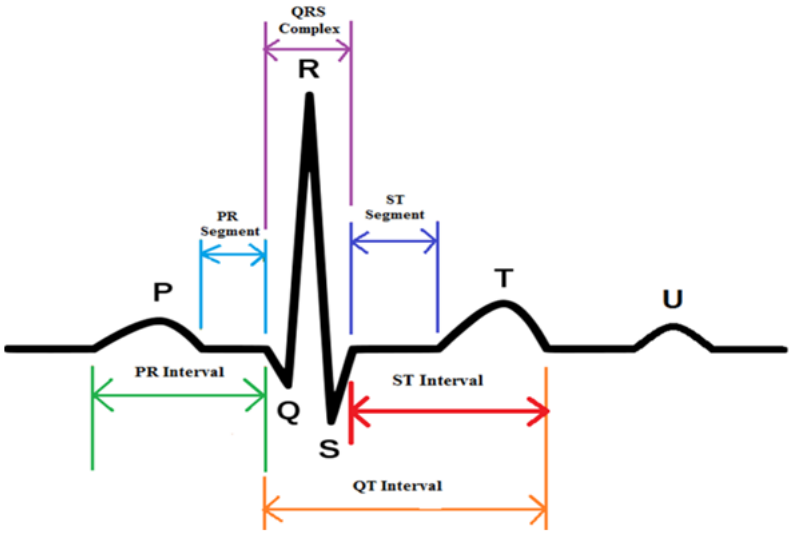


Figure 1.3: ECG waveform with characteristic waves and intervals. From left to right P wave, QRS complex, T wave and U wave.

The basic pattern of this electrical activity (Figure 1.3) comprises three waves, which are named P, QRS (a wave complex), and T. The first deflection is the P wave and represents depolarization of atrial muscle cells. It does not represent contraction of this muscle, nor does it represent firing of the sinoatrial node. The QRS complex represents depolarization of ventricular muscle cells; the Q portion is the initial downward deflection, the R portion is the initial upward deflection, and the S portion is the return to the baseline, the isoelectric point. After the depolarization, ventricular muscle repolarizes, and this event is great enough in amplitude to generate the T wave [28].

The standard 12-lead ECG is the most commonly used non-invasive tool for diagnosing electrical abnormalities, including AF. In the ECG, AF is characterized by rapid atrial activity that is irregular in timing and

morphology: discrete P waves are absent and replaced by an oscillating baseline that consists of low-amplitude fibrillatory f-waves. The characteristics of f-waves (shape, amplitude, and regularity) vary from patient to patient. The pattern of atrial activity could be similar to atrial flutter in some cases, mostly regular and with high amplitude f-waves, while in other cases it can have lower amplitude, be less regular, or both. Atrial rates detected from the ECG in AF vary between 240 and 540 beats per minute (bpm) with an average of 350 bpm, with changes in the presence of anti-fibrillatory drugs (slower rates) or acetylcholine (faster rates). During AF the ventricular heart rate results more irregular than in NSR, but less irregular than atrial activity because the atrioventricular node regulates the electrical trigger from atrial to ventricular tissue. The atrial activity is the most interesting for AF study and its waveforms are superimposed on the ventricular signals. Unfortunately, the atrial signals are much smaller (10 times lower in amplitude) than those related to ventricles (QRS complexes), thus the diagnostic applications based on P waves' analysis are compromised and computationally expensive. Furthermore, relying on body surface potential, the electrical image on the body surface is blurred with a low spatial definition in comparison to the potentials recorded inside the atria in an endocardial electrogram. There is substantial overlap between rates during atrial flutter and AF and so distinguishing between these two conditions is clinically important since the treatment may be different. During AF, the RR intervals, and hence the ventricular rate, is commonly irregular. However, ventricular activity of this sort is present not only in AF, but also in a variety of other arrhythmias, for example

multifocal atrial tachycardia, atrial flutter with variable atrioventricular (AV) block, frequent premature atrial complexes, and sinus arrhythmia. Conversely, AF may be present with a regular ventricular rate, as in the case of AV block with artificial ventricular pacing. Although irregular ventricular activity is commonly associated with AF, it often fails as the sole diagnostic criterion. In fact, the presence of AF has been shown to be under-recognized in paced patients, with important adverse clinical consequences since the RR intervals are regular [27]. The main disadvantage of the ECG is the low suitability for continuous long-term monitoring [18].

1.1.2 Intracardiac Electrogram

Intracardiac electrograms (IEGM) are recordings of cardiac potentials from electrodes directly in contact with the heart. Unipolar electrograms represent the potential difference between an "exploring electrode" in contact with the extracellular space of the active tissue, and "indifferent electrode" used as reference, which is at a distance from the heart. Ideally, this is placed at infinite distance, in practice, a chest patch, or, more commonly, the "Wilson central terminal" have been used as approximations of indifferent electrodes. Bipolar electrograms represent the potential difference between two closely spaced electrodes, and they produce a better signal-to-noise ratio than unipolar electrogram. Bipolar signals are calculated as the algebraic difference between the two unipolar electrograms at

the two sites, using the same reference [29]. IEGM records the electrical activity within the various chambers of the heart using multipolar electrodes placed inside the heart. While conventional electrocardiograms from the body surface are recorded at a paper speed of 25 mm/second, IEGMs are recorded at a much higher paper speed of 100 to 200 mm/second. EGMs allow clinicians to record the atrial activity directly in situ usually by passing catheter-guided electrodes through one of the major veins into the atria (endocardial signals). While ECG shows the electrical activity of all the heart, IEGM records the local wavefronts of depolarization and repolarization that pass below the electrode at the tip of the lead. The amplitude of the recorded electrical signal is 10 times greater than the amplitude of an ECG signal. The amplitude, frequency content and morphology of IEGMs depend on the location of the electrodes, the cardiac rhythm, posture, respiration and drugs. The amplitude of the atrial IEGM tends to be smaller and/or more spatially and temporally variable during AF than in NSR. Furthermore, the spectral coherence of IEGMs from two separate atrial sites is highly reduced during AF in comparison to other atrial rhythms. The analysis of IEGMs is used in diagnostic procedures as well as for guiding therapeutic interventions .The high level of accuracy, temporal and spatial detail is paid with the elevated invasivity of the recording system [27] [30].

1.1.3 Implantable Loop Recorders

Devices like Implanted Loop Recorders (ILRs) are typically inserted subcutaneously for ECG recording through closely spaced electrodes on their surface. Asymptomatic brady- and tachyarrhythmia events can been automatically record by ILRs and patients can manually signal symptomatic events using a handheld activator. ILRs cannot detect AF using P waves because their amplitude tends to be too small, but rely only on the analysis of RR intervals. Considering space, power and battery constraints, algorithms should be computationally simple and the recorded information must be condensed and summarized due to memory size restrictions. Furthermore, a big amount of repetitive or unimportant data can overwhelm the physician and hinder his analysis. Thus, this kind of device typically stores extremely detailed information only for a small subset of meaningful episodes, while general information is tabulated across all the episodes. In most implantable devices data is extracted during follow-up visits so the clinician can become aware of the presence of a new onset of AF after potentially 6 months from the time of the first AF occurrence. Some newer devices have wireless alert capabilities which automatically send an alarm to clinicians when an AF burden threshold is exceeded. This can dramatically improve anticoagulation therapy, reducing the risk of stroke in asymptomatic patients. ILR ECG is different from a near-field signal recorded by bipolar electrodes within the heart. The frequency content is lower like the amplitude of R waves and this, with the higher possibility of noise and artifacts, makes R wave sensing more challenging than with

intracardiac signals. On the other hand a subcutaneous device is less invasive than an intracardiac electrogram and causes less discomfort during everyday life compared to constant ECG monitoring [27] [30].

1.1.4 Photoplethysmogram

Photoplethysmography is a simple optical technology, it uses one or more light-emitting diodes to illuminate tissue with differing wavelengths. The intensity of the non absorbed light at each wavelength is measured by a photo-diode. The light absorption and transmission depends on the traveled light path, optical density of the tissue, volume of blood present in the tissue, and blood composition.

The Pulse Wave Form (PWF) of a PPG (Figure 1.4) consists of a pulsatile (ac) waveform, which is superimposed on a slowly varying (dc) baseline with lower frequency components. The ac component is attributed to cardiac synchronous changes in the blood volume with each heartbeat. The dc component is influenced by respiration, sympathetic nervous system activity, and thermoregulation. The systolic phase starts with a valley that marks the pulse wave begin (PWB) and ends with the pulse wave systolic peak (PWSP). The amplitude between this two points represents the pulse wave amplitude (PWA) whereas the time occurring between them is the Pulse Wave rise time (PWRT). During the diastolic phase, the small and downward deflection observed (Dicroctic Notch) corresponds to the closure of the aortic or pulmonic valve. A slight peak called diastolic peak

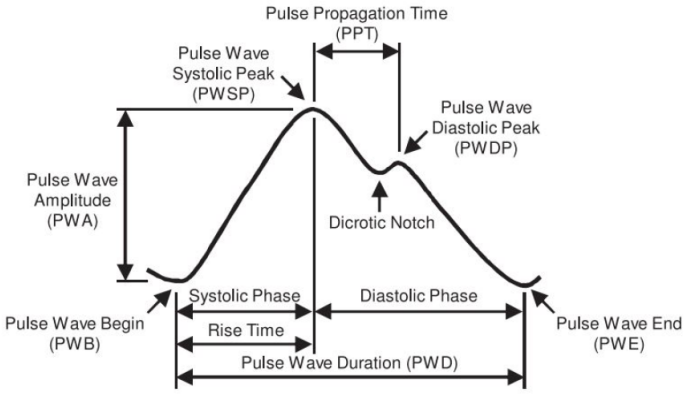


Figure 1.4: Pulse wave form of PPG with pulse wave characteristics.

(PWDP) follows the dicrotic notch. Finally, the pulse wave end (PWE) is marked by another valley at the end of the diastolic phase. Systolic phase (rise time) varies only in a narrow range inversely proportional to the heart rate compared to the pulse wave duration (PWD) [5].

Technology of photoplethismography allows a noninvasive measurement of variations of the blood volume over time. PPG has widespread clinical application, with the technology utilized in commercially available medical devices, for example in pulse oximeters, vascular diagnostics and digital beat-to-beat blood pressure measurement systems. The basic form of PPG technology requires only a few opto-electronic components (Figure 1.5): a light source to illuminate the tissue (e.g. skin), and a photodetector to measure the small variations in light intensity associated with changes in perfusion in the catchment volume. [31]

The PPG is accepted as a reliable source for measurements of peripheral arterial oxygen saturation (SpO_2) and pulse rate. Due to its technological

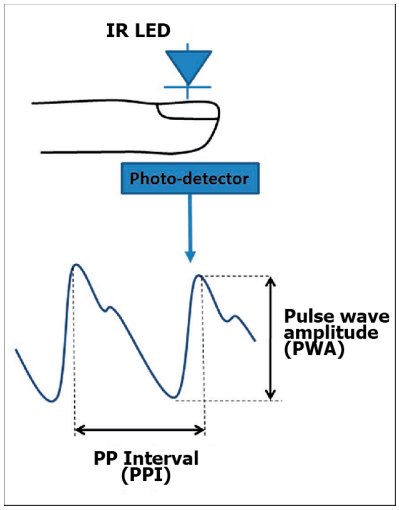


Figure 1.5: Photoplethysmography and photopletysmographic waveform. An Infrared LED illuminates the skin and the photo-detector measures changes in light absorption due to blood flow.

and practical advantages, PPG is becoming increasingly popular in wrist-worn devices for pulse rate detection. It is possible to apply a PPG sensor on the patient by using a dedicated wristband device; such device can be applied on the patient without causing discomfort even for prolonged acquisitions.

