

Characterization and Processing of Novel Neck Photoplethysmography Signals for Cardiorespiratory Monitoring

Irene María García López

November, 2020

Department of Electrical and Electronic Engineering
Imperial College London

Supervised by Prof. Esther Rodriguez-Villegas

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Electrical and Electronic Engineering
of the University of London and the Diploma of Imperial College

Abstract

Epilepsy is a neurological disorder causing serious brain seizures that severely affect the patients' quality of life. Sudden unexpected death in epilepsy (SUDEP), for which no evident disease reason is found after post-mortem examination, is a common cause of mortality. The mechanisms leading to SUDEP are uncertain, but, centrally mediated apneic respiratory dysfunction, inducing dangerous hypoxemia, plays a key role. Continuous physiological monitoring appears as the only reliable solution for SUDEP prevention. However, current seizure-detection systems do not show enough sensitivity and present a high number of intolerable false alarms. A wearable system capable of measuring several physiological signals from the same body location, could efficiently overcome these limitations. In this framework, a neck wearable apnea detection device (WADD), sensing airflow through tracheal sounds, was designed. Despite the promising performance, it is still necessary to integrate an oximeter sensor into the system, to measure oxygen saturation in blood (SpO_2) from neck photoplethysmography (PPG) signals, and hence, support the apnea detection decision.

The neck is a novel PPG measurement site that has not yet been thoroughly explored, due to numerous challenges. This research work aims to characterize neck PPG signals, in order to fully exploit this alternative pulse oximetry location, for precise cardiorespiratory biomarkers monitoring.

In this thesis, neck PPG signals were recorded, for the first time in literature, in a series of experiments under different artifacts and respiratory conditions. Morphological and spectral characteristics were analyzed in order to identify potential singularities of the signals. The most common neck PPG artifacts critically corrupting the signal quality, and other breathing states of interest, were thoroughly characterized in terms of the most discriminative features. An algorithm was further developed to differentiate artifacts from clean PPG signals. Both, the proposed characterization and classification model can be useful tools for researchers to denoise neck PPG signals and exploit them in a variety of clinical contexts. In addition to that, it was demonstrated that the neck also offered the possibility, unlike other body parts, to extract the Jugular Venous Pulse (JVP) non-invasively.

Overall, the thesis showed how the neck could be an optimum location for multi-modal monitoring in the context of diseases affecting respiration, since it not only allows the sensing of airflow related signals, but also, the breathing frequency component of the PPG appeared more prominent than in the standard finger location. In this context, this property enabled the extraction of relevant features to develop a promising algorithm for apnea detection in near-real time.

These findings could be of great importance for SUDEP prevention, facilitating the investigation of the mechanisms and risk factors associated to it, and ultimately reduce epilepsy mortality.

List of Publications

The following papers have been published as part of this work.

Journal papers

- I. Garcia-Lopez and E. Rodriguez-Villegas, “Characterization of artifact signals in neck photoplethysmography,” *IEEE Transactions on Biomedical Engineering*, 2020
- I. García-López and E. Rodriguez-Villegas, “Extracting the jugular venous pulse from anterior neck contact photoplethysmography,” *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020

Peer reviewed conference papers

- I. García-López, S. A. Imtiaz, and E. Rodriguez-Villegas, “Characterization study of neck photoplethysmography,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4355–4358, IEEE, 2018
- I. García-López, P. Sharma, and E. Rodriguez-Villegas, “Heart rate extraction from novel neck photoplethysmography signals,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6541–6544, IEEE, 2019

Acknowledgments

First of all, I would like to thank my supervisor Prof. Esther Rodriguez-Villegas for believing in my potential since the beginning and offering me the opportunity to pursue this PhD with her. I am extremely grateful for all the academic motivation and encouragement along the years, but also for the constant personal support. I have learnt a lot from this PhD experience, thanks to your valuable feedback and visionary ideas. Thank you for teaching me so many useful skills that have shaped me into the researcher I am today.

I would like to extend my thanks to all my labmates in the Wearable Technologies Lab: Sukhi, Piyush, Amit, Renard, Xuen, Anas, Mihal, Zhou and Stuart; for helping me with my research work and making this journey enjoyable. Especially, with the many cakes, pizza parties and escape rooms. I owe a special thanks to Wiesia for putting so much love in assisting with all the administrative paper work and baking such delicious banana bread to make everyone happy.

My PhD years would not have been the same without Marina, the sister I never had, and Pablo, my greatest friend and “second” supervisor. There are not enough words to express how lucky I feel to have shared so many priceless moments with you two. The bonding we have forged is so unique, that I can only call you family. Even more than that. You have been the craziest flatmates, best adventure companions and closest friends anyone could ask for. I thank you with all my heart for caring so much about me and persistently offering me your unconditional support. I wouldn’t have been able to accomplish this PhD without you doing it alongside me.

Similarly, I could not imagine these years in London without Patri, Lou and Mire. You have always been there for me in the most difficult moments, and have helped me stand up again. I will never forget all the good memories we have made together, dancing, hiking, going to concerts and trips. I am also very happy to have shared, with Adri, Axel, Juan, Joan and Nico, not only many lunch breaks, gym sessions and Union beers at Imperial, but also a long-lasting friendship.

I would also like to mention Cédric, who was there at the beginning, for encouraging me to start this PhD without fear, and supporting me closely, in the distance.

I consider myself also very lucky to have met other incredible people along the way. Especially the Brixton crew, who has brightened my last year of PhD during these strange times. I am specially

grateful to Hannah and Pieri, for their good vibes, caring and affection. I am also indebted with all the people forming Clapham Lacrosse, Spain Lacrosse and Swing Patrol, for helping me to keep my sanity during these years.

I am also very thankful to all my friends in Spain for always being there. Maria, Rafa, Jose and Lidia, you know very well what doing a PhD is, and I am glad I had you to encourage each other. Many thanks to Inés, Laura, Carol, Miri and Mela for consistently bringing some joy in the hardest times. I cannot forget to mention my cousin Pilar either, for sparkling my PhD with all her visits to London.

Last and most important, I would like to extend my most sincere gratitude to my parents, my brother and family for always encouraging me in the pursuit of all my goals. Words cannot express how grateful I am for your unconditional support, motivation and love. Your persistent help and reassurance means the world to me. I owe you everything. This thesis is for you.

Special thanks to “La Caixa” Foundation (LCF/BQ/EU15/10350027) and the European Research Council (ERC) (grant no. 724334), for funding my PhD research.

Declaration of Originality

I hereby confirm that the work presented in this thesis is my own, and that appropriate references have been used to denote the work of others.

Irene García López

November, 2020.

Copyright Statement

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Contents

Abstract	3
List of Publications	5
Acknowledgments	7
Declaration of Originality	9
Copyright Statement	11
List of Figures	18
List of Tables	19
Terms and Abbreviations	21
1 Introduction	25
1.1 Epilepsy mortality	25
1.2 SUDEP mechanisms and pathophysiology	26
1.3 Risk factors for SUDEP	27
1.4 Risk reduction and SUDEP prevention	28
1.5 Seizure monitoring devices for SUDEP prevention	29
1.5.1 Motor activity sensors	30
1.5.2 Audio detection	31
1.5.3 Cardiac-based detection	32
1.5.4 Respiratory changes detection	32
1.5.5 Electrodermal activity sensors	33
1.6 Motivation	34
1.7 Thesis Structure	36
1.8 Experiments outline	39
References	46
2 A Review on PPG Principles, Artifacts Cancellation and Apnea Detection	47
2.1 Photoplethysmography	47
2.1.1 Photoplethysmography principles	47
2.1.2 Applications: SpO ₂ and heart rate estimation	49
2.1.3 PPG sensors and measurement sites	51
2.1.4 PPG waveform	52
2.1.5 Changes in pulse shape	53
2.2 Overview of PPG artifacts cancellation methods	54
2.3 PPG artifacts detection and classification	56
2.3.1 Experimental protocol and datasets	56

2.3.2	Pulses segmentation and classification windows	60
2.3.3	Features extraction	61
2.3.4	Feature selection	62
2.3.5	Classification methods	62
2.3.6	Discussion	65
2.4	Automatic apnea detection using PPG sensors	66
2.4.1	Apnea types and definitions	66
2.4.2	Studies relevant to the current research	66
2.4.3	Datasets and cohorts	68
2.4.4	Epoch length	71
2.4.5	Signals and features exploited	71
2.4.5.1	SpO ₂ -based approaches	72
2.4.5.2	PPG signal features	73
2.4.6	Classification methods	74
2.4.7	Discussion	76
	References	85
3	Novel Neck PPG vs. Standard Finger PPG	87
3.1	Introduction	87
3.2	Experiment 1 procedures	88
3.2.1	Data acquisition set up	88
3.2.2	Experimental protocol	89
3.3	Comparison of morphological pulse characteristics	91
3.3.1	Average waveforms characteristics extraction and analysis	91
3.3.2	Results: Differences in average PPG waveforms morphological features	92
3.4	Spectral comparison under different breathing conditions	94
3.4.1	Finger and Neck PPG spectrograms differences	94
3.5	Heart rate extraction from novel neck photoplethysmography signals	97
3.5.1	Methods	97
3.5.1.1	Experimental protocol	97
3.5.1.2	Heart beat detection and HR estimation	97
3.5.1.3	Performance metrics	99
3.5.2	Results	99
3.5.2.1	HR estimation errors	99
3.5.2.2	Linear regression fit	100
3.5.2.3	Bland-Altman analysis	101
3.5.3	Discussion	101
3.6	SpO ₂ extraction with neck PPG	102
3.6.1	Experiment 2	102
3.6.1.1	Data acquisition sensors	102
3.6.1.2	Experimental protocol	102
3.6.2	Neck PPG signals processing for R ratio calculation	103
3.6.2.1	Artifacts removal	103
3.6.2.2	Signals conditioning	104
3.6.2.3	Ratio of ratios (R) extraction	104
3.6.2.4	R values processing	104
3.6.3	SpO ₂ estimation results	105
3.6.3.1	Linear regression	105
3.6.3.2	Leave-One-Subject-Out evaluation	105
3.6.4	Discussion	106
3.7	Discussion: Advantages and limitations of neck PPG	107
	References	110

4	Extracting the Jugular Venous Pulse from Anterior Neck Contact Photoplethysmography	111
4.1	Introduction	111
4.2	Methods	114
4.2.1	Experiment 3 - sensors and protocol	114
4.2.2	Sensing location	116
4.2.3	Neck JVP and finger PPG signals correlation analysis	117
4.2.4	Neck JVP waveforms annotations and averaging	117
4.3	Results	118
4.3.1	Anterior neck veins imaging	118
4.3.2	Ultrasound internal jugular vein measurements validation	119
4.3.3	Wavelet coherence between neck JVP and finger PPG signals	120
4.3.4	Neck JVP signals annotation and characteristic waves location	121
4.3.5	Timings between JVP a , c , v waves and ECG and PPG peaks	122
4.3.6	Average JVP waveforms for all subjects	125
4.4	Discussion	126
4.5	Conclusions	129
	References	131
5	Characterization of Artifact Signals in Neck Photoplethysmography	133
5.1	Introduction	133
5.2	Methods	134
5.2.1	Neck artifacts definition	135
5.2.1.1	Intrinsic	135
5.2.1.2	Environmental	136
5.2.1.3	External	136
5.2.2	Experiment 4 - data acquisition	137
5.2.2.1	Experimental set up	137
5.2.2.2	Experimental protocol	137
5.2.2.3	Recordings annotation	138
5.2.3	Features extraction	138
5.2.3.1	Time domain features extraction	139
5.2.3.2	Correlogram features	141
5.2.3.3	Frequency domain features extraction	141
5.2.4	Artifacts vs. Normal PPG statistical comparison	143
5.2.4.1	Statistical tests evaluation	143
5.2.4.2	Characterization significance matrix construction	143
5.2.4.3	Ranking of features for each artifact	144
5.2.4.4	mRMR feature selection	145
5.3	Results	145
5.3.1	Features ranking based on statistical results	145
5.3.2	Artifacts characterization	146
5.3.3	Characterization of additional respiratory states of interest	150
5.4	Discussion	151
5.5	Conclusion	153
	References	154
6	Artifacts Classification and Apnea Events Detection in Neck Photoplethysmography Signals	155
6.1	Introduction	155
6.2	Methods	156
6.2.1	Experimental protocol	156

6.2.2	Features extraction	157
6.2.2.1	Windows segmentation and labelling	157
6.2.2.2	Features	157
6.2.3	Classification pipeline	159
6.2.3.1	Data partition	160
6.2.3.2	Training	161
6.2.3.3	Features selection	162
6.2.3.4	Hyperparameters optimization	162
6.2.3.5	Performance metrics and model selection	163
6.2.4	Statistical evaluation of the classification results	163
6.3	Results	163
6.3.1	Classification results	163
6.3.1.1	Artifacts classification	164
6.3.1.2	Apnea classification	165
6.3.2	Statistical tests results:	165
6.3.2.1	Two-way ANOVA	165
6.3.2.2	Window length effect	166
6.3.2.3	Threshold of corruption (%) effect	167
6.3.3	Features selection results	168
6.4	Discussion	169
6.5	Conclusion	173
	References	176
7	Conclusions	177
7.1	Contributions	177
7.2	Future work	181
7.2.1	Areas of improvement of this PhD	181
7.2.2	Towards a neck PPG wearable apnea detection device	182
7.2.2.1	Artifacts algorithm validation	182
7.2.2.2	Apnea algorithm validation in patients	182
7.2.2.3	Low power implementation and WADD sensor fusion	182
7.2.2.4	Clinical trials	183
A	Ethics approval documents	185
A.1	Joint Research Compliance Office Approval Letter	186
A.2	Experimental Protocol	187
A.3	Participant Information Sheet	198
A.4	Informed Consent Form	201

List of Figures

2.1	Absorption spectra of oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb).	48
2.2	Light absorption by different tissues reflects the AC and DC components of the PPG signal.	48
2.3	Contact transmission and reflectance optical modalities to measure the PPG signal at the finger.	48
2.4	Diagram of the standard PPG waveform annotated with its main pulse characteristics.	53
2.5	Obstructive, Central and Mixed apnea recordings.	67
2.6	Annotated diagram of a SpO_2 desaturation with some common characteristics.	72
3.1	Pulse oximetry sensors used for data acquisition.	89
3.2	Polysomnography SOMNOscreen Plus monitoring system.	89
3.3	(a) Normal, (b) Breathing Fast, (c) Breathing Slow, (d) Breathing Apnea, recordings for one subject showing the finger PPG, neck PPG, ECG and respiratory channels. . .	90
3.4	Average finger and neck PPG pulse waveforms obtained for one participant.	91
3.5	Average pulse waveform and its second derivative, for diastolic and dicrotic notch detection.	92
3.6	Spectrograms of finger and neck PPG signals under different breathing conditions: normal breathing, fast breathing, slow breathing and apnea (breath-holding).	95
3.7	Substraction of the finger and neck PPG spectrograms for the different breathing conditions.	96
3.8	Inter-beat distances over time of a breathing normal recording, before and after spurious peaks correction.	98
3.9	(a-c). Linear regression of estimated HR_{neck} and HR_{finger} against HR_{ECG} . (b-d). Bland-Altman plots for all HR estimates.	100
3.10	Wearable neck PPG system used for SpO_2 calibration experiments.	103
3.11	SpO_2 (%) regression with calculated R values.	106
4.1	ECG, JVP and PPG signals traces.	112
4.2	Experimental setup.	115
4.3	Neck venous system anatomy. (a) Superficial veins schematic of the inferior anterior region of the neck. (b) Schematic of the venous tree connecting the external (EJV), internal (IJV) and anterior (AJVs) jugular veins to the right atrium of the heart. . . .	116
4.4	Anterior view of neck veins (subject 12).	118
4.5	Anterior neck veins vasculature imaging for all participants (n=20).	119
4.6	Validation of the proposed neck JVP-PPG signals with reference ultrasound (US-JVP) for three different subjects.	120
4.7	Recordings showing neck JVP and finger arterial PPG signals for all subjects (n=20). .	121
4.8	Wavelet coherence plot for neck inverted JVP and arterial finger PPG signals.	121
4.9	Manual a , c , v waves annotation of neck JVP signals with respect to ECG and PPG reference signals for 10 subjects.	123
4.10	Location of the annotated a , c , v , R and S peaks within a JVP cycle.	123

4.11	Boxplots showing the distributions of average time differences between characteristic peaks and waves from neck JVP, finger PPG and ECG (n=20).	125
4.12	Average JVP waveforms manually annotated for all participants (n=20).	126
5.1	Extraction of some frequency domain features from the spectrogram of a yawn recording.	142
5.2	Results of pairwise multiple comparisons performed after repeated measures one-way-ANOVA statistical test.	144
5.3	Ranking of features for each independent artifact and breathing state to discriminate against normal PPG, based on the p-values obtained from the characterization matrix	146
6.1	Classification pipeline. The predicted output classes for artifacts classification are: Artifacts/clean PPG; for apnea classification: Apnea/Normal PPG.	160
6.2	Data partitions for artifacts classification and apnea detection models.	161
6.3	Average classification results for the proposed artifacts and apnea classification algorithms, over the 30 randomization experiments.	164
6.4	Classification decision results of the best models for one head movement artifact and an apnea event.	166
6.5	Means plots of the classification performance metrics across different thresholds of corruption (%).	167
6.6	Features selection ranked by frequency of occurrence over the 30 randomization experiments for the best artifacts and apnea classification models.	169

List of Tables

2.1	Artifacts classification algorithms discriminating between good and bad quality PPG signals.	57
2.2	Apnea detection algorithms: SpO ₂ approaches.	69
2.3	Apnea detection algorithms using the PPG signal.	70
3.1	Wilcoxon Signed-Rank Test Pairwise Comparisons between Finger and Neck for the Different Features (*p<0.05)	93
3.2	Comparison of HR estimation performance between neck and finger PPG	100
4.1	Average time differences between characteristic peaks and waves from neck JVP, finger PPG, and ECG signals.	124
5.1	Classification of neck artifacts occurrence in sleep conditions: based on severity and frequency of incidence.	137
5.2	Experimental protocol: List of artifacts and breathing states tested with corresponding recording time sequences (in seconds)	138
5.3	Features definitions and formulas	140
5.4	Characterization of Artifacts and Breathing States of Interest	148
6.1	Average performance results ($\mu \pm \sigma$, n=30) for the best artifacts and apnea classification models	165
6.2	Comparison of artifacts classification results in the literature with our best (W=6s-Thd=20%) model	171
6.3	Comparison of apnea classification results in the literature with our best (W=10s-Thd=50%) model	172

Terms and Abbreviations

Term	Description
A1	Systolic Area of the PPG pulse
A2	Diastolic Area of the PPG pulse
AC	Alternating Current
ACC	Accuracy
AED	Anti-epileptic drugs
AHI	Apnea–Hypopnea Index
AI	Augmentation Index
AJV	anterior jugular veins
ANN	Artificial Neural Network
ANOVA	Analysis of variance
BMI	Body Mass Index
BP	Blood pressure
BPM	Beats-per-minute
CA	Carotid Artery
CFSA	cycle-by-cycle Fourier serial analysis
CNN	Convolutional Neural Network
CSA	Central Sleep Apnea
CTM50	Central Tendency Measure with radius 0.5
CVD	Cardio-vascular diagnosis
CVP	Central Venous Pressure
DAP	Decreases in Amplitude
DC	Direct Current
DWT	Discrete Wavelet Transform
ECG	Electrocardiography
EDA	Electrodermal activity
EEG	Electroencephalography
EJV	external jugular vein
EMG	Electromyography
F1	F1-score
FFT	Fast Fourier Transform

GTCS	Generalized tonic-clonic seizure
HR	Heart Rate
HRV	Heart Rate Variability
Hb	Deoxyhemoglobin
HbO₂	Oxyhemoglobin
ICA	Independent Component Analysis
ICU	Intensive Care Unit
IP	Impedance Photoplethysmography
IPA	Inflection point area
IQR	Interquartile range
IR	Infrared
JV	Jugular Vein
JVA	Jugular Venous Arch
JVP	Jugular Venous Pulse
LDA	Linear Discriminant Analysis
LED	Light-Emitting Diode
LMS	Least Mean Square
LOSO-CV	Leave-One-Subject-Out Cross-Validation
LVET	Left Ventricular Ejection Time
LoA	Limits of Agreement
MA	Moving Average
MAE	Mean Absolute Error
MAF	Moving Average Filter
ML	Machine Learning
MLP	Multilayer perceptron
MSA	Mixed Sleep Apnea
NMIFS	Normalized Mutual Information Feature Selection
NN	Neural Network
ODI	Oxygen Desaturation Index
OSA	Obstructive Sleep Apnea
PGES	Post-ictal Generalized Electroencephalographic Suppression
PPG	Photoplethysmography
PPT	Pulse Propagation time
PRSA	Phase Rectified Signal Averaging
PRV	Pulse Rate Variability
PSD	Power Spectral Density
PSG	Polysomnography
PTT	Pulse Transit Time
PTTV	PTT Variability
QDA	Quadratic Discriminant Analysis

R	Ratio of ratios
RA	Right Atrium
RBF	Radial Basis Function
RMSE	Root-Mean Square Error
RR	Respiratory Rate
SAHS	Sleep Apnea-Hypopnea Syndrome
SDAE	Standard Deviation Average Error
SDB	Sleep disordered breathing
SE	Sensitivity
SI	Stiffness Index
SP	Specificity
STFT	Short-time Fourier Transform
SUDEP	Sudden unexpected death in epilepsy
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SWT	Stationary Wavelet Transform
SpO₂	Peripheral saturation of oxygen in blood
TFS	Time-frequency spectrum
US	Ultrasound
VFCDM	Variable Frequency Complex Demodulation
WADD	Wearable Apnea Detection Device
cVCGI	Camera-based skin micro-motion Vibrocardiography Imaging
iPPG	Imaging Photoplethysmography
kNN	k-Nearest Neighbour
mRMR	minimum Redundancy Maximum Relevance

Chapter 1

Introduction

1.1 Epilepsy mortality

Epilepsy is a neurological condition that causes sudden bursts of electrical activity in the brain, called seizures. It affects more than 70 million people worldwide [1, 2], and the UK alone is home to 500,000 cases [3]. The symptoms vary depending on the type of seizure and the part of the brain that is affected, but they generally include uncontrollable convulsions, going blank, becoming stiff and even collapsing. Epilepsy can be a lifelong condition, with the highest incidence in infants and elderly age groups [4]. With the appropriate treatment, 70% of the patients could be seizure-free [5]. However, 80% of people with epilepsy live in low-income countries and do not have proper access to antiseizure treatments, nor to a complete diagnosis [6]. This greatly increases the premature mortality rate among epileptic patients.

Epilepsy-related deaths can be caused by status epilepticus, accidents and suicide [7]. In the UK, there are over 1000 deaths per year, which corresponds to the shocking number of 21 per week [8, 9, 10, 11]. However, in the majority of the cases, no obvious cause of death can be found post-mortem. This is described as sudden unexpected death in epilepsy (SUDEP), which often happens during sleep when less assistance is available. SUDEP is responsible for up to 9 deaths per 1000 patients each year [12], with the highest incidence in young adults (20-45 years old) [13]. It is very difficult to characterize such a phenomenon, but a widely adopted definition for SUDEP is: *“a sudden, unexpected, witnessed or not witnessed, non-traumatic and non-drowning death in patients with epilepsy, with or without evidence of a seizure and excluding documented status epilepticus, in which postmortem examination does not reveal a toxicological or anatomical cause of death”* [14].

1.2 SUDEP mechanisms and pathophysiology

The exact mechanisms leading to SUDEP are still unknown and unpredictable since very few cases have been possible to monitor. Indeed, the Mortality in Epilepsy Monitoring Units Study (MORTMUS) [15], was the first retrospective work evaluating the incidence and mechanisms of cardiorespiratory arrests encountered in 11 epileptic patients in the last moments preceding SUDEP. Generalized tonic-clonic seizures (GTCS), causing convulsions and severe alteration of the cardiorespiratory function, were found to trigger the physiological storm leading to death in all patients. A recurrent pattern indicated that many patients suffered from short tachycardia and tachypnea immediately after the seizure offset, which developed in a profound apnea and bradycardia. In some cases, patients died instantaneously from a terminal apnea and asystole, while in others, death occurred after a few intermittent respiratory efforts lasting up to 18min.

Despite this consistent postictal pattern, there is strong evidence suggesting that SUDEP pathophysiology results from an interplay between cardiovascular, respiratory and brainstem dysfunction, as well as autonomic dysregulation and arousal failure [16, 17]. The exact succession of events is poorly understood, but postmortem studies have provided insights into the potential causes, contributors or consequences of seizures, leading to death.

Cardiac dysfunction during seizures is a well-known phenomenon that results in arrhythmia of different types, such as tachycardia, bradycardia, asystole or atrioventricular conduction blocks [18]. Cardiac structural changes are commonly found in SUDEP cases, including myocyte hypertrophy and myocardial fibrosis abnormalities [19, 20]. Transient left ventricular dysfunction, in the form of Takotsubo cardiomyopathy [21] or neurogenic stunned myocardium [22], were also observed in living epilepsy patients. This was shown to be induced by GTCS or focal seizures, as a consequence of excessive sympathetic stimulation.

Respiratory dysfunction is also well-studied in the SUDEP literature, identifying seizure-related apneas as essential biomarkers [23]. The propagation of the seizure into the limbic areas of the brain controlling respiration results in the loss of respiratory drive. Prolonged periods of apneas can then dangerously disturb the oxygen supply to vital organs and lead to profound hypoxemia. Several studies have demonstrated that postictal apneic oxygen desaturations play a crucial role in SUDEP [24, 25, 26]. Even though hypoxemia is generally of central origin, upper airways obstruction and ictal laryngospasm can be other precipitating factors leading to asphyxiation in SUDEP, specially when patients are found in a prone position in bed [27, 28]. Post-mortem studies have similarly revealed pulmonary structural changes, following GTCS, that could be additional important contributors to postictal hypoxia [19]. Pulmonary edema, i.e. an accumulation of fluid in the lungs, is commonly found in autopsies [29, 20]. Several hypotheses point that a massive sympathetic discharge induced by the seizure, causes vascular

vasoconstriction and increases hydrostatic pressure [30]. This ultimately shifts an increased volume of blood from the systemic to the pulmonary circulation, generating pulmonary congestion. Others, support the idea that pulmonary edema is secondary to central respiratory inhibition leading to hypoxia [31]. Pulmonary edema was also postulated to be an epiphenomenon of prolonged seizures [32].

Postictal generalized electroencephalographic suppression (PGES), also appeared as an important SUDEP biomarker [33, 34, 35]. This global depression of the cortical electrical activity was widely present in all SUDEP patients evaluated in the MORTEMUS study [15]. PGES is significantly associated with GTCS and severe hypoxemic respiratory dysfunction [34, 36]. In fact, the longer the duration, the more profound the brainstem impairment, and hence the greater the risk of terminal apnea leading to SUDEP [33]. Other consequences of PGES are the loss of consciousness, postictal immobility, and the impairment of the arousal response. The lack of these recovery reflexes could be critical when patients with airway obstruction are in the prone position, not permitting their self-repositioning [37, 38].

Seizures, and in particular GTCS ones, often lead to irreversible brainstem dysfunction, causing autonomic system dysregulation. Besides cardio-respiratory impairment, this can be noticeable by a reduced heart rate variability (HRV) and changes in transpiration [39]. Indeed, research demonstrated that low HRV ratios during the awake state, and extremely high or extremely low HRV during the sleep to awake phases, were linked with SUDEP [40]. Some studies observed an increase in electrodermal activity (EDA), as a result of postictal sympathetic activation, which was directly correlated with a prolonged duration of PGES [41, 42, 43]. Although HRV and EDA abnormalities are far from being lethal, they can be used as potential biomarkers for autonomic dysfunction, which in turn is a surrogate of SUDEP.

To summarize, SUDEP can be considered as a multi-factorial phenomenon, in which several mechanisms come into play. There is not a unique hypothesis explaining it, but all the dysfunctions covered, could, individually or collectively, be responsible for the lethal outcome.

1.3 Risk factors for SUDEP

Based on SUDEP pathophysiology and epidemiology, various factors have been associated with an increased risk of SUDEP.

Static risk factors include young male gender [13], high frequency of tonic-clonic (convulsive) seizures (≥ 3 GTCS/year) [44], early onset of epilepsy before 16 years old and disease duration for at least 15 years [45, 46]. Mental health problems and intellectual disability were also shown to increase mortality [47, 48].

Genetic factors can also play a big role in SUDEP. These include mutations of genes that regu-

late sodium and potassium ion channels, which are widely present in the heart and the brain [40]. Channelopathies are typically correlated with cardiac comorbidities, which increase the likelihood of cardiac arrest in the postictal phase [40]. A well-studied example is the Dravet syndrome, an epileptic encephalopathy that predisposes patients to cardiac arrhythmias in addition to seizures [49]. Mutation of the SCN1A gene results in parasympathetic hyperactivity, which is reflected by decreased HRV and P wave dispersion [50].

Modifiable risk factors include the management of seizures and treatments, as well as the patients' lifestyle and environment. Having more than 13 seizures in the last year, not being seizure-free in the last 1-5 years and suffering 11-20 convulsive seizures in the last 3 months, are prospect determinants of SUDEP [51, 52]. Nocturnal seizures significantly augment the risk by threefold, partially due to the lack of surveillance [53]. Adherence to anti-epileptic drugs (AED) is essential to reduce the number of GTCS and hence prevent SUDEP. However, the prescription of an excessive number of AEDs, denoting poor seizure control, was also identified as a risk factor [45]. Sleep deprivation, alcohol, smoking or anxiolytics misuse are other lifestyle aspects that enhance seizure occurrence [47]. Sleeping in a prone position in combination with postictal immobility, make epileptic patients more susceptible to die of asphyxiation.

Even though static and genetic factors are hardly preventable, modifiable risks could be avoided by designing appropriate seizure reduction and intervention strategies.

1.4 Risk reduction and SUDEP prevention

The fatal consequences of GTCS demonstrate the necessity of taking preventive actions to reduce epileptic patients' mortality. SUDEP could be significantly prevented by different means [54].

First and simply, doctors and caregivers should properly inform epileptic patients of the risk factors leading to SUDEP, no matter how sensitive or distressful the conversations might be [55]. By improving the lifestyle and adherence to AED therapy, the management of the disease could be positively impacted. Besides these awareness conversations, there are some questionnaires, like the SUDEP and Seizure Safety Checklist [56] or the self-monitoring EpSMon app [57], that facilitate the risk assessment and the control of seizures over time. Adoption of these tools could help clinicians take informative decisions on the prescription of the most suitable treatment for each patient. However, at present there is little evidence of their use.

Anti-epileptic drugs are the first option for seizure freedom and reduction of SUDEP incidence [58]. These are, however, only effective in two-thirds of the population [5]. Drug-resistant patients at high risk are typically referred for resective epileptic surgery, which can greatly decrease the SUDEP rates if successful [59, 60]. Other palliative solutions, such as vagal nerve stimulation (VNS) [61] or responsive

neurostimulation [16] have also been proposed recently for patients that are not good candidates for surgery. Even though only a few studies have been published, these alternatives have provided satisfactory evidence to reduce the risk of SUDEP. Despite the efforts for seizure control and reduction, there are still many patients with frequent GTCS, that are not suitable for AED treatment or do not have access to invasive interventions.

At present, there are no effective solutions to SUDEP apart from preventive sleep supervision and constant physiological monitoring [62, 63, 64]. Indeed, many studies have demonstrated that nocturnal supervision by an individual or listening device greatly reduce the risk of SUDEP [62, 65]. Early assistance by nurses, caregivers, or relatives could be a lifesaver for epileptic patients found breathless or in a prone position, following a seizure. Peri-ictal interventions, including direct repositioning in bed, somatosensory stimulation or cardio-respiratory resuscitation could alleviate respiratory dysfunction, reduce PGES duration, and hence prevent SUDEP [66, 67]. Supplemental oxygen administration can also help in the recovery from early hypoxemia, by promoting arousal and stimulating respiration. However, it is controversial whether it is effective, especially in cases of central apnea and pulmonary edema [68].

Outside of epilepsy monitoring units, sleep supervision by room companions or frequent checks by relatives, could be tedious and overall unpractical on a night-to-night basis. Moreover, residential care settings or caregiver options when the patient lives alone, become very costly. To counteract these drawbacks and offer some autonomy to patients, anti suffocation lattice pillows have been proposed in the home environment to prevent asphyxiation [69]. However, very limited evidence supports their success in reducing SUDEP [68].

Continuous physiological monitoring, appears as a more reliable alternative, by scanning for biomarkers indicative of disease and automatically releasing alarms that warn of life-threateningly situations. An extended overview of the current state of the art of seizure monitoring devices for SUDEP prevention is presented as follows.

1.5 Seizure monitoring devices for SUDEP prevention

Video electroencephalography (v-EEG) is the gold standard for diagnosis of epilepsy and seizure-type characterization, due to its ability to detect most ictal events unequivocally [70]. This sensing modality measures, synchronously to video imaging, the brain electrical activity, by means of electrodes attached to the scalp or implanted invasively. Head-mounted EEG devices, do typically include a large number of bulky wired sensors, that cause discomfort to the user and can critically impair the quality of sleep. These cannot be easily placed in the home environment without trained personnel, making EEG unsuitable and unpractical for continuous monitoring on a night-to-night basis. Invasive EEG

electrodes are unobtrusive, but the risk of implantation is very high.

In this framework, non-EEG seizure detection devices, that are non-invasive, adapted for ambulatory monitoring and capable to trigger an alarm, are preferred. A wide variety of systems, exploiting different physiological parameters involved in SUDEP pathophysiology, exists [71, 72, 73]. These mainly focus on cardiovascular and respiratory changes, as well as on noise, electrodermal and motor activity.

1.5.1 Motor activity sensors

Motion-based seizure detection has been of primary interest in the literature, due to the simplicity of identifying convulsive movements that frequently arise in lethal GTCS. Different modalities have been explored to measure this characteristic signature.

Automated video approaches allow real-time image processing for quantitative movement analysis during seizures [74, 75]. The sleep activity monitor SAMi-3 (by the company SAMi Alert) exploits a wireless video camera coupled with infrared illuminators to automatically detect unusual movements during sleep and alert relatives [76]. A constant CCTV based epilepsy monitoring service is also offered by the company EpiView. An important issue of this is, however, the patient's privacy. Even though video detection is a practical contactless option that in some cases could even measure heart rate (HR) and respiration [77], it constraints the patient within the scope of the camera. Moreover, video-based systems are not safe as they tend to erroneously detect normal movements, or conversely, miss convulsions occluded by bedding.