Smartphones have become ensconced into our lives, and their utility is ever expanding. Once used only for communication, smartphones have come to replace the wrist watch, provide camera and navigation functions that allow easy access to the Internet. More recently, they have become powerful tools in monitoring our health. There are smartphone-enabled glucometers, blood pressure cuffs, oximeters, and even heart monitors. This present a challenge to physicians, namely the interpretation of diagnostic

information captured on smart phones, in particular cardiac arrhythmias [32].

Currently, the methods of using a smartphone to detect and monitor AF can be divided into two groups. The first group simply uses a downloadable application and hardware that already exists on modern smartphones, the camera and lamp, which transform the smartphone in a photopletysmogram [9]. The second group uses a pair of external electrodes, either built into the case or as a stand-alone -unit that communicates with an application downloaded to the phone, like Alivecor Kardia Mobile [33].

1.2 AF detection algorithms

In literature, there are many algorithms for the detection and analysis of AF. Several algorithms have been developed to detect AF which either rely on the absence of P-waves or are based on RR variability. Since there is no uniform depolarization of the atria during AF and consequently no discernible P-waves in the ECG, their absence has been utilized in the detection of AF. However, locating the P-wave fiducial point is very difficult because the low amplitude of the wave itself makes it susceptible to corruption by noise. The methods in the second category are based on RR interval dynamics and do not require identification of the P-wave.

It is possible divide the algorithms present in literature into two major classes:

- Methods based on atrial activity analysis;
- Methods based on ventricular response analysis.

The source of AF is in the atrial cells and heart conduction pathway, consequently, methods based on atrial activity can convey significant diagnostic information. However, in many recorded cardiac signals, the presence of the ventricular activity can disturb the atrial activity, because of his much larger amplitude, and also the presence of potential noise and artifacts superimposed can affect the atrial activity. Standard 12-lead ECG may not be enough to study atrial activity, because the number of electrodes is too small and their location is not ideal for AF detection. For atrial analysis a stable, high quality signal is required. It is difficult to obtain in real-time long-term recordings.

To conclude, atrial activity analysis, even if is highly effective, is not the most suitable approach for an automatic screening application, since simplicity of use, global acceptance and robustness to noisy tracings are important requirements. Methods based on the ventricular response, instead, are intended to capture the irregular, rapid nature of AF by exploiting the most prominent feature of cardiac signals, such as QRS peaks for ECG and systolic peaks for blood pressure signals. The ventricular response can be easily detected by non-invasive devices applied on the skin, making it a useful tool for long term monitoring and automatic screening [30]. Parameters evaluated in AF detection algorithms will be discussed in the next chapter.

Chapter 2

Materials and methods

In this chapter there is a description of the analyzed data, the algorithms used for characterization and classification of AF vs. NSR.

2.1 Database

2.1.1 Simulated database

The signals studied in this work are obtained from a PPG simulator is developed by Solosenko in collaboration with the Kaunas University of technology [6]. The simulator is based on a novel phenomenological model for simulating PPG signals in AF conditions. The model uses RR interval series as input for generating a PPG signal. The model accounts for the presence of premature beats by introducing amplitude and time scale factors which modify pulse width and amplitude, thus making it possible to simulate ectopic beats and certain rhythms such as bigeminy known to cause false alarms in RR interval-based AF detection.

The model consists of two main parts, namely, modeling of a single PPG pulse and concatenation of pulses into a connected signal. A PPG pulse is modeled as a linear combination of three functions: one log-normal waveform and two Gaussians. The log-normal function is defined as:

$$\phi_1(t; m, \sigma_1) = \begin{cases} \frac{1}{t\sqrt{2\pi\sigma_1^2}} e^{\frac{-(ln(t/m))^2}{2\sigma_1^2}} & t > 0 \\ 0 & t \leq 0 \end{cases} \tag{2.1}$$

where t is the time, m is a scale parameter, and σ_1^2 is a shape parameter. The Gaussian waveform is defined by:

$$\phi_i(t, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\frac{t^2}{2\sigma_i^2}} \tag{2.2}$$

where σ_i^2 is a width parameter. Then, the PPG pulse is modeled as a linear combination of weighted, time-shifted, and time-scaled versions of$\phi_1(t; m, \sigma_1)$, $\phi_2(t, \sigma_2)$, and $\phi_3(t, \sigma_3)$ in this way:

$$\phi(t;\theta) = w_1\phi_1(t-\tau_1; m, \sigma_1) + \sum_{2}^{3} w_i\phi_i(t-\tau_i;\sigma_i) + a \tag{2.3}$$

where a denots a DC offset. θ is a vector that includes all model parameters

$$\theta = [w_1, w_2, w_3, \tau_1, \tau_2, \tau_3, m, \sigma_1, \sigma_2, \sigma_3] \tag{2.4}$$

This parameters of the PPG are estimated by nonlinear least sqares fitting,

$$J(\theta) = \int_{-\infty}^{\infty} (y(t) - \phi(t;\theta)^2 dt) \tag{2.5}$$

$$\hat{\theta} = \arg\min_{\theta} J(\theta)\theta \tag{2.6}$$

where $\hat{\theta}$ is the vector minimizing the difference between PPG pulse template $y(t)$ and the model PPG $\phi(t;\theta)$. Prior to minimization each PPG is normalized to unit amplitude. Since pulse morphology varies considerably depending on factors such as age and medical condition, a set of template PPG pulses is employed.

A PPG pulse is composed of a systolic part and diastolic part, where the width of each part depends on the adjacent RR intervals. Another factor is the PPG pulse amplitude related to ventricular filling time and, accordingly, to the length of the RR interval. For example, in sinus rhythm, the ventricular filling time does not change much from beat to beat, leading

to negligible pulse amplitude variations. On the contrary, a premature beat causes diastole to be shorter, reducing the amplitude of the resulting pulse. The model proposed by [6] assumes that the amplitude of the premature pulse changes exponentially. Since a premature beat is followed by a compensatory pause, sufficiently long to fill the ventricles with extra blood, the subsequent pulse will have larger amplitude. The amplitude of a PPG pulse is assumed to be proportional to r_k, which denotes the RR interval preceding the kth pulse, unless the beat is premature when the relationship between the length of the current RR interval and the diastolic period can be characterized by an exponential function,

$$x(n) = \sum_{i=1}^{K} s_k(n - \delta_k) + v(n) \tag{2.7}$$

where K denotes the number of pulses in the connected signal. The noise $v(n)$ is generated by filtering white noise, where the filter is determined by the spectral properties of motion artifacts extracted from PPG signals in MIMIC database.

The simulated database was created with the PPG simulator [6]. PPG signals affected by AF and without AF have been generated. The database is composed by several signals which are different each other in terms of length, signal to noise ratio (SNR) and percentage of AF, it is possible to set the AF burden value, in order to generate signals completely in AF (AF burden=1) or signals in NSR (AF burden=0). The different length used in this study, in terms of number of RR interval, are 50, 100, 150,

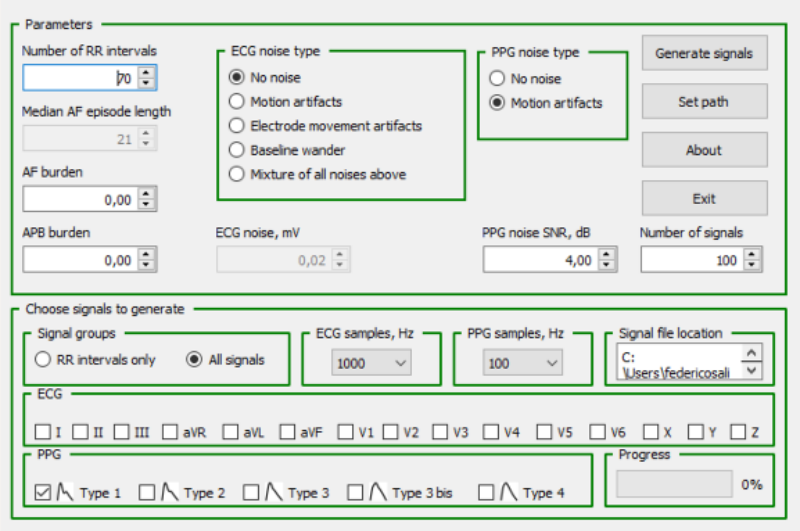


Figure 2.1: Graphical User Interface of PPG simulator [6].

200, 250, 300. The different SNR taken into account are 0, 1, 5, 10, 15, 20, 25 and 30; 0 db and 30 db are the limits set by the developer. The percentage of AF considered, for an easiest comprehension, are 0 and 100, in order to obtain signals in normal sinus rhythm and signals in atrial fibrillation rhythm. The sampling frequency is set to 100Hz. In Figure 2.1 it is shown the graphical user interface of the simulator. It is possible to set the number of RR intervals, median AF episode length, AF and APB burden, and the PPG noise SNR. To create NSR signal the parameter AF burden has to be set to 0, if this value is equal to 1, an AF signal is generated. It is also possible to generate different signals at the same time. In the lower part of the Figure 2.1 it is possible to modify the sampling frequency of the signal and set the pulse type.

2.1.2 Clinical database

The clinical database is populated by PPG signals recorded from 52 patients admitted to the Ospedale Maggiore Policlinico in Milan, Italy. 23 patients had persistent AF and 29 were healthy subjects (NSR group). All recordings were performed with the subject in a supine position, at rest. The subject was asked to stay as still as possible to reduce motion artifacts. The Empatica E4 wristband (Figure 2.2) was applied on the wrist of the non-dominant arm, with the main part of the device facing upward, in similar way to a regular wrist watch. Two minutes recording was acquired for each subject.

The Empatica E4 wristband is a wearable wireless device designed for continuous, real-time data acquisition in daily life. The device is equipped with sensors for the registration of different bio-signals: an electrodermal activity sensor, an infrared thermopile, a 3-axis accelerometer and a PPG

Figure 2.2: Empatica E4 wristband

sensor which measures the PPG signal. The PPG is sampled at 64Hz [26].

2.2 Peak detection and parameters computation

An algorithm for the detection of peaks and valleys is performed on each signal, in this way pulse-to-pulse intervals (PPi) are computed in order to study the signals. The algorithm performs a low pass averaging on the signal, useful to reduce noise and exploited in minima detection. The convoluted signal is lowest in zones which have very small values in a wide zone nearby, those points are to be identified as minima. First section of the algorithm identifies local maxima,then a second section identifies local minima on the low filtered signal. To be a minimum, a point must be smaller than its neighborhood and smaller than the signal half a beat earlier and half a beat later. For every minimum found in the low filtered signal, a search in their 0.1 seconds neighborhood for the lowest minima in the original signal (which will be considered as true minima) is performed. For each pair of minima, the algorithm transfers only those that:

- are closer than three seconds to the next one (otherwise there is likely a gap between the two, therefore the current peak might be right before a zero padded noisy zone);

- are not coincident with the next one (this is evidently a mistake)
- are more than 0.25 second far from the previous minimum (it is physiologically impossible to have two heart depolarizations closer than 0.25 of second, due to heart refractariety)

Only the local maxima which come right after a minimum point are saved. Whenever two peaks are too distant (PPi >1.5 seconds) the second one is been erased.

For each signal, 24 parameters have been computed and analyzed. Parameters studied are: PPi mean, PPi standard deviation, PPi rMSSD, PPi nrMSSD, ΔPPi standard deviation,Shannon entropy, sample entropy, five indexes of similarity between each pair of PPG pulses, pNN10, pNN20, pNN30, pNN40, pNN50, pNN60, pNN70, pNN80, pNN90, pNN100, Poincarè plot and Lorenz plot.

2.2.1 Time domain heart rate parameters

Time domain analysis of AF recordings usually includes the mean and standard deviation of normal-to-normal (SDNN) PP intervals, root mean-square differences of successive normal-to-normal PP intervals (rMSSD), and percentage of interval differences of successive normal-to-normal PP intervals greater than different interval of time in term of ms, the most used is pNN50 [27].

For a PP time series made of N intervals, the standard deviation is defined as:

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}|PP_i - \mu|^2} \tag{2.8}$$

where μ is the mean of the PP series,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} PP_i \tag{2.9}$$

SD and rMSSD are time-domain tools used to assess heart rate variability (HRV). The root mean square of the successive differences is defined as:

$$rMSSD = \sqrt{\frac{1}{N-1}\sum_{i=2}^{N}|PP_i - PP_{i-1}|^2} \tag{2.10}$$

rMSSD is usually normalized by the mean value of the RR time series (nrMSSD), because subjects have different mean heart rates,

$$nrMSSD = \frac{rMSSD}{\mu} \tag{2.11}$$

The percentage of interval differences of successive intervals greater than x ms (pNNx, with x=[10,20,..., 100]) is computed. It is defined as the mean number of times in which the change in successive normal sinus (NN) intervals exceeds x ms.

SD, nRMSSD and pNN are used to quantify beat-to-beat variability. Since AF exhibits higher variability than NSR, these metrics can be used to detect AF when higher than a previously fixed threshold [8] [9] [34] [35].

2.2.2 Entropy

Approximate entropy (ApEn), is a measure of signal randomness. It represents the likelihood that similar patterns of observations will not be followed by other similar observations. When there are many similar values in a time series ApEn will be small, for example in NSR; a less predictable and more complex process has a greater ApEn, as in AF. Given N points, the ApEn(m, r, N) is approximately equal to the negative average natural logarithm of the conditional probability that two sequences that are similar for m points remain similar, that is, within a tolerance r, at the next point. ApEn is defined as

$$ApEn(m, r, N) = \ln \left[\frac{C_N^m(r)}{C_N^{m+1}(r)} \right] \tag{2.12}$$

where C_N^m expresses the prevalence of repetitive patterns of length m in the series. The ApEn algorithm counts each sequence as matching itself, a practice carried over to avoid the occurrence of ln(0) in the calculations. This step has led to discussion of the bias of ApEn [36] [27]. Therefore, the sample entropy (SampEn), not counting self-matches, has been introduced. SampEn reduces the bias of ApEn by avoiding counting self-matches, it can be used on shorter series, it is more consistent than ApEn and it is easier to compute [37]. Given N points, SampEn(m, r, N) is the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point within a tolerance r, where self-matches are not included in calculating the probability.

Therefore a lower value of SampEn also indicates more self-similarity in the time series [37]. SampEn is computed as

$$SampEn(m, r, N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right] \tag{2.13}$$

where $A^m(r)$ is the total number of matches of length m+1 and $B^m(r)$ the total number of matches of length m [38].

Shannon entropy (ShE) is a common entropy definition in information theory. In information theory, entropy measures the amount of uncertainty of a random quantity [39] [40]. The entropy of a random variable X is defined to be

$$H(x) = -\sum_x P(x)\log_2 P(x) \tag{2.14}$$

where the sum is over all values x that the variable X can take, and P(x) is the probability of the value x occurring. In AF detection, ShE of the PP intervals is used to quantify the regularity of pattern in time series. The ShE of NSR is expected to be smaller to the one of AF or other cardiac arrythmias [9].

2.2.3 Shape analysis

To assess wave similarity, five indexes are computed (S1, S2, S3, S4 and S5). Each wave is represented as a point of the p-dimensional real space, the normalized waves are points belonging to the p-dimensional unitary sphere. The morphological dissimilarity between a pair of waves

is evaluated using the standard metric of the sphere to compute their distance

$$D_{ij} = \arccos(w_i^N \cdot w_j^N) \tag{2.15}$$

where w_i^N and w_j^N represent the ith and jth normalized waves, i.e.

$$w_i^N = \frac{w_i}{||w_i||} \tag{2.16}$$

and $(\cdot)$ denotes the scalar product. The measure of similarity is obtained by calculating the relative number of similar pairs of waves in the recording. The similarity depends on the threshold ϵ used in evaluating the similarity, if the distance between two waves is lower than ϵ, waves are considered similar. Value of ϵ used are the same proposed in [26], $\epsilon = [\pi/2, \pi/4, \pi/8, \pi/16, \pi/32]$.

2.2.4 Poincarè plot

In the Poincaré plot, each RR interval is plotted versus the preceding one; it should be noted that this type of plot is also referred to as the Lorenz plot. The pattern of such a plot can be inspected to distinguish AF from other supraventricular tachycardias such as atrial flutter where ventricular response is not as irregular as in AF. During sinus rhythm (SR), successive RR intervals are centered around the main diagonal forming an ellipsoid-like pattern, and each RR interval is strongly dependent on the preceding RR interval. During AF, the irregularity of RR intervals results

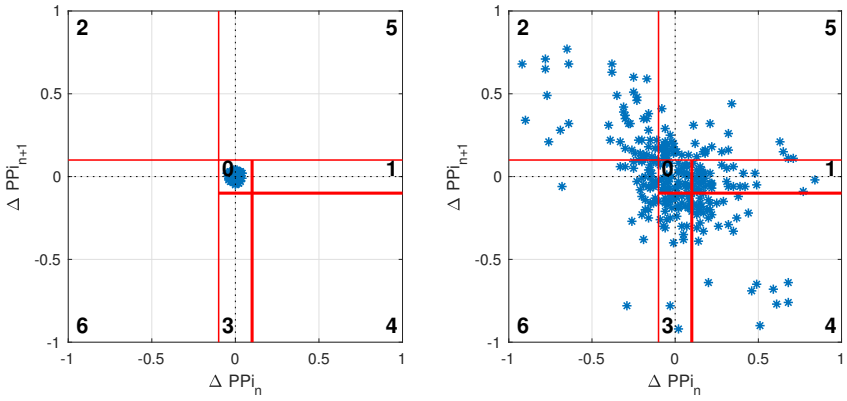


Figure 2.3: Examples of Poincarè plot in Chong et al. 2015, on the left Poincarè plot of a signal without AF, it is possible to see that all the points are in the zone 0; on the right Poincarè plot of a signal with AF, the points are in all the regions. Signals without noise generated with the simulator, with length of 300 beats.

in a widely scattered distribution which is representative of disorganized atrial activity combined with atrioventricular conduction properties (Figure 2.4).

Poincarè plot is used as a measure of organization of the heart beat, in [9] is used to characterize ΔPPi dynamics and to discriminate premature atrial (PAC) and ventricular (PVC) contractions episodes from those of AF and NSR. Poincarè plot is divided into seven regions, which represent permutations of all possible sequences "short" and "long" based on three consecutive pulse intervals. If the pattern obtained by the beats is irregular and with trajectories at all regions the signal is marked as AF, if all the trajectories stay in the zone 0 is NSR, see Figure 2.3.

The Lorenz plot is made plotting each ΔPP interval time against the following ΔPP interval time. The parameter Lor used in the analysis is

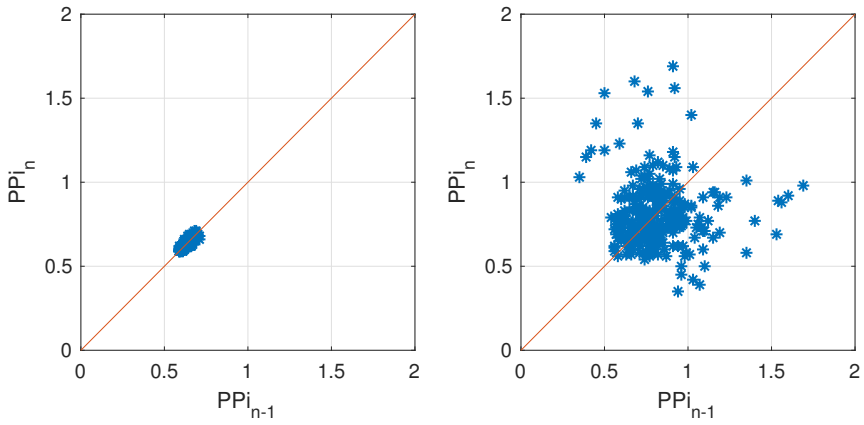


Figure 2.4: Examples of Poincarè plot, on the left Poincarè plot of a signal without AF, on the right Poincarè plot of a signal with AF. Signals without noise generated with the simulator, with length of 300 beats.

computed by the multiplication between the standard deviations for the transverse direction (Sd1) and the longitudinal direction (Sd2).

$$\text{Lor} = \text{Sd1} \cdot \text{Sd2}. \tag{2.17}$$

2.2.5 Turning Point Ratio

Turning Points Ratio measures the randomness of fluctuations within a data set and the degree of independence in a time series. It calculates the amount of turning points to the maximum number of possible turning points. Turning point is found if both the preceding and succeeding points are either greater or lower. Given 3 quantities a,b and c, b is a turning point if $a > b < c$, or $a < b > c$. It is expected in random data set of

arbitrary length N, that the number of possible turning points is $\frac{2N-4}{3}$, with a standard deviation of $\sqrt{\frac{16N-29}{90}}$ [41].

After the computation of the parameters cited, a t-test is performed to determine if there is a significant difference between the two groups of signals, AF and NSR.

2.3 Classification

A classification algorithm takes as input the matrix of observations F and a vector *label*. The matrix of observations has dimension NxM, where N is the number of signals and M the number of features. Since values of features extracted may differ by order of magnitudes, Z-score normalization was applied to transform each feature distribution so that it has zero mean and unitary standard deviation. The vector *label* has dimension Nx1, it has value 0 if the signal is NSR, and 1 if the signal is AF. In this study. for the classification linear SVMs are used.

SVMs are a family of separation methods for classification and regression developed in the context of statistical learning theory. They have been shown to achieve better performance in terms of accuracy with respect to other classifiers in several application domains, and to be efficiently scalable for large problems. A further important feature is concerned with the interpretation of the classification rules generated. Support vector machines identify a set of examples, called support vectors, which appear

to be the most representative observations for each target class. In a way, they play a more critical role than the other examples, since they define the position of the separating surface generated by the classifier in the attribute space [42].