Similarly, pressure sensors placed under the mattress, measuring the absence of motion or abnormal movements, did not show sufficient performance for SUDEP prevention [72]. Some commercial products, implementing under-mattress pressure mats in combination with other sensing modalities, are Aremco (by Aremco), Ep-iT (by Alert-iT), Emfit (by Emfit Ltd) and MP5 (by Medpage). The Emfit pressure mattress, for instance, was tested in several studies, showing variable accuracy in seizures detection from 30% [78] to 89% [79]. Despite the widespread adoption among epileptic patients due to its simplicity, this option is not very reliable.

Electromyography (EMG) devices measuring muscle electrical activity are promising approaches to detect the tonic phase of GTCS [80]. Some wearable FDA-cleared devices include the Sensing Portable sEMG Analysis Characterization system (SPEAC, by Brain Sentinel) and the Epileptic seizure Detector Developed by IctalCare (EDDI, by IctalCare). The Eddi patch, attached to the biceps, demonstrated good performance in detecting GTCS with a sensitivity of 93.8% and false alarm rate (FAR) of 0.67 per day [81]. Similarly, the SPEAC portable system was able to detect all GTCS and reported a FAR of 1.44 per 24h in a prospective multi-centre phase III trial [82]. However,

a disadvantage of EMG electrodes is the potential skin irritation, which in the SPEAC study led to adverse events in 28% of the subjects.

Accelerometry-based seizure detection is probably the most widely used in SUDEP prevention devices. Accelerometers (ACM), gyroscopes and magnetometers, measuring acceleration, orientation, and position; are low-cost sensors that can be easily integrated into miniaturized systems for ambulatory monitoring. These are capable of identifying when the patient is in the prone position or can detect movement during seizures. ACM are typically exploited, on their own or in combination with other sensing modalities, in smartwatches and bracelets. Some ACM-based bracelet systems include the EpiLert (by BioLert) and the Epi-Care Free (by Danish Care). These demonstrated good sensitivities of 91% [83] and 89.7% [84] respectively, with FAR of 0.11/day [83] and 0.2/day [84].

Some seizure monitoring apps, intended to be installed in existing fitness tracker devices like the Apple Watch, have also been proposed. Some examples include the SmartWatch Inspire (by Smart Monitor), SeizAlarm (by SeizAlarm), EpiWatch (by Johns Hopkins University), or My Medic Watch (by My Medic Watch Pty Ltd). These mainly make use of ACM and cardiac sensing sensors present in smartwatches to detect repetitive shaking motion, falls and abnormal heart rates. These apps require the coupling with a smartphone, to notify the designated caregivers or relatives of a dangerous seizure event. The patient's location is then shared for prompt intervention. An additional feature of the EpiWatch is the display of a cognitive test on the screen, to evaluate the patient's responsiveness. Despite the attractiveness of these user-friendly apps, a study testing the SmartWatch Inspired in a pediatric cohort for various seizure types reported that only 16% of the total seizures and 31% of the GTCS were detected [85]. This is indeed the overall disadvantage of motion-based seizure detection systems. They mainly target GTCS involving convulsions and are unable to accurately identify other types of seizures.

1.5.2 Audio detection

Some audio systems have also been proposed to measure noises like seizure vocalizations and screams, lip biting or bed noises. Some under-mattress commercial products incorporating audio recording include the MP5 (by Medpage) [86, 87], Aremco (by Aremco), Ep-It (by Alert-It) and the video-based SAMi-3 (by SAMI Alert) sleep activity monitor. Audio-based systems offer the advantage of being low-cost and ensure the patients' comfort, but are generally characterized by poor performance and a high number of false positives. To illustrate, the MP5 device placed between the bed and the mattress, showed a sensitivity of 63% and a precision (PPV) of 3% in the detection of 8 tonic-clonic seizures [86].

Seizure detection approaches based on physiological measurements tend to be more accurate and reliable.

1.5.3 Cardiac-based detection

Heart rate monitoring sensors are widely used in seizure detection, as they can provide useful information regarding cardiac dysfunction in SUDEP. These mainly include electrocardiography (ECG) and photoplethysmography (PPG) modalities, nevertheless the Emfit system (by Emfit Ltd) also claims to exploit ballistocardiography.

A miniaturized ECG wireless system, attached to the arm and measuring cardiac electrical activity with electrodes placed on the chest, was developed by Massé *et al.* [88]. The integration of ACM to this prototype resulted in an acceptable sensitivity (SE) in the range of 71-87%, and a high FAR of 2.3–5.7 per night, when tested against different types of seizures [89]. This system evolved into the commercial upper arm bracelet NightWatch (by LivAssured), which implemented a PPG sensor, this time instead, to monitor heart rate changes [90]. The seizure detection performance of this device exploiting HR and movements features, resulted in a reduced FAR of 0.03 per night, but a PPV value of 49%, in a long term prospective study [91]. Another commercial product similar to the preliminary prototype is the ProGuardian system (by Cyberonics) consisting of a chest-worn ECG system with ACM. Seizure detection performance resulted in a SE larger than 80% but a high number of 2 false alarms per night [92]. Cardiac-worn sensors are also found coupled to EEG in the smart-clothing Neuronate product (by Bioserenity) under current development [93].

However, the most common way of measuring heart rate changes is through portable wristbands incorporating PPG sensors. Some commercially available products are the PulseGuard (by PulseGuard International Ltd) [94], Pulse Companion (by Alert-iT), Brio System (by Possum) and the E4 (by Empatica) bracelets. The latter is only intended for research and does not include any embedded seizure detection algorithm. However, PPG algorithms trained on data collected by the E4 have reported very low sensitivity of 32% in the detection of frontotemporal lobe seizures [95].

1.5.4 Respiratory changes detection

Respiratory monitoring is probably the most important SUDEP prevention strategy, since central respiratory arrests, hypoxemia and terminal apneas are usually the primary triggers of the physiological cascade of lethal events. Measurement of chest and abdominal movements denoting respiratory efforts, can provide useful information about ventilation. Respiratory inductance plethysmography (RIP) straps around the torso, EMG sensors on the intercostal muscles or ACM on the chest or placed in the bed, are some sensing options to record breathing. However, RIP bands, for example, cause discomfort to patients and tend to move considerably during sleep. Several under-mattress multi-modal commercial monitors already mentioned (Aremco, Ep-iT and Emfit), also measure respiratory movements by means of bed-mounted sensors. However, their performance is questionable [78].

Tracking oxygen saturation in blood (SpO_2) could be a more reliable alternative, however, no pulse oximeter-based systems have been marketed to date for epileptic seizure detection. This is probably due to the delayed response of SpO_2 desaturations, which could increase the latency of apnea detection and hence prevent prompt intervention. However, some research studies measuring SpO_2 changes with an oximetry finger cuff coupled to a wrist-worn control unit [96, 97], showed an accuracy of 86% and a specificity of 73%. This seizure detection algorithm combined SpO_2 , HR and EDA features [97].

A miniaturized wearable breathing sensing system, placed externally over the trachea, was developed in our laboratory to acoustically sense airflow and instantaneously detect apnea events [98]. A later version of this prototype, the Wearable Apnea Detection Device (WADD), was tested in a pilot study in comparison to a standard polysomnography system [99]. A sensitivity of 88.6% and a specificity of 99.6% in detecting apnea events showed promising applicability for SUDEP prevention.

1.5.5 Electrodermal activity sensors

Changes in the autonomic nervous system preceding SUDEP, are evident through increased transpiration. Seizure-detection devices exploiting this aspect, mainly focus on measuring changes in skin conductance, i.e. electrodermal activity (EDA). The stand-alone alert system Embrace (by Empatica) implements EDA and ACM sensors in a wrist-band for GTCS detection. The identification of biomarkers by the embedded machine learning algorithm, triggers an alarm in the mobile app, which in turn alerts caregivers for assistance. The Empatica company has iteratively improved the Embrace system by testing it in wider cohorts and outpatient settings [100]. This has led to very high sensitivities in the range 92-100% and reduced FAR from 2 down to 0.2-1/day. Despite the promising results, EDA changes appear slower than cardio-respiratory changes during seizures [97], which could be a disadvantage for prompt detection.

In summary, despite the general efforts to explore different non-EEG sensing modalities for seizure detection, several review papers argue that the effectiveness of current devices in reducing SUDEP mortality is still limited and needs further improvement [54, 72, 101]. In fact, many devices have only been tested in a very small number of seizures, arising in a reduced cohort of participants, which were monitored in highly controlled inpatient settings [72, 101]. Other commercial systems, are under ongoing development and have not even published any performance results yet. Moreover, current systems mainly focus on identifying movements characterizing GTCS, but lack the ability to detect non-convulsive focal seizures [101, 102]. Even though the sensitivity is very promising for certain devices, it is still very variable across different modalities (2-100%) [71]. Another critical issue is the excessive FAR that is extremely disruptive for patients and intolerable for caretakers. Even in the best-

case scenarios, for which near excellent performance is obtained, the absolute number of false alarms still translates into several per week, which is unacceptable if medical intervention is required [101]. Alarm fatigue would then severely impact the user’s adherence. Multimodal signal acquisition systems might be the future research direction for false alarm suppression [73, 100].

1.6 Motivation

To overcome present limitations, there is the necessity for a user-friendly wearable technology for automatic long-term sleep monitoring. This non-invasive wearable sensor should be comfortable and durable. Ideally, it would sense as many physiological parameters as possible from just one location in the body. It must provide at the same time reliable signals, in order to extract accurate biomarkers for the correct identification of potentially dangerous situations. Specifically, since cardiovascular effects are not the major cause triggering SUDEP, but death is most often due to pulmonary dysfunction, it would be ideal to focus on recording respiratory signals. In addition, the number of false positives should be minimized as much as possible, otherwise an excessive number of false alarms would quickly result in patients’ non-compliance.

Given these specifications, the Wearable Apnea Detection Device (WADD), sensing acoustic breathing signals at the neck, was developed in our laboratory, as already mentioned. Despite the successful trade-off of this wearable solution between comfort and accuracy, the sensitivity of 88.6% could still be critical for epileptic patients. Even though airflow is the key physiological signal to track apneic respiratory arrests, it is very risky to uniquely rely on a single sensor for apnea detection in SUDEP. Therefore, in order to avoid neglecting life-threatening apnea events, it would be suitable to integrate an oximeter sensor into the system to support the decision. This would offer the possibility to simultaneously track oxygen saturation ($\text{SpO}_2\%$) levels and HR, besides just sensing acoustic sounds from the respiratory system [98].

However, the neck has never been thoroughly evaluated as a measurement site for pulse oximetry since it could appear sub-optimal for sensor attachment and patient’s aesthetics. Indeed, the gold standard measurement site is the finger as it provides a stable clip-type attachment and is well perfused by a large bed of capillaries. Other body locations such as the wrist, earlobe and forehead have alternatively been explored in the literature, but none of these typical options offer access to airflow measurements. This anatomical rationale demonstrates why studying neck PPG can have a great impact in the detection of critical terminal apneas in the context of SUDEP. But also for a wide number of respiratory monitoring applications such as sleep apnea, neck PPG could be particularly beneficial to minimize the number of sensors and improve patients comfort.

Preliminary proof-of-concept neck systems have only focused on obtaining the neck PPG signal and SpO₂% [103], as well as determining the most stable position for signal acquisition [104]. But, in order to properly assess its feasibility and utility, a complete characterization of this novel PPG measurement site is necessary. Even though it is a very attractive location for multipurpose monitoring, the neck presents numerous challenges. Being an unexplored sensing position, a completely new variety of artifacts corrupting the PPG signal might arise. Degradation in the quality of the signal could lead to an erroneous estimation of HR and SpO₂% levels. This could, in turn, hinder the detection of apnea events, or increase the number of false alarms. To ensure, good PPG signals quality, accurate readings, false alarms suppression, and hence patients safety; an in-depth study of the unknown sources of noise artifacts in neck PPG must be performed first, to further develop an algorithm to detect and remove them.

As a consequence, artifacts-free PPG signals could enable reliable estimation of SpO₂%, which in fusion with the existing WADD technology, could enhance the detection of apneas. This added feature to the wearable system could ultimately have a significant impact in reducing SUDEP mortality, by automatically alerting of dangerous cardiorespiratory dysfunction. Similarly, this wearable system could enable reliable constant physiological monitoring in the home environment, and hence help investigate the mechanisms and risk factors of SUDEP. This would in turn greatly improve the disease management, treatment and prevention strategies.

This PhD thesis is framed in the scope of this challenge. The general research goal is to characterize the novel neck PPG signals, for subsequent processing and ultimate exploitation in cardiorespiratory monitoring applications. This fundamental aim can be further subdivided into a list of research sub-objectives:

- What are the singularities, advantages and limitations of neck PPG signals compared to the gold standard finger PPG?
- Is neck PPG a suitable pulse oximetry location for apnea detection and other potential cardiorespiratory applications?
- What are the sources of interference specific to the novel neck pulse oximetry region and how can they be characterized?
- How can neck PPG artifact signals be detected for future removal?
- Is it feasible to reliably detect apnea events in near real-time with neck PPG signals for SUDEP prevention? If so, how?

1.7 Thesis Structure

This thesis is organised into six major chapters. [Chapter 2](#) is a theoretical chapter that covers the background and literature review to put readers into context. [Chapter 3](#) and [Chapter 4](#) are two technical chapters studying the singularities and potential cardiorespiratory applications of novel neck PPG signals. [Chapter 5](#) and [Chapter 6](#) are two other technical chapters focusing on the processing of neck PPG signals for artifacts classification and apnea detection.

The main contributions of each chapter are summarized below.

Chapter 2 - A Review on PPG principles, Artifacts Cancellation and Apnea Detection presents the fundamentals on photoplethysmography, as well as a literature review on PPG noise cancellation, artifact classification and apnea detection.

First, the general background on pulse oximetry is reviewed. The optical foundations of PPG measurement are introduced, covering the two sensing modes: transmission and reflection. The main morphological traits of single PPG pulses are described and correlated to the corresponding cardiac physiological events. The physiological factors modifying the standard shape of the signal are subsequently analyzed. A review of existing PPG measurement sites in the literature is also presented, as well as the formulation of the two vital signs parameters commonly extracted in PPG: HR and SpO₂%. A brief overview of artifacts cancellation techniques in the literature follows. An in depth review of artifacts classification methods exploiting machine learning techniques is extended. At last, a review of automatic apnea detection algorithms, uniquely using PPG sensors, is covered in this chapter.

Chapter 3 - Novel Neck PPG vs. Standard Finger PPG The finger is the gold standard measurement site in pulse oximetry. This chapter presents a comparison between finger and neck photoplethysmography (PPG) in order to assess the potential and limitations of this, non-conventionally used body site, for application in pulse oximetry.

PPG signals were recorded at both sites in a series of experiments, in which healthy participants were asked to breathe at different respiratory rhythms, such like normal, fast, slow and breath-holding (apnea). The differences in average PPG waveforms, frequency content, as well as in oxygen saturation (SpO₂%) and heart rate (HR) estimation, were inspected. In terms of pulse waveform statistical analysis, the results showed significant differences in the average PPG pulse shape for different contour features, such as a higher diastolic or dicrotic notch amplitudes. In terms of cardiac frequency extraction, the HR could be successfully estimated from neck signals with high correlation to reference ECG. The feasibility of extracting SpO₂% from novel neck PPG signals was also proven. A linear regression model was fitted in order to calibrate the novel measurement site against the ground truth. Leave-One-Subject-Out validation results showed that SpO₂% could be reliably extracted. Spectrograms

under different breathing conditions revealed that the respiratory frequency was more prominent in neck PPG than in finger, which has a great potential for respiratory rate (RR) extraction. The results of this chapter, are very promising for the suitability of the neck as, not only, an alternative pulse oximetry location, but also, as a multi-modal site for cardio-respiratory diseases monitoring. However, some of the main limitations found for neck PPG, included the sensitivity to artifacts, and the uncovering of the optimal signal acquisition location. In fact, when moving the sensor away from the suprasternal notch, the venous signal could be identified, as further studied in [Chapter 4](#).

Chapter 4 - Extracting the Jugular Venous Pulse from Anterior Neck Contact Photoplethysmography. In this chapter, a novel approach to measure the jugular venous pressure using neck contact PPG is presented. Several claims supporting the feasibility of measuring this hardly accessed physiological signal, are formulated.

The jugular venous pulse (JVP) is the reference physiological signal used to detect right atrial and central venous pressure (CVP) abnormalities in cardio-vascular diseases (CVDs) diagnosis. Invasive central venous line catheterization is the gold standard method to extract it reliably. However, due to all the risks it entails, novel non-invasive approaches are needed. In this study, we demonstrate, for the first time, that reflectance photoplethysmography (PPG) can be used for extracting the JVP from the anterior jugular veins, in a contact manner. Neck JVP-PPG signals were recorded from 20 healthy participants, together with reference ECG and arterial finger PPG signals for validation. B-mode ultrasound imaging of the internal jugular vein also proved the validity of the proposed method. The results show that it is possible to identify the characteristic a , c , v pressure waves in the novel signals, and confirm their cardiac-cycle timings in consistency with established cardiac physiology. This work is of great significance for the future of CVDs diagnosis, as it has the potential to reduce the risks associated with conventional catheterization and enable continuous non-invasive point-of-care monitoring of CVP.

Chapter 5 - Characterization of Artifact Signals in Neck Photoplethysmography focuses on the characterization of neck PPG artifacts through an exhaustive study of the distinctive features discriminating noise sources from clean PPG.

Neck PPG signals are very susceptible to artifacts which greatly compromise their quality. But the extent of this, is going to depend on the nature of the artifacts and the strength of the sensed signal, both of which are location dependent. This chapter presents for the first time the characterization of artifacts affecting neck PPG signals. The study of two other respiratory states of interest, breathing slow and apnea, is also covered. Neck PPG data was recorded from 19 participants, who simulated these two respiratory rhythms and performed ten different activities to deliberately introduce common artifacts. 41 PPG features were extracted and statistically analyzed to investigate which ones

showed the greatest ability to differentiate normal PPG from each artifact. A customized minimum Redundancy Maximum Relevance (mRMR) feature selection approach was implemented, to select the top 10 features with the highest potential. Artifacts caused by Swallowing, Yawning and Coughing exhibited larger Spectral Entropy, Average Power and smaller Spectral Kurtosis, than normal PPG. Head movement artifacts, also demonstrated highly disordered and noisy frequency spectra, and were characterized by having larger and irregular time domain features. In addition, the analysis showed that the studied respiratory states, such as apneas, were also distinguishable from sources of interference. These findings are important for the development of PPG denoising algorithms and subsequent obtention of biomarkers; or alternatively for applications where the events of interest are the artifacts themselves.

Chapter 6 - Artifacts Classification and Apnea Events Detection in Neck Photoplethysmography Signals In order to fully exploit novel neck PPG for reliable biomarkers extraction and apneas monitoring, this chapter aims at developing, an artifacts classification and an apnea detection algorithms. Neck PPG signals were recorded in simulated noise and breathing conditions. 51 features from the time, correlogram and frequency domains were extracted, together with some envelope-derived characteristics. Two support vector machine (SVM) classifiers with a radial basis function (RBF) kernel were trained for different window lengths ($W = 4, 5, 6, 7, 8, 10s$) and thresholds of corruption ($Thd = 20\%, 30\%, 40\%, 50\%$). Chi-square tests coupled with sequential feature selection, were applied to fit the model using Leave-one-subject-out cross-validation. The classification process was repeated 30 times with different randomizations of the data. A two-way ANOVA and Tukey's post-hoc comparisons, were used to statistically test how the different windows and thresholds affected the performance. For artifacts classification, the maximum performance was obtained for the parameters combination of $[W=6s-Thd=20\%]$, and for apnea detection, the model $[W=10s-Thd=50\%]$ maximized all the performance metrics significantly. The findings of this proof of concept, are of great importance to design the future filtering strategy of neck PPG signals, and to facilitate a prompt detection of apneic events. This could have a significant impact in real-time applications, in which extremely high accuracy is necessary but cannot be obtained from just one sensing modality; such as in the prevention of sudden unexpected death in epilepsy (SUDEP).

Chapter 7 - Conclusions summarizes the major contributions of this thesis and describes the future directions to extend this PhD work.

1.8 Experiments outline

In order to study and explore different aspects of the novel neck PPG site, a total of 4 independent data recording experiments were performed during this PhD. This experimental work was approved by the Local Ethics Committee of Imperial College London (ICREC ref.: 18IC4358), and written informed consent was obtained from all subjects. All data was anonymized and stored in password-secured computers. Relevant ethics documents can be found in [Appendix A](#).

Between 10 to 20 healthy subjects, with an heterogeneous variety of skin tones, were recruited at university for each experiment. Due to the specific pool of participants, the age group was 20-30 years old, with a normal BMI of around 23kg, an overall males/females ratio of 3:1 and a majority of white skin tones. Although a wider range of participants would have been ideal to ensure a broader generalization of the results, this cohort was deemed suitable for a first proof-of-concept study of neck PPG. Moreover, since being a male is one of the factors increasing the risk of SUDEP, the gender imbalance did not cause any inconvenient.

The experimental procedures are presented within each corresponding chapter. However, for the sake of clarity, a brief overview of the purpose of each study and the outcomes of the recorded data, are outlined as follows:

Experiment 1. is presented in [Chapter 3](#), Section 3.2. It consists in recording neck PPG signals at the neck for the first time in supine position and compare them with finger PPG signals acquired simultaneously. Neck PPG signals were obtained under different respiratory states (spontaneous breathing, fast breathing, slow controlled breathing and breath-holds) to explore the pulse morphology (Section 3.3) and spectral (Section 3.4) characteristics of novel neck PPG signals. This dataset was also used in Section 3.5 to demonstrate that it is possible to reliably extract the heart rate from neck PPG signals.

Experiment 2. (in [Chapter 3](#), Section 3.6.1.1) The aim of this experiment was to record raw neck PPG signals during SpO₂ desaturations. This was necessary for the calibration of a wearable neck prototype sensor, and to demonstrate the feasibility of reliably estimating SpO₂% levels from the novel PPG site.

Experiment 3. in [Chapter 4](#), consisted in exploring the vein vasculature of the neck and to demonstrate, for the first time in literature, that it is possible to record the Jugular Venous Pulse (JVP) from the anterior veins using neck contact PPG.

Experiment 4. in [Chapter 5](#), aims at recording neck PPG signals under different artifacts conditions and breathing states of interest. This data is exploited to characterize the different artifacts in terms of time and frequency features. The same dataset is subsequently used in [Chapter 6](#) to train and test an artifacts and apnea classification models.

References

- [1] A. Singh and S. Trevick, “The epidemiology of global epilepsy,” *Neurologic clinics*, vol. 34, no. 4, pp. 837–847, 2016.
- [2] A. K. Ngugi, C. Bottomley, I. Kleinschmidt, J. W. Sander, and C. R. Newton, “Estimation of the burden of active and life-time epilepsy: a meta-analytic approach,” *Epilepsia*, vol. 51, no. 5, pp. 883–890, 2010.
- [3] Epilepsy Society UK, “Epilepsy facts and myths,” <https://www.epilepsysociety.org.uk/facts-and-statistics>, (Accessed: 09/2020).
- [4] R. D. Thijs, R. Surges, T. J. O’Brien, and J. W. Sander, “Epilepsy in adults,” *The Lancet*, vol. 393, no. 10172, pp. 689–701, 2019.
- [5] M. Brodie, S. Barry, G. Bamagous, J. Norrie, and P. Kwan, “Patterns of treatment response in newly diagnosed epilepsy,” *Neurology*, vol. 78, no. 20, pp. 1548–1554, 2012.
- [6] World Health Organization (WHO), “Epilepsy,” <https://www.who.int/news-room/fact-sheets/detail/epilepsy>, June 2019.
- [7] O. Devinsky, T. Spruill, D. Thurman, and D. Friedman, “Recognizing and preventing epilepsy-related mortality: a call for action,” *Neurology*, vol. 86, no. 8, pp. 779–786, 2016.
- [8] Epilepsy Research UK, “Epilepsy statistics,” <https://epilepsyresearch.org.uk/about-epilepsy/epilepsy-statistics/>, (Accessed: 09/2020).
- [9] Office for National Statistics, “Mortality statistics - underlying cause, sex and age,” <https://www.nomisweb.co.uk/query/construct/summary.asp?reset=yes&mode=construct&dataset=161&version=0&anal=1&initset=>, 2019.
- [10] National Records for Scotland, “Vital events reference tables 2018. section 6: Deaths - causes,” <https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/general-publications/vital-events-reference-tables/2018/section-6-death-causes>, 2019.
- [11] Northern Ireland Statistics and Research Agency (NISRA), “Registrar general annual report 2018 deaths,” <https://www.nisra.gov.uk/publications/registrar-general-annual-report-2018-cause-death>, 2019.
- [12] S. Duncan and M. J. Brodie, “Sudden unexpected death in epilepsy,” *Epilepsy & Behavior*, vol. 21, no. 4, pp. 344–351, 2011.
- [13] D. J. Thurman, D. C. Hesdorffer, and J. A. French, “Sudden unexpected death in epilepsy: assessing the public health burden,” *Epilepsia*, vol. 55, no. 10, pp. 1479–1485, 2014.
- [14] L. Nashef, “Sudden unexpected death in epilepsy: terminology and definitions,” *Epilepsia*, vol. 38, no. s11, pp. S6–S8, 1997.
- [15] P. Ryvlin, L. Nashef, S. D. Lhatoo, L. M. Bateman, J. Bird, A. Bleasel, P. Boon, A. Crespel, B. A. Dworetzky, H. Høgenhaven *et al.*, “Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (mortemus): a retrospective study,” *The Lancet Neurology*, vol. 12, no. 10, pp. 966–977, 2013.
- [16] O. Devinsky, D. C. Hesdorffer, D. J. Thurman, S. Lhatoo, and G. Richerson, “Sudden unexpected death in epilepsy: epidemiology, mechanisms, and prevention,” *The Lancet Neurology*, vol. 15, no. 10, pp. 1075–1088, 2016.

-
- [17] A. D. Elmali, N. Bebek, and B. Baykan, "Let's talk sudep," *Archives of Neuropsychiatry*, vol. 56, no. 4, p. 292, 2019.
 - [18] M. van der Lende, R. Surges, J. W. Sander, and R. D. Thijs, "Cardiac arrhythmias during or after epileptic seizures," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 87, no. 1, pp. 69–74, 2016.
 - [19] F. A. Nascimento, Z. H. Tseng, C. Palmiere, J. J. Maleszewski, T. Shiomi, A. McCrillis, and O. Devinsky, "Pulmonary and cardiac pathology in sudden unexpected death in epilepsy (sudep)," *Epilepsy & Behavior*, vol. 73, pp. 119–125, 2017.
 - [20] L. Zhuo, Y. Zhang, H. R. Zielke, B. Levine, X. Zhang, L. Chang, D. Fowler, and L. Li, "Sudden unexpected death in epilepsy: evaluation of forensic autopsy cases," *Forensic science international*, vol. 223, no. 1-3, pp. 171–175, 2012.
 - [21] M. Dupuis, K. Van Rijkevorsel, F. Evrard, N. Dubuisson, F. Dupuis, and P. Van Robays, "Takotsubo syndrome (tks): a possible mechanism of sudden unexplained death in epilepsy (sudep)," *Seizure*, vol. 21, no. 1, pp. 51–54, 2012.
 - [22] S. Al-Najafi and H. Rosman, "Seizure-induced myocardial stunning: a possible cardiac link to sudden unexpected death in epilepsy (sudep)," *Seizure-European Journal of Epilepsy*, vol. 24, pp. 137–139, 2015.
 - [23] N. Lacuey, B. Zonjy, J. P. Hampson, M. S. Rani, A. Zaremba, R. K. Sainju, B. K. Gehlbach, S. Schuele, D. Friedman, O. Devinsky *et al.*, "The incidence and significance of periictal apnea in epileptic seizures," *Epilepsia*, vol. 59, no. 3, pp. 573–582, 2018.
 - [24] L. Vilella, N. Lacuey, J. P. Hampson, M. S. Rani, R. K. Sainju, D. Friedman, M. Nei, K. Strohl, C. Scott, B. K. Gehlbach *et al.*, "Postconvulsive central apnea as a biomarker for sudden unexpected death in epilepsy (sudep)," *Neurology*, vol. 92, no. 3, pp. e171–e182, 2019.
 - [25] L. Nashef, F. Walker, P. Allen, J. Sander, S. Shorvon, and D. Fish, "Apnoea and bradycardia during epileptic seizures: relation to sudden death in epilepsy." *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 60, no. 3, pp. 297–300, 1996.
 - [26] A. S. Blum, J. R. Ives, A. L. Goldberger, I. C. Al-Aweel, K. Krishnamurthy, F. W. Drislane, and D. L. Schomer, "Oxygen desaturations triggered by partial seizures: implications for cardiopulmonary instability in epilepsy," *Epilepsia*, vol. 41, no. 5, pp. 536–541, 2000.
 - [27] J. Tavee and H. Morris III, "Severe postictal laryngospasm as a potential mechanism for sudden unexpected death in epilepsy: a near-miss in an emu," *Epilepsia*, vol. 49, no. 12, pp. 2113–2117, 2008.
 - [28] M. Stewart, R. Kollmar, K. Nakase, J. Silverman, K. Sundaram, R. Orman, and J. Lazar, "Obstructive apnea due to laryngospasm links ictal to postictal events in sudep cases and offers practical biomarkers for review of past cases and prevention of new ones," *Epilepsia*, vol. 58, no. 6, pp. e87–e90, 2017.
 - [29] J. E. Leestma, T. Walczak, J. R. Hughes, M. B. Kalelkar, and S. S. Teas, "A prospective study on sudden unexpected death in epilepsy," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 26, no. 2, pp. 195–203, 1989.
 - [30] C. F. Terrence, G. R. Rao, and J. A. Perper, "Neurogenic pulmonary edema in unexpected, unexplained death of epileptic patients," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 9, no. 5, pp. 458–464, 1981.

- [31] M. Pezzella, P. Striano, C. Ciampa, L. Errichiello, P. Penza, and S. Striano, "Severe pulmonary congestion in a near miss at the first seizure: further evidence for respiratory dysfunction in sudden unexpected death in epilepsy," *Epilepsy & Behavior*, vol. 14, no. 4, pp. 701–702, 2009.
- [32] J. D. Kennedy, K. A. Hardin, P. Parikh, C.-S. Li, and M. Seyal, "Pulmonary edema following generalized tonic clonic seizures is directly associated with seizure duration," *Seizure*, vol. 27, pp. 19–24, 2015.
- [33] S. D. Lhatoo, H. J. Faulkner, K. Dembny, K. Trippick, C. Johnson, and J. M. Bird, "An electroclinical case-control study of sudden unexpected death in epilepsy," *Annals of neurology*, vol. 68, no. 6, pp. 787–796, 2010.
- [34] R. Surges, A. Strzelczyk, C. A. Scott, M. C. Walker, and J. W. Sander, "Postictal generalized electroencephalographic suppression is associated with generalized seizures," *Epilepsy & behavior*, vol. 21, no. 3, pp. 271–274, 2011.
- [35] Y. Tang, D. An, Y. Xiao, R. Niu, X. Tong, W. Liu, L. Zhao, Q. Gong, and D. Zhou, "Cortical thinning in epilepsy patients with postictal generalized electroencephalography suppression," *European journal of neurology*, vol. 26, no. 1, pp. 191–197, 2019.
- [36] V. Alexandre, B. Mercedes, L. Valton, L. Maillard, F. Bartolomei, W. Szurhaj, E. Hirsch, C. Marchal, F. Chassoux, J. Petit *et al.*, "Risk factors of postictal generalized eeg suppression in generalized convulsive seizures," *Neurology*, vol. 85, no. 18, pp. 1598–1603, 2015.
- [37] J. Kuo, W. Zhao, C.-S. Li, J. D. Kennedy, and M. Seyal, "Postictal immobility and generalized eeg suppression are associated with the severity of respiratory dysfunction," *Epilepsia*, vol. 57, no. 3, pp. 412–417, 2016.
- [38] J. A. Liebenthal, S. Wu, S. Rose, J. S. Ebersole, and J. X. Tao, "Association of prone position with sudden unexpected death in epilepsy," *Neurology*, vol. 84, no. 7, pp. 703–709, 2015.
- [39] N. Barot and M. Nei, "Autonomic aspects of sudden unexpected death in epilepsy (sudep)," *Clinical Autonomic Research*, vol. 29, no. 2, pp. 151–160, 2019.
- [40] K. A. Myers, L. E. Bello-Espinosa, J. D. Symonds, S. M. Zuberi, R. Clegg, L. G. Sadleir, J. Buchhalter, and I. E. Scheffer, "Heart rate variability in epilepsy: a potential biomarker of sudden unexpected death in epilepsy risk," *Epilepsia*, vol. 59, no. 7, pp. 1372–1380, 2018.
- [41] R. W. Picard, M. Migliorini, C. Caborni, F. Onorati, G. Regalia, D. Friedman, and O. Devinsky, "Wrist sensor reveals sympathetic hyperactivity and hypoventilation before probable sudep," *Neurology*, vol. 89, no. 6, pp. 633–635, 2017.
- [42] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 786–794, 2010.
- [43] R. A. Sarkis, S. Thome-Souza, M.-Z. Poh, N. Llewellyn, J. Klehm, J. R. Madsen, R. Picard, P. B. Pennell, B. A. Dworetzky, T. Loddenkemper *et al.*, "Autonomic changes following generalized tonic clonic seizures: an analysis of adult and pediatric patients with epilepsy," *Epilepsy research*, vol. 115, pp. 113–118, 2015.
- [44] L. Nilsson, B. Farahmand, P. Persson, I. Thiblin, and T. Tomson, "Risk factors for sudden unexpected death in epilepsy: a case control study," *The Lancet*, vol. 353, no. 9156, pp. 888–893, 1999.