For feature selection, a sequential forward floating search (SFFS) algorithm [7] is used to identify the best subset of features differentiating the two conditions. Briefly, starting from the empty set of features, the feature x_i that maximizes the objective function ($Y_k 1 X_i$) when combined with the features Y_k that have already been selected, is added. After this forward step, SFFS performs backward steps as long as the objective function increases. A backward step consists in removing from Y_k the feature that makes the objective function increase.

The quality of each model is assessed through train-validation-test split (55% of the data is used as training set, 25% as validation set, 20% as test set), repeated 100 times (bootstrapping). Training and validation sets are used to choose the best model, the test set to test the model on never-seen data. The data split took into account the outcome (responder/non-responder).

2.4 AF Detection algorithms

In this work, different algorithms are used to detect AF in the simulated database. It has been used different algorithms: Lee et al. 2012 [8],

Chong et al. 2015 [9], Petrenas et al. 2015 [10] and Solosenko et al. 2019 [43].

2.4.1 Lee et al. 2012

Lee developed a comprehensive iPhone application for collection of pulsatile time series followed by real-time detection of AF using: rMSSD, ShE and SampEn. The iPhone App was developed using the Objective-C programming language. For the pulsatile acquisition, the camera of an iPhone 4S was placed on either the index or middle finger of partecipants for two minutes prior to (see Figure 2.5), and immediately after, cardioversion. The iPhone 4S videos were recorded and pre-processed resampling the signal, then a peak detection algorithm has been performed. The signal was divided in 64-beat segment and then normalized rMSSD, ShE and SampEn were computed.

Figure 2.5: Camera placement for the pulsatile acquisition in [8]

The condition for AF detection is based on each threshold value of the three statistical methods TH_{Rm}, TH_{SE} and TH_{SA}: rmSSD normalized has to pass the threshold, rMSSD/mean $\geq TH_{Rm}$, and at least one of ShE or SampEn has to be higher than the corresponding threshold, ShE $\geq TH_{SE}$ and/or SampEn $\geq TH_{SA}$. Thresholds' values found, using the MIT-BIF AF and NSR databases, are: TH_{Rm}=0.130, TH_{SE}=0.76 and TH_{SA}=1.3. With iPhone 4S data and combination of all the three methods, the detection accuracy is 0.9951 with TH_{Rm}=0.115, TH_{SE}=0.55 and TH_{SA}=0.76.

2.4.2 Chong et al. 2015

Chong proposed an algorithm, used in arrhythmia discrimination using a a smartphone, which can differentiates between NSR, AF, PVC and PAC. The algorithm combines rmSSD of successive RR differences and ShE with Poincarè plot (or TPR) and pulse rise and fall times to increase the sensitivity of AF discrimination and add new capabilities of PVC and PAC identification. The 88 participants of the study were instructed to place their index or middle finger on a standard phone camera with the flash turned on for 2 min. Videos of their fingertip blood flow intensity were pre-processed extracting the region with best signal quality from the green signal, which is the band signal with the best signal fidelity. Pulse beat-to-beat detection was performed. The signal was divided in 60-beat segment. The arrythmia discrimination algorithm is detailed in the

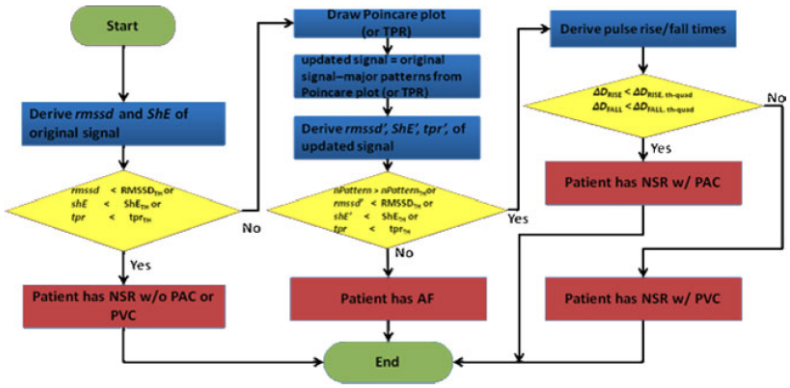


Figure 2.6: Flowchart of arrythmia discrimination algorithm proposed by Chong [9].

flowchart in Figure 2.6. After the computation of the statistical methods, the parameters are compared with their thresholds. If they are less than their thresholds, the pulsatile time series is classified as NSR without PAC or PVC. Otherwise, the algorithm goes to next step and check if the pulsatile time series is AF, PAC or PVC studying Poincarè plot and TPR.

Since the signals studied are completely in NSR or AF, the discrimination that I made in the replication of the algorithm is: if the values of rMMSD, ShE and TPR are less than the thresholds the pulsatile time series is classified as NSR, otherwise as AF. The optimal rMSSD, ShE and TPR threshold values are derived as 0.09275, 0.3800 and 0.3600. The method detects NSR with accuracy of 0.9626.

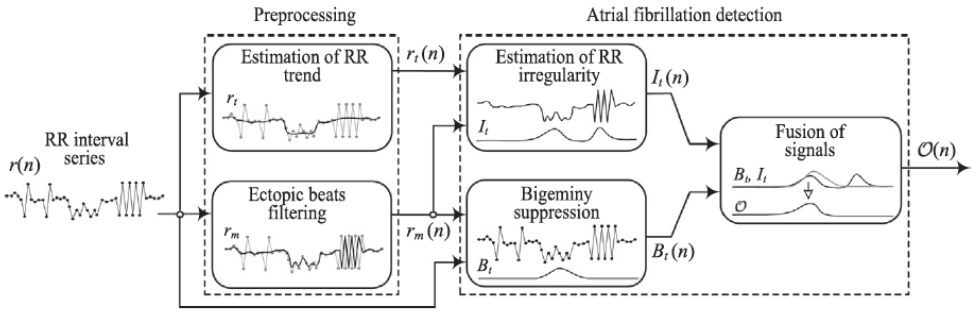


Figure 2.7: Flowchart of low-complexity detection of AF in continuous long-term monitoring algorithm proposed by Petrenas [10].

2.4.3 Petrenas et al. 2015

The study of Petrenas describes an AF detector whose structure is well-adapted both for detection of subclinical AF episodes and for implementation in a battery-powered device for use in continuous long-term monitoring applications. A key aspect for achieving these two properties is the use of an 8-beat sliding window, which thus is much shorter than the n-beat window used in most existing AF detectors, must used is 128-beat sliding window. The building blocks of the proposed detector include ectopic beat filtering, bigemini suppression, characterization of RR interval irregularity, and signal fusion. With one design parameter (α), the performance can be tuned to put more emphasis on avoiding false alarms due to non-AF arrhythmias or more emphasis on detecting brief AF episodes. The results show that the detector performs really good on the MIT-BIH AF database, with high sensitivity and specificity (97.1% and 98.3%, respectively). The detector can be implemented with just a few arithmetical operations (8 multiplications, 2 divisions and 45 addition/subtractions)

and it does not require a large memory buffer thanks to the short window. In the preprocessing, the simple 3-point median filter is used for reducing the influence of ectopic beats in the RR series. The filter is defined by

$$r_m(n) = \text{median}\{r(n-1), r(n), r(n+1)\}, \tag{2.18}$$

where *r(n)* is the length of the nth RR interval (in seconds). Median filtering is also useful for rejecting outlier RR intervals due to missed QRS complexes. An exponential averager is perform to track the "trend" in the RR interval series, and it is defined by

$$r_t(n) = r_t(n-1) + \alpha(r(n) - r_t(n+1)), \tag{2.19}$$

where α ($0<\alpha<1$) determines the degree of smoothing, the lowpass cutoff frequency. Since the exponential averager in 2.19 Techniques of filtering, like forward-backward filtering, have to be used to achieve linear (null) phase. For RR interval irregularity, in a sliding detection window of length N (N=8), located at time n the number of all pairwise RR interval combinations differing more than γ seconds (γ=0.03s) is determined, and normalized with its maximum value $N(N-1)/2$,

$$M(n) = \frac{2}{N(N-1)} \sum_{j=0}^{N-1} \sum_{k=j+1}^{N} H\big(\big|r(n-j) - r(n-k)\big| - \gamma\big), \tag{2.20}$$

where $H(\cdot)$ is the Heaviside step function and $0 \leq M(n) \leq 1$. $M(n)$ is based on the same principle as is sample entropy estimation. The ratio between $M_t(n)$, a smoothed version of $M(n)$ obtained by exponential

averaging, and the RR interval trend $r_t(n)$ provides a feature of RR irregularity,

$$I_t(n) = \frac{M_t(n)}{r_t(n)}, \tag{2.21}$$

where the division by $r_t(n)$ is to emphasize RR irregularity at higher heart rates. $I_t(n)$ is close to 0 for regular rhythms since the difference between pairs of RR intervals is usually smaller than γ, whereas $I_t(n)$ approaches 1 during AF. It is well-known that bigeminy can be incorrectly interpreted as AF when the detection is based on the RR. To discriminate bigeminy it is useful study introduce another measure of the RR irregularity, defined as

$$B(n) = \left(\frac{\sum_{j=0}^{N-1} r_m(n-j)}{\sum_{j=0}^{N-1} r(n-j)} - 1 \right)^2, \tag{2.22}$$

where N is an even-valued integer. For bigeminy and NSR 2.22 is approximately 0 because $r_m(n)$ and $r(n)$ are quite similar. For AF the variation in $r_m(n)$ is lower than in $r(n)$ because of the median filtering, and thus $B_t(n)$, which results from averaging filter of $B(n)$, will increase indicating RR irregularity. Signal fusion is employed to produce the decision function $O(n)$: $O(n)$ is equal to $B_t(n)$ if $B_t(n)$ does not exceed a fixed threshold δ, otherwise is equal to $I_t(n)$,

$$O(n) = \begin{cases} I_t(n) & B_t(n) \geq \delta \\ B_t(n) & B_t(n) < \delta \end{cases} \tag{2.23}$$

where δ in the study is equal to 2×10^{-4}. AF is detected whenever $O(n)$ exceeds η, a fixed threshold set equal to 0.725.

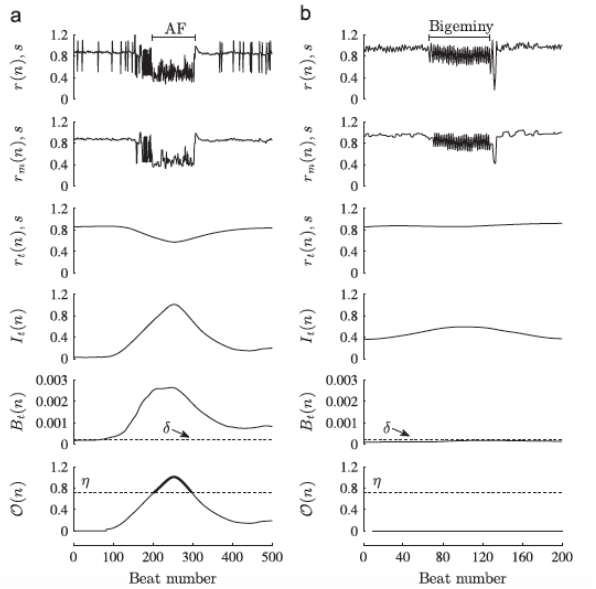


Figure 2.8: The output of each building block of the algorithm proposed by Petrenas [10] for (a) an AF episode and (b) bigeminy.

All the algorithms explained are performed on the database, in order to make a comparison with the classification made and study the accuracy, TPR and TNR of each AF detection algorithm.

Chapter 3

Results

In this chapter the results of the work are presented. In the first section the comparison between AF and NSR signals is illustrated. The second section shows the results of classification, in particular accuracy, TPR and TNR. The last section compares the classification algorithms.

3.1 Characterization of AF vs. NSR

The parameters obtained are studied for signals in AF and signals in NSR. Boxplots are used to display the distribution of data, they are based on a five number summary ("minimum", first quartile (Q1), median, third quartile (Q3), and "maximum"). Figure 3.1, Figure 3.3 and Figure 3.2 show the comparisons of parameters for signals with length of 300 RR intervals between the worst scenario, signal with SNR=0, and the best one, SNR=30. T-test is performed to determine if there are significant parameters differences between NSR and AF.

Figure 3.1 compares ShE, SampEn, rMMSD and Lorenz plot standard deviation values of NSR and AF signals with SNR=0 and SNR=30. Figure 3.2 compares similarity indexes values of NSR and AF signals with SNR=0 and SNR=30. Figure 3.3 compares the percentage of interval differences of successive intervals greater than x ms (pNNx, with x=[10,20,. . . , 100]) values of NSR and AF signals with SNR=0 and SNR=30. For signals with RR length=300 and SNR=0, SampEn, rMSSD, Lorenz plot standard deviation, S1, S2, pNN10, pNN20, pNN30, pNN40, pNN50, pNN60, pNN80, pNN90, pNN100 values are significantly different ($p<0.05$) between NSR and AF. For signals with RR length=300 and SNR=30, all the parameters values are significantly different ($p<0.05$) between NSR and AF.

To display the results of the t-test colormaps are used. A binary color classification is made, if *p-value* of the t-test is significant (*p-value*<0.05) the corresponding block is purple, otherwise is yellow.

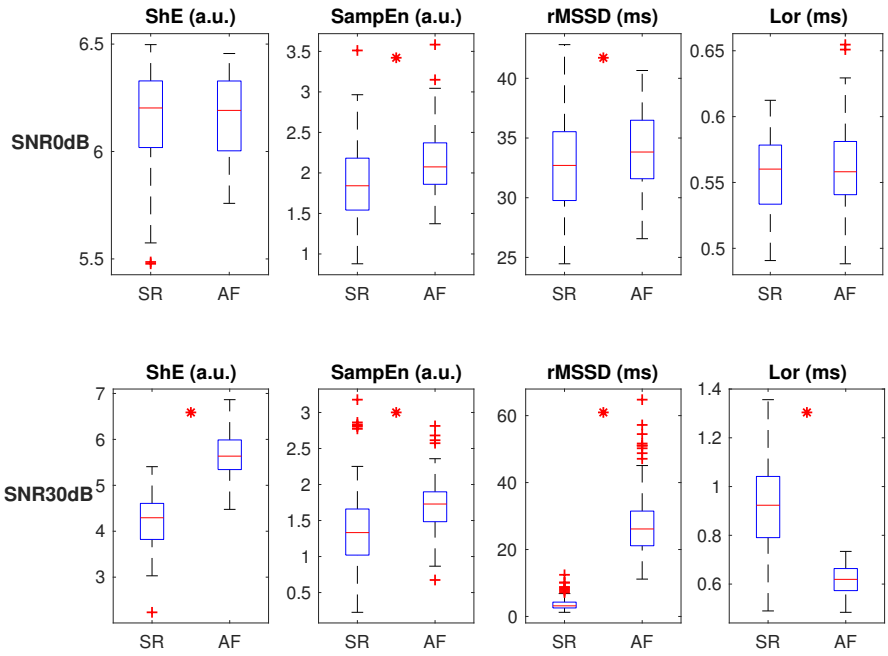


Figure 3.1: Shannon Entropy, Sample Entropy, rMSSD and Lorenz Plot SD comparison for signals with RR length=300 and SNR=0, in the upper part of the figure, and with RR length=300 and SNR=300, in the lower part of the figure. $*$ in the boxplot means that the p-value resulting from the t-test is less than 0.05. a.u=arbitrary unit.

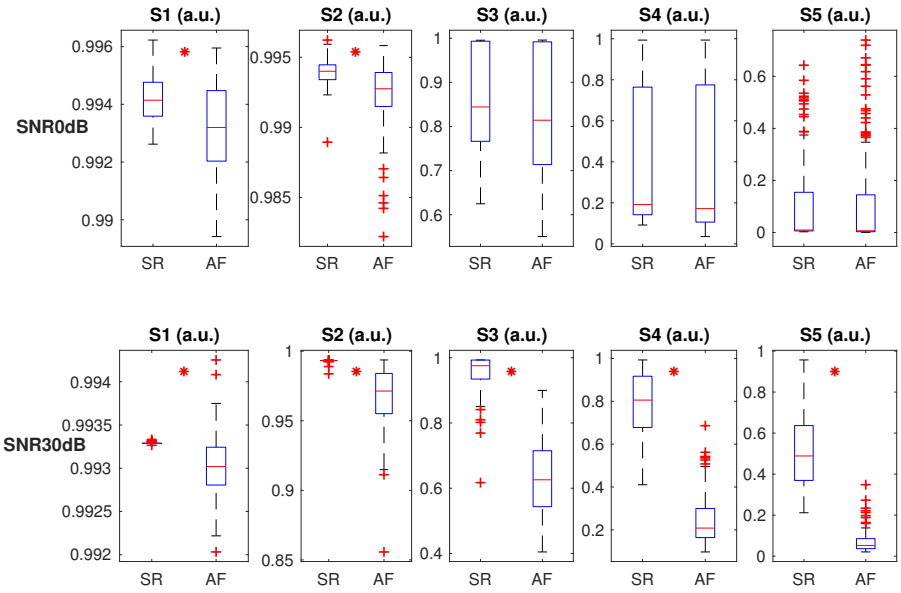


Figure 3.2: Similarity indexes comparison for signals with RR length=300 and SNR=0, in the upper part of the figure, and with RR length=300 and SNR=300, in the lower part of the figure. $*$ in the boxplot means that the p-value resulting from the t-test is less than 0.05. a.u=arbitrary unit.

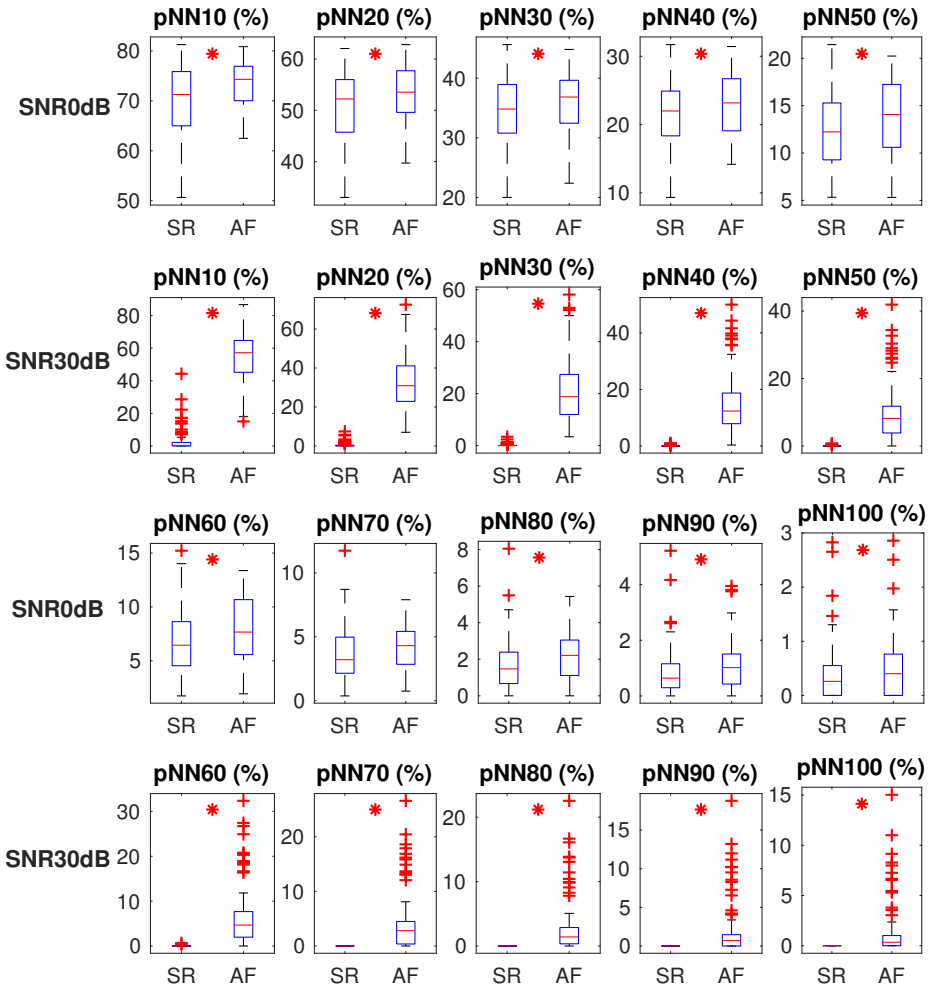


Figure 3.3: pNN comparison for signals with RR length=300 and SNR=0, in the upper part of the figure, and with RR length=300 and SNR=300, in the lower part of the figure. $*$ in the boxplot means that the p-value resulting from the t-test is less than 0.05.

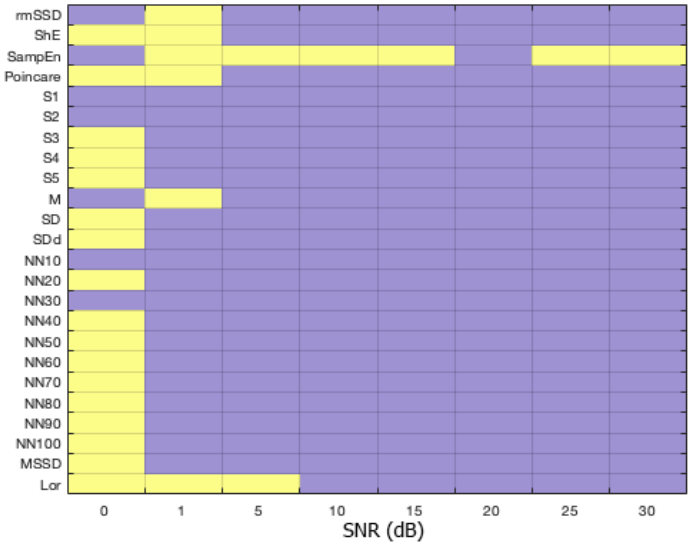


Figure 3.4: *p-value* resulting from t-test signals with RR length=50 and different SNR. Blocks in yellow represent *p-value*$\geq$0.05, blocks in purple represent *p-value*$<$0.05

Figure 3.4 shows the *p-values* colormap obtained for signals with RR length=50 and different SNR. It is possible to notice that parameters are significantly different for NSR and AF for SNR values bigger than 1. With SNR=0 there is no significant different except for seven parameters, nRMSSD, SampEn, S1, S2, M, NN10 and NN40. The t-test shows also that there is no significant SampEn difference for NSR and AF for almost all SNR values.