-
- [45] D. C. Hesdorffer, T. Tomson, E. Benn, J. W. Sander, L. Nilsson, Y. Langan, T. S. Walczak, E. Beghi, M. J. Brodie, A. Hauser *et al.*, “Combined analysis of risk factors for sudep,” *Epilepsia*, vol. 52, no. 6, pp. 1150–1159, 2011.
 - [46] T. Tomson, R. Surges, R. Delamont, S. Haywood, and D. C. Hesdorffer, “Who to target in sudden unexpected death in epilepsy prevention and how? risk factors, biomarkers, and intervention study designs,” *Epilepsia*, vol. 57, pp. 4–16, 2016.
 - [47] L. Ridsdale, J. Charlton, M. Ashworth, M. P. Richardson, and M. C. Gulliford, “Epilepsy mortality and risk factors for death in epilepsy: a population-based study,” *British Journal of General Practice*, vol. 61, no. 586, pp. e271–e278, 2011.
 - [48] R. Kiani, F. Tyrer, A. Jesu, S. Bhaumik, S. Gangavati, G. Walker, S. Kazmi, and M. Barrett, “Mortality from sudden unexpected death in epilepsy (sudep) in a cohort of adults with intellectual disability,” *Journal of Intellectual Disability Research*, vol. 58, no. 6, pp. 508–520, 2014.
 - [49] C. R. Frasier, H. Zhang, J. Offord, L. T. Dang, D. S. Auerbach, H. Shi, C. Chen, A. M. Goldman, L. L. Eckhardt, V. J. Bezzerides *et al.*, “Channelopathy as a sudep biomarker in dravet syndrome patient-derived cardiac myocytes,” *Stem cell reports*, vol. 11, no. 3, pp. 626–634, 2018.
 - [50] Y. Ergul, B. Ekici, B. Tatli, K. Nisli, and M. Ozmen, “Qt and p wave dispersion and heart rate variability in patients with dravet syndrome,” *Acta Neurologica Belgica*, vol. 113, no. 2, pp. 161–166, 2013.
 - [51] C. Harden, T. Tomson, D. Gloss, J. Buchhalter, J. H. Cross, E. Donner, J. A. French, A. Gil-Nagel, D. C. Hesdorffer, W. H. Smithson *et al.*, “Practice guideline summary: sudden unexpected death in epilepsy incidence rates and risk factors: report of the guideline development, dissemination, and implementation subcommittee of the american academy of neurology and the american epilepsy society,” *Epilepsy currents*, vol. 17, no. 3, pp. 180–187, 2017.
 - [52] C. M. DeGiorgio, D. Markovic, R. Mazumder, and B. D. Moseley, “Ranking the leading risk factors for sudden unexpected death in epilepsy,” *Frontiers in neurology*, vol. 8, p. 473, 2017.
 - [53] R. J. Lamberts, R. D. Thijs, A. Laffan, Y. Langan, and J. W. Sander, “Sudden unexpected death in epilepsy: people with nocturnal seizures may be at highest risk,” *Epilepsia*, vol. 53, no. 2, pp. 253–257, 2012.
 - [54] M. J. Maguire, C. F. Jackson, A. G. Marson, and S. J. Nevitt, “Treatments for the prevention of sudden unexpected death in epilepsy (sudep),” *Cochrane Database of Systematic Reviews*, no. 7, 2016.
 - [55] R. Shankar, W. Henley, C. Boland, R. Laugharne, B. N. McLean, C. Newman, J. Hanna, S. Ashby, M. C. Walker, and J. W. Sander, “Decreasing the risk of sudden unexpected death in epilepsy: structured communication of risk factors for premature mortality in people with epilepsy,” *European journal of neurology*, vol. 25, no. 9, pp. 1121–1127, 2018.
 - [56] R. Shankar, D. Cox, V. Jaliha, S. Brown, J. Hanna, and B. McLean, “Sudden unexpected death in epilepsy (sudep): development of a safety checklist,” *Seizure*, vol. 22, no. 10, pp. 812–817, 2013.
 - [57] C. Newman, R. Shankar, J. Hanna, B. McLean, A. Osland, C. Milligan, A. Ball, C. Jory, and M. Walker, “Developing an evidence-based epilepsy risk assessment ehealth solution: from concept to market,” *JMIR research protocols*, vol. 5, no. 2, p. e82, 2016.

- [58] P. Ryvlin, M. Cucherat, and S. Rheims, “Risk of sudden unexpected death in epilepsy in patients given adjunctive antiepileptic treatment for refractory seizures: a meta-analysis of placebo-controlled randomised trials,” *The Lancet Neurology*, vol. 10, no. 11, pp. 961–968, 2011.
- [59] M. R. Sperling, S. Barshow, M. Nei, and A. A. Asadi-Pooya, “A reappraisal of mortality after epilepsy surgery,” *Neurology*, vol. 86, no. 21, pp. 1938–1944, 2016.
- [60] M. R. Sperling, H. Feldman, J. Kinman, J. D. Liporace, and M. J. O’Connor, “Seizure control and mortality in epilepsy,” *Annals of neurology*, vol. 46, no. 1, pp. 45–50, 1999.
- [61] C. A. Granbichler, L. Nashef, R. Selway, and C. E. Polkey, “Mortality and sudep in epilepsy patients treated with vagus nerve stimulation,” *Epilepsia*, vol. 56, no. 2, pp. 291–296, 2015.
- [62] Y. Langan, L. Nashef, and J. Sander, “Case-control study of sudep,” *Neurology*, vol. 64, no. 7, pp. 1131–1133, 2005.
- [63] L. Watkins and R. Shankar, “Reducing the risk of sudden unexpected death in epilepsy (sudep),” *Current treatment options in neurology*, vol. 20, no. 10, p. 40, 2018.
- [64] F. Rugg-Gunn, “The role of devices in managing risk,” *Epilepsy & Behavior*, vol. 103, p. 106456, 2020.
- [65] M. van der Lende, D. C. Hesdorffer, J. W. Sander, and R. D. Thijs, “Nocturnal supervision and sudep risk at different epilepsy care settings,” *Neurology*, vol. 91, no. 16, pp. e1508–e1518, 2018.
- [66] M. Seyal, L. M. Bateman, and C.-S. Li, “Impact of periictal interventions on respiratory dysfunction, postictal eeg suppression, and postictal immobility,” *Epilepsia*, vol. 54, no. 2, pp. 377–382, 2013.
- [67] S. Wu, N. P. Issa, S. L. Rose, A. Ali, and J. X. Tao, “Impact of periictal nurse interventions on postictal generalized eeg suppression in generalized convulsive seizures,” *Epilepsy & Behavior*, vol. 58, pp. 22–25, 2016.
- [68] F. Rugg-Gunn, J. Duncan, H. Hjalgrim, M. Seyal, and L. Bateman, “From unwitnessed fatality to witnessed rescue: Nonpharmacologic interventions in sudden unexpected death in epilepsy,” *Epilepsia*, vol. 57, pp. 26–34, 2016.
- [69] P. G. Catcheside, A. A. Mohtar, and K. J. Reynolds, “Airflow resistance and co2 rebreathing properties of anti-asphyxia pillows designed for epilepsy,” *Seizure*, vol. 23, no. 6, pp. 462–467, 2014.
- [70] G. D. Cascino, “Video-eeg monitoring in adults,” *Epilepsia*, vol. 43, pp. 80–93, 2002.
- [71] A. Van de Vel, K. Cuppens, B. Bonroy, M. Milosevic, K. Jansen, S. Van Huffel, B. Vanrumste, P. Cras, L. Lagae, and B. Ceulemans, “Non-eeg seizure detection systems and potential sudep prevention: state of the art: review and update,” *Seizure*, vol. 41, pp. 141–153, 2016.
- [72] C. Jory, R. Shankar, D. Coker, B. McLean, J. Hanna, and C. Newman, “Safe and sound? a systematic literature review of seizure detection methods for personal use,” *Seizure*, vol. 36, pp. 4–15, 2016.
- [73] E. G. Gutierrez, N. E. Crone, J. Y. Kang, Y. I. Carmenate, and G. L. Krauss, “Strategies for non-eeg seizure detection and timing for alerting and interventions with tonic-clonic seizures,” *Epilepsia*, vol. 59, pp. 36–41, 2018.
- [74] J. Rémi, J. P. S. Cunha, C. Vollmar, Ö. B. Topçuoğlu, A. Meier, S. Ulowetz, P. Beleza, and S. Noachtar, “Quantitative movement analysis differentiates focal seizures characterized by automatism,” *Epilepsy & Behavior*, vol. 20, no. 4, pp. 642–647, 2011.

-
- [75] G. M. K. Ntonfo, G. Ferrari, R. Raheli, and F. Pisani, “Low-complexity image processing for real-time detection of neonatal clonic seizures,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 375–382, 2012.
 - [76] SAMI Alert, “Sami-3,” <https://www.samialert.com/sami-3>.
 - [77] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Advancements in noncontact, multiparameter physiological measurements using a webcam,” *IEEE transactions on biomedical engineering*, vol. 58, no. 1, pp. 7–11, 2010.
 - [78] K. V. Poppel, S. P. Fulton, A. McGregor, M. Ellis, A. Patters, and J. Wheless, “Prospective study of the emfit movement monitor,” *Journal of child neurology*, vol. 28, no. 11, pp. 1434–1436, 2013.
 - [79] A. P. Narechania, I. I. Garić, I. Sen-Gupta, M. P. Macken, E. E. Gerard, and S. U. Schuele, “Assessment of a quasi-piezoelectric mattress monitor as a detection system for generalized convulsions,” *Epilepsy & Behavior*, vol. 28, no. 2, pp. 172–176, 2013.
 - [80] S. Beniczky, I. Conradsen, and P. Wolf, “Detection of convulsive seizures using surface electromyography,” *Epilepsia*, vol. 59, pp. 23–29, 2018.
 - [81] S. Beniczky, I. Conradsen, O. Henning, M. Fabricius, and P. Wolf, “Automated real-time detection of tonic-clonic seizures using a wearable emg device,” *Neurology*, vol. 90, no. 5, pp. e428–e434, 2018.
 - [82] J. J. Halford, M. R. Sperling, D. R. Nair, D. J. Dlugos, W. O. Tatum, J. Harvey, J. A. French, J. R. Pollard, E. Faught, K. H. Noe *et al.*, “Detection of generalized tonic-clonic seizures using surface electromyographic monitoring,” *Epilepsia*, vol. 58, no. 11, pp. 1861–1869, 2017.
 - [83] U. Kramer, S. Kipervasser, A. Shlitner, and R. Kuzniecky, “A novel portable seizure detection alarm system: preliminary results,” *Journal of Clinical Neurophysiology*, vol. 28, no. 1, pp. 36–38, 2011.
 - [84] S. Beniczky, T. Polster, T. W. Kjaer, and H. Hjalgrim, “Detection of generalized tonic-clonic seizures by a wireless wrist accelerometer: a prospective, multicenter study,” *Epilepsia*, vol. 54, no. 4, pp. e58–e61, 2013.
 - [85] A. L. Patterson, B. Mudigoudar, S. Fulton, A. McGregor, K. Van Poppel, M. C. Wheless, L. Brooks, and J. W. Wheless, “Smartwatch by smartmonitor: assessment of seizure detection efficacy for various seizure types in children, a large prospective single-center study,” *Pediatric neurology*, vol. 53, no. 4, pp. 309–311, 2015.
 - [86] C. Carlson, V. Arnedo, M. Cahill, and O. Devinsky, “Detecting nocturnal convulsions: efficacy of the mp5 monitor,” *Seizure*, vol. 18, no. 3, pp. 225–227, 2009.
 - [87] S. Fulton, K. V. Poppel, A. McGregor, M. Ellis, A. Patters, and J. Wheless, “Prospective study of 2 bed alarms for detection of nocturnal seizures,” *Journal of child neurology*, vol. 28, no. 11, pp. 1430–1433, 2013.
 - [88] F. Massé, M. V. Bussel, A. Serteyn, J. Arends, and J. Penders, “Miniaturized wireless ecg monitor for real-time detection of epileptic seizures,” *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 12, no. 4, pp. 1–21, 2013.
 - [89] J. van Andel, C. Ungureanu, J. Arends, F. Tan, J. Van Dijk, G. Petkov, S. Kalitzin, T. Gutter, A. de Weerd, B. Vledder *et al.*, “Multimodal, automated detection of nocturnal motor seizures at home: Is a reliable seizure detector feasible?” *Epilepsia open*, vol. 2, no. 4, pp. 424–431, 2017.

- [90] LivAssured, “Nightwatch,” <https://www.nightwatchepilepsy.com/nightwatch>.
- [91] J. Arends, R. D. Thijs, T. Gutter, C. Ungureanu, P. Cluitmans, J. Van Dijk, J. van Andel, F. Tan, A. de Weerd, B. Vledder *et al.*, “Multimodal nocturnal seizure detection in a residential care setting: A long-term prospective trial,” *Neurology*, vol. 91, no. 21, pp. e2010–e2019, 2018.
- [92] S. Sabesan and R. Sankar, “Improving long-term management of epilepsy using a wearable multimodal seizure detection system,” *Epilepsy & Behavior*, vol. 46, pp. 56–57, 2015.
- [93] Bioserenity, “Neuronaute,” <https://www.bioserenity.com/en/neuro/>.
- [94] PulseGuard International Ltd, “Pulseguard,” <https://pulseguard.org/product/pulseguard/>.
- [95] K. Vandecasteele, T. De Cooman, Y. Gu, E. Cleeren, K. Claes, W. V. Paesschen, S. V. Huffel, and B. Hunyadi, “Automated epileptic seizure detection based on wearable ecg and ppg in a hospital environment,” *Sensors*, vol. 17, no. 10, p. 2338, 2017.
- [96] D. Cogan, M. Nourani, J. Harvey, and V. Nagaraddi, “Epileptic seizure detection using wristworn biosensors,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 5086–5089.
- [97] D. Cogan, J. Birjandtalab, M. Nourani, J. Harvey, and V. Nagaraddi, “Multi-biosignal analysis for epileptic seizure monitoring,” *International journal of neural systems*, vol. 27, no. 01, p. 1650031, 2017.
- [98] P. Corbishley and E. Rodríguez-Villegas, “Breathing detection: towards a miniaturized, wearable, battery-operated monitoring system,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 196–204, 2008.
- [99] E. Rodríguez-Villegas, G. Chen, J. Radcliffe, and J. Duncan, “A pilot study of a wearable apnoea detection device,” *BMJ open*, vol. 4, no. 10, p. e005299, 2014.
- [100] G. Regalia, F. Onorati, M. Lai, C. Caborni, and R. W. Picard, “Multimodal wrist-worn devices for seizure detection and advancing research: focus on the empatica wristbands,” *Epilepsy research*, vol. 153, pp. 79–82, 2019.
- [101] J. W. Lee, “Real-time non-ecg convulsive seizure detection devices: they work; now what?” *Epilepsy currents*, vol. 18, no. 3, pp. 164–166, 2018.
- [102] P. Ryvlin, L. Cammoun, I. Hubbard, F. Ravey, S. Beniczky, and D. Atienza, “Noninvasive detection of focal seizures in ambulatory patients,” *Epilepsia*, 2020.
- [103] M. Peng, S. A. Imtiaz, and E. Rodríguez-Villegas, “Pulse oximetry in the neck—a proof of concept,” in *EMBC*. IEEE, 2017, pp. 877–880.
- [104] Y. Zhong, Y. Pan, L. Zhang, and K.-T. Cheng, “A wearable signal acquisition system for physiological signs including throat ppg,” in *Engineering in Medicine and Biology Society (EMBC), 2016 38th Annual International Conference of the IEEE*. IEEE, 2016, pp. 603–606.

Chapter 2

A Review on PPG Principles, Artifacts Cancellation and Apnea Detection

This chapter introduces all the background necessary to understand the research work behind this thesis. It starts with an introduction to the fundamentals of PPG technology and its applications. Then, the basic characteristics of the signal and the physiology behind are covered. Finally, the state of the art of artifacts removal techniques and algorithms for apnea detection is reviewed.

2.1 Photoplethysmography

2.1.1 Photoplethysmography principles

Photoplethysmography (PPG) is a non-invasive, low-cost, optical technique measuring changes in blood volume at the skin surface [1]. Pulse oximeter sensors are used to extract the PPG signal relying on the difference in light absorption properties of the microvascular bed of tissues. In fact, as it can be observed in Figure 2.1, oxyhemoglobin absorbs more infra-red (940 nm) and less red light (660nm) in comparison to deoxyhemoglobin, which allows fluctuations of oxygenated arterial blood to be measured [2]. According to this principle, a PPG sensing device consists of a red and an infrared light-emitting diode (LED) that illuminate the tissue, and a photodetector that senses light intensity variations after tissue penetration. The resulting PPG waveform shows two principal constituents (Figure 2.2): An AC pulsatile component with frequency of approximately 1 Hz, or 60 beats-per-min (BPM), reflects the arterial pulsations at each heartbeat; and a DC component, accounts for the light absorption of non-pulsatile arterial, venous blood and other non-vascular tissues [3].

Two main types of photoplethysmography exist: contact and non-contact. The contact modality includes transmission and reflective PPG (Figure 2.3). In transmission PPG, generally used in fingertip commercial pulse oximetry due to the easy attachment and rich capillarity, the light emitted by the

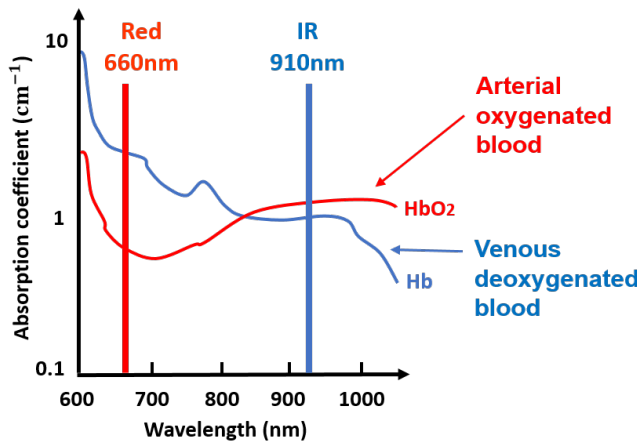


Figure 2.1: Absorption spectra of oxy-hemoglobin (HbO_2) and deoxyhemoglobin (Hb).

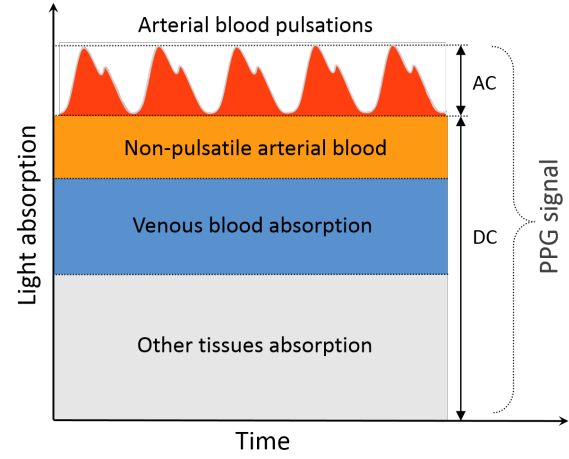


Figure 2.2: Light absorption by different tissues reflects the AC and DC components of the PPG signal.

LEDs is transmitted through the finger before it reaches the photodetector. Reflective PPG can be used in a wider range of measurements sites, for which the photodetector cannot be typically placed directly opposite to the light source, but alongside to it; to detect the light reflected from superficial vessels. However, despite this advantage, reflective PPG is more prone to the uncoupling of the sensor with the skin, making this modality more susceptible to detachment and ambient light interference.

The second modality is imaging-PPG (IPPG). This non-contact peripheral blood perfusion measurement type uses a remote light source and a camera to detect blood pulsations in multiple areas of the skin simultaneously [4].

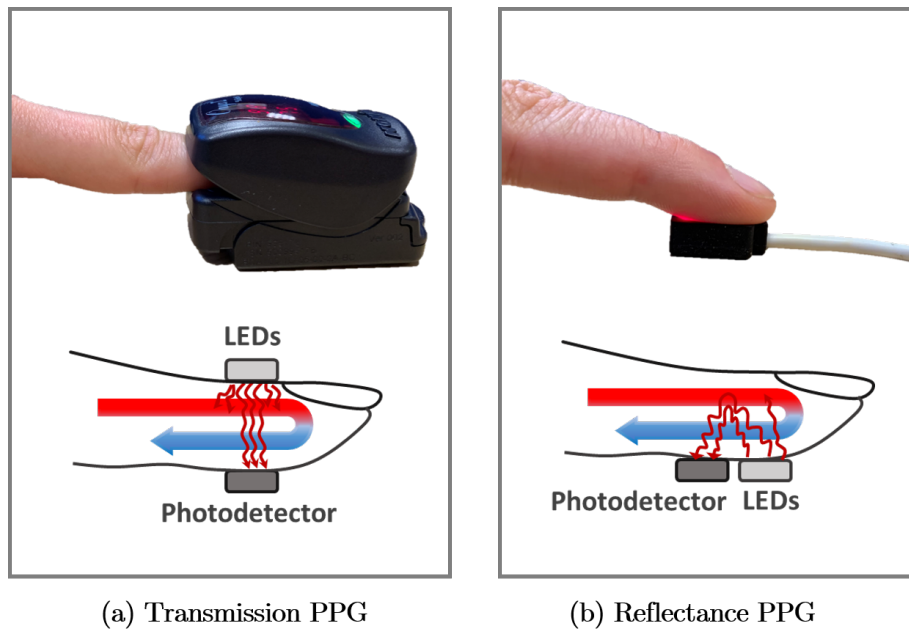


Figure 2.3: Contact transmission and reflectance optical modalities to measure the PPG signal at the finger.

2.1.2 Applications: SpO₂ and heart rate estimation

Pulse oximeter devices can be used to monitor an individual's general health status. They are essential in outpatient clinics, inpatient wards, intensive care units and operating theaters [5]. The applications of PPG are numerous. Vital cardio-respiratory parameters such as peripheral oxygen saturation (SpO₂), heart rate (HR), respiratory rate (RR), blood pressure or cardiac output, can be extracted from the PPG signal. The vascular and autonomic function can be similarly assessed [6]. In this thesis, I focus on the two most important biomarkers usually measured by pulse oximeters: SpO₂% and HR.

Heart rate: Since the PPG signal carries arterial blood pulsations information, the instantaneous heart rate can be calculated as follows:

$$HR \text{ (BPM)} = \frac{60}{T_{p-p}} \quad (2.1)$$

where T_{p-p} is the average peak-to-peak time interval between two PPG pulses in seconds. A factor of 60 outputs the result in the standard units of beats-per-minute (BPM). Normal heart rate in healthy humans ranges from 60 to 100 BPM at rest. Artifacts can disrupt the periodicity between pulses and alter the measure of T_{p-p} , resulting in an inaccurate heart rate estimation.

Peripheral oxygen saturation: Hemoglobin, is a gas transport protein present in red blood cells, that binds to oxygen and delivers it to the organs. Different forms of hemoglobin exist, that can be grouped in two major categories depending on their ability to bind to oxygen: functional and non-functional. Functional hemoglobins, such as oxyhemoglobin (HbO₂) and deoxyhemoglobin (Hb), are the only ones capable of carrying oxygen. Non-functional ones, including carboxyhemoglobin (HbCO) or methemoglobin (MetHb), cannot participate in respiratory function. Arterial blood oxygen saturation (SaO₂) is defined as the relative amount of oxygenated blood over total blood (oxygenated and de-oxygenated), such that:

$$SaO_2 = \frac{HbO_2}{HbO_2 + Hb + HbCO + MetHb + Hb_{other}} \quad (2.2)$$

where, HbO_2 and Hb represent the concentrations of functional hemoglobins; and $HbCO$, $MetHb$ and Hb_{other} the concentrations of other non-functional hemoglobins, incapable of binding to oxygen.

Pulse oximetry measures the peripheral oxygen saturation SpO₂, as a surrogate measure of SaO₂, assuming no non-functional hemoglobins, such that:

$$SpO_2 = \frac{HbO_2}{HbO_2 + Hb} \quad (2.3)$$

The derivation of SpO_2 from optical pulse oximetry measurements, is based on the principles of spectrophotometry [7]. The Beer–Lambert law states that the attenuation of light, i.e. the absorbance (A), through a non-scattering media is directly proportional to the concentration of the absorbing substances and the optical path length (L), such that:

$$A(\lambda) = L \sum_i \varepsilon_i(\lambda) c_i \quad (2.4)$$

where, $\varepsilon_i(\lambda)$ represents the extinction coefficient for each molecular absorber i at a specific wavelength λ , and c_i the corresponding concentrations.

The absorbance is likewise defined as the negative logarithm of light transmission T , which is calculated with the ratio of incident (I_{in}) and output (I_{out}) light intensities. This can be mathematically described as:

$$A = -\log(T) = -\log(I_{out}/I_{in}) \quad (2.5)$$

Pulse oximetry exclusively focuses on arterial *changes* in blood volume. *Constant* absorbance terms of non-pulsatile tissues (skin, bone, veins etc.), which do not change over time, will not have an effect in the calculation of SpO_2 . The two primary light absorbers hence are HbO_2 and Hb . This implies that the blood absorbance equation can be rewritten as:

$$A(\lambda) = (\varepsilon_{HbO_2}(\lambda) c_{HbO_2} + \varepsilon_{Hb}(\lambda) c_{Hb}) \cdot L \quad (2.6)$$

Using equation 2.3, the hemoglobin concentrations can be expressed as a function of SpO_2 as: $c_{HbO_2} = SpO_2(c_{HbO_2} + c_{Hb})$ and $c_{Hb} = c_{HbO_2}(1 - SpO_2)/SpO_2$. By rearranging equation 2.6, this gives:

$$A(\lambda) = (SpO_2 \cdot \varepsilon_{HbO_2}(\lambda) + (1 - SpO_2) \cdot \varepsilon_{Hb}(\lambda)) \cdot (c_{HbO_2} + c_{Hb}) \cdot L \quad (2.7)$$

The *change* in absorbance of blood pulsations at a particular wavelength λ , becomes:

$$\Delta A(\lambda) = -\frac{d \log(I_{out}(t)/I_{in})}{dt} = (SpO_2 \cdot \varepsilon_{HbO_2}(\lambda) + (1 - SpO_2) \cdot \varepsilon_{Hb}(\lambda)) \cdot \Delta Hb_a \cdot L \approx \frac{AC}{DC} \Big|_{\lambda} \quad (2.8)$$

where Hb_a denotes the time varying hemoglobin concentration in arterial blood. Equation 2.8 results from calculating the time derivative of the light intensity sensed by the photodiode, that can be approximated to the AC/DC ratio of the PPG signal at wavelength λ .

Using a monochromatic light would not be sufficient to differentiate the contributions of HbO_2 and Hb in the overall absorbance. This is why two wavelengths $\lambda_1 = 660nm$ (red) and $\lambda_2 = 940nm$ (infra-red) are required to compute the ratio of ratios (R) of pulsatile and non-pulsatile absorbance,

such that:

$$R \approx \frac{(AC/DC)_{\lambda_1}}{(AC/DC)_{\lambda_2}} = \frac{(SpO_2 \cdot \varepsilon_{HbO_2}(\lambda_1) + (1 - SpO_2) \cdot \varepsilon_{Hb}(\lambda_1)) \cdot \Delta Hb_a \cdot L}{(SpO_2 \cdot \varepsilon_{HbO_2}(\lambda_2) + (1 - SpO_2) \cdot \varepsilon_{Hb}(\lambda_2)) \cdot \Delta Hb_a \cdot L} \quad (2.9)$$

By taking the assumption that the pulsatile changes in blood concentration ΔHb_a are equal for both wavelengths, this can be solved for SpO_2 :

$$SpO_2 = \frac{\varepsilon_{Hb}(\lambda_1) - \varepsilon_{Hb}(\lambda_2) \cdot R}{\varepsilon_{Hb}(\lambda_1) - \varepsilon_{HbO_2}(\lambda_1) + (\varepsilon_{HbO_2}(\lambda_2) - \varepsilon_{Hb}(\lambda_2)) \cdot R} \quad (2.10)$$

This theoretical derivation of the peripheral oxygen saturation, SpO_2 , is not always precise, as not all the Beer-Lambert's law assumptions can be met. Therefore, in commercial pulse oximetry SpO_2 is generally determined by empirical calibration of the sensor. A generally adopted model, linearly approximates SpO_2 to the absorption ratio (R) at the red (λ_1) and infrared (λ_2) wavelengths with two empirical constants a and b , as shown in Equation 2.11 [8]. The accuracy of SpO_2 measurements for every specific device, is directly dependent on the calibration of these constants against some reference SpO_2 values.

$$SpO_2 \approx a - bR = a - b \frac{I_{AC}(\lambda_1)/I_{DC}(\lambda_1)}{I_{AC}(\lambda_2)/I_{DC}(\lambda_2)} \quad (2.11)$$

where, I_{AC} and I_{DC} correspond to the light intensities amplitudes of the AC and DC components of the PPG signal. Human healthy blood oxygen levels range from 95 to 100 %. In cases where artifacts of high amplitude perturb the coupling between sensor and skin, unusual AC and DC changes in light intensity sensed by the photodetector will directly alter the final SpO_2 value. This undesirable outcome would result in inaccurate SpO_2 measurements. In the worst scenario, unreliable readings could cause false alarms, alerting of non-existent hypoxia events, when saturation levels drop below 95%.

2.1.3 PPG sensors and measurement sites

The gold standard measurement location in pulse oximetry is the finger. The majority of commercial pulse oximeters in the market focus on this positioning, as it provides large capillarity, good perfusion, and is an accessible body part. However, in the past years, a broad variety of body locations have also been examined. Novel body sites encompass toes [1], ring finger [9], brachia [10], wrist [11], forehead [12], in-ear [13], oesophagus and bowel [14]. Successful wearable devices have demonstrated the feasibility of using alternative monitoring sites for pulse oximetry. Some examples include a magnetic earring placed at the earlobe [15], a wrist-watch PPG sensor array [16], earphones integrated PPG sensors [17] or forehead wireless-based reflectance sensors [18].

To my knowledge, only two studies exploring the neck as an innovative measurement site for

photoplethysmography exist. The first one, tested the anterior neck region to define the most stable PPG positioning as the mid-throat, and the best respiration rate extraction location as the lower throat, with an error of less than 5% [19]. Despite the novelty, the proposed prototype system was very cumbersome. It was composed of a slave PPG sensor board placed on the throat, that was connected via a wire to a master board placed on the chest. Such a system, is very uncomfortable for long term monitoring and unsuitable for portability.

The second attempt of neck PPG signal measurement was developed in our group [20]. The advanced wireless wearable prototype showed the feasibility of neck reflectance PPG to measure oxygen saturation (SpO_2) with an accuracy of 98.6%. Although some efforts of this proof of concept were invested towards the identification of sources of noise affecting neck PPG, an exhaustive characterization of the artifacts is still required, to accurately define the most effective signal processing strategy to deal with these sources of noise.

2.1.4 PPG waveform

The AC component of the PPG signal reflects the arterial blood volume changes resulting from cardiac synchronous pulsations. Each PPG pulse of the AC component, corresponds to a single heart beat. As it can be observed in Figure 2.4, the waveform is characterized by the systolic and diastolic phases. The systolic phase starts in the first valley or pulse onset, and ends at the point where the pulse wave reaches its maximum, i.e. the systolic peak. This time interval is denoted as rise time. The height of the pulse from its onset until the systolic peak is referred as the pulse wave amplitude. The diastolic phase starts right after the systolic peak and ends at the offset of the pulse. The pulse width or the total pulse wave duration are defined as the time interval between the onset and the offset. Another distinctive characteristic of the PPG waveform is the second maximum called the diastolic peak. The valley formed between the systolic and diastolic peaks is named the dicrotic notch. Finally, the interval between the two peaks is denoted as the Pulse Propagation Time (PPT).

The physiology behind this particular morphology can be simply explained. When ventricular contraction of the heart occurs during systole, a big amplitude pressure wave propagates towards the periphery. This corresponds to the systolic peak. However, due to the complexity of the systemic circulation in the central arterial tree, different arteries bifurcate and the decrease in their diameter generates new reflected pressure waves that are transmitted in a delayed manner to the peripheral vasculature [21]. This occurs during early systole and gives origin to the diastolic peak. The PPT corresponds to the reflection time, i.e., the time delay between the systolic wave and the second reflected wave.

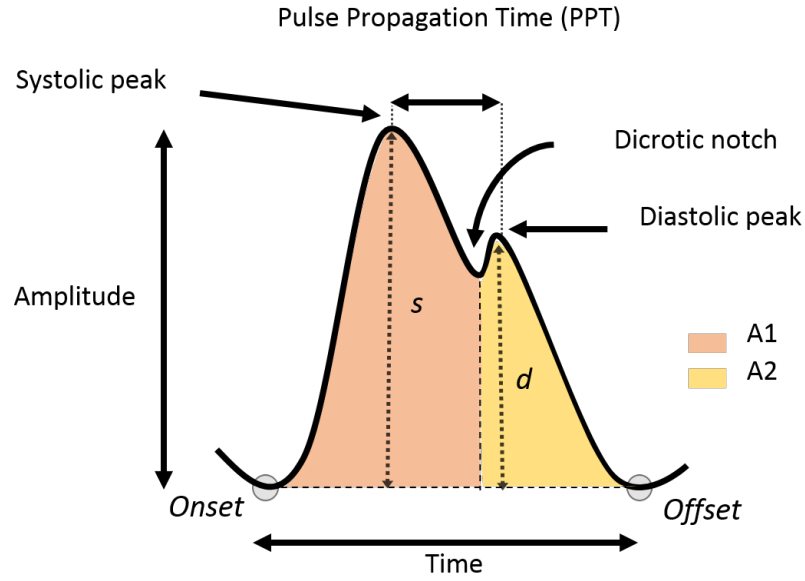


Figure 2.4: Diagram of the standard PPG waveform annotated with its main pulse characteristics.

2.1.5 Changes in pulse shape

Different factors can modify the morphology of the standard PPG pulse shape.

Physiological factors: There are some subject specific factors which confer some unique morphological characteristics to the PPG signal. These are also the principal reason explaining the huge variability in PPG shapes among and intra-subjects.

- *Medical conditions:* Age and cardiovascular diseases (e.g. hypertension, atherosclerosis etc.), result in changes in the vascular properties of the arterial tree, and hence, directly alter the pulse shape of the PPG waveform [1, 22]. Indeed, arterial stiffening entails an increase in the pulse wave velocity, leading to the occurrence of the reflected wave during late systole [21]. This means that the diastolic reflected wave travels faster than normal and overlaps with the systolic wave, preventing the distinction between them in the PPG pulse [23]. The PPT is therefore reduced, and the diastolic peak can apparently disappear [24]. Pulse rise time was also observed to increase with age [24].
- *Autonomic system regulation:* The PPG waveform is modulated by the sympathetic nervous system and thermoregulation activity [25]. High temperatures and other vasodilator agents, can increase the PPG amplitude and decrease the height of the dicrotic notch, as a result of an apparent loss of the reflected wave [21]. In this case, the PPG pulse waveform seems to undergo a triangulation. In addition, the DC baseline of the PPG signal is modulated by the respiration frequency [1, 26, 27].