Figure 3.4 shows the *p-values* colormap obtained for signals with RR length=300 and different SNR. It is possible to notice that parameters are significantly different for NSR and AF for SNR values bigger than 5. With SNR=1 there is no significant different except for five parameters, SampEn, S1, S2, M and NN10.

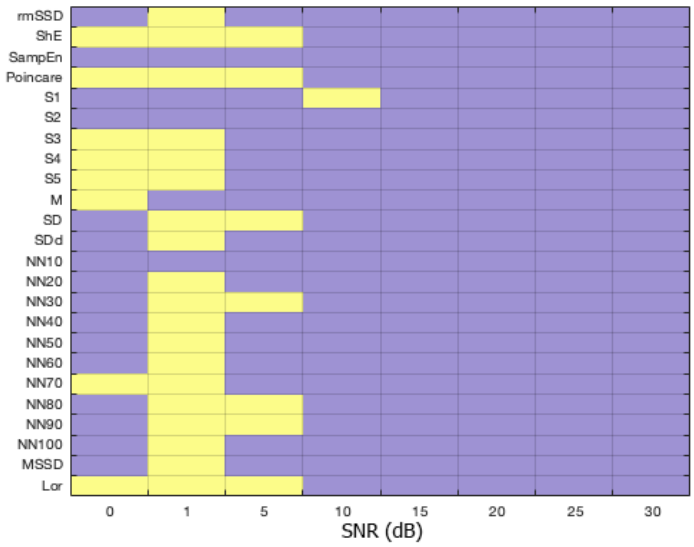


Figure 3.5: *p-value* resulting from t-test signals with RR length=300 and different SNR. Blocks in yellow represent *p-value*$\geq$0.05, blocks in purple represent *p-value*$<$0.05

Figure 3.6 shows the *p-values* colormap obtained for signals with different RR intervals length and SNR=0. The parameters SampEn, S1 and S2 are significant different for NSR and AF for each length. For RR length=250 and RR length=300 parameters are significantly different for NSR and AF.

Figure 3.7 shows the *p-values* colormap obtained for signals with different RR intervals length and SNR=30. The parameters are significantly different for NSR and AF for all the values of RR length. The only two yellow block in the colormap are SampEn for RR length=50 and M for RR Length=200. In Figure 2 validation and test accuracy

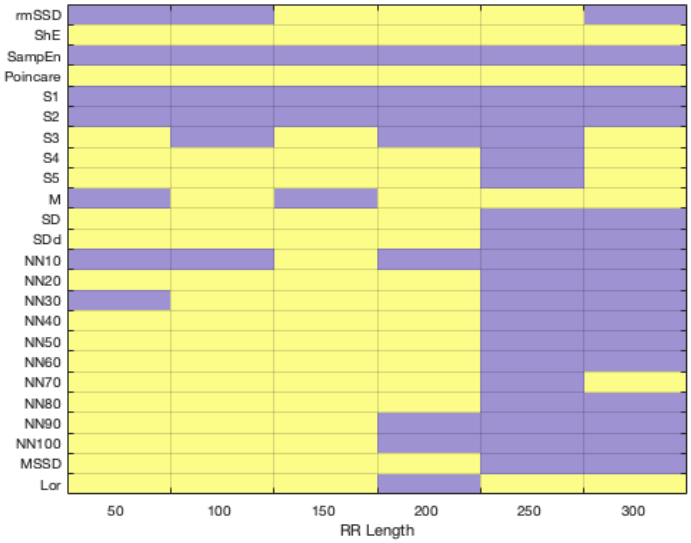


Figure 3.6: *p-value* resulting from t-test signals with different RR length and SNR=0. Blocks in yellow represent *p-value*$\geq$0.05, blocks in purple represent *p-value*$<$0.05

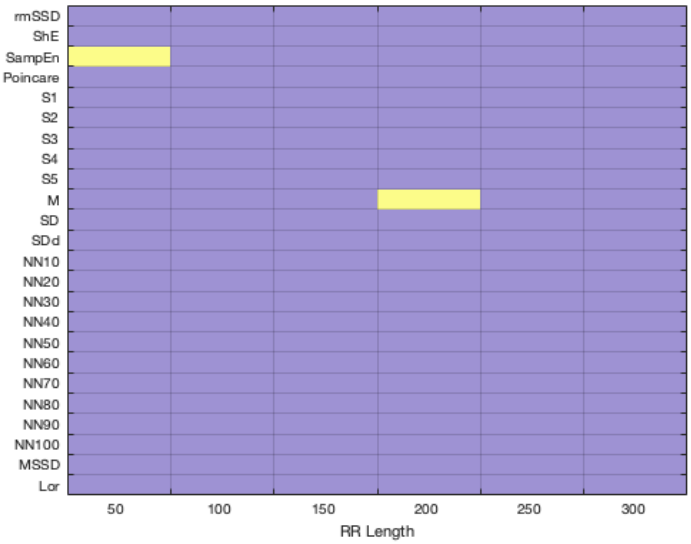


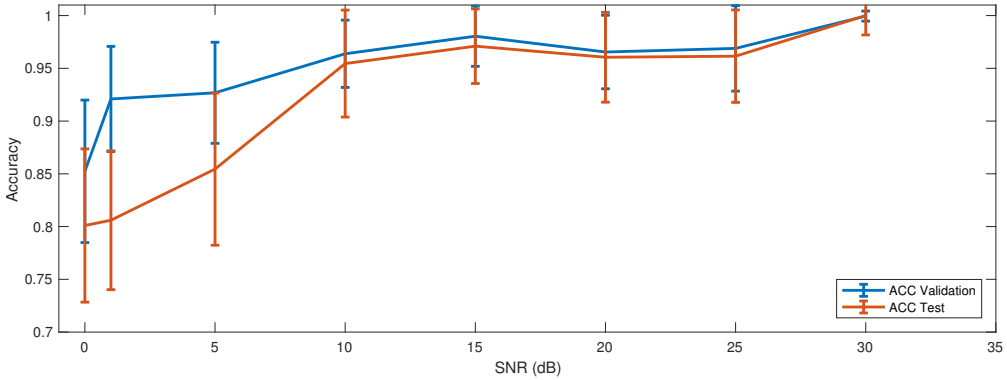
Figure 3.7: *p-value* resulting from t-test signals with different RR length and SNR=30. Blocks in yellow represent *p-value*$\geq$0.05, blocks in purple represent *p-value*$<$0.05

3.2 Classification

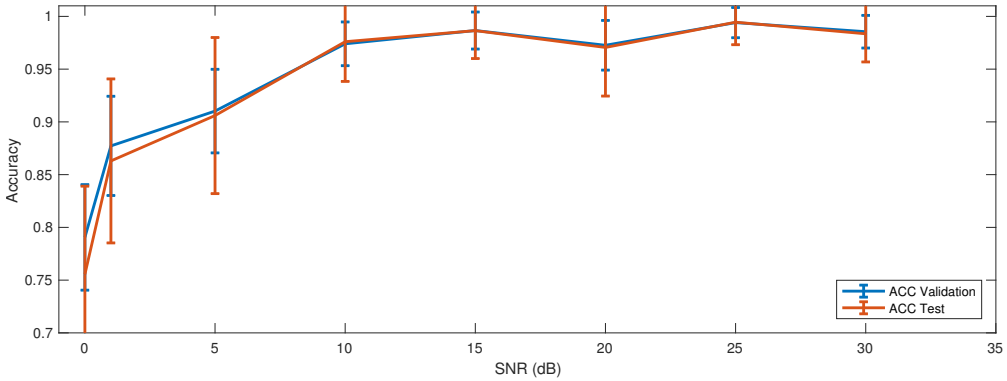
3.2.1 Simulated data

In this section results of classification in simulated PPG signals are illustrated. Figure 3.8 shows validation accuracy and test accuracy for signals with RR length=50, 100 and 300 in function of SNR . It can be observed for RR length = 50, that the accuracy increases since SNR=10, after this point it is almost constant; the validation accuracy is higher than test accuracy for SNR values lower than 10, then they have almost the same value. For RR length=100, the accuracy increases since SNR=10, after this point it is almost constant; the validation accuracy and test accuracy have almost the same value. For RR length=300, passing from SNR=0 to SNR=1 causes a decrease in accuracy and then it increases since SNR=10, after this point it is almost constant; the validation accuracy and test accuracy have the same value.

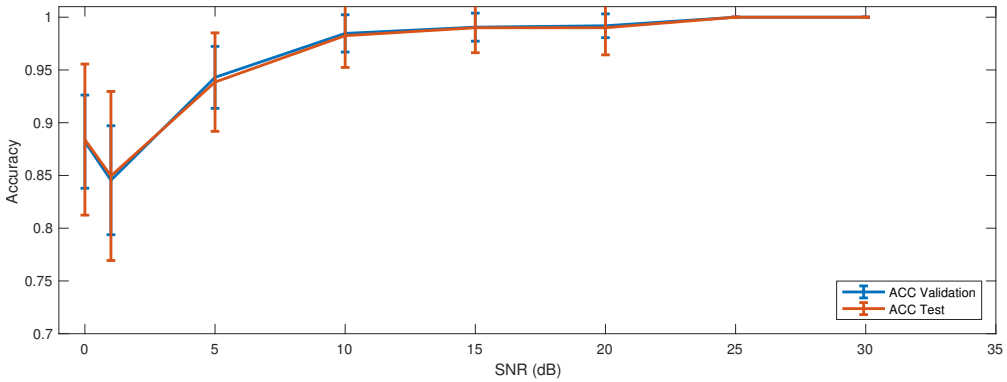
Figure 3.9 shows true positive rate (TPR) and true negative rate (TNR) for signals with RR length = 50, 100 and 300 in function of SNR. In Figure 3.9(a) it can be observed, for RR length = 50, that the TPR passing from SNR=0 to SNR=1, decreases and then it increases since SNR=10, after this point it is almost constant. For RR length=100, the TPR increases since SNR=10, after this point it is almost constant. For RR length=300, passing from SNR=0 to SNR=1 causes a decrease in TPR and then it increases since SNR=10, after this point it is almost constant.



(a)



(b)



(c)

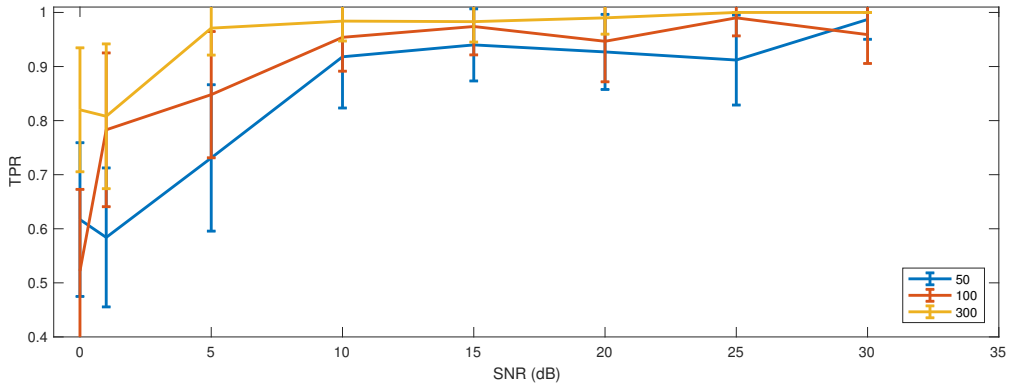
Figure 3.8: Validation and test accuracy for classification algorithm in signals with RR length=50 (a), RR length=100 (b) and RR length=300

For SNR≥10, the TPR for each RR length is over 0.91 ($TPR10_{50}$=0.918, $TPR10_{100}$=0.954, $TPR10_{300}$=0.984).