- *Measurement site:* Different body parts have different capillarity and tissue characteristics, showing variations in the reflected light optical path. As a consequence, not only the DC baseline but also the AC component can be impacted. In fact, peripheral areas have a larger vascular bed, resulting in turn in a higher PPG amplitude. The variation in perfusion volume is also greater. Moreover, depending on the anatomical site, different reflection waves can be transmitted through the complex arterial tree with different amplitude and transit times (PTT) [28]. As a result, the PPG pulse contour can be affected, with changes in the diastolic peak height and occurrence time [24, 29].

These physiological factors are observed in the photoplethysmography waveform. They do not usually disturb the quality of the signal, but they have to be taken into account when interpreting the PPG pulse contour, specially when studying a novel measurement site.

Artifacts: Artifacts are noise components that do not naturally belong to the PPG signal itself. They corrupt the PPG signal by distorting the time and frequency characteristics with varying severity. As a consequence, they result in inaccurate measurements and false alarms during clinical monitoring [30]. Some artifacts can be efficiently filtered when they do not alter the fundamental frequency of the PPG signal. But in general, artifacts can severely disturb the frequency bands of interest, making it challenging to restore the clean PPG signal. Furthermore, artifacts can be very variable depending on the measurement site, the pulse oximetry sensor mode, or the subject wearing the sensor. For instance, the neck will be susceptible to artifacts typical of the trachea, such as swallowing, coughing or talking, while these will not have an impact in other body parts. Likewise, the sensor attachment can be less stable for the reflective PPG mode. The differences in anatomical characteristics among individuals can also influence the amplitude and duration of the artifacts produced.

Despite the uniqueness, it is still possible to characterize artifacts with general features when some of the sources of variability, such as the measurement site, can be fixed to narrow the problem. In the case of finger pulse oximetry, PPG artifacts are well known and recurrent. This facilitates the application of common filtering techniques as analyzed in the following sections.

2.2 Overview of PPG artifacts cancellation methods

The most common sources of noise PPG artifacts have been identified as motion, powerline interference of ambient light, displacement disturbances of the sensor, and heterogeneous optical properties of tissue [31]. In the literature, an ample range of signal processing techniques have been proposed for artifact cancellation in PPG signals. Special attention has been paid to in-band low frequency motion artifacts, which are usually considered the most problematic ones.

One of the simplest filtering techniques explored is the **Moving Average Filter** (MAF). By sequentially averaging point by point of the PPG waveform with values of the neighbouring data points, baseline drift can be successfully removed [32]. Despite the simplicity, this method is not very effective because it cannot deal with sudden or big amplitude noise, and the PPG signal is usually degraded. To solve this issue, a Periodic Moving Average Filter (PMAF) was proposed [33], where the novelty of segmenting the signal into periods and carrying out the averaging of points across periods as well, eliminated artifacts without deteriorating the signal.

Some other studies have focused on **Fourier series analysis**. A cycle-by-cycle Fourier serial analysis (CFSA) demonstrated that clean PPG signals can be reconstructed from the Fourier series coefficients, extracted from noisy PPG at each cycle [34].

Another noise cancellation technique that not only captures the frequency information of the PPG signal but also the temporal resolution is **Wavelet denoising**. Noise reduction consisted on, first, decomposing the PPG corrupted signals using the stationary wavelet transform (SWT); and then reconstructing clean PPG with the inverse SWT and wavelet transform modulus maxima (WTMM) [35]. Other studies [32, 36] reported promising performance as well, by implementing the discrete wavelet transform (DWT) with the Daubechies (db4) wavelet. Nevertheless, the major drawbacks are that PPG signals are not completely restored, and a phase shift is susceptible to be introduced during reconstruction [37].

One of the most popular trends in motion artifact reduction is **Adaptive filtering**. The superiority of this method relies on the filter coefficients that are iteratively adjusted, to estimate the noise at each iteration, and ultimately subtract it from the corrupted PPG. To this end, a reference noise signal is required and it is generated either by direct measurement of motion with accelerometers sensors [15, 38, 39], or by a synthetic noise generator [40]. A least mean square (LMS) algorithm is then typically implemented to recursively minimize the error cost function between the true and estimated noise components, for a precise update of the filter coefficients. Although optimal filters are very powerful, some studies argue that, as wavelets, they could introduce phase shifts in the PPG signal [37]. Also, some works discuss that a correlation might not always exist between accelerometers signals and motion artifacts [40].

Other studies have exploited the **Independent Component Analysis** (ICA) method to separate artifacts sources from clean PPG. To illustrate, the combination of ICA and a preprocessor, consisting of a block interleaving with a low pass filter, demonstrated to reduce the noise significantly [41]. Moreover, clean PPG extraction from contaminated PPG signals was also successful when combining ICA and an adaptive filter [42]. Although ICA assumes that all the sources constituting noisy PPG are independent, there might exist a correlation between the arterial flow and the movements of the

patient, turning this method invalid.

Finally, as an alternative to previous techniques that assume the existence of an inherent noise underlying the whole PPG signal, other studies have opted to follow a **Motion Artifact Detection-Reduction** strategy, where only contaminated PPG detected segments are filtered [43, 44, 45, 46]. Since this technique will be adopted in this research work, an extended literature review is presented in next section.

2.3 PPG artifacts detection and classification

The next subsections compare the most relevant PPG artifacts detection studies in the recent years. A summary of each approach is presented in Table 2.1. To discriminate between clean PPG and noise artifacts, various algorithms usually followed the same steps, with slight changes in the methodology. First an experimental protocol is designed or a database is selected to obtain the set of PPG recordings to be analyzed. Second, a segmentation of the signals and corresponding labelling into corrupted or clean categories is performed. Subsequently, unique features of the PPG waveform are extracted to be inputted in a classification model for pulse quality assessment. When the features subset is too large, some studies also include a feature selection step before the classification process. Finally, once contaminated PPG segments have been identified, they can be directly rejected, filtered or reconstructed implementing the methods discussed in section 2.2.

2.3.1 Experimental protocol and datasets

In the literature, some research groups used specific datasets to test the performance of their noise detection algorithms. Others, opted to record their own data in a series of experiments where specific artifacts can be simulated. But overall, the general characteristics of the employed datasets, can be summarized as:

- **Measurement site:** The most common PPG modality was transmission finger pulse oximetry. Some of the reviewed studies also used reflectance PPG in other body locations such as finger [63], wrist [57, 60, 61], and forehead [43, 55]. It is worth highlighting the sensing modality used in [58, 59], which used the camera and flashlight of a smartphone to extract finger PPG signals.
- **Datasets size:** In controlled experiments, where healthy participants were recruited for data acquisition, cohorts usually range from 10 to 20 subjects [51, 57, 63]. Subjects are typically asked to simulate defined artifacts at specific times. Larger cohorts are found in studies where the dataset is constructed from one or several annotated databases [54, 56]. This is important to validate the algorithms generalizability in a wider population and in different clinical contexts.

Table 2.1: ARTIFACTS CLASSIFICATION ALGORITHMS DISCRIMINATING BETWEEN GOOD AND BAD QUALITY PPG SIGNALS.

Author	Sensor Site	Subjects	Window	Type of corruption	Features	Classifier	ACC (%)	SE (%)	SP (%)	F1 (%)
[45] Krishnan	Finger reflectance	10 healthy	Short frames of equal length	Finger, wrist & elbow movements in controlled experiments	High order statistical features from: time, freq, bi-spectral domains	Neyman-Pearson rules for each feature+global decision fusion	91	-	-	-
[47] Selvaraj	Finger, ear & forehead	24 healthy	1min (10s sliding window)	Involuntary movements in clinical setting + voluntary movements in experiments	Kurtosis & Shannon Entropy	Decision rules for each feature + decision fusion	88.8	86.9	98.3	-
[48] Sukor	Reflection finger	13 healthy	Pulse-by-pulse	8 types of finger and wrist movements	Time morphology characteristics	Decision lists	83 \pm 11	89 \pm 10	77 \pm 19	-
[49] Li	Finger	MIMIC II database [50] (105 patients)	6 s	ICU data during arrhythmia alerts	4 pulse template matching signal quality indexes	MLP	95.2	99	80.6	-
[43] Chong	Reflectance finger & forehead	20 healthy subjects	3, 5, 7*, 9, 11 s	Daily activities & lab controlled movements & simulated motion	4 standard deviation time features	Linear SVM + major voting	93.9	94.3	92.4	-
[51] Couceiro	Finger	8 healthy & 7 CVD patients	Beat-to-beat	11 hand and body movements in controlled experiments	26 time and period domain	SVM	87.5 \pm 0.6	78.4 \pm 1.2	94.4 \pm 0.6	-

* Best classification window, for which the performance is reported.

ACC: Accuracy, SE: Sensitivity, SP: Specificity, F1: F1-score, ICU: Intensive Care Unit, CVD: Cardiovascular disease, MLP: Multilayer perceptron, SVM: Support Vector Machine.

	Author	Sensor Site	Subjects	Window	Type of corruption	Features	Classifier	ACC (%)	SE (%)	SP (%)	F1 (%)
[52]	Orphanidou	Finger	19 hospital patients	10 s	Ambulatory monitoring in hospital	HR, peak-to-peak intervals, template matching correlation coefficient	Decision list	-	91	95	-
[53]	Cherif	Finger	Same dataset as in [48]	Pulse-by-pulse	Finger movements induced for 20s	Derivative of the correlation coefficient between each beat & a pulse template	Non-parametric random distortion testing (adaptive thresholding)	83 ± 8	84 ± 16	83 ± 12	-
[54]	Elgendi	Finger	40 emergency responders	Window length tested from 1s to 30s. → 2-5s optimal window for skewness feature	Rest periods before and after exercise (x3). Random stitching of PPG segments	8 SQIs from time, statistical and freq domains	LDA, SVM, QDA & Mahalanobis distance	-	-	-	79.1*
[55]	Dao	Finger & forehead	Several datasets from laboratory experiments + emergency patients	8s with 50% overlap	Controlled motion artifacts + Treadmill exercises + Involuntary movements in emergency unit	Time-frequency spectrum features	TifMA: → SVM + traceback strategy → HR tracking usability metric	92 ± 2	93 ± 2	90 ± 3	-
[56]	Fischer	Finger	69 subjects from three databases	Beat-to-beat	ICU, sleep & ergometry laboratory data	Time domain analysis of pulse waveform	3 decision lists (threshold based)	98.3	99.6	90.5	99
[57]	Pradhan	Empatica E4 wristband	15 healthy	10s	24h daily activities	Time, freq & correlogram	Naïve Bayes	-	98.3	98.3	-

* Average performance for all classifiers for the best Skewness feature.

ACC: Accuracy, SE: Sensitivity, SP: Specificity, F1: F1-score, SQI: Signal Quality Index, LDA: Linear Discriminant Analysis, SVM: Support Vector Machine, QDA: Quadratic Discriminant Analysis, ICU: Intensive Care Unit, HR: Heart Rate.

	Author	Sensor Site	Subjects	Window	Type of corruption	Features	Classifier	ACC (%)	SE (%)	SP (%)	F1 (%)
[58]	Poh	Finger smartphone	Several databases: outpatient clinic & ICU patients	17s	Data with motion artifacts, AF and ectopic beats	Self-learned	DCNN	-	97	100	-
[59]	Tabei	Finger smartphone	35 healthy & 5 AF patients	7s	Hand & fingertip movement artifacts + AF patients, no movements.	Time, statistical and time-freq	→ PNN <i>personalized</i>	98 ± 2.0	92.6 ± 6.5	99.7 ± 0.9	-
							→ PNN <i>generalized</i>	89.9 ± 9	84.2 ± 14	93.6 ± 7	-
[60]	Vandecasteele	Empatica E4 wristband	17 epilepsy patients	7 s	24h recordings. <i>Automatic MNA labelling using ECG.</i>	Time & freq domains	LS-SVM	90.2 ± 3	85.5 ± 7	92.3 ± 4	-
[61]	Pereira	Finger & wrist	13 stroke patients & 3764 ICU patients	30s	Continuous hospital recordings, including motion artifacts and AF events	42 features from time and spectral domains	SVM	90.5 ± 0.5	91.3 ± 1.0	88.8 ± 0.6	92.9 ± 0.4
[62]	Goh	Finger	4 datasets: healthy & ICU patients	5s	Controlled laboratory movements + Continuous monitoring ICU patients	Self-learned	CNN	-	96.6	91.2	-

ACC: Accuracy, SE: Sensitivity, SP: Specificity, F1: F1-score, AF: Atrial Fibrillation, CNN: Convolutional Neural Network, DCNN: Deep CNN, PNN: Probabilistic Neural Network, SVM: Support Vector Machine, LS-SVM: Least Squares SVM.

Combination of several publicly available datasets is specially necessary for the development of deep learning models [58, 62].

- **Artifacts types:** The bad quality pulses to analyze were of different origins. Some resulted from involuntary movements of daily activities, and others were simulated in controlled laboratory experiments. They include for example, finger, hand or arm movements, and displacement of the sensor probe [48]. More realistic data incorporates cases of patients with cardiovascular diseases (CVD) [51, 59, 61], epilepsy [60], being admitted to the emergency department [55], in intensive care units (ICU) suffering from arrhythmias [49, 56, 58, 61], or healthy emergency responders after heat-stress conditions [54]. The aim of these studies was to include beats of unusual signal quality, that were not of induced artifacts origin.
- **Time recordings:** The recordings duration of the controlled experiments was commonly of 60 seconds, with the induced artifacts occurring in the middle interval (20-40s) [43, 48, 51]. This type of recordings were usually preceded by 5min of rest. In a particular study [54], in order to increase variability, 60s recordings were synthetically generated through random stitching of three 20s segments, from distinct subjects and various signal qualities.

2.3.2 Pulses segmentation and classification windows

In a first step, recordings need to be segmented for feature extraction. Some studies focused on a beat-to-beat segmentation to extract information from single PPG pulses. For that, characteristic points of the pulse contour, such as the pulse onset, offset and systolic peak, needed to be determined. Different segmentation techniques were applied in the literature to find these. First in a pre-processing step, the PPG signal was typically filtered with a 4th order band-pass Butterworth filter (0.5-12Hz), to remove the DC and low frequency components. Then, some algorithms used the lower envelope, with a minimum filter, to detect the troughs of the PPG pulses [48]. Similarly, the minima, marking the start of the segmented pulses, could be identified after applying a moving-average-filter to the original signal [53]. Systolic peaks were subsequently identified as the maximum between the two troughs of the pulse. Other techniques rely on the PPG signal derivatives, and more specifically use the maximum of the 3rd derivative previous to the maximum of the 1st derivative, to determine the onset of the pulse [51]. Finally, another simple but effective method to find peaks and valleys, was adaptive thresholding with an amplitude correction factor, that amended undesirable diastolic peaks [56].

Other studies, did not concentrate in such a small beat-to-beat scale, but rather, in time windows. These did not require complex segmentation techniques, since calculated features were simply averaged over subsequent time frames [43, 49, 52, 54, 57, 59]. For instance, Dao *et al.* implemented a sliding

window of 8s with 50% overlap [55]. Elgendi *et.al* [54], optimized the window size from 1s to 30s in steps of 1s, and found that windows of 2-5s were optimal for quasi-real time assessment of pulses quality using the skewness statistical feature. Other time and frequency features required longer windows, and were extracted using the whole duration (60s) of the recordings. Chong *et al.* [43], also tested different classification windows (3, 5, 7, 9 & 11s), and identified the 7s one as the most optimal for discriminating artifacts from clean PPG fragments.

2.3.3 Features extraction

The discriminatory capability of artifacts detection algorithms mainly relies on the features extracted from the PPG signal, being able to subtly differentiate between corrupted and clean PPG segments. This is why, feature extraction is one of the most important steps in the classification process. Different feature sets have been proposed in different works. However, the two main categories are: time and frequency domain features. Some exceptions to these are period domain characteristics, such as height, length, width and area of the predominant peaks in the spectrum, obtained by applying the Period Short Fourier Transform (PD-STFT) [51]; or the quadratic phase coupling (QPC) extracted from bi-spectral analysis [45, 63].

In the time domain subgroup, absolute and relative pulse morphological characteristics prevail. To illustrate, the simplest PPG traits extracted are: pulse amplitude [48, 51, 56, 59], pulse length [48, 51, 56], peak distance [51], peak height and trough depth difference [48, 51], and rise time [56]. Other contour metrics such as pulse skewness and kurtosis have also been demonstrated to have great noise discriminative capability [45, 51, 54, 63]. Alternatively, other studies focused on choosing features reflecting the variability of corrupted PPG segments in the form of the standard deviation of peak-to-peak interval and amplitude [43, 59], systolic and diastolic ratio, and the mean standard deviation of the pulse shape [43].

Some characteristics cannot be directly extracted from PPG signals, but need a pulse template to be built from the average of consecutive pulses, in a time window. Signal quality indices can then be calculated based on the pulse-by-pulse comparison with that template. Some examples include the template matching and the dynamic time warping correlation coefficients [49, 52], or the derivative of the correlation coefficient [53].

In contrast, some features require larger time windows compared to beat-to-beat characteristics. To illustrate, some signal quality indices are relative power, zero-crossing rate, and signal to noise ratio (SNR) [54]. Hjorth parameters, which are commonly used in electroencephalography (EEG), have also been implemented in a PPG artifacts detector [64].

In the frequency domain, it is worth highlighting the use of power spectrum values at specific

frequencies (5,7,9,13 Hz), the number of median-crossings of the instantaneous frequency [57]; the relative power [54], and the frequency kurtosis features [45, 63]. Other spectral characteristics include spectral entropy, total power, and non-linear features derived from the Pointcare plot [61].

Dao *et al.* [55] proposed to derive the time-frequency spectrum (TFS) of the PPG using the variable frequency complex demodulation (VFCDM) technique, and extract features from it. These included the residual noise power, the projected frequency modulation difference and the HR difference, between the frequency modulation and the time domain calculation.

In order to avoid pulses segmentation and manual feature engineering, recent efforts have focused on training deep learning models that self-learn the features by error backpropagation [62, 65].

2.3.4 Feature selection

The success of artifacts detection algorithms relies on the selection of the most relevant features to maximize the performance of the classification process. Training on a large set of features is prone to introduce high variability in the classifier and increase the computational effort. Therefore, it is crucial to detect any possible linear or non-linear interactions among extracted features to not only diminish redundancy in the feature set, but also to enhance the discrimination and generalization capability of the classifier.

Some studies implement feature selection techniques to reduce the number of extracted features. The minimum redundancy maximum relevance (mRMR) filter method was applied in [66]. It consists on ranking the most relevant features using mutual information metrics. A later version of mRMR, is the Normalized Mutual Information Feature Selection (NMIFS) algorithm. The normalized Mutual Information is used here as the measure of redundancy. This method implemented in [51], has the advantage to not require any manual parameters tweaking. In another study [57], the Bhattacharya distance, a measure of the separation between two distributions of data, was used to reduce the features subset. Vandecasteele *et al.*, implemented backwards wrapper feature selection using a leave-one-patient-out cross-validation approach to find the most optimal set that maximized the F1-score.

The majority of studies in the literature do not include any feature selection step in their algorithms, since the number of features is already small enough. Such an optimal subset of features was probably obtained by trial and error of previous experience, or domain knowledge.

2.3.5 Classification methods

Once features of the PPG waveform are extracted and properly selected, they can be input in a classifier for pulse quality assessment. Here, a comparison of the main classification approaches found in the PPG literature is presented.

The most elementary type of classification model found was decision lists. Sukor *et al.* [48], proposed a threshold-based classifier, with pulse resolution, based on waveform morphology analysis to classify whether a pulse was of acceptable quality or not. This algorithm achieved an average sensitivity (SE), specificity (SP) and accuracy (ACC), of 89%, 77% and 83%, respectively. Cherif *et al.* [53], argued that this method cannot be implemented in a stand alone system since the template and thresholds are predefined. In contrast, they proposed Random Distortion Testing (RDT), a robust non-parametric hypothesis testing, relying on a single adaptive thresholding decision, instead of 6 manually fixed thresholds. Despite the efforts, the classification of a pulse as an artifact or not, yielded similar outcomes as previously obtained by Sukor *et al.*, with mean 84% (SE), 83% (SP) and 83% (ACC). Another decision lists algorithm, proposed by Orphanidou *et al.* [52], based on the correlation coefficient of each PPG pulse with an adaptive template, achieved higher sensitivity (91%) and specificity (95%) in discriminating good and bad quality PPG segments. Fischer *et al.* outperformed these results with 99.6% (SE), 90.5% (SP), 98.3% (ACC) and precision of 98.5%, using 3 decision lists based on contour analysis [56]. This novel embedded algorithm, differentiating artifacts and reliable pulses, achieved segmentation and beat-to-beat artifact detection in real-time, and it could run in devices with limited memory and computational power. However, even though this model was validated with three different clinical databases, it can be criticized that it uses too many manually defined thresholds and normative ranges (13), and therefore, lacks generality.

In the same line, Krishnan *et al.* [45, 63] applied a Neyman-Pearson detection rule to individual high order statistics features and then fused the individual measures through the Varshney-Chair rule to further enhance the artifact detection capability. This combination resulted in a sensitivity of 91% and a specificity of 94%. However, the reduced number of repetitions of each artifact in the dataset, added to the small number of subjects, questioned the validity of this method. In addition, a simple probabilistic naive Bayes classifier, suggested a better performance in discriminating clean from corrupted PPG with 98.3% (SE) and 98.3% (SP) [57]. For complete validation of this approach, it would be interesting to analyse the reliability of the selected features in patients suffering arrhythmias for example. This medical condition, characterized by an increased heart rate variability, would probably negatively impact the class separability potential of autocorrelation features.

More advanced classifiers with the ability to model non-linear decision boundaries and showing robustness against overfitting, included Support Vector Machines (SVM). To illustrate, time and period domain characteristics input in a SVM classifier demonstrated good performance with 84.3% (SE), 91.5% (SP), 88.5% (ACC) [51]. Nevertheless, in this work, the discriminative capability of the selected features is suspected to be dependent on the particular context of differentiating between a CVD group of subjects and healthy controls. Vandecasteele *et al.* [60], trained a Least-Squares

SVM (LS-SVM) with data from 17 epileptic patients recorded with an E4 Empatica (Empatica Inc.) wristband, for a duration of 24h. The motion artifacts detection algorithm implemented time and frequency PPG features, as well as accelerometer characteristics, and achieved an ACC of 90.23%, SE of 85.5% and SP of 92.3%. Chong *et al.* [43], proposed to incorporate a major voting of neighbouring PPG segments approach, to a SVM classifier, to enhance the classification decision. This algorithm outperformed previous results with average 94.3% (SE), 92.4% (SP), 93.9% (ACC) [43]. Dao *et al.* [55] also combined an SVM classifier with a traceback strategy in the first stage of their TifMA algorithm, to differentiate motion artifacts segments from clean PPG data and locate their onset and offset times with 1s precision. For that, time-frequency characteristics from the TFS obtained through VFCDM, were exploited. In a second stage, the algorithm assessed whether the identified corrupted PPG segments, were usable for HR extraction or not with average 92% (ACC), 93%(SE) and 90% (SP).

Some works, trained several classifiers and evaluated their performance for model selection. El-gendi [54] tested four classification models: Mahalanobis distance, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Support Vector Machine (SVM). The aim was to find a single optimal signal quality index for PPG signals able to discriminate between excellent, acceptable and unfit pulses with high accuracy. Skewness showed the best overall performance in differentiating excellent pulses from acceptable (F1=86%), unfit (F1=79.1%) and both together (F1=87.2%), with respective optimal window sizes of 3s, 5s and 2s. Pereira *et al.* [61], also tested three different classifiers: an SVM, a K-nearest neighbors, and a decision tree, to investigate which one resulted in the best performance. The SVM model, trained with time and frequency features, was the most successful one in differentiating between good and bad quality signal fragments (ACC=90.5%, SE=91.3%, SP=88.8%).

Artificial neural networks (ANN) were also proposed in the literature. These consist of a set of connected nodes organized in different layers, that use non-linear activation functions to discriminate between corrupted and clean PPG signals. A type of feed-forward ANN, a Multilayer Perceptron (MLP), with 10 nodes and 6 input features, achieved a performance of 99.0% (SE), 80.6% (SP) and 95.2% (ACC) in the classification of PPG pulses with different signal qualities [49]. Cherif *et al.* [53], argued that this approach lacked generality since the weights were specific to the dataset. A probabilistic neural network (PNN), was proposed by Tabei *et al.* [59], to detect motion artifacts using two different training approaches. On one hand, a generalized PNN, was trained on the clean and corrupted PPG segments from all subjects, except the test subject. On the other hand, a personalized PNN was trained on the clean data of the test subject and the artifacts data of the rest of the subjects. Both methods, were tested in one test subject at a time, including clean and corrupted PPG segments. The results showed that the personalized PNN improved the performance compared to the generalized

approach, with average 98 ± 2 % (ACC), 92.6 ± 6.5 % (SE) and 99.7 ± 0.9 % (SP). Despite the promising findings, a limitation of this personalized method is that the PNN must be retrained, prior to being implemented on every new user.

Finally, deep learning convolutional neural networks (CNN) were also explored in the recent years. Goh *et al.* [62] proposed a one-dimensional CNN architecture for motion artifact detection in PPG, trained on different datasets including laboratory controlled experiments and ICU data. This 13-layers deep learning model demonstrated the ability to successfully distinguish artifacts from clean PPG segments with an overall 94.5% (ACC), 96.6% (SE) and 91.2% (SP). Similar performance was obtained, with a deep CNN (DCNN), with six densely connected blocks and a total of 201 layers proposed by Poh *et al.* [58]. This DCNN model was trained to distinguish among 4 different classes including: artifacts, synus rythm clean PPG, ectopic beats, and atrial fibrillation segments. The sensitivity and specificity of noise-corrupted PPG segments detection was 97% and 100%, respectively. Despite the high discriminatory power of deep learning, CNNs are very sensitive to weights initialization and rely on large-scale manually annotated PPG datasets. These are indeed, difficult, time-consuming and costly to obtain.

2.3.6 Discussion

Some of the most relevant papers in PPG artifacts classification have been reviewed in this section. In these studies, many efforts have being put into designing an ideal real-time algorithm with pulse-by-pulse resolution, showing high generality to PPG signals of different qualities, and at the same time achieving a high discriminative performance. However, current approaches still do not completely fulfill all these requirements at the same time.

First, increasing the algorithms generality could adversely affect the detection performance. Indeed, it would critically limit the features specificity to the classification task, since their discrimination capability is reliant on each particular artifacts context. Likewise, in many works, the architecture of the algorithm appeared strongly dependent on the dataset. For instance, many decision lists relied on manually defined thresholds, and the personalized PNN needed to be trained on every specific subject to be successful.

Despite some efforts to validate the proposed algorithms in various publicly availability datasets, including motion artifacts, realistic ambulatory conditions, or ICU patients suffering from atrial fibrillations; developing a general model that can be used in many different clinical contexts indistinctively, seems a difficult task to be accomplished in the near future. A standardized large-scale database, with annotated PPG signals from all types of medical conditions and environments, would be required.

Another important aspect that is currently limited in the field is the real-time implementation.

Only one algorithm [56] of the reviewed studies, was successfully executed in real time in a micro-controller. Nevertheless, this approach was solely based on short time beat-to-beat window features, which restricted the inclusion of features with high reliability in detecting changes over longer periods of time. This is for example the case of Hjorth parameters [64], that require a larger window size than a single beat.

Another issue is the number of features and complexity of the algorithm that directly impacts the computational power. Some papers have proposed large features subsets and more complex classification models. Despite the higher classification performance, these might have the drawback of being computationally expensive in contrast to smaller features subsets and simpler decision lists.

The artifact detection work developed as part of this thesis tries to tackle these issues, in order to find a trade off between generality, features specificity, computational efficiency and high performance for successful real-time implementation.

2.4 Automatic apnea detection using PPG sensors

2.4.1 Apnea types and definitions

An apnea is a temporary interruption of breathing typically occurring during sleep, that can lead to the desaturation of oxygen in blood circulation. Three types of apnea events exists, as shown in Figure 2.5, depending on the cause of respiratory cessation [67]. Obstructive apnea manifests when the upper airways are blocked preventing the passage of air. However, thoracic and abdominal respiratory efforts still occur. When the obstruction is only partial, allowing some shallow breathing, it is called hypopnea. Central apnea is another type, that is characterized by both the absence of airflow and the cessation of respiratory movements. It typically originates when the brain, is unable to send the appropriate neurological messages to the muscles that control respiration. Mixed apnea is a combination of both and occurs when the airways obstruction follows a central respiratory pause.

2.4.2 Studies relevant to the current research

Unfortunately at present, due to the difficulty of witnessing and monitoring SUDEP, there are no available oximetry-based approaches for real-time seizure-related apnea detection in epilepsy [68]. The only research work using oximetry for seizure detection, searches for simple decreases in SpO_2 in combination with HR and EDA changes [69]. However, sleep apnea-hypopnea syndrome (SAHS) has been widely studied and there are many publicly available polysomnography (PSG) databases. This has promoted the development of many automatic apnea detection algorithms in the past years [70, 71, 72]. These exploit a wide variety of sensors including: oximetry, ECG, sounds and respiratory

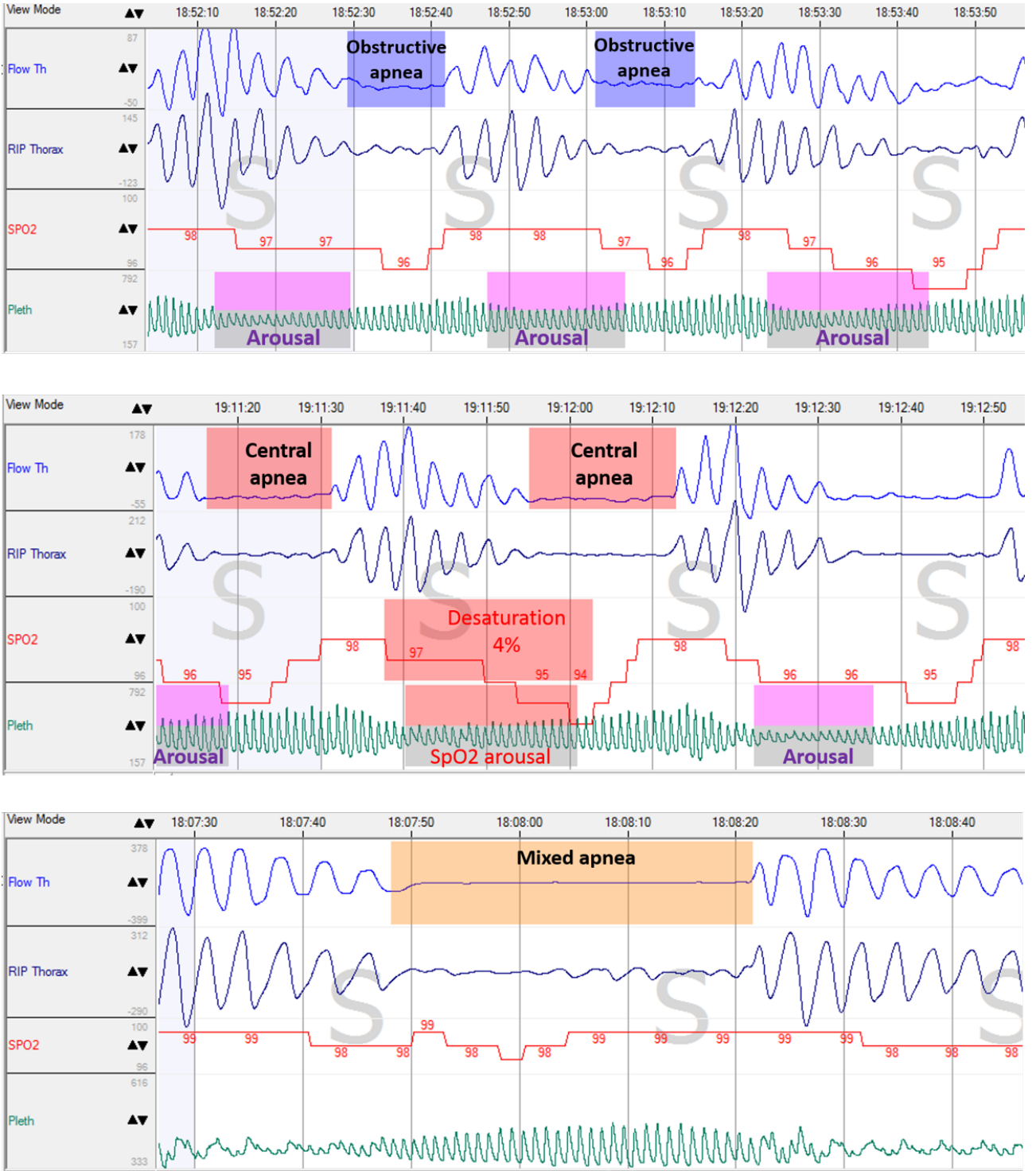


Figure 2.5: Obstructive, Central and Mixed apnea recordings. Four different channels are shown: flow thermistor, respiratory inductance plethysmography (RIP) at the thorax, SpO₂ and PPG signals. During the three apnea cases, the flow is clearly interrupted. During obstructive apneas, the RIP thorax trace still shows some small oscillations, indicating the presence of respiratory efforts. During central apneas, no RIP activity can be observed. In the mixed apnea case, the RIP channel oscillations only appear at the end.

signals, as well as multiple combinations of these [66, 64]. However, only oximetry approaches using PPG sensors, are relevant to the present research work. That is why in the following subsections, only PPG and SpO₂-derived techniques are considered.

In the context of SUDEP prevention, where continuous physiological monitoring is critical, works that have focused on subject-based classification for sleep apnea diagnosis are irrelevant. Here the focus is on the papers in the literature that propose either epoch-based apnea detection or direct SpO₂ desaturations detection. These algorithms classified each individual epoch or desaturation event as apneic or normal.

In the next subsections, the methodology followed by the most relevant papers performing apnea detection with single PPG sensors is presented. We only reviewed algorithms that were potentially suitable for online processing. Summaries of the main aspects of the proposed SpO₂ and PPG approaches are presented in Table 2.2 and Table 2.3, respectively. The papers are listed, in ascending order, by year.

2.4.3 Datasets and cohorts

There are many PSG databases publicly available that have been extensively used by the different research works in the literature to train and test algorithms for PPG-based apnea events detection. The majority of these included subjects suspected to have sleep apnea. These had different degrees of SAHS severity, diagnosed based on the apnea-hypopnea index (AHI). This index, measuring the total number of respiratory events occurring per hour of sleep, divided SAHS patients into one of the four categories: normal ($AHI < 5$), mild ($5 \leq AHI < 15$), moderate ($15 \leq AHI < 30$) or severe ($AHI \geq 30$).

The most commonly used databses in SpO₂-based approaches, were:

- *The Apnea-ECG Physionet database.* This consists of 8 full PSG night recordings including the SpO₂ channel, with half of the cohort diagnosed with severe SAHS and the rest as controls. This dataset was used by several studies [74, 75, 77, 81, 82]. Annotations were made minute-by-minute.
- *The University College Dublin (UCD) Sleep Apnea database.* It also collected by Physionet and included 25 subjects with average AHI of $24.1 \pm 20.3s$. Annotations consisted of onset times and duration of obstructive, central, mixed and hypopneas respiratory events..
- *The dataset from the Gran Canaria University Hospital (HuGCDN2008).* It included 70 patients suspected to have sleep apnea and was annotated every 30s, but many works re-labeled the recordings every minute to facilitate the comparison with other databases [81, 82, 80].
- *The Sleep Heart Health Study (SHHS) dataset.* This contains recordings from 5793 participants undergoing full night PSG from different sleep laboratories. An uniform distribution of subjects into the four AHI categories was ensured [77]. Desaturations are labeled as being linked to an apnea or not.

Table 2.2: APNEA DETECTION ALGORITHMS: SPO₂ APPROACHES.

	Author	Epoch length	Features	Classifier	Datasets	ACC	SE	SP	PPV
[73]	Burgos	1 min	ODI & TSA	Bagging ADTree	Apnea-ECG	93.03	92.35	93.52	-
[74]	Almazaydeh	Desaturation	ODI3, Δ index, CTM50	NN	Apnea-ECG	93.3	87.5	100	-
[75]	Mostafa	1 min	Time, freq and wavelet	ANN + GA	Apnea-ECG	97.7	96.5	98.5	-
[76]	Pathinarupothi	1 min	Self-learned	LSTM-RNN	Apnea-ECG	95.5	92.9	-	99.2
[77]	Deviaene	Desaturation	Time, desaturation severity and quasi- periodicity	Random Forest	UZ Leuven, Apnea-ECG, SHHS	82.8	64.3	88.6	-
[78]	Deviaene	1 min	SpO ₂ as in [77] + PPG time and freq	LS-SVM	UZ Leuven	83.4	73.7	86.6	64.8
[79]	Jung	1 min	3 characteristic points in the desaturation	Decision lists	Seoul National University Hospital	91	83	89	-
[80]	Mendonça	5 min (1 min overlap)	Time & freq	Logistic regression	HuGCDN2008	88	80	91	-
[81]	Mostafa	1, 3, 5 min (1 min overlap)	Self-learned	CNN + GA	Apnea-ECG (1min)*	94.24	92.04	95.78	-
					UCD (3min)*	85.79	60.38	93.9	-
					HuGCDN2008 (5min)*	89.32	74.75	94.44	-
[82]	Mostafa	1, 3, 5 min (1 min overlap)	Self-learned	CNN + GBO	HuGCDN2008 (5min)*	88.49	73.64	93.8	-
					Apnea-ECG (3min)*	95.14	92.36	97.08	-

*Best epoch length performance

ODI: oxygen desaturation index, TSA: time spent in apnea, CMT50: the central tendency measure with radius 0.5, NN: Neural Network, LS-SVM: Least-Squares Support Vector Machine, LSTM-RNN: Long Short-Term Memory Recurrent Neural Network, CNN: Convolutional NN, GA: Genetic Algorithm, GBO: Greedy Based Optimization.

Table 2.3: APNEA DETECTION ALGORITHMS USING THE PPG SIGNAL.

	Author	Subjects	Epoch length	Features	Classifier	ACC	SE	SP	PPV
[83]	Knorr-Chung	20 sedated (17 intubated / 3 spontaneous breathing)	5 s	Time & freq domains	NN (obstructive apnea/normal)	75.4	91.6	84.7	85.9
[84]	Pradhapan	40 healthy subjects (breath-holding experiments)	15s <i>selected</i> segments	PPG time (RI & SI) + Power ratio between the LF (0.04-0.4Hz) and HF (0.15-0.4Hz)	SVM	97.22	94.62	100	-
[85]	Gaurav	16 healthy subjects (breath-holding experiments)	Continuous	Area under power spectrum (0.05-0.5Hz) of PPG-derived respiratory signal	Wavelet decomposition + Threshold	<i>Under 10 s detection*</i>			
[86]	Lázaro	PSG recordings from 21 children (10 with OSA)	Continuous detection / Performance over 1h fragments	PRV freq features from 4 windows [before, during, post-apnea (5s) & global (20s)]	LDA (obstructive apnea/normal)	70.37	81.82	68.57	-
[87]	Uçar	5 PSG subjects (3 sleep apnea patients)	10 - 30 s (<i>variable</i>)	Time domain pulse characteristics	MLFFNN	97.07	98	96	-
[88]	Bozkur	10 OSA patients	>10 s (<i>variable</i>)	Time and freq from PPG & PRV	Ensemble classifier (apnea/normal)	95	93	96	-
[89]	Papini	SOMNIA (469 PSG patients) & HealthBed datasets (33 healthy)	30 s	PPG-derived: PRV + respiratory features	Deep learning model with stacked convolutional blocks	86	39	94	51

*No performance metrics were reported for this study. Only the time taken to detect an apneic event was specified per subject.

OSA: Obstructive Sleep Apnea, PSG: polysomnography, PRV: pulse rate variability, LF: low frequency, HF: high frequency, RI: reflection index, SI: stiffness index, : decreases in amplitude of PPG signal, NN: Neural Network, MLFFNN: Multilayer feedforward NN, LDA: Linear Discriminant Analysis, SVM: Support Vector Machine.

- The *UZ Leuven dataset* included the PSG of 100 patients with moderate and severe SAHS ($AHI \geq 15$). Desaturations were also annotated based on whether they were linked to an apnea or not.

Some works, used several of these public databases to test their models generalizability. Others instead, used their own recorded datasets. For instance, Jung *et al.* [79] collected 230 PSG recordings at the Seoul National University Hospital, of which 46 did not suffer from SAHS and the rest were characterized by different AHI indexes.

As it can be observed in Table 2.3, studies exploiting PPG signals typically used their own datasets, including healthy and SAHS subjects. The cohorts sizes are smaller than in SpO₂ databases, ranging from 5 to 40 subjects. Data collection varied from anesthetized patients [83] to controlled breath-holding experiments [84, 85], but also included PSG recordings from suspected obstructive sleep apnea (OSA) patients [86, 87]. Overall, the classification decision focused on differentiating apnea PPG events from non-apneic PPG fragments, without making further distinctions between central, obstructive, mixed or hypoapneas events. However, the SOMNIA dataset used in [89], included all of these, which enabled to assess the algorithm sensitivity for each particular respiratory event.

2.4.4 Epoch length

Full night PSG recordings were labeled by experts with various durations in different studies. The most common epoch length found in the SpO₂ literature was 1 min. Many studies that utilized publicly available databases, such as the Apnea-ECG database, were indeed limited by this annotation window of the manual scoring. Mostafa *et al.* also explored larger epochs of 1min, 3min and 5 min duration, with 1min window overlap [81, 82].

PPG studies instead, used more instantaneous windows of 5s [83], 15s [84], and variably between 10 to 30s [87, 88]. However, none of these works implemented it in an epoch-by-epoch running window manner, but instead manually segmented the most representative PPG fragments in the recordings, showing either normal breathing or apneic patterns for classification. This non-fully automated process is non-realistic and does not prove the method's validity for online processing.

Lázaro *et al.* continuously detected events, but reported the detection performance over 1 hour fragments [86]. The algorithm proposed by Papini *et al.* [89], was the only one performing epoch-by-epoch classification with a window of 30s, relying on the PPG signal.

2.4.5 Signals and features exploited

Pulse oximeter sensors enable the detection of apneas in a twofold manner. Either by using the raw PPG signal, or the SpO₂-derived time-series. In some rare cases, a combination of both is implemented.

2.4.5.1 SpO₂-based approaches

The most established strategy for apnea detection using pulse oximetry sensors is to directly exploit features from the SpO₂ signal. More specifically, works in the literature focus on the characteristics of SpO₂ desaturations. For that, drops in oxygen saturation need to be first detected. Deviaene *et al.* [77] proposed to detect SpO₂ desaturations over the whole recording by smoothing the SpO₂ time series with a moving average (MA) and then identifying characteristic points using the first derivative. Jung *et al.* [79] also used 3 specific points to determine the desaturation: the first one (A) was located when the oxygen dropped by 1% at the beginning of the desaturation; the second (B) was placed at the deepest point of the desaturation with a further decrease of 3%; the last point (C) was based 1% below A or 3% above B on the resaturation. The duration between A and C had to be $\geq 10s$ and $\leq 90s$ for the desaturations to be considered as an apnea.

Several approaches, quantifying patterns in pulse oximetry, have been presented in the literature [90]. Typical desaturations characteristics include the desaturation's duration and depth, the oxygen desaturation index quantifying the depth of the drop in oxygen levels by 2%, 3%, 4% (ODI2, ODI3, ODI4), and the desaturation area under these thresholds [77, 75]. Some of these characteristics are illustrated in Figure 2.6. Time series analysis characteristics such as the mean, variance, maximum, minimum, skewness, kurtosis, and deviations from these were also found in [77, 75]. The cumulative time spent in apnea below a certain SpO₂ threshold value such as $\leq 90\%$ (TSA90) was also an important trait considered in [73]. Other statistical features explored were the delta index (Δ) and the central tendency measure with radius 0.5 (CTM50) [74, 75].

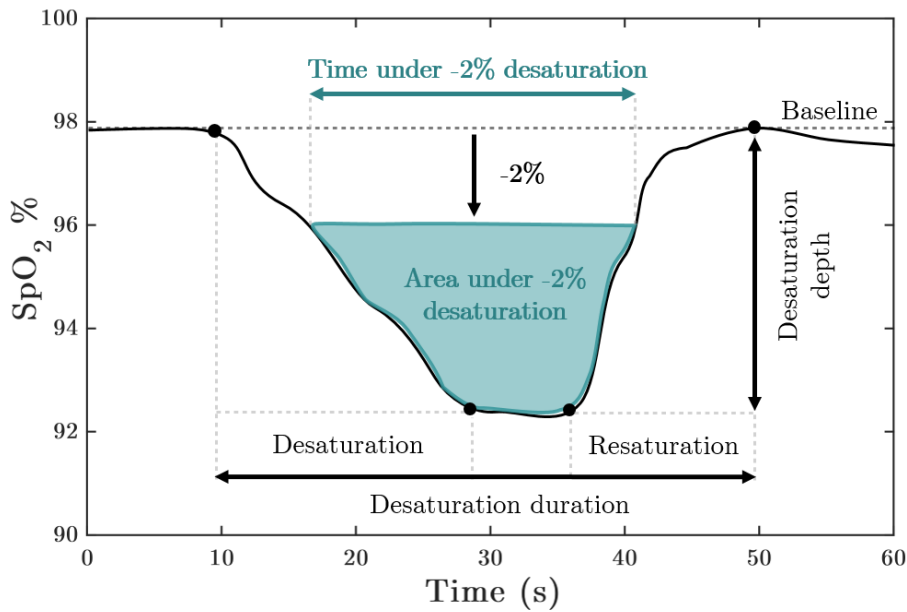


Figure 2.6: Annotated diagram of a SpO₂ desaturation with some common characteristics. Features such as desaturation depth, desaturation duration, area and time under the -2% threshold are depicted.

Deviaene *et al.* [77] extracted some features from the quasi-periodicity domain using the Phase Rectified Signal Averaging (PRSA) and auto-correlation methods. For instance, the upwards and downwards slopes of the PRSA, and the first auto-correlation peak together with its relative amplitude, were some aspects studied.

Some studies [80, 75], proposed some frequency domain features as the output of a 20 equally spaced filter bank, covering the whole frequency band of a 5 min epoch. The bandwidth of each filter was of 1.25 Hz.

Mostafa *et al.* [75] also explored some wavelet features from the time-frequency domain. By applying the db3 mother wavelet for wavelet decomposition, approximation and detail coefficients were extracted. The Shannon entropy and standard deviation of these for the 6 decomposition levels were also derived.

2.4.5.2 PPG signal features

Less attention has been paid in the literature to features directly extracted from the PPG signal. However, since not all the apneic events entail significant oxygen desaturations in adults [77], and this is even more frequent in children, researchers needed to take advantage of some unique apneic PPG traits in order to avoid missing critical respiratory events.

Time domain features were commonly computed from the PPG signal. For instance, PPG amplitude characteristics, rise time and beat-to-beat intervals were explored in [83, 87, 88], as well as the variability of some of these (pulse rate, amplitude and width variability) [78]. Some studies, proposed morphological pulse traits such as the pulse width, the ratio of the systolic (A1) and diastolic (A2) areas [87], the reflection index or the stiffness index [84]. Several statistical features including the mean, standard deviation, skewness, kurtosis, central moments, average curve length, Hjorth parameters; were also found in the literature [87, 88].

Frequency domain features mainly focused on the low (0.04-0.15 Hz), medium (0.09-0.15 Hz) and high (0.15-0.6 Hz) frequency bands of the PPG power spectrum [83, 88]. They included the frequency peaks and power values in these bands.

Some studies derived some additional physiological signals from the raw PPG, such as the pulse rate variability (PRV) or the respiratory signal. Additional features were computed from these surrogates for apnea detection. Lázaro *et al.* [86] focused on detecting decreases in amplitude of the PPG signal (DAP) that were previously shown to be correlated with apnea [64]. This vasoconstriction response triggered by the sympathetic system is hypothetically caused to generate an arousal after respiratory cessation. However, since not all DAPs in the PPG signal are always related to apneic events, PRV features were additionally extracted based on the inverse of the instantaneous

PPG pulses interval. PRV characteristics included the mean and the variance, as well as the power in the very low (0.003-0.04Hz), low (0.04-0.15 Hz), and high (0.15-0.5 Hz) frequency bands. In another study [85], the respiratory signal was extracted from the PPG by wavelet decomposition using the db9 mother wavelet. The area under the power spectrum of this surrogate signal, in the band between 0.05-0.5Hz was continuously calculated and compared to an adaptive average for apnea detection. Papini *et al.* included both PRV and respiratory activity derived features from the PPG signal. Some PRV characteristics included time domain statistics, the arousal probability, multi-scale entropy, Hilbert transform or frequency domain features among others. For the respiratory ones: time, frequency peaks, amplitude and visibility graphs analysis were likewise analyzed.

2.4.6 Classification methods

Various apnea classification techniques were explored in the literature, from simple threshold-based decisions lists to complex deep learning algorithms. The most distinctive algorithms are presented as follows.

Jung *et al.* [79] automatically detected apneic desaturations using the SpO₂ signal. To classify whether these resulted from apneic events or not, they passed the coordinates of three characteristic points of the desaturation into simple decision lists. The minute-by-minute classification of apneic events resulted in an average ACC of 91%, SE of 83% and SP of 89%. This research claimed that their algorithm was near real-time, however they first detected the slow response of the SpO₂ and then they located the original apneic event by subtracting 25s from the onset of the desaturation. This delay of 25s could potentially be acceptable in diagnostic applications. However, it is questionable to consider this algorithm suitable, for real-time SUDEP monitoring.

A simple Linear regression (LR) classifier analyzing a total of 22 statistical and frequency-based SpO₂ features was implemented by Mendonça *et al.* [80] in a wireless device for minute-by-minute apnea detection. A sequential forward feature selection process computed the most optimal feature set. This achieved an acceptable performance of ACC of 88%, SE of 80% and SP of 91%.

Another algorithm proposed by Burgos *et al.* [73] also claimed to perform real-time apnea detection on a PDA. Yet again, a minute-by-minute epoch-based detection, as annotated in the Apnea-ECG database, can be argued to not be real-time. In this study, the ODI2, 3, 4 and TSA95, 90, 85, 80, 70 indexes, were input into a Bagging of decision trees (ADTree). The performance obtained (ACC=93.03%, SE=92.35%, SP=93.52%), was very promising for an apnea monitoring system capable of transmitting physiological data online via Bluetooth into a mobile device for processing.

Lázaro *et al.* [86] exploited PRV derived features from DAPs hypothetically correlated with apnea events. A total of 34 statistical and frequency characteristics from four different windows preceding,

following and spanning the detected DAP events, were fed into a LDA classifier. The DAP event classification as related to an apneic event or not, showed a performance of ACC=70.37%, SE=81.82% and SP=68.57%.

Deviaene *et al.* [77] extracted time domain, statistical, desaturation severity and quasi-periodicity features from detected SpO₂ desaturations, and tested different ML classifiers. A random forest classifier showed the best average performance (ACC=82.8%, SE=64.3%, SP=88.6%, PPV=64.2%) in identifying whether the desaturations were caused by an apneic event or not. However the sensitivity of this algorithm, was very poor due to the fact that some apneic events were not always followed by oxygen desaturations. In order to avoid missing apneas, the same authors exploited the PPG signal to compliment the SpO₂ detection, in a follow up study [78]. The inclusion of additional time and frequency PPG features in a least squares SVM classifier (LS-SVM) resulted in a an improvement of the sensitivity (SE=73.7%). However the precision still remained very low (PPV=64.8%). Another study [84], feeding a small subset of time and frequency PPG features into a linear SVM classifier obtained very good results (ACC=97.22%, SE=94.62%, SP=100%). However, these might not be very realistic, since the experiments were carried out in healthy participants performing breath-holding exercises to simulate apneas.

Some studies exploited ANN for SpO₂-based apnea classification [74, 75]. In order to select an optimal subset of SpO₂ features from the time, frequency and wavelet domains; a Genetic Algorithm (GA) was additionally applied in [75]. Almazaydeh *et al.* [74], on the contrary only employed a total of three desaturation features (ODI3, Δ index and CTM50). Both studies used the Physionet Apnea-ECG dataset for validation, resulting in very high performance values (ACC=97.7% [75] and ACC=93.3% [74]).

Other NN approaches exploiting features from the PPG signal were also proposed. Knorr-Chung *et al.* [83], extracted some time and frequency domain features from normal and obstructive apneas PPG segments observed during general anesthesia with intubation pocedure. The trained ANN demonstrated very high sensitivity (SE=91.8%), yet debatable, since the most representative obstructive PPG segments were chosen per subject. Another study including only 5 participants, derived 34 morphological and statistical pulse characteristics, as well as Hjorth parameters [87]. These were input in a multilayer feedforward NN (MLFFNN). The results obtained were good (ACC=97.07%, SE=98%, SP=96%), but questionable, due to the inconsistent epoch length used in the range [10-30s], and the reduced number of participants (a total of 5, with only 3 sleep apnea patients). A similar issue was observed in another work by the same authors [88], in which a combination of time and frequency features from the PPG and PRV-derived signals were employed. An ensemble classifier consisting of a kNN, MLFFNN, probabilistic NN and SVM, attained doubtful successful performance. However

again, the cohort was only composed of 10 subjects with different epoch lengths.

In the past years, deep learning models have gained more attention in the field of apnea detection [72]. Their main objective is to replace the feature engineering stage in traditional ML techniques, by an automatic learning of features from raw data. For that, Mostafa *et al.* proposed to input the SpO₂ time series directly into a convolutional network (CNN) in 1D [81, 82]. In order to find the best CNN structure and select all the complex hyperparameters, they implemented a non-dominated Genetical Algorithm (GA) [81], and a weighted-topology transfer, greedy based optimization (GBO) [82]. In both studies, they tested three epoch lengths of 1, 3 and 5min duration, and various available datasets. However, despite the large computational times invested in these optimizations of 25-38 days [81] and 8-30 hours [82] respectively; the results did not show such a significant improvement compared to other simpler ML models [75, 73]. Unless the slight increase in performance is strictly crucial for specific applications, these complex and time-consuming deep learning methods are at the moment not worth the effort. Among the unsupervised feature deep learning models, Pathinarupothi *et al.* [76] also proposed a long short-term memory recurrent neural network (LSTM-RNN) with 60 neurons in the input layer (one per each second of the SpO₂ epoch) and 32 memory blocks in the hidden layer. The good performance achieved in the Apnea-ECG database is comparable to other ML and CNN works (ACC=95.5%, SE=92.9%, PPV=99.2%).

Papini *et al.* [89] developed a deep learning model with stacked convolutional blocks that performed 30s epoch-based classification, of apneic respiratory events from different origins (obstructive, central, mixed and hypopnea). Instead of adopting a self-learned features strategy as other deep learning works, they employed HRV and respiratory PPG-derived features. The performance resulted in very poor sensitivity and precision (ACC=86%, SE=39%, PPV=51%, PR AUC=0.47). This can be explained by the differences in sensitivities among different types of apnea, being hypopnea the lowest SE=37%. The high number of false positives was probably due to the high percentage of epochs corrupted with movement artifacts.

2.4.7 Discussion

The current-state-of-the-art of automatic apnea detection algorithms making use of pulse oximeter sensors has been reviewed, in terms of datasets, classification windows, extracted features and classifiers. Overall, even though there are common efforts to maximize detection performance by extensively engineering new features from the PPG and SpO₂ signals, and devising complex classifiers; there are still some inherent limitations that prevent from directly applying these solutions for apnea detection in SUDEP prevention.

Datasets and annotations Validation on several publicly available databases is essential to ensure that classification results are reproducible. Despite some attempts, it is still challenging to assess and compare the performance of the proposed classification models. This is due to the huge variability in the number and type of participants undergoing the PSG studies in different datasets. The same algorithms tested in several databases showed indeed very distinct performance. Subjects with higher AHI are likely to suffer more frequent and severe apneic desaturations than subjects with mild SAHS. This can greatly bias the classifier when all the subjects are included equally, instead of balancing the training set with the same number of apneic and non-apneic desaturations per subject.

In addition, the Physionet Apnea-ECG consistently stands out with the greatest performance no matter the algorithm. This can be explained by the 1-min-epoch annotations, that facilitates correct classification. In fact, even if a 1-minute segment contains more than one apnea, and the algorithm only identifies one, the minute would still be counted as a true positive. This demonstrates that larger epoch lengths might not be very reliable, specially for non-diagnostic online applications.

Classification Window Epoch length selection is critical to ensure good performance and reliability of the classification method. It has been observed that among different studies the epoch length varies significantly. It is therefore very difficult to agree on a suitable duration equally valid for all clinical applications. Diagnostic automatic algorithms have the flexibility of exploring larger windows, whereas real-time applications are limited by more instantaneous ones. So a trade-off in epoch length should be found between real-time implementation and accurate detection.

Inconvenients of single SpO₂ approaches SpO₂-based approaches have been widely studied in the literature and have shown good performance in classifying apneic versus normal epochs/events. However, the start of the SpO₂ desaturation is delayed by 20-40s with respect to the start of the apneic event [67], seriously limiting the implementation in real-time. Even though algorithms detecting SpO₂ desaturations or DAPs claim that they can seek the apneic event in the previous 25s [79], that might be too late to set an alarm in life-threatening SUDEP situations.

In addition, some apneic events are not always followed by oxygen desaturations. In these cases single SpO₂-based approaches will not be able to successfully detect these respiratory events [77]. This shows the importance of using the available PPG signal simultaneously, in order to maximize the pool of biomarkers available to make a decision on critical apneas.

Combination with respiratory signals Even if current state-of-the-art PPG sensors have a great potential to classify apneic and normal events, they still lack the capability to properly distinguish hypoapneas. The combination of oximetry with additional sensors could improve the performance.

Specially, respiratory signals obtained through nasal airflow, tracheal sounds, or thoracic and abdominal efforts, could successfully complement oximetry approaches in providing meaningful information to support the classification decision [91].

Classifier type Supervised machine learning classifiers have been proven to be very successful in detecting apneas. Yet, some disadvantages are the tedious feature engineering stage, that relies on a good understanding of the physiology behind, which in turn requires a certain level of expertise in the field. However, in the case of apnea, the well established PPG pattern makes it easy to model it. It could be argued though, that features extraction could result in infinite combinations in order to maximize performance. But still, one of the greatest advantages of ML methods is the interpretability of the model, such as depicting the most discriminative features. This could be of great importance in the understanding of SUDEP mechanisms.

The main gap in the literature of apnea detection is deep learning algorithms. These have become more popular in the recent years, but still need to be further explored and developed, as there is big room for improvement. Specially in the direction of optimizing the degree of complexity while maintaining the outstanding performance. Indeed, the high number of hyperparameters entails very long training times, which restrict the cross-validation of the models and limit them to be retrained if additional data becomes available. Yet, these deep learning approaches are very promising, we always have to keep in mind the end-application. Efficient hardware implementations should be of primary focus in situations where computational power is a constraint.

Conclusion The major challenge in apnea detection hence is: to develop a robust system that incorporates an efficient automatic algorithm capable of detecting apneic events accurately in real-time. Shorter epochs are necessary to reduce the delays in detection, and ensure a more instantaneous and reliable online processing. The performance of the proposed models ($\sim 90\%$) is promising, but still not sufficient to avoid an excessive number of false alarms. Current algorithms only have the ability to classify apneas, but cannot distinguish between other types of respiratory events. That is why, measurement of airflow with additional sensing modalities could be helpful to differentiate, for example, central and obstructive apneas. As well, the combination of SpO_2 and PPG features could enhance the overall performance.

References

- [1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.
- [2] J. E. Sinex, "Pulse oximetry: principles and limitations," *The American journal of emergency medicine*, vol. 17, no. 1, pp. 59–66, 1999.
- [3] E. D. Chan, M. M. Chan, and M. M. Chan, "Pulse oximetry: understanding its basic principles facilitates appreciation of its limitations," *Respiratory medicine*, vol. 107, no. 6, pp. 789–799, 2013.
- [4] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 463–477, 2016.
- [5] A. Jubran, "Pulse oximetry," *Critical Care*, vol. 19, no. 1, p. 272, 2015.
- [6] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current cardiology reviews*, vol. 8, no. 1, pp. 14–25, 2012.
- [7] P. D. Mannheim, "The light–tissue interaction of pulse oximetry," *Anesthesia & Analgesia*, vol. 105, no. 6, pp. S10–S17, 2007.
- [8] M. Nitzan, A. Romem, and R. Koppel, "Pulse oximetry: fundamentals and technology update," *Medical Devices (Auckland, NZ)*, vol. 7, p. 231, 2014.
- [9] S. Rhee, B.-H. Yang, and H. H. Asada, "Artifact-resistant power-efficient design of finger-ring plethysmographic sensors," *IEEE transactions on biomedical engineering*, vol. 48, no. 7, pp. 795–805, 2001.
- [10] M. Maguire, T. Ward, C. Markham, D. O'Shea, and L. Kevin, "A comparative study in the use of brachial photoplethysmography and the qrs complex as timing references in determination of pulse transit time," in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 1. IEEE, 2001, pp. 215–218.
- [11] A.-M. Tăuțan, A. Young, E. Wentink, and F. Wieringa, "Characterization and reduction of motion artifacts in photoplethysmographic signals from a wrist-worn device," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE, 2015, pp. 6146–6149.
- [12] S. Sugino, N. Kanaya, M. Mizuuchi, M. Nakayama, and A. Namiki, "Forehead is as sensitive as finger pulse oximetry during general anesthesia," *Canadian Journal of Anesthesia*, vol. 51, no. 5, pp. 432–436, 2004.
- [13] H. J. Davies, I. Williams, N. S. Peters, and D. P. Mandic, "In-ear spo2: A tool for wearable, unobtrusive monitoring of hypoxaemia in covid-19," *medRxiv*, 2020.
- [14] J. P. Phillips, P. A. Kyriacou, D. P. Jones, K. H. Shelley, and R. M. Langford, "Pulse oximetry and photoplethysmographic waveform analysis of the esophagus and bowel," *Current opinion in anesthesiology*, vol. 21, no. 6, pp. 779–783, 2008.
- [15] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 786–794, 2010.
- [16] Y. Lee, H. Shin, J. Jo, and Y.-K. Lee, "Development of a wristwatch-type ppg array sensor module," in *Consumer Electronics-Berlin (ICCE-Berlin), 2011 IEEE International Conference on*. IEEE, 2011, pp. 168–171.

- [17] M.-Z. Poh, K. Kim, A. Goessling, N. Swenson, and R. Picard, "Cardiovascular monitoring using earphones and a mobile device," *IEEE Pervasive Computing*, vol. 11, no. 4, pp. 18–26, 2012.
- [18] W. Johnston and Y. Mendelson, "Extracting breathing rate information from a wearable reflectance pulse oximeter sensor," in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, vol. 2. IEEE, 2004, pp. 5388–5391.
- [19] Y. Zhong, Y. Pan, L. Zhang, and K.-T. Cheng, "A wearable signal acquisition system for physiological signs including throat ppg," in *Engineering in Medicine and Biology Society (EMBC), 2016 38th Annual International Conference of the IEEE*. IEEE, 2016, pp. 603–606.
- [20] M. Peng, S. A. Imtiaz, and E. Rodriguez-Villegas, "Pulse oximetry in the neck - a proof of concept," in *Engineering in Medicine and Biology Society (EMBC), 39th Annual International Conference of the IEEE*. IEEE, 2017.
- [21] S. C. Millasseau, J. M. Ritter, K. Takazawa, and P. J. Chowienczyk, "Contour analysis of the photoplethysmographic pulse measured at the finger," *Journal of hypertension*, vol. 24, no. 8, pp. 1449–1456, 2006.
- [22] J. Allen and A. Murray, "Age-related changes in the characteristics of the photoplethysmographic pulse shape at various body sites," *Physiological measurement*, vol. 24, no. 2, p. 297, 2003.
- [23] T. R. Dawber, H. E. THomas Jr, and P. M. McNamara, "Characteristics of the dicrotic notch of the arterial pulse wave in coronary heart disease," *Angiology*, vol. 24, no. 4, pp. 244–255, 1973.
- [24] J. Allen and A. Murray, "Age-related changes in peripheral pulse timing characteristics at the ears, fingers and toes," *Journal of human hypertension*, vol. 16, no. 10, pp. 711–717, 2002.
- [25] J. Allen, J. R. Frame, and A. Murray, "Microvascular blood flow and skin temperature changes in the fingers following a deep inspiratory gasp," *Physiological measurement*, vol. 23, no. 2, p. 365, 2002.
- [26] H. Liu, J. Allen, D. Zheng, and F. Chen, "Recent development of respiratory rate measurement technologies," *Physiological measurement*, vol. 40, no. 7, p. 07TR01, 2019.
- [27] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton, R. Beale, and P. J. Watkinson, "An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram," *Physiological measurement*, vol. 37, no. 4, p. 610, 2016.
- [28] J. Allen and A. Murray, "Variability of photoplethysmography peripheral pulse measurements at the ears, thumbs and toes," *IEE Proceedings-Science, Measurement and Technology*, vol. 147, no. 6, pp. 403–407, 2000.
- [29] V. Hartmann, H. Liu, F. Chen, Q. Qiu, S. Hughes, and D. Zheng, "Quantitative comparison of photoplethysmographic waveform characteristics: effect of measurement site," *Frontiers in physiology*, vol. 10, p. 198, 2019.
- [30] M. T. Petterson, V. L. Begnoche, and J. M. Graybeal, "The effect of motion on pulse oximetry and its clinical significance," *Anesthesia & Analgesia*, vol. 105, no. 6, pp. S78–S84, 2007.
- [31] T. Tamura, Y. Maeda, M. Sekine, and M. Yoshida, "Wearable photoplethysmographic sensors—past and present," *Electronics*, vol. 3, no. 2, pp. 282–302, 2014.
- [32] A. K. Bhoi, S. Sarkar, P. Mishra, and G. Savita, "Pre-processing of ppg signal with performance based methods," *International Journal of Computer Application*, vol. 4, no. 2, pp. 251–256, 2012.

-
- [33] H.-W. Lee, J.-W. Lee, W.-G. Jung, and G.-K. Lee, "The periodic moving average filter for removing motion artifacts from ppg signals," *International journal of control automation and systems*, vol. 5, no. 6, p. 701, 2007.
 - [34] K. A. Reddy, B. George, and V. J. Kumar, "Use of fourier series analysis for motion artifact reduction and data compression of photoplethysmographic signals," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 5, pp. 1706–1711, 2009.
 - [35] C. Lee and Y. Zhang, "Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach," in *Biomedical Engineering, 2003. IEEE EMBS Asian-Pacific Conference on*. IEEE, 2003, pp. 194–195.
 - [36] G. Joseph, A. Joseph, G. Titus, R. M. Thomas, and D. Jose, "Photoplethysmogram (ppg) signal analysis and wavelet de-noising," in *Emerging Research Areas: Magnetism, Machines and Drives (AICERA/iCMMD), 2014 Annual International Conference on*. IEEE, 2014, pp. 1–5.
 - [37] J. Y. A. Foo, "Comparison of wavelet transformation and adaptive filtering in restoring artefact-induced time-related measurement," *Biomedical signal processing and control*, vol. 1, no. 1, pp. 93–98, 2006.
 - [38] H. H. Asada, H.-H. Jiang, and P. Gibbs, "Active noise cancellation using mems accelerometers for motion-tolerant wearable bio-sensors," in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, vol. 1. IEEE, 2004, pp. 2157–2160.
 - [39] H. Han, M.-J. Kim, and J. Kim, "Development of real-time motion artifact reduction algorithm for a wearable photoplethysmography," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 1538–1541.
 - [40] R. Yousefi, M. Nourani, and I. Panahi, "Adaptive cancellation of motion artifact in wearable biosensors," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 2004–2008.
 - [41] B. S. Kim and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE transactions on biomedical engineering*, vol. 53, no. 3, pp. 566–568, 2006.
 - [42] F. Peng, Z. Zhang, X. Gou, H. Liu, and W. Wang, "Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter," *Biomedical engineering online*, vol. 13, no. 1, p. 50, 2014.
 - [43] J. W. Chong, D. K. Dao, S. Salehizadeh, D. D. McManus, C. E. Darling, K. H. Chon, and Y. Mendelson, "Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection–reduction approach. part i: motion and noise artifact detection," *Annals of biomedical engineering*, vol. 42, no. 11, pp. 2238–2250, 2014.
 - [44] S. Salehizadeh, D. K. Dao, J. W. Chong, D. McManus, C. Darling, Y. Mendelson, and K. H. Chon, "Photoplethysmograph signal reconstruction based on a novel motion artifact detection-reduction approach. part ii: Motion and noise artifact removal," *Annals of biomedical engineering*, vol. 42, no. 11, pp. 2251–2263, 2014.
 - [45] R. Krishnan, B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE transactions on biomedical engineering*, vol. 57, no. 8, pp. 1867–1876, 2010.