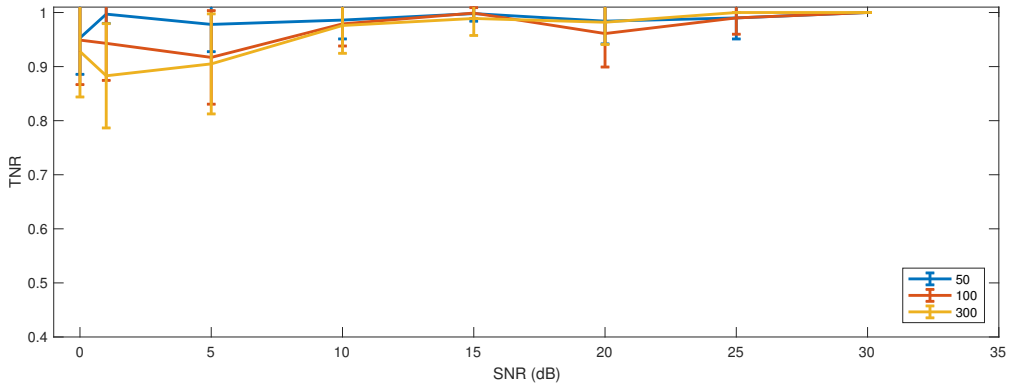
In Figure 3.9(b) it can be observed, for RR length = 50, that the TNR passing from SNR=0 to SNR=1, decreases and then it increases since SNR=10, after this point it is almost constant. For RR length = 100, TNR decreases passing from SNR=0 to SNR=5, then it increases and becomes constant, except for a small deflection in SNR=20. For RR length=300, passing from SNR=0 to SNR=1 causes a decrease in TNR and then it increases since SNR=10, after this point it is almost constant. The TNR for each RR length for SNR≥10 is over 0.96.

3.2.2 Clinical data

In this section results of classification in clinical signals are illustrated. In Figure 3.10 validation accuracy and test accuracy in function of the number of features selected (n) are shown. It is possible to observe that validation accuracy is almost higher than test accuracy. The number of features n which maximize the validation accuracy is 2, after this value the accuracy decreases. In Figure 3.11(a) TPR in function of the number of features selected is shown. TPR increases with number of features since n=4, then it decreases until n=19 where it increases in n=20, after this value it decreases.



(a)



(b)

Figure 3.9: TPR and TNR of classification for different lengths of signal: 50, 100 and 300.

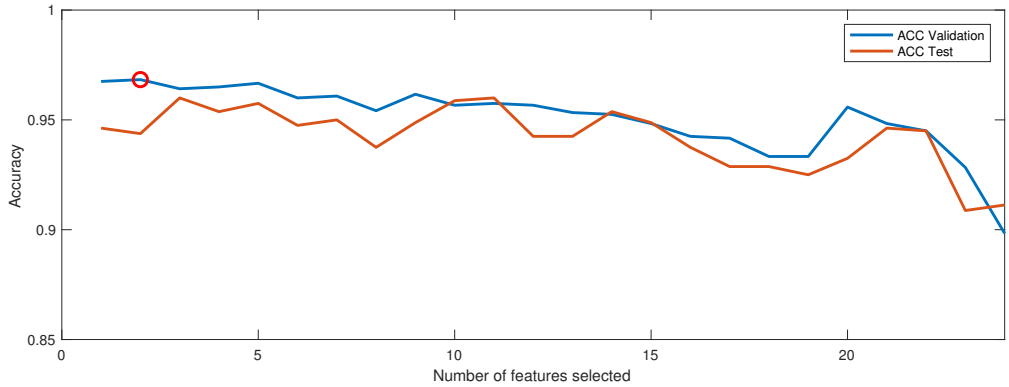
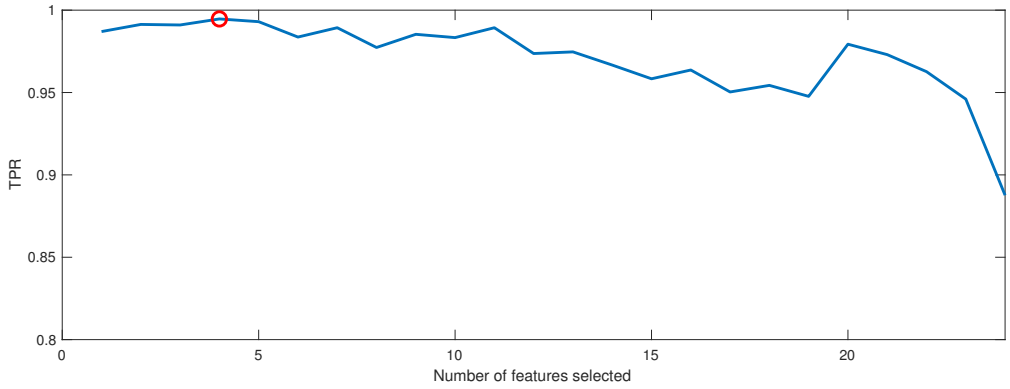
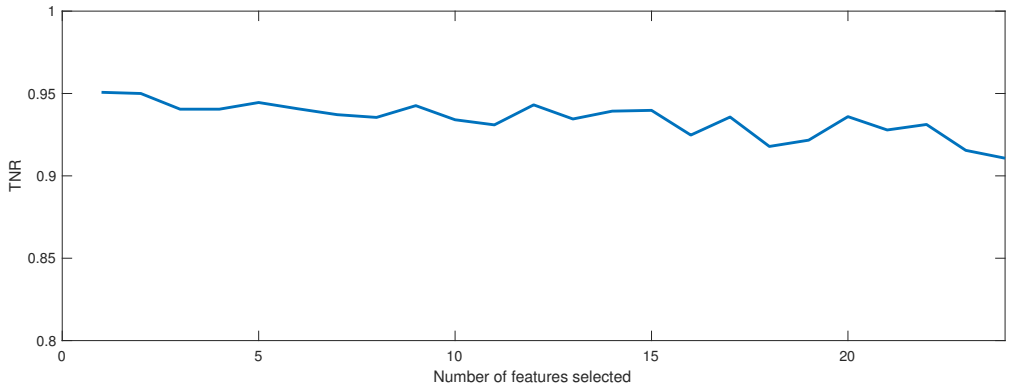


Figure 3.10: Validation accuracy and test accuracy of classification for different number of features in clinical data.

In Figure 3.11(b) TPR in function of the number of features selected is shown. TPR have almost the same value for each number of features n, it decreases with higher n.



(a)



(b)

Figure 3.11: TPR and TNR of classification for different number of features in clinical data.

3.3 Comparison of AF Algorithms

In this section results of comparison between the classification algorithm (A) and existing AF detection algorithms (Petrenas, Lee and Chong) are

presented. In Table 3.1 and Table 3.2 the results are computed considering, for the existing algorithms, AF detection right when at least one episode of AF is detected, in case of AF signals, and when no one episode of AF is detected, in case of NSR signals. Table 3.1 shows accuracy, TPR and TNR for the different algorithms in function of SNR in signals with RR length = 100. The accuracy for each algorithm increases with the SNR until SNR=25, then it decreases except for Petrenas. TPR has optimal values for each existing algorithm, the worst cases are in Chong's algorithm with SNR=20 and SNR=30 ($TPR20_{Chong} = TPR30_{Chong} = 0.88$). TNR increases with SNR for Lee and Chong. In Petrenas TNR increase until SNR=25 and then it decreases. For the classification algorithm the TPR increase with the SNR since SNR=25 and then it decreases, TNR has optimal value even with low SNR values.

SNR	**Accuracy**				**TPR**				**TNR**			
	A	P	L	C	A	P	L	C	A	P	L	C
0	0.7555	0.495	0.497	0.5	0.522	0.99	1	0.98	0.9490	0	0	0.02
1	0.8630	0.5	0.5	0.52	0.7830	1	1	0.98	0.9430	0	0	0.06
5	0.9060	0.525	0.51	0.55	0.8480	1	1	1	0.9170	0.05	0.02	0.1
10	0.9760	0.66	0.625	0.695	0.9540	0.99	1	0.95	0.9790	0.33	0.25	0.44
15	0.9865	0.905	0.84	0.83	0.9740	1	0.99	0.93	0.999	0.81	0.69	0.73
20	0.9706	0.955	0.94	0.875	0.9467	0.99	0.98	0.88	0.9611	0.92	0.9	0.87
25	0.9945	0.955	0.985	0.92	0.99	1	1	0.91	0.9919	0.99	0.97	0.93
30	0.9835	0.965	0.975	0.915	0.9590	0.98	0.99	0.88	1	0.95	0.96	0.95

Table 3.1: Accuracy, TPR and TNR for signals with RR length=100 and different SNR computed for the classification algorithm (A), Petrenas (P), Lee (L) and Chong (C).

Table 3.2 shows accuracy, TPR and TNR for the different algorithms in function of SNR in signals with RR length = 300. The accuracy for each algorithm increases with the SNR. TPR has optimal values for each existing algorithm, for the classification algorithm the TPR increase with

SNR	Accuracy				TPR				TNR			
	A	P	L	C	A	P	L	C	A	P	L	C
0	0.884	0.5	0.5	0.5	0.82	1	1	1	0.9237	0	0	0
1	0.8495	0.5	0.5	0.5	0.808	1	1	1	0.8730	0	0	0
5	0.9385	0.5	0.5	0.505	0.971	1	1	1	0.9130	0	0	0.01
10	0.9825	0.52	0.505	0.565	0.984	1	0.99	0.99	0.9774	0.04	0.02	0.14
15	0.99	0.75	0.7	0.76	0.983	1	1	1	0.9907	0.5	0.4	0.52
20	0.99	0.86	0.895	0.88	0.99	1	1	0.99	0.9926	0.72	0.79	0.77
25	1	0.88	0.94	0.91	1	1	1	0.98	1	0.76	0.88	0.84
30	1	0.945	0.975	0.945	1	1	1	0.98	1	0.89	0.95	0.91

Table 3.2: Accuracy, TPR and TNR for signals with RR length=300 and different SNR computed for the classification algorithm (A), Petrenas (P), Lee (L) and Chong (C).

the SNR. TNR increases with SNR for each algorithm, for the classification algorithm TNR has optimal values even with low SNR values.

Tabel 3.3 shows accuracy, TPR and TNR for the different algorithms in signals from clinical database. The algorithm with an higher accuracy is the classification algorithm, then Petrenas, Chong and Lee. TPR values are high and similar to 1 in each algorithm. TNR values are low for each existing algorithm, but for the classification algorithm the value is 0.913.

	Accuracy	**TPR**	**TNR**
Petrenas	0.673	1	0.414
Lee	0.577	1	0.241
Chong	0.596	0.957	0.310
Classification A	0.958	0.995	0.913

Table 3.3: Accuracy, TPR and TNR for each algorithm, for clinical data

Chapter 4

Discussion and conclusions

In this chapter, the results obtained in this work are discussed, evaluating the efficiency of the classification algorithms. In the last part limits of this work and future developments are presented.

4.1 Discussion and conclusions

The PPG signal, considering the large quantity of devices which are able to record and thanks to its easy recording method, can give an important help in the screening and detection of AF and other arrythmias. Nowadays smartphones and smartwatches are part of the daily life of all people making easier the health monitoring, including the cardiac health.

This study aims at evaluating the potential of a set of PPG-derived measures to discriminate between AF and NSR. In signals with high value of SNR the distribution of the parameters is always significantly different in AF and NSR, quite the opposite, with low values of SNR the distributions are more similar. This is likely to be due to the noise that worsens the peak detection. Focusing on the duration of the signal, signals with RR length=50 have more parameters that are significantly different than signals with RR length=300, in which the t-test better performs for SNR=0 than SNR=1.

The classification algorithm has good results, the best results (i.e., for signals with RR length = 100, accuracy=0.9825 for SNR=10db, accuracy = 0.99 for SNR=15db, 20db and accuracy=1 for SNR=25db and 30db) are obtained with values of SNR$\geq$10 db. It suggests that noise, in particular motion artifacts, really affects the detection of AF episodes. Comparing the results for the classification algorithm A in Table 3.1 and Table 3.2, is possible to observe that longer signals are better classified, with higher value of accuracy and TPR. The classification algorithm works better

than the others even with low values of SNR. High values of TPR and TNR indicate that the discrimination between the two rhythms is well performed and it reduces the number of false alarms.

The classification algorithm classifies well also clinical data, being the number of features that maximizing the accuracy 2. If the purpose is to maximize the TPR, in order to detect a major number of AF episodes, the number of features increases to 4. As reported in Table 3.3 the algorithm work better than the other taken in consideration, particularly for the right classification of the NSR.

In the end, it is possible to conclude that the significant difference between AF and NSR signals depends more on SNR than length of the signal and that the classification algorithm works better than the others considered in this study, with high level of accuracy and TPR even with a small number of feature selected, this is demonstrated with generated signals and also with clinical data.

4.2 Limitation of the study and future developments

The main limit of this study are generated signals used to evaluate the classification algorithms, since the data are generated completely affected by AF or completely NSR. It would be useful study signals that are more

close to the reality, using clinical databases. It may be interesting to consider the length of the AF episode detected and compares it with the real episode. Another limitation of this study is the use of Petrenas algorithm instead of Solosenko algorithm [43], because it is an improvement of the first one but it has been just published and I did not have the time to reproduce it. With these consideration, future development could be a study with only clinical data in which also the length of the episodes detected is compared.

Given the spread and impact on the quality of life of heart diseases, another development could be a creation of a classification algorithms for different type of arrythmias, not focused only AF and NSR, but which can investigate also the others.

Bibliography

[1] D.-g. Fu. Cardiac arrhythmias: Diagnosis, symptoms, and treatments. *Cell biochemistry and biophysics* 73.2 (2015), pp. 291–296.

[2] P. Zimetbaum. Atrial Fibrillation. *Annals of Internal Medicine* 166.5 (Mar. 2017), ITC33–ITC48. ISSN: 0003-4819. DOI: 10.7326/AITC201703070. eprint: https://annals.org/acp/content_public/journal/aim/936088/aitc201703070.pdf. URL: https://doi.org/10.7326/AITC201703070.

[3] G. Boriani et al. Asymptomatic atrial fibrillation: clinical correlates, management, and outcomes in the EORP-AF Pilot General Registry. *The American journal of medicine* 128.5 (2015), pp. 509–518.

[4] F. Pistoia, S. Sacco, C. Tiseo, D. Degan, R. Ornello, and A. Carolei. The epidemiology of atrial fibrillation and stroke. *Cardiology clinics* 34.2 (2016), pp. 255–268.

[5] C. Fischer, B. Dömer, T. Wibmer, and T. Penzel. An algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms. *IEEE journal of biomedical and health informatics* 21.2 (2016), pp. 372–381.

[6] A. Sološenko, A. Petrėnas, V. Marozas, and L. Sörnmo. Modeling of the photoplethysmogram during atrial fibrillation. *Computers in biology and medicine* 81 (2017), pp. 130–138.

[7] P. Heuristics. *Intelligent Search Strategies for Computer Problem Solving*. 1984.

[8] J. Lee, B. A. Reyes, D. D. McManus, O. Maitas, and K. H. Chon. Atrial fibrillation detection using an iPhone 4S. *IEEE Transactions on Biomedical Engineering* 60.1 (2012), pp. 203–206.

[9] J. W. Chong, N. Esa, D. D. McManus, and K. H. Chon. Arrhythmia discrimination using a smart phone. *IEEE journal of biomedical and health informatics* 19.3 (2015), pp. 815–824.

[10] A. Petrėnas, V. Marozas, and L. Sörnmo. Low-complexity detection of atrial fibrillation in continuous long-term monitoring. *Computers in biology and medicine* 65 (2015), pp. 184–191.

[11] S. P. Johnsen, L. W. Dalby, T. Täckström, J. Olsen, and A. Fraschke. Cost of illness of atrial fibrillation: a nationwide study of societal impact. *BMC health services research* 17.1 (2017), p. 714.

[12] M. Z. Berisso et al. The cost of atrial fibrillation in Italy: a five-year analysis of healthcare expenditure in the general population. From

the Italian Survey of Atrial Fibrillation Management (ISAF) study. *Eur Rev Med Pharmacol Sci* 21.1 (2017), pp. 175–183.

[13] F. Rahman, G. F. Kwan, and E. J. Benjamin. Global epidemiology of atrial fibrillation. *Nature Reviews Cardiology* 11.11 (2014), p. 639.

[14] S. S. Chugh et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* 129.8 (2014), pp. 837–847.

[15] V. Fuster et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 guidelines for the management of patients with atrial fibrillation) developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Journal of the American College of Cardiology* 48.4 (2006), e149–e246.

[16] A. D. Margulescu and L. Mont. Persistent atrial fibrillation vs paroxysmal atrial fibrillation: differences in management. *Expert Review of Cardiovascular Therapy* 15.8 (2017), pp. 601–618.

[17] C. Rugarli et al. *Medicina interna sistematica*. Elsevier, 2010.

[18] S. Z. Rosero, V. Kutyifa, B. Olshansky, and W. Zareba. Ambulatory ECG monitoring in atrial fibrillation management. *Progress in cardiovascular diseases* 56.2 (2013), pp. 143–152.

[19] A. Tarniceriu et al. "The accuracy of atrial fibrillation detection from wrist photoplethysmography. a study on post-operative patients". *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2018, pp. 1–4.

[20] R. Couceiro, P. Carvalho, J. Henriques, M. Antunes, M. Harris, and J. Habetha. "Detection of atrial fibrillation using model-based ECG analysis". *2008 19th International Conference on Pattern Recognition*. IEEE. 2008, pp. 1–5.

[21] S. Babaeizadeh, R. E. Gregg, E. D. Helfenbein, J. M. Lindauer, and S. H. Zhou. Improvements in atrial fibrillation detection for real-time monitoring. *Journal of electrocardiology* 42.6 (2009), pp. 522–526.

[22] M. Grond et al. Improved detection of silent atrial fibrillation using 72-hour Holter ECG in patients with ischemic stroke: a prospective multicenter cohort study. *Stroke* 44.12 (2013), pp. 3357–3364.

[23] W. Cai et al. Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network. *Computers in Biology and Medicine* (2019), p. 103378.

[24] R. S.C. P. Vaizurs, R. Sankar, and F. Leonelli. "Atrial fibrillation source identification". *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2011, pp. 4398–4401.

[25] C. Israel et al. Detection of atrial fibrillation in patients with embolic stroke of undetermined source by prolonged monitoring with

implantable loop recorders. *Thrombosis and haemostasis* 117.10 (2017), pp. 1962–1969.

[26] V. D. Corino, R. Laureanti, L. Ferranti, G. Scarpini, F. Lombardi, and L. T. Mainardi. Detection of atrial fibrillation episodes using a wristband device. *Physiological measurement* 38.5 (2017), p. 787.

[27] L. Mainardi, L. Sörnmo, and S. Cerutti. Understanding atrial fibrillation: The signal processing contribution, part ii. *Synthesis Lectures on Biomedical Engineering* 3.1 (2008), pp. 1–139.

[28] D. E. Becker. Fundamentals of electrocardiography interpretation. *Anesthesia progress* 53.2 (2006), pp. 53–64.

[29] J. J. Goldberger and J. Ng. *Practical signal and image processing in clinical cardiology*. Springer, 2010.

[30] L. FERRANTI and R. LAUREANTI. Atrial fibrillation detection in PPG signal recorded through a wristband device (2015).

[31] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement* 28.3 (2007), R1.

[32] R. D. White and G. Flaker. Smartphone-based Arrhythmia Detection: Should we encourage patients to use the ECG in their pocket? *Journal of atrial fibrillation* 9.6 (2017).

[33] M. R. Gropler, A. S. Dalal, G. F. Van Hare, and J. N. A. Silva. Can smartphone wireless ECGs be used to accurately assess ECG intervals in pediatrics? A comparison of mobile health monitoring to standard 12-lead ECG. *PloS one* 13.9 (2018), e0204403.

[34] T. Schäck, Y. S. Harb, M. Muma, and A. M. Zoubir. "Computationally efficient algorithm for photoplethysmography-based atrial fibrillation detection using smartphones". *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2017, pp. 104–108.

[35] A. Tarniceriu et al. "Detection of beat-to-beat intervals from wrist photoplethysmography in patients with sinus rhythm and atrial fibrillation after surgery". *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE. 2018, pp. 133–136.

[36] S. M. Pincus. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences* 88.6 (1991), pp. 2297–2301.

[37] J. S. Richman and J. R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology* 278.6 (2000), H2039–H2049.

[38] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 283.3 (2002), R789–R797.

[39] S.-M. Shan et al. "Reliable PPG-based algorithm in atrial fibrillation detection". *2016 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE. 2016, pp. 340–343.

[40] Z. Ghahramani. Information theory. *Encyclopedia of Cognitive Science* (2006).

[41] K. Horoba et al. "Recognition of Atrial Fibrilation Episodes in Heart Rate Variability Signals Using a Machine Learning Approach". *2019 MIXDES-26th International Conference" Mixed Design of Integrated Circuits and Systems"*. IEEE. 2019, pp. 419–424.

[42] C. Vercellis. *Business intelligence: data mining and optimization for decision making*. Wiley Online Library, 2009.

[43] A. Sološenko, A. Petrėnas, B. Paliakaitė, L. Sörnmo, and V. Marozas. Detection of atrial fibrillation using a wrist-worn device. *Physiological measurement* 40.2 (2019), p. 025003.