- [46] B. Tarvirdizadeh, A. Golgouneh, F. Tajdari, and E. Khodabakhshi, "A novel online method for identifying motion artifact and photoplethysmography signal reconstruction using artificial neural networks and adaptive neuro-fuzzy inference system," *Neural Computing and Applications*, vol. 32, no. 8, pp. 3549–3566, 2020.
- [47] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in photoplethysmogram," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 4972–4975.
- [48] J. A. Sukor, S. Redmond, and N. Lovell, "Signal quality measures for pulse oximetry through waveform morphology analysis," *Physiological measurement*, vol. 32, no. 3, p. 369, 2011.
- [49] Q. Li and G. Clifford, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiological measurement*, vol. 33, no. 9, p. 1491, 2012.
- [50] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L.-W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, "Multiparameter intelligent monitoring in intensive care ii (mimic-ii): a public-access intensive care unit database," *Critical care medicine*, vol. 39, no. 5, p. 952, 2011.
- [51] R. Couceiro, P. Carvalho, R. P. Paiva, J. Henriques, and J. Muehlsteff, "Detection of motion artifact patterns in photoplethysmographic signals based on time and period domain analysis," *Physiological measurement*, vol. 35, no. 12, p. 2369, 2014.
- [52] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, "Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring," *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 832–838, 2014.
- [53] S. Cherif, D. Pastor, Q.-T. Nguyen, and E. L'Her, "Detection of artifacts on photoplethysmography signals using random distortion testing," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 6214–6217.
- [54] M. Elgendi, "Optimal signal quality index for photoplethysmogram signals," *Bioengineering*, vol. 3, no. 4, p. 21, 2016.
- [55] D. Dao, S. M. Salehizadeh, Y. Noh, J. W. Chong, C. H. Cho, D. McManus, C. E. Darling, Y. Mendelson, and K. H. Chon, "A robust motion artifact detection algorithm for accurate detection of heart rates from photoplethysmographic signals using time–frequency spectral features," *IEEE journal of biomedical and health informatics*, vol. 21, no. 5, pp. 1242–1253, 2016.
- [56] 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*, vol. 21, no. 2, pp. 372–381, 2017.
- [57] N. Pradhan, S. Rajan, A. Adler, and C. Redpath, "Classification of the quality of wristband-based photoplethysmography signals," in *Medical Measurements and Applications (MeMeA), 2017 IEEE International Symposium on*. IEEE, 2017, pp. 269–274.
- [58] M.-Z. Poh, Y. C. Poh, P.-H. Chan, C.-K. Wong, L. Pun, W. W.-C. Leung, Y.-F. Wong, M. M.-Y. Wong, D. W.-S. Chu, and C.-W. Siu, "Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms," *Heart*, vol. 104, no. 23, pp. 1921–1928, 2018.
- [59] F. Tabei, R. Kumar, T. N. Phan, D. D. McManus, and J. W. Chong, "A novel personalized motion and noise artifact (mna) detection method for smartphone photoplethysmograph (ppg) signals," *IEEE Access*, vol. 6, pp. 60 498–60 512, 2018.

- [60] K. Vandecasteele, J. Lázaro, E. Cleeren, K. Claes, W. Van Paesschen, S. Van Huffel, and B. Hunyadi, "Artifact detection of wrist photoplethysmograph signals." in *BIOSIGNALS*, 2018, pp. 182–189.
- [61] T. Pereira, K. Gadhomi, M. Ma, X. Liu, R. Xiao, R. A. Colorado, K. J. Keenan, K. Meisel, and X. Hu, "A supervised approach to robust photoplethysmography quality assessment," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 3, pp. 649–657, 2019.
- [62] C.-H. Goh, L. K. Tan, N. H. Lovell, S.-C. Ng, M. P. Tan, and E. Lim, "Robust ppg motion artifact detection using a 1-d convolution neural network," *Computer Methods and Programs in Biomedicine*, p. 105596, 2020.
- [63] R. Krishnan, B. Natarajan, and S. Warren, "Analysis and detection of motion artifact in photoplethysmographic data using higher order statistics," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2008, pp. 613–616.
- [64] E. Gil, J. M. Vergara, and P. Laguna, "Detection of decreases in the amplitude fluctuation of pulse photoplethysmography signal as indication of obstructive sleep apnea syndrome in children," *Biomedical Signal Processing and Control*, vol. 3, no. 3, pp. 267–277, 2008.
- [65] J. Torres-Soto and E. A. Ashley, "Multi-task deep learning for cardiac rhythm detection in wearable devices," *NPJ digital medicine*, vol. 3, no. 1, pp. 1–8, 2020.
- [66] V. Monasterio, F. Burgess, and G. D. Clifford, "Robust classification of neonatal apnoea-related desaturations," *Physiological measurement*, vol. 33, no. 9, p. 1503, 2012.
- [67] V. Moret-Bonillo, D. Alvarez-Estévez, A. Fernández-Leal, and E. Hernández-Pereira, "Intelligent approach for analysis of respiratory signals and oxygen saturation in the sleep apnea/hypopnea syndrome," *The open medical informatics journal*, vol. 8, p. 1, 2014.
- [68] F. Rugg-Gunn, "The role of devices in managing risk," *Epilepsy & Behavior*, vol. 103, p. 106456, 2020.
- [69] D. Cogan, J. Birjandtalab, M. Nourani, J. Harvey, and V. Nagaraddi, "Multi-biosignal analysis for epileptic seizure monitoring," *International journal of neural systems*, vol. 27, no. 01, p. 1650031, 2017.
- [70] F. Mendonca, S. S. Mostafa, A. G. Ravelo-García, F. Morgado-Dias, and T. Penzel, "A review of obstructive sleep apnea detection approaches," *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 825–837, 2018.
- [71] N. Pombo, N. Garcia, and K. Bousson, "Classification techniques on computerized systems to predict and/or to detect apnea: A systematic review," *Computer methods and programs in biomedicine*, vol. 140, pp. 265–274, 2017.
- [72] S. S. Mostafa, F. Mendonça, A. G. Ravelo-García, and F. Morgado-Dias, "A systematic review of detecting sleep apnea using deep learning," *Sensors*, vol. 19, no. 22, p. 4934, 2019.
- [73] A. Burgos, A. Goñi, A. Illarramendi, and J. Bermúdez, "Real-time detection of apneas on a pda," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, pp. 995–1002, 2009.
- [74] L. Almazaydeh, M. Faezipour, and K. Elleithy, "A neural network system for detection of obstructive sleep apnea through spo2 signal features," *International Journal of Advanced Computer Science and Applications*, vol. 3, 2012.
- [75] S. S. Mostafa, J. P. Carvalho, F. Morgado-Dias, and A. Ravelo-García, "Optimization of sleep apnea detection using spo2 and ann," in *2017 XXVI international conference on information, communication and automation technologies (ICAT)*. IEEE, 2017, pp. 1–6.

- [76] R. K. Pathinarupothi, E. S. Rangan, E. Gopalakrishnan, R. Vinaykumar, K. Soman *et al.*, “Single sensor techniques for sleep apnea diagnosis using deep learning,” in *2017 IEEE international conference on healthcare informatics (ICHI)*. IEEE, 2017, pp. 524–529.
- [77] M. Deviaene, D. Testelmans, B. Buyse, P. Borzée, S. Van Huffel, and C. Varon, “Automatic screening of sleep apnea patients based on the SpO₂ signal,” *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 607–617, 2018.
- [78] M. Deviaene, J. Lázaro, D. Huysmans, D. Testelmans, B. Buyse, S. Van Huffel, and C. Varon, “Sleep apnea detection using pulse photoplethysmography,” in *2018 Computing in Cardiology Conference (CinC)*, vol. 45. IEEE, 2018, pp. 1–4.
- [79] D. W. Jung, S. H. Hwang, J. G. Cho, B. H. Choi, H. J. Baek, Y. J. Lee, D.-U. Jeong, K. S. Park *et al.*, “Real-time automatic apneic event detection using nocturnal pulse oximetry,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 706–712, 2017.
- [80] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, “An oximetry based wireless device for sleep apnea detection,” *Sensors*, vol. 20, no. 3, p. 888, 2020.
- [81] S. S. Mostafa, F. Mendonça, A. G. Ravelo-García, G. G. Juliá-Serdá, and F. Morgado-Dias, “Multi-objective hyperparameter optimization of convolutional neural network for obstructive sleep apnea detection,” *IEEE Access*, vol. 8, pp. 129 586–129 599, 2020.
- [82] S. S. Mostafa, D. Baptista, A. G. Ravelo-García, G. Juliá-Serdá, and F. Morgado-Dias, “Greedy based convolutional neural network optimization for detecting apnea,” *Computer Methods and Programs in Biomedicine*, vol. 197, p. 105640, 2020.
- [83] B. R. Knorr-Chung, S. P. McGrath, and G. T. Blike, “Identifying airway obstructions using photoplethysmography (PPG),” *Journal of clinical monitoring and computing*, vol. 22, no. 2, pp. 95–101, 2008.
- [84] P. Pradhapan, M. Swaminathan, H. K. S. V. Mohan, and N. Sriraam, “Identification of apnea during respiratory monitoring using support vector machine classifier: a pilot study,” *Journal of clinical monitoring and computing*, vol. 27, no. 2, pp. 179–185, 2013.
- [85] G. Gaurav, S. Mohanasankar, and V. J. Kumar, “Apnea sensing using photoplethysmography,” in *2013 Seventh International Conference on Sensing Technology (ICST)*. IEEE, 2013, pp. 285–288.
- [86] J. Lázaro, E. Gil, J. M. Vergara, and P. Laguna, “Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children,” *IEEE journal of biomedical and health informatics*, vol. 18, no. 1, pp. 240–246, 2013.
- [87] M. K. Uçar, M. R. Bozkurt, C. Bilgin, and K. Polat, “Automatic detection of respiratory arrests in osa patients using ppg and machine learning techniques,” *Neural Computing and Applications*, vol. 28, no. 10, pp. 2931–2945, 2017.
- [88] M. R. Bozkurt, M. K. Uçar, F. Bozkurt, and C. Bilgin, “In obstructive sleep apnea patients, automatic determination of respiratory arrests by photoplethysmography signal and heart rate variability,” *Australasian Physical & Engineering Sciences in Medicine*, vol. 42, no. 4, pp. 959–979, 2019.
- [89] G. B. Papini, P. Fonseca, M. M. van Gilst, J. W. Bergmans, R. Vullings, and S. Overeem, “Wearable monitoring of sleep-disordered breathing: estimation of the apnea–hypopnea index using wrist-worn reflective photoplethysmography,” *Scientific reports*, vol. 10, no. 1, pp. 1–15, 2020.

- [90] P. I. Terrill, “A review of approaches for analysing obstructive sleep apnoea-related patterns in pulse oximetry data,” *Respirology*, vol. 25, no. 5, pp. 475–485, 2020.
- [91] M. Uddin, C. Chow, and S. Su, “Classification methods to detect sleep apnea in adults based on respiratory and oximetry signals: a systematic review,” *Physiological measurement*, vol. 39, no. 3, p. 03TR01, 2018.

Chapter 3

Novel Neck PPG vs. Standard Finger PPG

3.1 Introduction

In clinical settings, the finger is considered the gold standard measurement site for pulse oximetry due to the easy sensor attachment and its rich capillarity. It is important, however, to investigate other sites for recording the PPG on the body as it can free the hands from measurement devices. With the recent explosion of wearable tracker systems, wrist PPG sensors have become the most popular alternative for consumer grade heart rate (HR) monitoring during daily and physical activities [1]. Despite their promising application in ambulatory environments ([2, 3, 4, 5]), in stationary conditions such as the ones this research focuses on, frequent arm motion could still corrupt the PPG signal more regularly than more steady PPG sensors placed in the head.

In this context, head measurement sites are preferred. The earlobe and forehead are already, some prevalent locations for sensors positioning when the patient's hands are unavailable (e.g. wounds, burns, amputation, surgery) [6]. However, the neck could provide great benefits over these conventional PPG sites in the context of some diseases for which additional physiological biomarkers are desired to be recorded simultaneously with the same wearable system. Unlike other body parts, the neck region offers the possibility to measure tracheal sounds that can be exploited for respiratory diseases diagnosis [7, 8, 9, 10]. Specifically, for apnea detection, the neck is a unique location for cardio-respiratory multi-modal signal acquisition [7].

Even though the neck could be a favourable new alternative for clinical PPG, this novel site has not yet been thoroughly studied. The only works attempting to use it, focused on showing the feasibility of extracting the PPG signal from this site [11], and on determining the most stable position for signal acquisition [12]. None of these studies examined the actual characteristics of the signal. This is

necessary, however, not only to correctly interpret and potentially extract new disease indicators, but also to develop targeted signal processing techniques to eliminate artifacts that could possibly disrupt the PPG signal.

This chapter aims to compare finger and neck PPG signals in order to characterize the novel measurement site. The comparison starts from the most basic PPG unit, the pulse waveform, then moves into the spectral analysis of the novel signals, and builds up into the extraction of higher-level physiological signals that can be extracted during long-term PPG monitoring. These include the respiration, the cardiac pulsatile function and the blood oxygen saturation (SpO₂). As a summary, a discussion on the potentials and limitations of the novel neck PPG site is presented, for its application in pulse oximetry.

3.2 Experiment 1 procedures

Data recorded in this experiment was used for analysis in Sections 3.3, 3.4 and 3.5.

3.2.1 Data acquisition set up

PPG, ECG and respiratory signals were acquired from a total of 9 healthy participants in a relaxed supine position on a bed to simulate sleep conditions, such that both heart and respiration rate slowed down. The cohort was composed of 5 males and 4 females, with average age of 24 ± 3 years old, and BMI of $22.6 \pm 3.7 \text{ kg/m}^2$. The study was approved by the Local Ethics Committee of Imperial College London (ICREC ref.: 18IC4358), and written informed consent was obtained from all subjects.

Three different measurement systems were used for data acquisition. Neck PPG signals were acquired by a reflectance pulse oximeter sensor (8000R, Nonin) attached externally at the suprasternal space with an adhesive, and connected to a processing module (Xpod, Nonin). Standard finger PPG signals were obtained from a synchronized transmission pulse oximeter (Onyx II 9560, Nonin) with Bluetooth connectivity, placed on the left hand. These two pulse oximeter sensors can be observed in Figure 3.1. Signals were sampled at a rate of 75Hz, and visualized in the computer in real-time.

Reference cardiac and respiratory signals were also obtained with a portable polysomnography (PSG) system (SOMNOscreen Plus, SOMNOmedics), shown in Figure 3.2. This consisted of: 2-leads ECG, flow thermistor sensor, nasal pressure cannula, chest and abdomen impedance plethysmography (IP) bands, and a finger transmission PPG sensor placed on the right hand. The PSG system also allowed the possibility of adding markers, which were used for posterior synchronization with the Nonin PPG sensors. In order to verify that all signals were correctly aligned in time for further processing, the correlation between finger PPG signals, from the Nonin and SOMNO systems, was obtained. If the initial synchronization did not maximize this correlation, a re-alignment was performed accordingly.

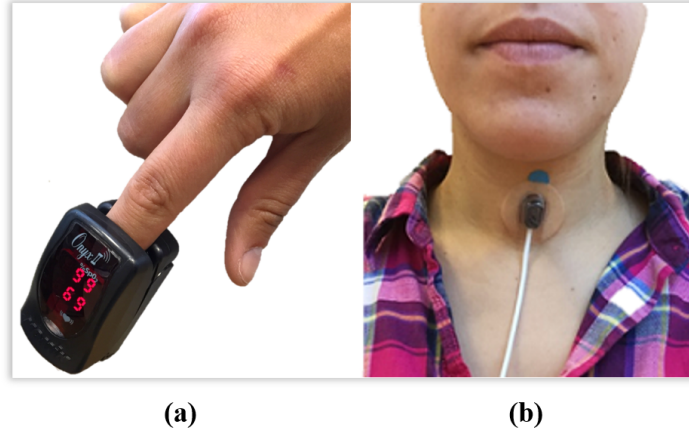


Figure 3.1: *Pulse oximetry sensors used for data acquisition. (a) Finger transmission (Onyx II 9560, Nonin). (b) Neck reflectance (8000R, Nonin).*

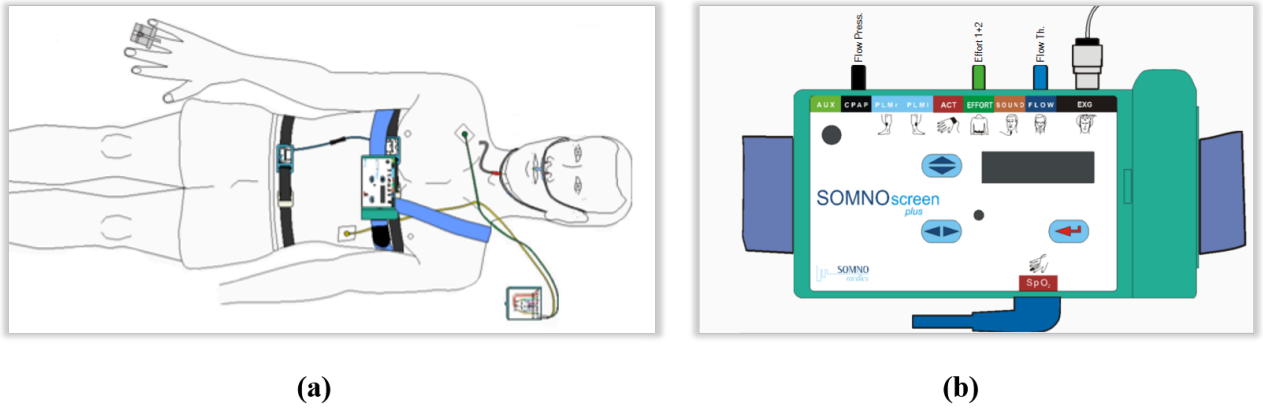


Figure 3.2: *Polysomnography SOMNOscreen Plus monitoring system. (a) 2-leads ECG, flow thermistor sensor, nasal pressure cannula, chest and abdomen impedance plethysmography (IP) bands, and finger transmission PPG sensors. (b) SOMNOscreen data acquisition unit.*

3.2.2 Experimental protocol

Four data recordings of 140s were acquired for each participant at rest, while breathing at four different respiratory frequencies. Figure 3.3 shows the different recordings for the neck and finger PPG, ECG, and respiratory channels. In the first recording (a), the participants were asked to relax and breath at their normal spontaneous pace. The following three epochs were similarly recorded at rest on a bed, but different respiration frequency modulations of around 20s-30s duration were introduced. These included fast breathing (b), slow breathing (c) and breath-holding simulated apnea events (d). For that, at specific moments during each recording session, I instructed the participants to either increment their breathing pace (b), reduce it with controlled breaths (c), or stop breathing (d). These changes in respiration frequency were marked online, and subsequently verified with the respiratory bands signals as depicted by the yellow horizontal arrows in Fig 3.3. It is worth noting how these alterations immediately affect the neck PPG signals. All signals were normalized in the range of $[-1,1]$.

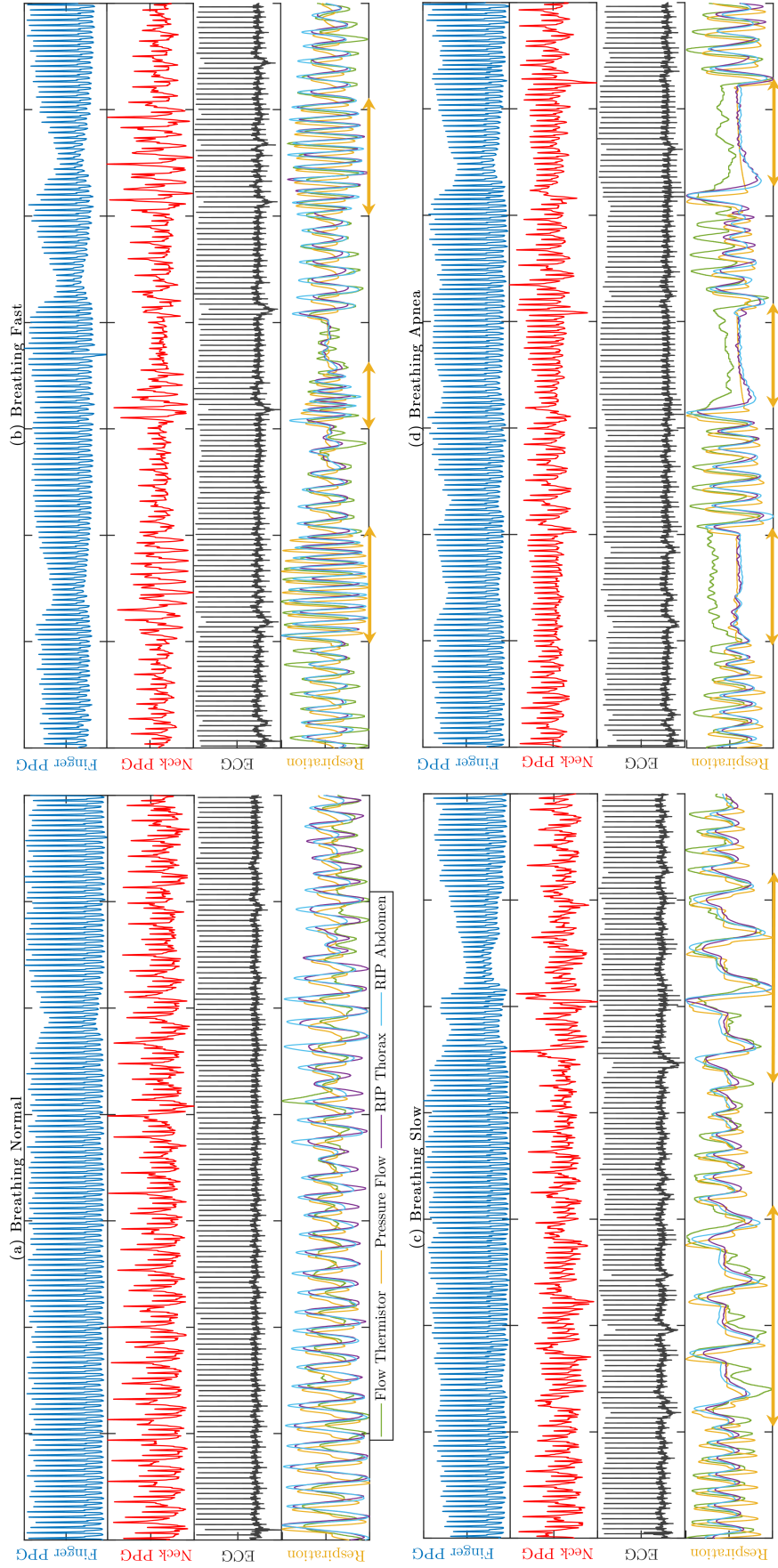


Figure 3.3: (a) Normal, (b) Breathing Fast, (c) Breathing Slow, (d) Breathing Apnea, recordings for one subject showing the finger PPG, neck PPG, ECG and respiratory channels. Yellow horizontal arrows indicate the time interval of respiratory frequency modulation.

3.3 Comparison of morphological pulse characteristics

The contents of the following sections are an edited version of the research published in:

I. García-López, S. A. Imtiaz, and E. Rodríguez-Villegas, “Characterization study of neck photoplethysmography,” in 40th Annual International Conference of the IEEE (EMBC). *IEEE*, 2018, pp. 4355–4358 © IEEE

3.3.1 Average waveforms characteristics extraction and analysis

For this analysis, only the first Breathing Normal recording (Figure 3.3 (a)) was used. The neck and finger PPG channels were automatically segmented to extract each individual PPG pulse and obtain an average pulse shape. The neck PPG waveform was statistically compared to the gold standard finger PPG one by means of time domain morphological features, as detailed in the next subsections.

Automated pulse segmentation and waveform averaging: The PPG signal was first filtered to remove the DC and low frequency components, with a fourth order high pass Butterworth filter with 0.7Hz cut off frequency. In order to detect the onsets of each pulse, the lower peak envelope of the PPG signal was determined using spline interpolation over local maxima separated by a time window of 0.667s. The maximum between successive troughs was selected as the systolic peak. Wrongly segmented pulses with a duration greater than the mean pulse duration by three standard deviations were discarded. Subsequently, a normalization in time and amplitude was performed for each of the remaining pulses. Pulses were aligned by calculating a delay shift for each pulse that maximized the overall correlation among all pulses. The first approximation of the mean pulse waveform was then obtained by averaging point by point. The 60 beats with the shortest Euclidean distance to the mean waveform were then selected. Taking into consideration this final set of good quality pulses, the final finger and neck average PPG pulse waveforms were extracted for each subject, as shown in Fig. 3.4.

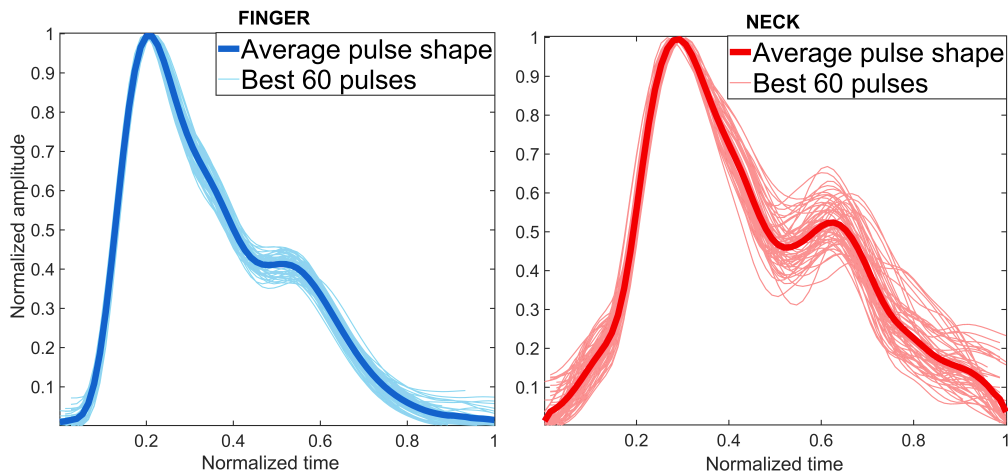


Figure 3.4: Average finger and neck PPG pulse waveforms obtained for one participant.

To an untrained eye, these two average pulse shapes might seem similar, however, the diastolic peak appears more prominent for the neck waveform, as well as the systolic peak is wider. However other more subtle features not observable to the naked eye might also differ between both sites. This is why the extraction of more complex features for further statistical analysis is necessary.

Feature extraction: As shown in Fig. 3.5, in order to obtain the morphological characteristics of the average normal pulse waveforms, the systolic peak was detected with the maximum of the mean pulse, the dicrotic notch with the second maximum of the second derivative, and the diastolic peak with the minima right after this maximum. The different features extracted for finger and neck PPG waveforms comparison, are described in Table 3.1.

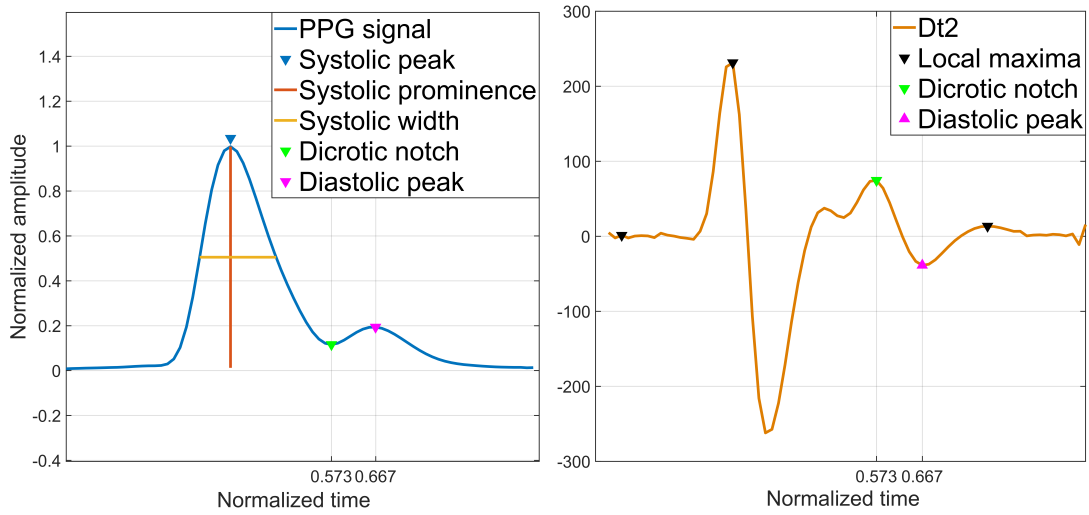


Figure 3.5: Average pulse waveform and its second derivative, for diastolic and dicrotic notch detection.

Statistical tests: The non-parametric Wilcoxon signed-rank test was used for the pairwise comparisons of the features, because the number of samples was too small and the population was not normally distributed. This tested the null hypothesis that the pulse features from finger and neck PPG signals come from continuous distributions with equal medians, against the alternative. Median values of each feature, as well as the interquartile range (IQR) variability, were also reported in Table 3.1.

3.3.2 Results: Differences in average PPG waveforms morphological features

Based on the results from the Wilcoxon signed-rank test, the neck PPG average pulse was significantly different from the finger standard pulse waveform for different features. This can be seen in bold in the comparison of median values in Table 3.1. Based on these, the neck waveform shows a significantly lower systolic peak prominence, pulse skewness and augmentation Index (AI) than the finger one. The neck pulse also exhibits a significantly greater systolic peak width, diastolic peak amplitude, pulse area, IPA, systolic area (A1), diastolic area (A2) and dicrotic notch amplitude than the finger

Table 3.1: WILCOXON SIGNED-RANK TEST PAIRWISE COMPARISONS BETWEEN FINGER AND NECK FOR THE DIFFERENT FEATURES (* $p < 0.05$)

<i>Features</i>	<i>Finger</i>		<i>Neck</i>		<i>Definition</i>
	Median	IQR	Median	IQR	
Systolic Peak Prominence*	0.985	0.008	0.937	0.074	measured by how much the systolic peak stood out due to its intrinsic height and its location relative to the diastolic peak [14].
Systolic Peak Time	0.267	0.107	0.327	0.140	time at which the systolic peak occurred in the normalized time of 1 second.
Systolic Peak Width*	0.242	0.087	0.324	0.120	computed as the distance between the points to the left and right of the systolic peak measured at half prominence height [14].
Diastolic Peak Amplitude*	0.293	0.125	0.500	0.149	corresponded to the normalized height of the second peak in the PPG pulse waveform.
Diastolic Peak Prominence	0.035	0.072	0.026	0.101	measured by how much the diastolic peak stood out due to its intrinsic height and its location relative to the systolic peak [14].
Diastolic Peak Time	0.620	0.080	0.647	0.093	time at which the diastolic peak occurred in the normalized time of 1 second.
Diastolic Peak Width	0.076	0.086	0.036	0.101	computed as the distance between the points to the left and right of the diastolic peak measured at half prominence height [14].
Pulse Propagation Time (PPT)	0.333	0.047	0.307	0.047	time difference between the occurrence of the systolic and diastolic peaks.
Pulse Kurtosis	2.694	1.234	2.092	0.660	referred to the degree of peakedness of the pulse [15].
Pulse Skewness*	0.898	0.600	0.482	0.476	referred to the degree of symmetry of the pulse [15].
Augmentation Index (AI)*	0.706	0.123	0.496	0.149	ratio between the difference in heights of the systolic (s) and diastolic (d) peaks in relation to the systolic peak height [16]: $AI = (s - d)/s$.
Stiffness Index (SI)	5.250	0.648	5.674	0.879	indicates arterial distensibility, can be defined as the ratio between height (h) and PPT [17]: $SI = h/PPT$.
Total Pulse Area*	0.321	0.081	0.419	0.072	the area under the pulse curve from onset to offset.
Systolic Area (A1)*	0.246	0.061	0.304	0.051	area under the pulse curve from the onset until the dicrotic notch.
Diastolic Area (A2)*	0.066	0.028	0.109	0.039	area under the pulse curve from the dicrotic notch until the offset.
Inflection Point Area (IPA)*	0.262	0.092	0.365	0.246	ratio between the diastolic and systolic areas: $IPA = A2/A1$. It is an indicator of pulse wave reflection [16].
Dicrotic Notch Amplitude*	0.285	0.200	0.496	0.128	normalized height of the inflection point between the systolic and diastolic peaks.
Dicrotic Notch Time	0.533	0.087	0.547	0.127	time at which the dicrotic notch occurred in the normalized time of 1 second.

PPG. Overall, the neck PPG pulse is less skewed than the finger PPG pulse. This is corroborated by the fact that the systolic peak prominence is lower since the diastolic notch and diastolic peak heights are greater. As a consequence, A1 and A2 have increased, increasing the total area of the pulse.

A possible interpretation for these changes in the PPG pulse contour is that, depending on the anatomical site, the reflection wave giving rise to the diastolic peak will be transmitted via different possible paths through the complex arterial tree. Thus, due to the proximity of the neck to the body trunk, the reflection wave arrives with greater amplitude since it has travelled across less bifurcations than when it travels to the periphery (finger).

3.4 Spectral comparison under different breathing conditions

3.4.1 Finger and Neck PPG spectrograms differences

Finger and neck PPG spectrograms for normal, fast, slow and apnea breathing conditions were computed using the squared magnitude of the Short-Time Fourier Transform (STFT). These can be observed in Fig. 3.6. A window of 15 s and an overlap of 90% were chosen to ensure respiratory rates as low as 0.067Hz could be captured with good resolution. The spectrograms were compared in the range from 0 to 1.5Hz, since the major distinctions appeared in the low-frequency respiratory bands. Normalized spectrograms of each sensor were subtracted to inspect the differences in power between them, as shown in Fig. 3.7.

In the case of finger PPG, no matter under which breathing condition, the spectrograms in Fig. 3.6 showed approximately the same spectral components. A high concentration of power at the PPG fundamental frequency of 1Hz revealed the AC arterial pulsations. A reduced power density at low frequencies ($< 0.28\text{Hz}$) denoted the PPG DC component modulated by respiration [18].

In contrast for neck PPG, spectral differences could be clearly noticed among various breathing situations. For normal breathing in Fig. 3.6, the principal power band at 1Hz found in the neck PPG spectrogram coincided with the fundamental frequency already observed in finger PPG. However, two other low frequency bands, 0.28Hz and 0.56Hz, stood out equally with high power ($>70 \text{ dB/Hz}$). These corresponded to the fundamental frequency and the first harmonic of the respiration, which presented larger spectral amplitude in the neck than in finger signals, as highlighted by Fig. 3.7. A possible explanation to this is the proximity of the neck to the thoracic cage that makes it more sensitive to respiratory movements. This finding is of great importance, as the prominence of the neck PPG respiratory spectral component could facilitate the extraction of the respiratory rate. During fast breathing in Fig. 3.6, the neck PPG spectrograms displayed two high power bands at 0.55Hz & 1.1Hz, which correspond to the fundamental (F_0) and first harmonic (F_1) of the fast respiratory rhythm.

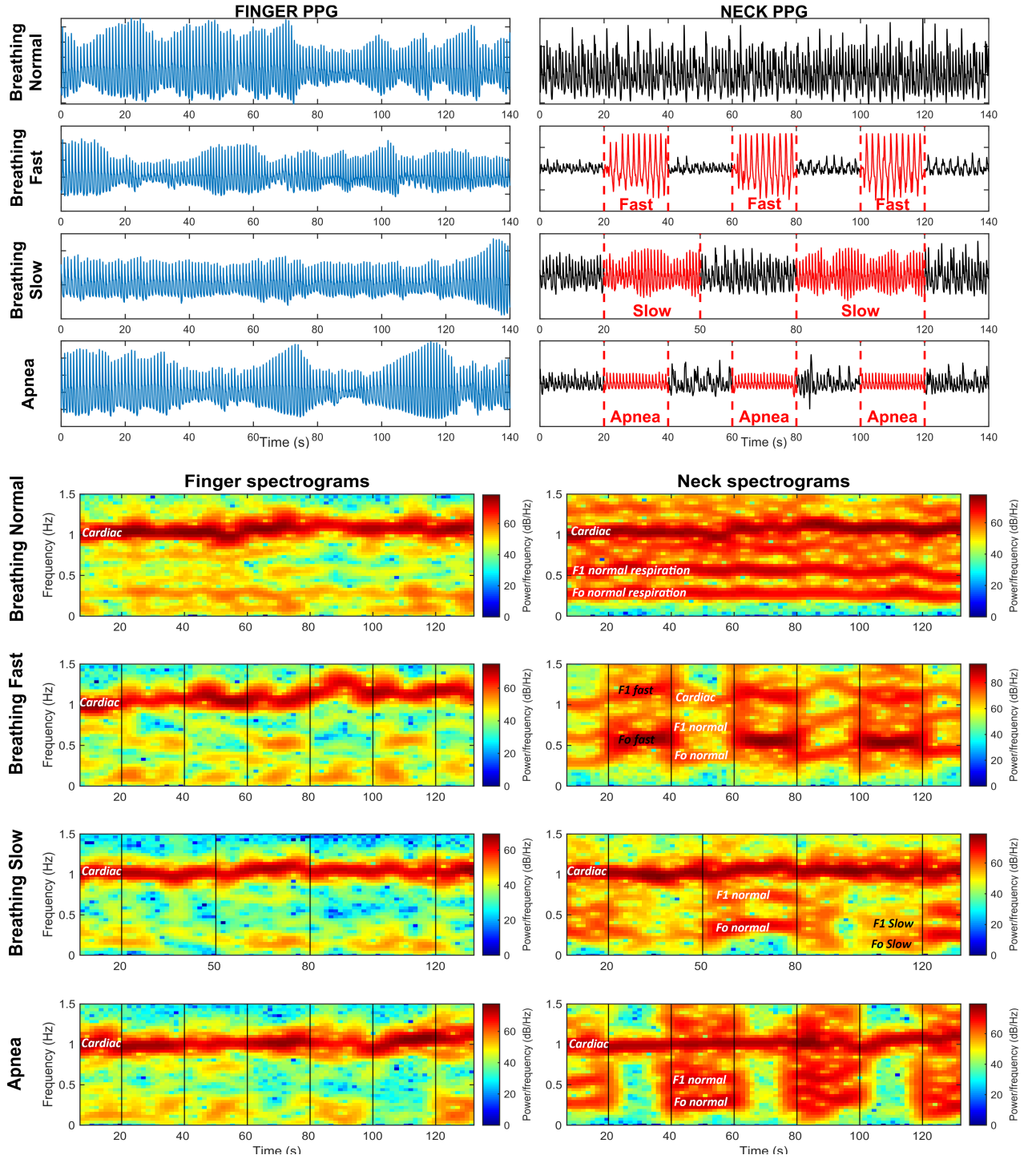


Figure 3.6: Spectrograms of finger and neck PPG signals under different breathing conditions: normal breathing, fast breathing, slow breathing and apnea (breath-holding). The top panels show the PPG signals for both measurement sites. Finger PPG signals are in blue. Neck PPG signals appear in black when breathing at normal pace and in red, when breathing was altered with various respiratory modulations. On the bottom panels, the corresponding spectrograms are displayed. Cardiac: indicates the pulsatile PPG frequency band, Fo & F1: fundamental and first harmonic respiratory frequency bands (for normal, fast and slow breathing)

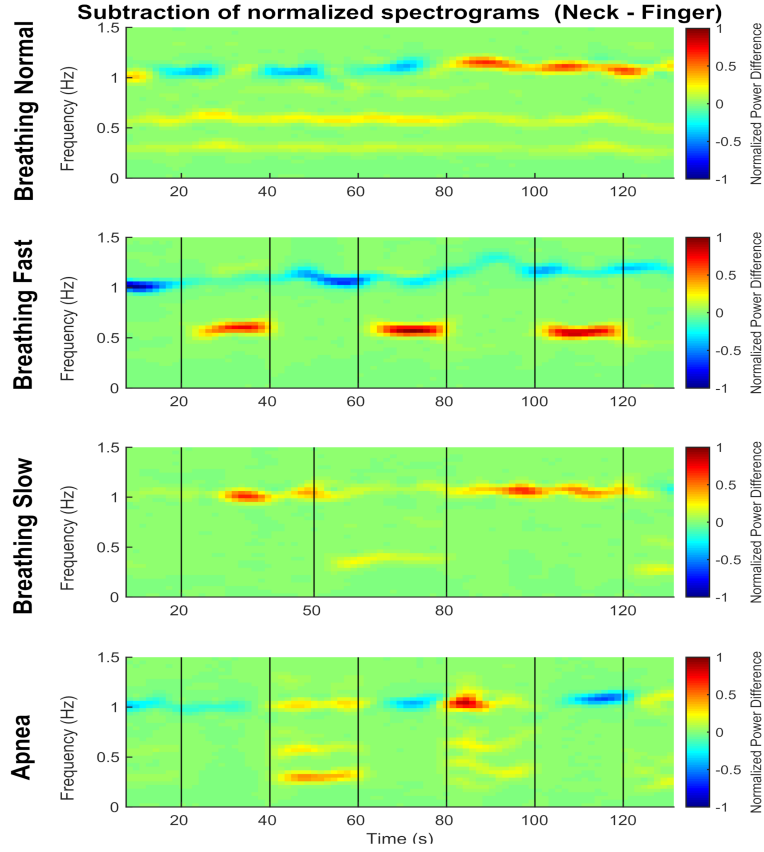


Figure 3.7: Subtraction of the finger and neck PPG spectrograms for the different breathing conditions. The colormap represents the difference between the neck and finger normalized powers in the range of $[-1,1]$. Positive power values, in hot colours, indicate that the neck PPG power is larger than the finger PPG (Neck>Finger), and the contrary applies for the negative values in cold colours (Neck<Finger).

In Fig. 3.7, the main differences between the finger and neck spectrograms were the respiratory band at 0.55 Hz, which was larger for the neck; and the 1 Hz one for which finger PPG dominated. This could be explained by the predominance of the fast breathing fundamental frequency and its harmonics that overlapped with the PPG pulsatile component (1 Hz) of normal PPG. This could be a limitation for the novel measurement site of the neck, as new filtering techniques would be needed.

In the case of slow breathing, the respiratory frequency of neck PPG signals slightly decreased to ~ 0.2 Hz as seen in Fig. 3.6. The power was significantly reduced compared to the normal rhythm, probably because the breathing was stably controlled. When apnea events were simulated, the two respiratory power bands completely disappeared from the neck spectrogram, as a result of the extremely low or inexistent respiratory amplitude. Meanwhile in the finger PPG spectrogram, a poorly localized respiratory band still remained present. This unique property of the neck PPG measurement site could help identify the interruption of respiration, and hence facilitate apnea monitoring. The potential of neck PPG for apnea detection is therefore twofold, as, not only the decrease in SpO_2 levels could indicate hypoxia, but also, the absence of respiration could be detected by examining the power spectral content.

3.5 Heart rate extraction from novel neck photoplethysmography signals

The contents of this section are an edited version of the research published in:

I. García-López, P. Sharma, and E. Rodríguez-Villegas, “Heart rate extraction from novel neck photoplethysmography signals,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2019, pp. 6541–6544 © IEEE

PPG signals are commonly used in clinical and ambulatory monitoring for tracking heart rate (HR), mostly in consumer products, since PPG provides an easy to attach, low cost, and cable free alternative to traditional electrocardiography (ECG) [20].

In this section, we compare the accuracy of extracting HR from neck PPG against reference ECG. HR estimation from finger PPG is similarly evaluated, in order to assess the potential of neck PPG with respect to the preferred location in pulse oximetry.

3.5.1 Methods

3.5.1.1 Experimental protocol

The same recordings as in section 3.2.2 were used here for HR extraction. The variations in respiratory rate were of interest as, according to the cardio-respiratory coupling, HR varies accordingly. This allowed to test the response of neck PPG in a wider range, and therefore simulate more realistic situations.

3.5.1.2 Heart beat detection and HR estimation

HR was calculated for neck PPG, finger PPG and ECG channels in windows of 8s with 6s overlap, to simulate real-time readings updates every 2s. Since the total recording duration was of 140s, 67 windows were used. PPG signals were pre-processed with a 4th order high pass Butterworth filter with 0.7Hz cut off frequency, in order to remove the DC and low frequency components. Systolic peaks of PPG pulses and R peaks of ECG signals were first detected in MATLAB 2018b using the *findpeaks* command. The minimum peak distance parameter was set to 0.6s, to ensure sufficient time resolution for large HRs (up to 100 BPM). A minimum peak amplitude threshold of -0.15 (a.u.) was also established. The cardiac frequency HR_w estimate for each 8s window $w = 1, 2, \dots, 67$, was calculated in beats-per-minute (BPM), as:

$$HR_w = \frac{60}{RR_w} \quad (3.1)$$

where, RR_w represents the average time interval between successive peaks in window w , and was computed as:

$$RR_w = \frac{1}{N} \sum_{i=1}^N \Delta t_{peaks}(i) = \frac{1}{N} \sum_{i=1}^N t_{peak}(i+1) - t_{peak}(i) \quad (3.2)$$

with N representing the total number of inter-beat time differences Δt_{peaks} , in window w .

A correction step was included before final HR_w storage, in order to amend spurious peaks that could negatively affect the precision of RR_w . For that, an adaptive average ($\mu_{\Delta t_{peaks}}$) and standard deviation ($\sigma_{\Delta t_{peaks}}$) of precedent inter-beat distances, were calculated and updated after each processed window. Each peak-to-peak time difference in the evaluated window, was compared against an adaptive threshold such that:

$$|\Delta t_{peaks}(i) - \mu_{\Delta t_{peaks}}| > 2.5\sigma_{\Delta t_{peaks}} \quad (3.3)$$

If the condition was true, the abnormal peak-to-peak difference $\Delta t_{peaks}(i)$ was discarded, and not taken into account for the updated calculation of RR_w , and ultimate HR_w estimation. Figure 3.8 shows the neck PPG Δt_{peaks} traces before (in gray) and after (in red) the correction for a normal breathing recording over time. The true ECG Δt_{peaks} distances were also plotted for reference. It is noticeable how the correction reduced the distance between the neck PPG Δt_{peaks} and the reference beat-to-beat ECG intervals.

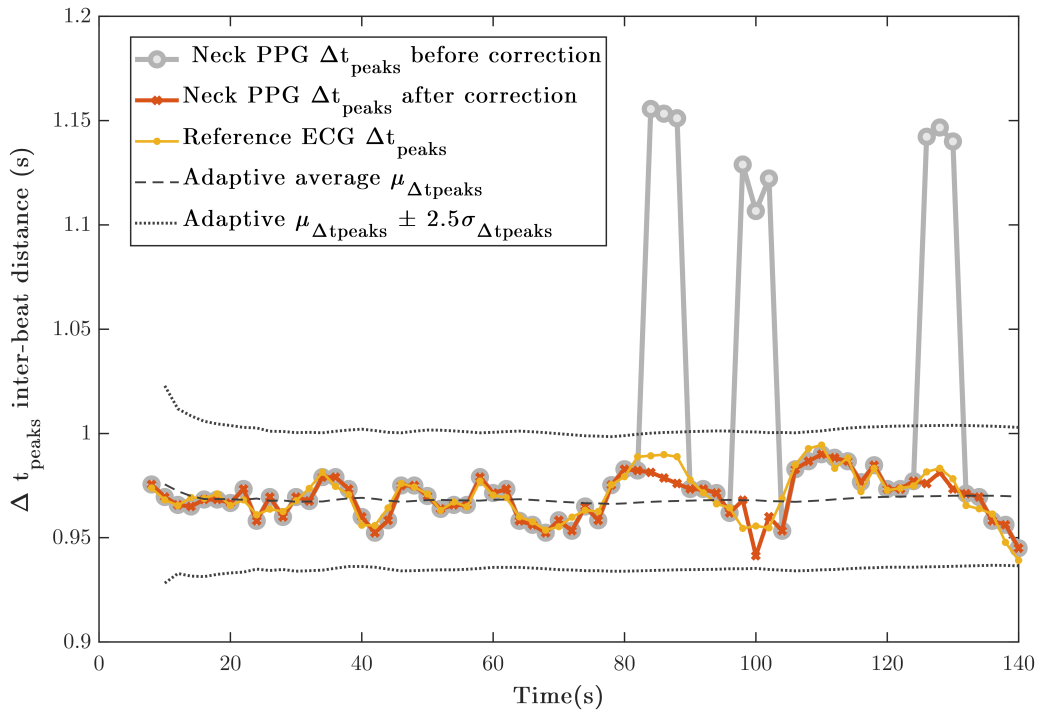


Figure 3.8: *Inter-beat distances over time of a breathing normal recording, before and after spurious peaks correction.*

3.5.1.3 Performance metrics

In order to evaluate the performance of HR extraction from PPG signals, several indices were used. Mean absolute error (MAE), standard deviation absolute error (SDAE) and root-mean-square error (RMSE) were calculated for each subject recording, and subsequently averaged out for the whole cohort. Additionally, HR_{neck} or HR_{finger} values were scattered against reference, HR_{ECG} , to find the best fitting using linear regression. Pearson's correlation coefficient (R) was also computed. Strongly positive (or negative) linear relationships would result in values close to 1 (or -1), whereas absence of correlation would output values in the proximity to 0.

These metrics could only assess the distance and linear relationship between the two sets of PPG estimated and true ECG HR_w observations. But, in order to adequately evaluate the degree of agreement between two quantitative methods measuring the same variable, a Bland-Altman analysis is commonly used [21]. For that, mean-differences of the two measurement methods were plotted against reference values to evaluate the bias. Limits of agreement (LoA) were also constructed to define the interval containing 95% ($\pm 1.96SD$) of the paired differences. In our case, HR_w paired differences, $(HR_{neck} - HR_{ecg})$ and $(HR_{finger} - HR_{ecg})$, were graphically visualized against true ECG HR_{ECG} values in Bland-Altman plots.

3.5.2 Results

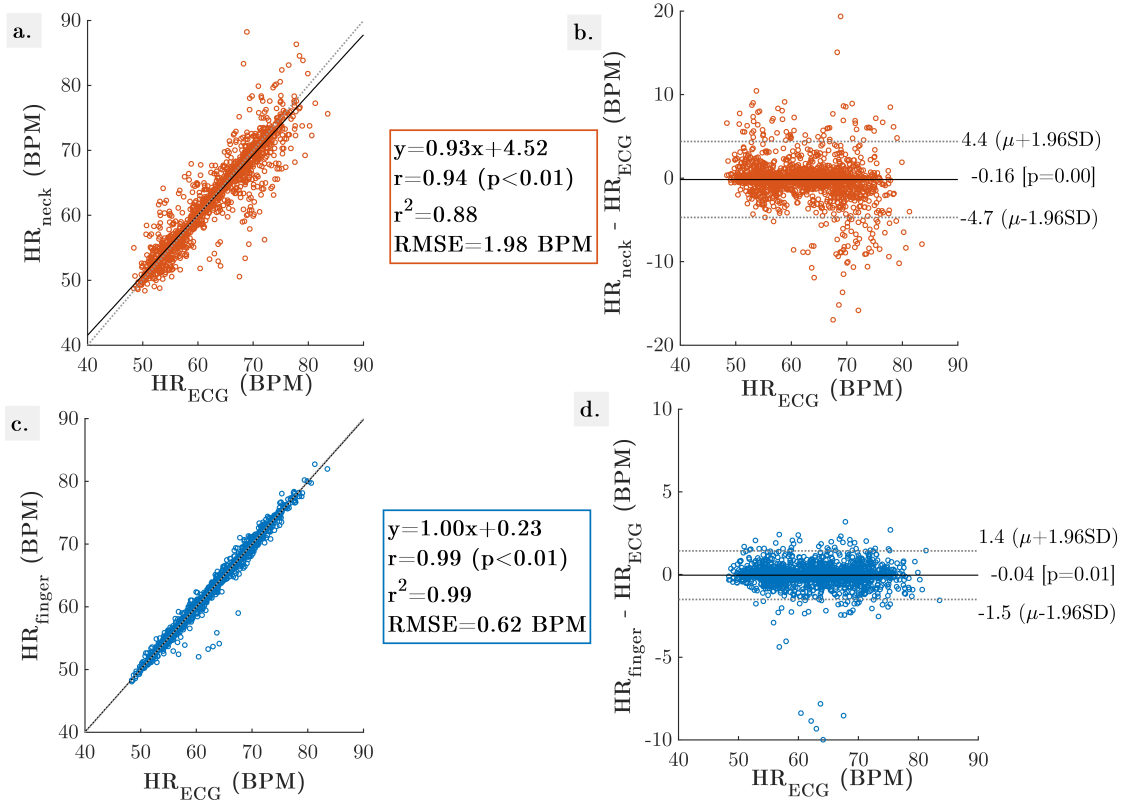
Cardiac frequency estimation from the novel PPG measurement site of the neck was tested and compared against the gold standard ECG method. For the sake of having a reference range of typical errors in conventional pulse oximetry, finger PPG signals were also used to extract HR and were compared to ECG values. A total of 31 recordings of 140s were evaluated, and 2077 HR estimates were extracted from each of the sensing modes.

3.5.2.1 HR estimation errors

Table 3.2 shows the HR performance between each PPG sensor modality and the ECG reference signals, in terms of errors and correlation coefficient. As it can be observed, HR estimation for the neck had in average a 1.22 BPM MAE, 1.54 BPM SDAE and 1.98 BPM RMSE. Pearson's correlation coefficient was very high ($R=0.94$) showing strong correlation between the estimated HR_{neck} and true HR_{ECG} values. For conventional finger PPG, MAE, SDAE and RMSE errors were even smaller with values of 0.38 BPM, 0.48 BPM and 0.62 BPM respectively. The correlation coefficient was also very close to 1 ($R=0.99$).

Table 3.2: COMPARISON OF HR ESTIMATION PERFORMANCE BETWEEN NECK AND FINGER PPG

PPG sensor	MAE (BPM)	SDAE (BPM)	RMSE (BPM)	R (correlation)
<i>Neck</i>	1.22	1.54	1.98	0.94 ($p < 0.01$)
<i>Finger</i>	0.38	0.48	0.62	0.99 ($p < 0.01$)

**Figure 3.9:** (a-c). Linear regression of estimated HR_{neck} and HR_{finger} against HR_{ECG} . (b-d). Bland-Altman plots for all HR estimates.

3.5.2.2 Linear regression fit

HR estimates from finger and neck PPG sensors, were linearly fitted against true HR values of the reference ECG channel. The resulting linear models can be graphically observed in Figures 3.9 (a-c), with their corresponding slope-intercept equations. In both cases, the predictor variables (HR_{neck} and HR_{finger}) suggested a strong positive linear relationship with the ground truth HR_{ECG} , as all data points seem to lie on the diagonal straight line with slope coefficient of 1.

This was pretty much the same for the finger model, where the slope coefficient equalled exactly 1 and the intercept had a very small value of 0.23 BPM. For the neck, the linear fit was very close to ideal, but with a slightly higher intercept of 4.52 BPM, probably due to a larger spread of the data. The coefficients of determination for neck and finger, $r^2 = 0.88$ and $r^2 = 0.99$ respectively, also suggested that the proportion of variance that each variable had in common with the ECG ground truth was very high.

3.5.2.3 Bland-Altman analysis

Bland-Altman graphical analysis for the paired differences of $HR_{neck} - HR_{ECG}$ and $HR_{finger} - HR_{ECG}$ are presented in Figures 3.9 (b-d). Novel neck PPG and reference ECG heart rates showed a good agreement with a bias of -0.16 BPM and 95% LoA of (-4.7, 4.4). The percentage of values found to lie beyond $\pm 1.96SD$ from the mean difference was 7%. The conventional finger PPG and reference ECG cardiac frequencies demonstrated a bias of -0.04 BPM and 95% LoA (-1.5, 1.4), with 2.3% of paired differences scores beyond $\pm 1.96SD$.

3.5.3 Discussion

In this study, HR was extracted for the first time from novel neck PPG signals, at four different respiratory paces, in supine position. The developed algorithm precisely detected cardiac pulses relying on a mean inter-beat distances adaptive threshold for spurious peaks correction. HR was obtained in windows of 8s with 6s overlapping, by the inverse of peak-to-peak time differences multiplied by a factor of 60. The accuracy of HR_{neck} extraction was evaluated with simultaneous ground truth ECG values. Average errors for all subjects and respiratory conditions, revealed very small differences between both techniques in the order of magnitude of ~ 1 BPM. This precision is very promising for accurate HR estimation. In addition, a linear regression model and Pearson's correlation coefficient demonstrated the strong linear relationship between neck PPG and ECG HR estimates. Ultimately, a Bland-Altman analysis showed a good agreement between both techniques, with very small bias between mean paired differences of HR observations, and 95% LoA of (-4.7, 4.4). These findings show the potential of neck PPG to reliably extract cardiac frequency with good accuracy.

When juxtaposing these results to the ones obtained for conventional finger PPG, HR_{finger} presented a reduced error and even higher similitude to ECG overall. Although this analysis was only included for comparative purposes within the same PPG modality, it is worth pointing out that, finger PPG is the widely established gold standard location in pulse oximetry to extract physiological parameters, whereas this work is still at the stage of a proof of concept. Further work is required to test the proposed HR_{neck} algorithm in different sleep positions since this study was limited to supine controlled conditions as a first approach.

This work confirmed the suitability of the neck as an alternative body site for HR estimation based on PPG measurements. Future work should focus on estimating HR_{neck} in the presence of artifacts to ensure its applicability in real life long term monitoring. These advances could have a significant impact in the development of a unique neck wearable sensor incorporating multiple sensing modalities for multi-purpose cardio-respiratory applications.

3.6 SpO₂ extraction with neck PPG

The findings presented in this section were obtained in collaboration with other members from my research group. My contributions did not include the development of the wearable sensor, the recording experiments or the synchronization of the data. But, I jointly worked on the design of the experimental protocol, as well as on the processing and analysis of the data.

The goal of this section is to demonstrate the feasibility of extracting SpO₂% from neck PPG, as an alternative to finger pulse oximetry.

3.6.1 Experiment 2

3.6.1.1 Data acquisition sensors

Due to the fact that no pulse oximeter consumer product currently exists for the neck region, a general reflectance Nonin PPG sensor (8000R, Nonin), typically implemented in forehead applications, was used in previous experiments presented in this chapter. However, different body parts have dissimilar capillarity and tissue characteristics that result in variations of the reflected light optical path. This means that pulse oximeter sensors cannot be used interchangeably for various body locations, but need to be particularly calibrated for each measurement site instead. For this reason, raw PPG signals obtained from the neck had to be further processed for a precise SpO₂ estimation. However, since the Nonin sensor did not provide *both* the infra-red (IR) and red PPG channels *required* for SpO₂ calculation, a neck PPG wearable sensor prototype was used here.

The custom wearable system can be observed in Figure 3.10. It consisted of a reflective optical sensor (MAX30102, MAXIM integrated) for PPG extraction, and an accelerometer to provide body position and help with motion detection. Acquired data was transmitted at 75 Hz via Bluetooth to a mobile device for storage. The wearable PPG sensor was placed at the suprasternal notch of the neck. Additionally, a finger PPG sensor from the portable polysomnography system (SOMNOscreen, SOMNOmedics) was used as reference for SpO₂ calibration. Two respiration impedance bands, were also employed to track DC respiration baseline variations.

An Ultrabreathe® respiratory trainer device, mimicking high altitude oxygen conditions, was safely used to slightly reduce the amount of oxygen intake required for SpO₂ calibration.

3.6.1.2 Experimental protocol

A total of 9 participants (6 males and 3 females) with average age of 28 ± 3 years old, BMI of $23.73 \pm 2.9 \text{ kg/m}^2$, took part in the experiments. The study was approved by the Local Ethics Committee of Imperial College London (ICREC ref.: 18IC4358), and written informed consent was obtained from



Figure 3.10: *Wearable neck PPG system used for SpO_2 calibration experiments.*

all subjects.

Participants were asked to lie down in supine position on a bed and rest for 1min while breathing spontaneously at their normal rhythm. After this, they had to put the Ultrabreath in their mouth and breath through it for another minute. This progressively reduced their oxygen intake to enable the measurement of SpO_2 desaturations. When finished, they were allowed to breath normally again without the device, for as long as their HR returned to baseline. This process was repeated three times per subject.

Following the experiments, the data from the different sensors was synchronized with the help of markers and the reference respiratory bands.

3.6.2 Neck PPG signals processing for R ratio calculation

As already explained in section 2.1.2, the most optimal linear regression fit between the reference SpO_2 values and the calculated ratio of ratios (R) must be sought for SpO_2 calibration at the neck. In order to obtain the R values, the red and infra-red PPG channels were first processed for a reliable extraction of the AC and DC components. The different steps followed to calculate the AC/DC ratio at each wavelength and the final R value are explained in the next subsections.

3.6.2.1 Artifacts removal

Initially, the accelerometer sensor was used to discard motion artifacts. For that, the magnitude of the x, y and z output values was calculated and a threshold was set to cutoff all the signal fragments above it. These sections were labeled as artifacts and discarded for further processing.

3.6.2.2 Signals conditioning

A template denoising function was implemented to remove high frequency noise. The procedure consisted on quantifying the bandwidth of interest using the ground truth finger PPG signal, and then, filtering the inputted red and IR neck signals.

3.6.2.3 Ratio of ratios (R) extraction

In order to calculate the ratio R as detailed in equation 2.11, the AC and DC components of both the red and IR channels had to be inferred. Since the SpO₂ readings were continuously output by the reference pulse oximeter as a 4-beats average, our calculated R ratio had to match the same window. For that, individual heartbeats were identified using the built-in MATLAB function *findpeaks*. This enabled the location of each pulse index for subsequent averaging in later stages of the algorithm.

DC: A moving average was implemented to smooth the DC of the PPG signal, and hence obtain a respiration-free DC baseline signal. The DC value was averaged every 4-beats.

AC: The AC was computed as the amplitude difference between the upper and lower envelopes of the signal over a 4 pulses moving window.

AC/DC ratio: Once the AC and DC values were obtained, their ratio was computed.

R ratio: By combining the AC/DC ratios at each wavelength, the ratio of ratios was estimated as:

$$R = \frac{AC_{red}/DC_{red}}{AC_{IR}/DC_{IR}} \quad (3.4)$$

3.6.2.4 R values processing

In order to ensure an homogeneous distribution of values and avoid bias from potential outliers in the regression analysis, the extracted R ratios were processed as follows:

Plateaus selection: Since the SpO₂ measurements were susceptible to quick fluctuations during experimental oxygen desaturations, the sensor's manufacturer (MAXIM integrated) suggested to localize stable data plateaus, and only use the corresponding R values from these segments for regression [22]. These guidelines were implemented in an algorithm that detected when the first derivative of the SpO₂ time-series became zero (plateaus), and discarded the rest of the data fragments.

Outliers removal: In order to remove outliers and hence ensure a robust regression fitting, a loop through each SpO₂ percentage (%) level selected the R observations within one standard deviation of the mean ($R_{obs} \in \mu \pm \sigma$). Due to the variability among R values for different subjects, this step ensured the most representative observations were used for calibration.

K-nearest neighbours selection: After the signals conditioning and processing stages, and given that each subject was able to desaturate up to a distinct depth, not all SpO₂ (%) levels had the same number of samples. Indeed the higher saturation levels spanning 95-100% concentrated the higher number of points, leaving the lower levels underrepresented. This unbalanced distribution accross the SpO₂ range could severely impact the accurate estimation of the regression model. In addition, this issue was aggravated by the discretization of SpO₂ values. In order to give the same weight to all SpO₂ (%) levels for fitting, a k-nearest neighbours (knn) search algorithm was implemented. By uniquely choosing the k closest points to the mean at each SpO₂ (%) percentage, the number of observations was balanced. The common number of neighbours used for all percentages (%) was the number of points in the lowest level (92%), which was the minimum number of samples over all.

3.6.3 SpO₂ estimation results

3.6.3.1 Linear regression

Figure 3.11 shows the scatter of the reference SpO₂ values against the corresponding processed R ratios for all subjects. The observations within one standard deviation ($R_{obs} \in \mu \pm \sigma$) can be observed in blue and the k-nearest neighbours surrounding the mean in yellow. A transparency factor of $\alpha = 0.01$ and 0.2 respectively, were used to highlight the unbalanced concentration of observations. A linear model was fitted to the data by least squares approximation, i.e. by minimizing the squared distances between each observation and the regression line. The resulting linear function $SpO_2 = 103.63 - 7.14R$, provided the neck calibration coefficients $a = 103.63$ and $b = 7.14$ as indicated in equation 2.11. The coefficient of determination (R^2), demonstrated that the proposed model significantly ($p < 0.001$) explained 63% of the SpO₂ variance.

3.6.3.2 Leave-One-Subject-Out evaluation

Once the calibration of the neck PPG measurement site was obtained, the ability of this novel pulse oximetry location to estimate SpO₂ values was evaluated. A Leave-One-Subject-Out cross-validation (LOSO-CV) strategy was applied to assess the performance of the proposed linear model. For that, one test subject was removed from the whole set iteratively, to repeat the linear regression fitting with the rest of the subjects data. The calculated R values of the test subject were then inputted in the new calibrated model to derive the corresponding SpO₂ level. The root-mean-square error (RMSE) was calculated to quantify the absolute prediction error for each test subject j , such as:

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (3.5)$$

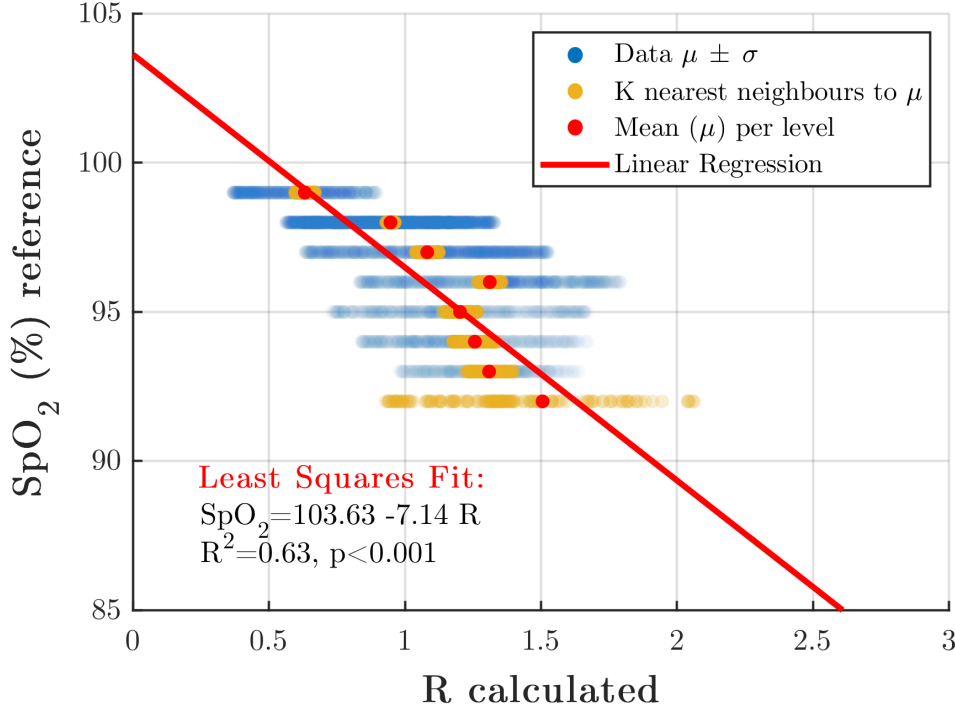


Figure 3.11: SpO_2 (%) regression with calculated R values.

where, $(y_i - \hat{y}_i)^2$ represents the residuals, i.e. how close the test observations are to the model's predicted values; and N refers to the total number of points evaluated.

The overall RMSE (ORMSE) was calculated by averaging the individual $RMSE_j$ for all test subjects. A low value of $ORMSE = 1.79\%$ indicated that the proposed model accurately predicted the SpO_2 response. These results successfully show that the neck could be a promising pulse oximetry location, capable of estimating SpO_2 (%) reliably.

3.6.4 Discussion

In this section, we demonstrated the feasibility of extracting $\text{SpO}_2\%$ from novel neck PPG signals. A fitted linear regression model, proposed an adequate set of coefficients for the SpO_2 calibration at this novel PPG site. A LOSO-CV approach, validated the prediction of test SpO_2 values, with a very low absolute error of $ORMSE = 1.79\%$. According to the U.S. Food and Drug Administration (FDA), the ORMSE for reflective PPG sensors must be below 3.5% for commercialization [23, 24]. This demonstrates the promising value of neck PPG for its future adoption as an alternative pulse oximetry location.

Despite the great significance of these findings, this proof-of-concept must be validated in a larger cohort of participants with various ages, gender, and skin tones, for the linear model to be able to generalize over a wider population. In the same line, the linear regression fit could be similarly improved by further increasing the range of SpO_2 (%) values. Indeed by replacing the Ultrabreath respira-

tor with an oxygen delivery system, the participants could breath an air mix with specific oxygen concentrations through a gas mask. Starting with a rich oxygen saturation to ensure $\text{SpO}_2=100\%$ measurements, the air could be gradually diluted to decrease the oxygen levels in steps of 5% desaturations, down to 70%, as recommended [23, 24]. This protocol, could also solve the problem of having an unbalanced number of ($\text{SpO}_2\text{-R}$) measures per (%) level in the scatter plot, as all subjects would desaturate for the same amount of time and to the same depth. However, this set up requires medical equipment and trained experts to comply with clinical safety measures.

In terms of the processing pipeline, the extraction of the AC and DC components was limited by the presence of noise, even though the participants were lying down at rest. Accelerometer thresholds, simple denoising filters and outliers removal techniques helped with the processing. However, more elaborate automatic signal processing algorithms need to be developed in future work to ensure reliable SpO_2 prediction. Specifically, it is very likely that, in real world monitoring situations, the quality of neck PPG signals worsens as a consequence of motion artifacts corruption.

3.7 Discussion: Advantages and limitations of neck PPG

In this chapter, the novel neck PPG measurement site was compared against standard finger pulse oximetry, in terms of morphological characteristics of the PPG pulse, and spectral frequency components at various respiratory rhythms. The comparison was expanded into the extraction of two physiological parameters typically measured by pulse oximeters, the HR and SpO_2 . The feasibility of extracting HR and SpO_2 from neck PPG was proven, demonstrating its potential as an alternative pulse oximetry location.

Based on these findings, the advantages and limitations of this novel PPG site are discussed here.

The neck is a unique multi-purpose signal acquisition site particularly attractive for respiratory diseases monitoring [10, 7]. The positioning of the PPG sensor on the suprasternal notch allows the recording of tracheal sounds, that could provide respiratory flow information as well as the measurement of tidal volumes [9, 7]. This is specially promising for the detection of apneas since respiratory pauses are characterized by airflow reduction and SpO_2 desaturations. No other pulse oximetry site has the potential to simultaneously record these key physiological signals, using the same wearable device and from a single location. For sleep monitoring applications, this would ultimately reduce the excessive number of sensors currently required in polysomnography.

In addition to being an optimal region to measure respiration through other sensor modalities, it was demonstrated in this chapter that, the neck PPG signal itself presents a more prominent respiration component with respect to conventional finger PPG. The easy identification of the respiratory fundamental frequency and harmonics, could facilitate the extraction of the respiratory rate in dif-

ferent breathing conditions. In the specific case of apneas, respiratory cessation could be detected instantaneously, without having to wait the typical delay of 20-40s to observe an SpO₂ desaturation. This opens up a huge number of possibilities to engineer novel ways of detecting apneic events directly from the PPG signal, without the surrogate SpO₂ signal. In fact, the value of neck PPG is two-fold, as respiratory pauses could be detected by the absence of the respiratory spectral band; or by analyzing conventional SpO₂ desaturations.

The study of the morphological contour characteristics of the average PPG waveform presented at the beginning of this chapter, set the foundations of how a standard neck PPG pulse looks like. This is of great importance to avoid confusing significant neck PPG traits, that might differ from standard finger PPG; with abnormal cardiovascular diseases (CVD) biomarkers.

Neck PPG could also provide interesting new insights into the head vasculature function and hence potentially be used, for example, to reduce the risks of stroke. The morphology of the neck PPG waveform could contribute to the creation of new rapid biophysical markers of the aging process or diseases. For example it has the potential to identify carotid artery blocks and bulges, typical in carotid occlusive disease and extracranial carotid aneurysms respectively. Neck PPG could similarly be a useful tool in clinical diagnosis, to monitor the stiffening of vessels in the head vasculature caused by arteriosclerosis. Other cardiovascular applications for which neck PPG has shown to be promising are the extraction of Left Ventricular Ejection Time (LVET) [25] and the location of Takayasu's arteritis (TA) vessels inflammatory disease in the arterial tree [26].

At present, the novel neck PPG site has several limitations. Despite the promising results in extracting HR and SpO₂ comparably to the gold standard finger PPG, the proposed algorithms need to be tested in the presence of motion artifacts and in a larger cohort of participants. In particular, the SpO₂ calibration has to be expanded to a wider range of SpO₂ (%) values, in controlled hospital conditions, to be suitable for apnea disorders diagnosis.

One of the main disadvantages of the neck region is that it is susceptible to new and unknown artifacts of various sources, that are not typically found in other pulse oximetry sites. Some examples include vibrations or autonomic reflexes originating in the trachea, as well as head movements. Moreover, the high prominence of the respiratory signal could be a great shortcoming during fast breathing, as the first harmonic completely hinders the AC pulsatile component of interest. These demonstrate the necessity of extensively studying the sources of interference in novel neck PPG, to design adequate filtering strategies to clean up the signal.

In addition, neck PPG can only be implemented in the reflective PPG mode that is more complex than the transmission one. In the latter, light travels on a straight line and interacts with all the tissues found in that direction. However, in reflective PPG, light is reflected at different depths under

the skin, before being sensed by the photo-diode. The huge variability in fat content and underlying system of vessels and veins at the neck, makes it very tricky to collect the neck PPG signal under the same conditions for all participants. Indeed, anatomical characteristics of the neck vasculature are much more variable than simple finger or earlobe capillarity. This could make it challenging to find a standard optimal signal acquisition region valid for the whole population.

The suprasternal notch was generally suitable for arterial PPG measurement. However, during various experiments performed in this chapter, a completely different signal appeared sometimes if the sensor was positioned slightly outside of this region. Indeed, when the PPG sensor was located on top of a visible vein, the venous pulse could be measured instead. This could be of interest for other jugular venous pulse (JVP) applications as further explored in next Chapter.

References

- [1] Y. Zhang, R. G. Weaver, B. Armstrong, S. Burkart, S. Zhang, and M. W. Beets, "Validity of wrist-worn photoplethysmography devices to measure heart rate: A systematic review and meta-analysis," *Journal of sports sciences*, vol. 38, no. 17, pp. 2021–2034, 2020.
- [2] Z. Zhang, Z. Pi, and B. Liu, "Troika: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on biomedical engineering*, vol. 62, no. 2, pp. 522–531, 2014.
- [3] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE transactions on biomedical engineering*, vol. 62, no. 8, pp. 1902–1910, 2015.
- [4] M. B. Mashhadi, E. Asadi, M. Eskandari, S. Kiani, and F. Marvasti, "Heart rate tracking using wrist-type photoplethysmographic (ppg) signals during physical exercise with simultaneous accelerometry," *IEEE Signal Processing Letters*, vol. 23, no. 2, pp. 227–231, 2015.
- [5] D. Biswas, N. Simues-Capela, C. Van Hoof, and N. Van Helleputte, "Heart rate estimation from wrist-worn photoplethysmography: A review," *IEEE Sensors Journal*, 2019.
- [6] D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care," *International journal of biosensors & bioelectronics*, vol. 4, no. 4, p. 195, 2018.
- [7] E. Rodriguez-Villegas, G. Chen, J. Radcliffe, and J. Duncan, "A pilot study of a wearable apnoea detection device," *BMJ open*, vol. 4, no. 10, 2014.
- [8] P. Corbishley and E. Rodríguez-Villegas, "Breathing detection: towards a miniaturized, wearable, battery-operated monitoring system," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 196–204, 2008.
- [9] G. Chen, I. de la Cruz, and E. Rodriguez-Villegas, "Automatic lung tidal volumes estimation from tracheal sounds," 2014, pp. 1497–1500.
- [10] B.-S. Lin and B.-S. Lin, "Automatic wheezing detection using speech recognition technique," *Journal of Medical and Biological Engineering*, vol. 36, no. 4, pp. 545–554, 2016.

- [11] M. Peng, S. A. Imtiaz, and E. Rodriguez-Villegas, "Pulse oximetry in the neck-a proof of concept," in *EMBC*. IEEE, 2017, pp. 877–880.
- [12] Y. Zhong, Y. Pan, L. Zhang, and K.-T. Cheng, "A wearable signal acquisition system for physiological signs including throat ppg," in *EMBC*. IEEE, 2016, pp. 603–606.
- [13] I. García-López, S. A. Imtiaz, and E. Rodriguez-Villegas, "Characterization study of neck photoplethysmography," in *40th Annual International Conference of the IEEE (EMBC)*. IEEE, 2018, pp. 4355–4358.
- [14] Mathworks. Findpeaks. <https://uk.mathworks.com/help/signal/ref/findpeaks.html>, note = (Accessed: 2018-01-24).
- [15] R. Couceiro, P. Carvalho, R. P. Paiva, J. Henriques, and J. Muehlsteff, "Detection of motion artifact patterns in photoplethysmographic signals based on time and period domain analysis," *Physiological measurement*, vol. 35, no. 12, p. 2369, 2014.
- [16] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current cardiology reviews*, vol. 8, no. 1, pp. 14–25, 2012.
- [17] S. C. Millasseau, J. M. Ritter, K. Takazawa, and P. J. Chowienczyk, "Contour analysis of the photoplethysmographic pulse measured at the finger," *Journal of hypertension*, vol. 24, no. 8, pp. 1449–1456, 2006.
- [18] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.
- [19] I. García-López, P. Sharma, and E. Rodriguez-Villegas, "Heart rate extraction from novel neck photoplethysmography signals," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 6541–6544.
- [20] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 463–477, 2016.
- [21] D. G. Altman and J. M. Bland, "Measurement in medicine: the analysis of method comparison studies," *The statistician*, pp. 307–317, 1983.
- [22] Maxim integrated, "Guidelines for SpO₂ measurement using the MAXIM® MAX32664 sensor hub," Accessed 2020. [Online]. Available: <https://www.maximintegrated.com/en/design/technical-documents/app-notes/6/6845.html>
- [23] U.S. Food and Drugs Administration (FDA), "pulse oximeters - premarket notification submissions [510(k)s]: Guidance for industry and food and drug administration staff."
- [24] "Medical electrical equipment — Part 2-61: Particular requirements for basic safety and essential performance of pulse oximeter equipment," International Organization for Standardization, Geneva, CH, Standard, 2017.
- [25] S.-H. Liu, J.-J. Wang, C.-H. Su, and D.-C. Cheng, "Improvement of left ventricular ejection time measurement in the impedance cardiography combined with the reflection photoplethysmography," *Sensors*, vol. 18, no. 9, 2018.
- [26] S. Lakshmanan, D. Chatterjee, and M. Muniyandi, "Noninvasive assistive method to diagnose arterial disease-takayasu's arteritis," in *Computational Vision and Bio Inspired Computing*. Springer, 2018, pp. 384–398.

Chapter 4

Extracting the Jugular Venous Pulse from Anterior Neck Contact Photoplethysmography

The research presented within this chapter is an edited version of the previously published study:

I. García-López and E. Rodríguez-Villegas, “Extracting the jugular venous pulse from anterior neck contact photoplethysmography,” Scientific Reports, vol. 10, no. 1, pp. 1–12, 2020

4.1 Introduction

Cardiovascular diseases (CVDs) are listed as the principal cause of mortality worldwide by the World Health Organization (WHO). Accounting for 17.9 million deaths annually, they represent 31% of global deceases [2]. In Europe, CVDs claim 3.9 million deaths per year with an estimated total cost of 210 billion euros to the European Union [3]. In the United States, 92.1 million people are currently living with some variety of CVD [4]. CVDs result in an impaired blood supply to the different organs of the body. They include vascular disorders, involving blocked or damaged blood vessels (e.g., coronary, peripheral or cerebrovascular arterial disease); and cardiac disorders, resulting from heart contractility dysfunction (e.g. cardiomyopathy, heart failure, cardiac dysrhythmias).

Non-invasive diagnostic and monitoring methods for CVDs include the assessment of cardiac electrical activity with electrocardiography (ECG), examination of blood arterial pulsations with photoplethysmography (PPG) or ballistocardiography, and scrutiny of the vasculature’s function by measuring arterial blood pressure (BP). Most of these physiological signals provide information about cardiac pump efficiency and blood delivery function. However, a significant number of CVDs affecting the

right side of the heart cannot be diagnosed with these, since they result in central venous pressure (CVP) abnormal values that cannot be inferred from the arterial pulse. In these cases, the jugular venous pulse (JVP), is the standard physiological signal commonly measured for diagnosis [5]. Examples of CVDs that can be diagnosed by identifying JVP abnormalities are: hypovolaemia, tricuspid stenosis and regurgitation, constrictive pericarditis, atrio-ventricular block, atrial fibrillation, or cardiac tamponade [6].

The JVP is considered as a manometer for right atrial pressure and CVP, because the superior vena cava extends all the way from the entrance of the right atrium up to the internal jugular vein at the neck [7]. Due to the quasi-absence of bifurcations between the first and the latter, the pressure at the right atrium is transmitted following a straight route to the internal jugular vein. Moreover, during diastole, while the tricuspid valve is open, the right ventricle is directly connected in series with the right atrium, superior vena cava and ultimately with the jugular vein. This results in differentiating characteristics of the JVP signal when there is an abnormal behaviour of the right ventricle too. The typical JVP tracing is presented in Figure 4.1, together with the typical ECG and PPG signals for cardiac time reference purposes.

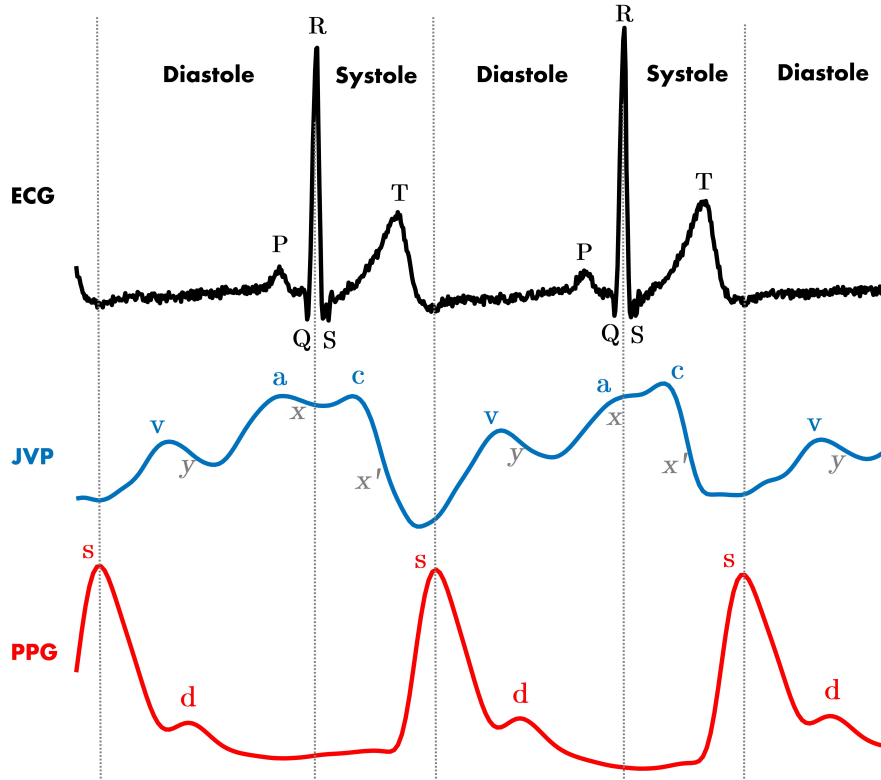


Figure 4.1: ECG, JVP and PPG signals traces. Dashed vertical lines separate the two phases of the cardiac cycle: systole and diastole. Each physiological signal is annotated with the corresponding significant waves. For ECG: *P* corresponds to atrial depolarization, *QRS* complex to ventricular depolarization and *T* to ventricular repolarization. For JVP: the *a, c, v* ascents are annotated in blue and the *x, x'* and *y* descents in gray. For PPG: *s* indicates the systolic peak and *d* the diastolic peak.

The JVP waveform is formed by three ascents (a , c and v waves) and three descents (x , x' and y), which respectively represent the different events of the cardiac cycle in terms of pressure variations. The (a) wave, occurring right after the P peak of the ECG, represents atrial contraction and is followed by the (x) descent that indicates relaxation of the atrium and closure of the tricuspid valve. Subsequently, after the QRS complex of the ECG, the (c) wave shows right ventricular contraction as a result of the bulging of the tricuspid valve during ejection, and it is followed by the drop in pressure (x'). Finally, after ventricular repolarization (T peak of ECG), the (v) wave reflects the maximum pressure attained from blood filling of the right atrium before the tricuspid valve opens again. This last event results in the fall of pressure during rapid ventricular filling, represented by the (y) wave; and the whole cycle starts again [8]. The typical arterial PPG waveform onset happens approximately at the start of the diastolic phase, but it can occur earlier or later in time, specific to the person, as it depends on the pulse transit time to the periphery.

The most broadly used technique to measure the jugular venous pulse (JVP) is central venous line catheterization. This consists in surgically introducing a catheter into the right internal jugular vein, all the way to the right atrium and sometimes even advancing it down the pulmonary artery [9]. This invasive method is not routinely performed in primary care due to all the risks it entails, but instead, it is only used in acute CVD patients at the ICU [10]. Some undesirable complications that justify this decision include: pneumothorax, carotid artery puncture, pulmonary infarction, arrhythmias or catheter infection [11].

The alternative non-invasive methods that eliminate the risks of such invasive technique, look for JVP waveform abnormalities by indirectly observing the internal pressure of the heart at the more accessible site of the external neck. This can be done, for example, by simple visual examination of the jugular vein blood column height in the triangle formed by the sternocleidomastoid muscles and the clavicle [12]. However, this option requires trained personnel and is usually very subjective, making its accuracy questionable. Alternatively, a study also showed the feasibility of extracting the JVP using ultrasound (US) B-mode of the jugular vein cross-sectional area [13]. Although this method showed potential, a stable probe-skin contact is crucial and US equipment remains very expensive.

Non-invasive remote optical techniques that try to counteract the disadvantages of the current state-of-the-art have been explored recently. Some initial efforts visualising the neck area with a camera system equipped with a near-IR light, were able to identify neck pulsations in the triangle of the sternocleidomastoid muscles underlying the right jugular vein [14]. However, this proof of concept was unable to distinguish whether the pulsations were of carotid or jugular origin, to prove their hypothesis. A different approach, based on a video imaging photoplethysmographic system, demonstrated that, with the subject still in supine position it was feasible to extract the JVP waveform from the neck [15].

Another recent study [16], showed that camera-based skin micro-motion vibrocardiography (cVCGI) with subjects in recumbent-to-supine position could sense skin-displacement caused by jugular vein pulsations. Similarly, Lam Po Tang *et al.* [17] non-contact alternative was based on using a commodity camera and a subpixel image registration algorithm to identify the same jugular skin displacements.

However, despite the promising proposed methods in the literature, all these video remote diagnostic systems constraint the patient to a specific body position facing the camera in front of them. This severely restricts the patients from moving or changing position during monitoring, making it unsuitable for longer term signal acquisition. In addition, in the clinical setting, they all require training of hospital personnel to ensure correct set up and equipment calibration.

This chapter shows, for the first time, that it is possible to extract the JVP from the anterior jugular veins by using contact photoplethysmography (PPG). The advantage of this is that PPG is an easy and cheap physiological sensing modality, and thus, a system developed using this could potentially overcome the shortcomings of all other existing techniques. The fundamentals insights provided by this study are intended to set the foundations for the future design of a small, light and user-friendly wearable sensor, which would be able to record continuously for long periods of time, without constraining the position of the patient in bed. This technique thus, has the potential to be a breakthrough for physiological monitoring of CVDs, mostly in outpatient clinics, domestic environments, and low resourced settings.

4.2 Methods

4.2.1 Experiment 3 - sensors and protocol

Signals were recorded from 20 subjects (15 males and 5 females), with average age 27 ± 4 years old, height of 1.75 ± 0.09 m and weight of 74.2 ± 13.1 kg. The study was approved by the Local Ethics Committee of Imperial College London (ICREC ref.: 18IC4358), and written informed consent was obtained from all subjects. Informed consent was likely provided from participants whose images were displayed in this work, for both study participation and publication of identifying information/images in an online open-access publication. All experiments in this work were performed in accordance with the Declaration of Helsinki. The subjects were asked to lie down on a bed, since it was hypothesized (and confirmed) that JVP pulsations could not be properly sensed if participants were seated [16]. This is due to the fact that when the body is lying down, the venous return does not have to counteract the effect of gravity to pump blood back to the heart from the lower extremities, and blood is pumped more easily to the head. In fact, neck venous pressure is elevated, enhancing the pulsatility of the JVP pulse at the neck.

One reflectance pulse oximeter sensor (8000R, Nonin) coupled to an OEM processing module (Xpod, Nonin), was located on the anterior area of the neck. A transmission finger pulse oximeter (Onyx II 9560, Nonin) with Bluetooth connectivity, was located on the index finger of the left hand, and used as the ground truth. In order to validate and identify more precisely the JVP characteristic (a, c, v) waves and their corresponding timings, 2-lead electrocardiography (ECG) sensors from a polysomnography system (SOMNOscreenplus, SOMNOMedics), were also used as reference. A PPG sensor connected to the polysomnography system and placed on the index finger of the right hand, was used to synchronize with the two Nonin PPG sensors.

In order to further confirm the relative position of the reflectance neck PPG sensor, with respect to the neck veins, an IR vein camera (Infrared Vein Finder, Z-imaging) connected via USB to a computer, was placed 30cm above the participant's head. The mounted imaging system, that was used to visualize the underlying venous system of the anterior neck, consisted of a 1080P IR sensitive lens surrounded by IR LEDs (800-1000nm) to illuminate the tissue.

Figure 4.2 shows the experimental setup, with one of the subjects lying down in supine position on a bed with all the attached sensors and imaging system. The IR vein camera online image and Nonin PPG sensors recordings could be visualized in real-time on the screen to verify the venous anatomy of the neck area. Signals were acquired for 60 seconds duration for each subject. After data acquisition, ECG recordings were imported into the computer for processing. Recordings were synchronized in time by finding the maximum correlation between both fingers PPG sensors' signals.

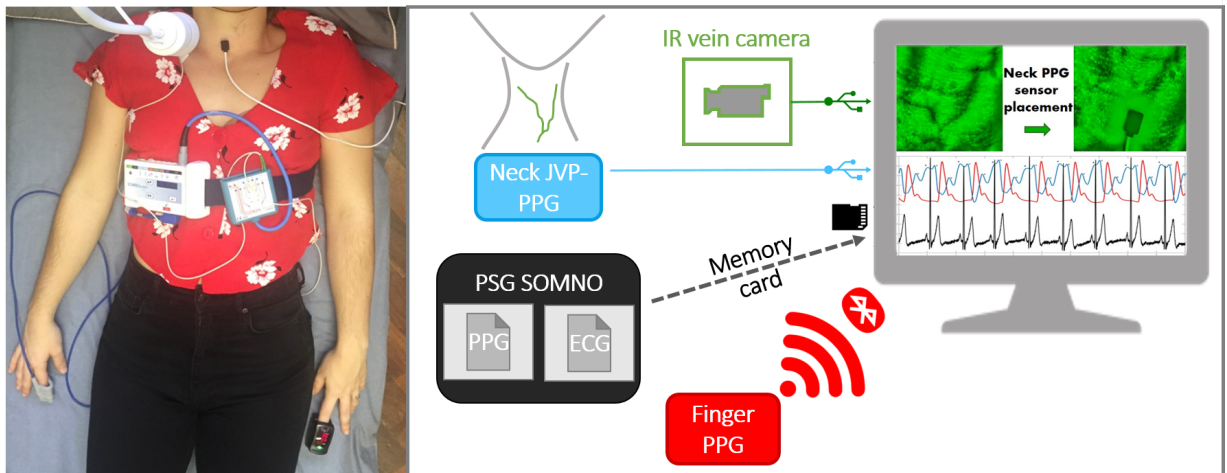


Figure 4.2: Experimental setup. A participant lying down with two transmission pulse oximeters on each hand and one reflectance PPG sensor on the neck. ECG signals were simultaneously recorded using a portable polysomnography system (SOMNOMedics). The vein imaging camera was placed 30cm on top of the participant's neck to visualize the neck vasculature. Online vein imaging, neck JVP-PPG signals and ground truth finger PPG, were observable in real-time on the screen of the computer. ECG and PPG signals acquired in parallel with the polysomnography SOMNO system were loaded in the computer, synchronized and processed after the data acquisition session was finished.

A diagnostic ultrasound (US) system (Sonix RP, Ultrasonix) coupled with a linear transducer probe (14-5 MHz) was also used to obtain transverse B-mode images and videos of the internal jugular vein in order to validate our PPG-based jugular measurements. Neck JVP-PPG recordings were acquired in parallel with the US probe placed on the right side of the neck for 6.5 s. Using the 3.2 cm cross-section measurement tool of the US system, the jugular vein wall distension across the imaged internal jugular vein was measured in the transverse plane over time. The output 2D topographic graph (of 6.5s duration) was further processed to segment the jugular vein cavity area and calculate the cross-sectional diameter at each instant of time. For that, all the pixels inside the segmented jugular walls were summed and scaled in terms of distance (cm). The resulting variations in diameter of the jugular cross-section represent the reference ultrasound JVP waveform for comparison with our neck JVP-PPG signals.

4.2.2 Sensing location

In order to extract the JVP from the anterior neck, the contact reflectance PPG sensor was placed at the middle inferior region. As it can be observed in Figure 4.3(a), this area is highly perfused by the anterior jugular veins (AJVs) that combine into the jugular venous arch (JVA) [18, 19]. Figure 4.3(b) shows how the JVA drains into the subclavian vein (or occasionally into the external jugular vein (EJV)), which directly joins the superior vena cava further down, which in turn is ultimately connected to the right atrium of the heart. This venous configuration linking the neck jugular veins in almost a straight path with the right side of the heart, allows RA pressure changes to be easily transmitted to the AJVs and JVA, in the form of the JVP.

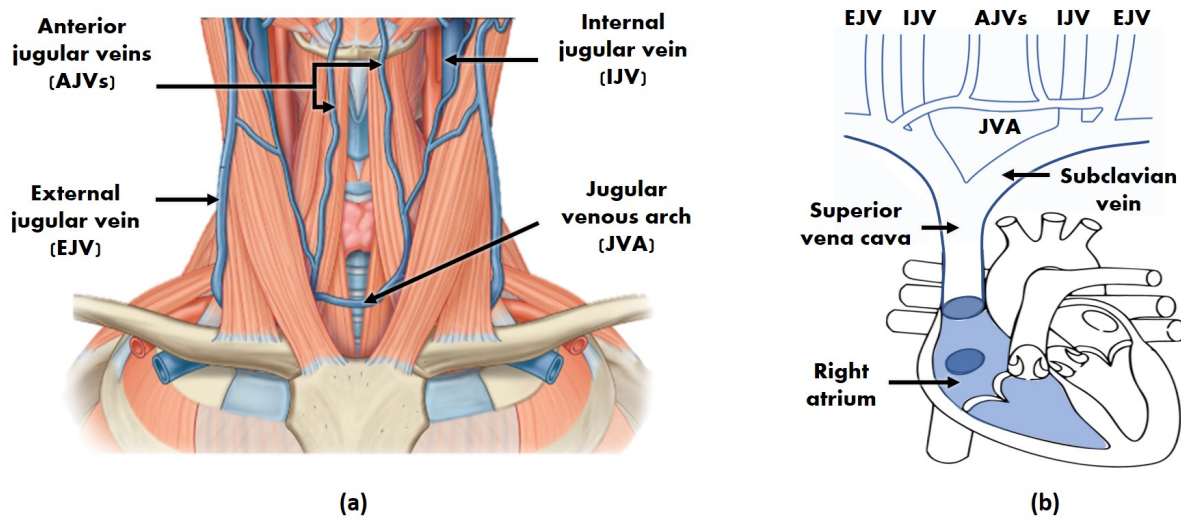


Figure 4.3: Neck venous system anatomy. (a) Superficial veins schematic of the inferior anterior region of the neck (adapted from [18]). (b) Schematic of the venous tree connecting the external (EJV), internal (IJV) and anterior (AJVs) jugular veins to the right atrium of the heart.

But the JVP is only noticeable in this area of interest, when the participant's body is in recumbent or supine positions [16]. For this reason, experiments were carried out with participants lying down. Blood then concentrates more in the central venous compartment as a result of being less influenced by the gravitational pull, and hence, the JVP can be better observed in the neck venous system.

Another key feature that makes the frontal neck suitable for reflectance contact PPG sensing is the absence of thick tissues preventing the penetration of light. Neither the sternocleidomastoid nor the platysma superficial muscle cover the central lower area along the midline. On the contrary, the AJVs appear exposed in front of the infrahyoid muscles, and accessible superficially as they pierce the investing fascia [18]. Moreover, the skin thickness of the anterior neck is narrower than for the anterior-lateral regions [20]. In addition, the inferior part, is the thinnest compared to the middle and superior neck regions. These anatomical properties facilitate the light to easily reach the subcutaneous plexus.

4.2.3 Neck JVP and finger PPG signals correlation analysis

In order to measure the correlation in the time-frequency plane between the neck JVP and finger ground-truth PPG signals the magnitude-squared wavelet coherence was used. It was computed using the MATLAB *wcoherence()*, based on the analytic Morlet wavelet over logarithmic scales. For two signals x and y , it is defined as:

$$WCoherence_{xy}(f) = \frac{|C_{xy}(f)|^2}{C_{xx}(f)C_{yy}(f)} = \frac{|s(C_x^*(a, b)C_y(a, b))|^2}{s(|C_y(a, b)|^2)s(|C_y(a, b)|^2)} \quad (4.1)$$

where, $C_{xy}(f)$ denotes the wavelet cross-spectrum between x and y , and $C_{xx}(f)$ and $C_{yy}(f)$ the wavelet auto-spectra of x and y respectively. $C_x(a, b)$ and $C_y(a, b)$ refer to the continuous wavelet transforms of x and y at scales a and positions b . s represents the smoothing operator in time and scale, and $*$ the complex conjugate. The magnitude-squared wavelet coherence values range from 0 to 1, i.e. from low to high correlation. The phase of the wavelet cross-spectrum values were also extracted to inspect the relative lag between the neck JVP and finger PPG signals.

4.2.4 Neck JVP waveforms annotations and averaging

For each subject, signals segments of 5 seconds duration were selected. JVP signals were manually annotated by marking the characteristic (a, c, v) waves, as well as the onset of each JVP pulse corresponding to the trough before each v wave. The main peaks of the ECG (R) and arterial finger PPG (S) were also identified for reference.

Average time differences between ECG, PPG peaks and JVP waves were calculated for 5-cycles. Each JVP cycle (i.e. v - v interval) was defined as the duration between two subsequent onsets (O),

and was used to normalize the distances between waves. This was done similarly to Lam Po Tang *et al.* [17], for the sake of comparison.

Using the whole duration of the recording (60 s), average JVP pulse shapes were obtained for each subject. For that, the automatic algorithm previously presented in [21], was applied to perform segmentation, alignment and averaging with a quality correction stage, of JVP pulses. The average waveforms were annotated, after being normalized in amplitude and in time (by the average pulse duration), for inter-subjects comparison.

4.3 Results

4.3.1 Anterior neck veins imaging

Figure 4.4 shows the anterior view of the venous system at the lower neck, visible to the naked eye (a) and by means of the IR camera system (b-c). As it can be observed in the leftmost panel 4.4(a), veins are very evident already in the selected participant with very pale thin skin. It is not always that obvious for participants with darker skin tones, which is why the IR imaging system was used to confirm the underlying venous anatomy. With the IR camera it was possible to visualize with a low resolution (and hence in some subjects better than in others) the pair of anterior jugular veins (AJVs) descending parallel to the midline of the neck and connecting at the jugular venous arch at the inferior part, as shown in Figure 4.4(b). The rightmost panel 4.4(c), demonstrates a possible longitudinal positioning of the reflectance PPG sensor on top of the left AJV, that enables extraction of the desired neck JVP-PPG signal.

Vein vasculature imaging, at the anterior neck, varies significantly among subjects as it can be observed in Fig. 4.5. Differences in anatomy and fat content in the neck, affect the identification of veins that are located more or less superficially. For some subjects (4, 5, 7, 9, 12, 13, 19); broad venous ramifications are clearly noticeable as intertwined black curvilinear lines. AJVs connecting into the

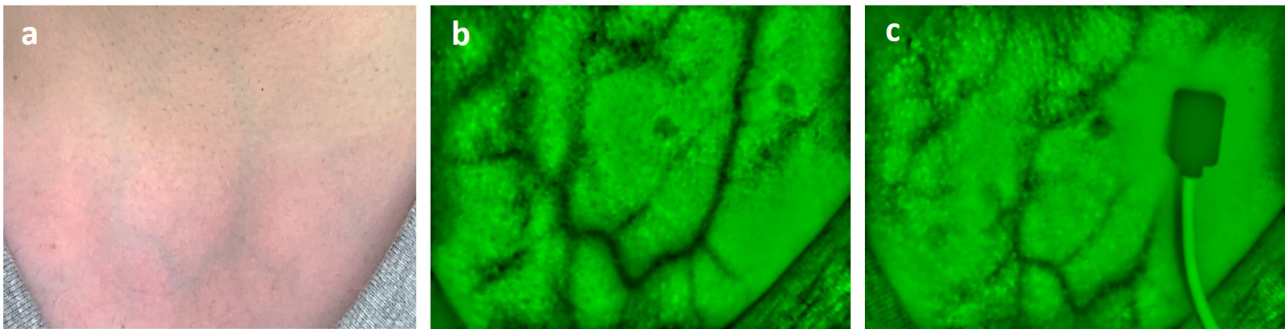


Figure 4.4: Anterior view of neck veins (subject 12). (a) Image of the anterior lower neck. Veins are already visible to the naked eye. (b) Vein camera image of anterior jugular veins and jugular venous arch. (c) Reflectance PPG sensor placed on top of the left anterior jugular vein for recordings.

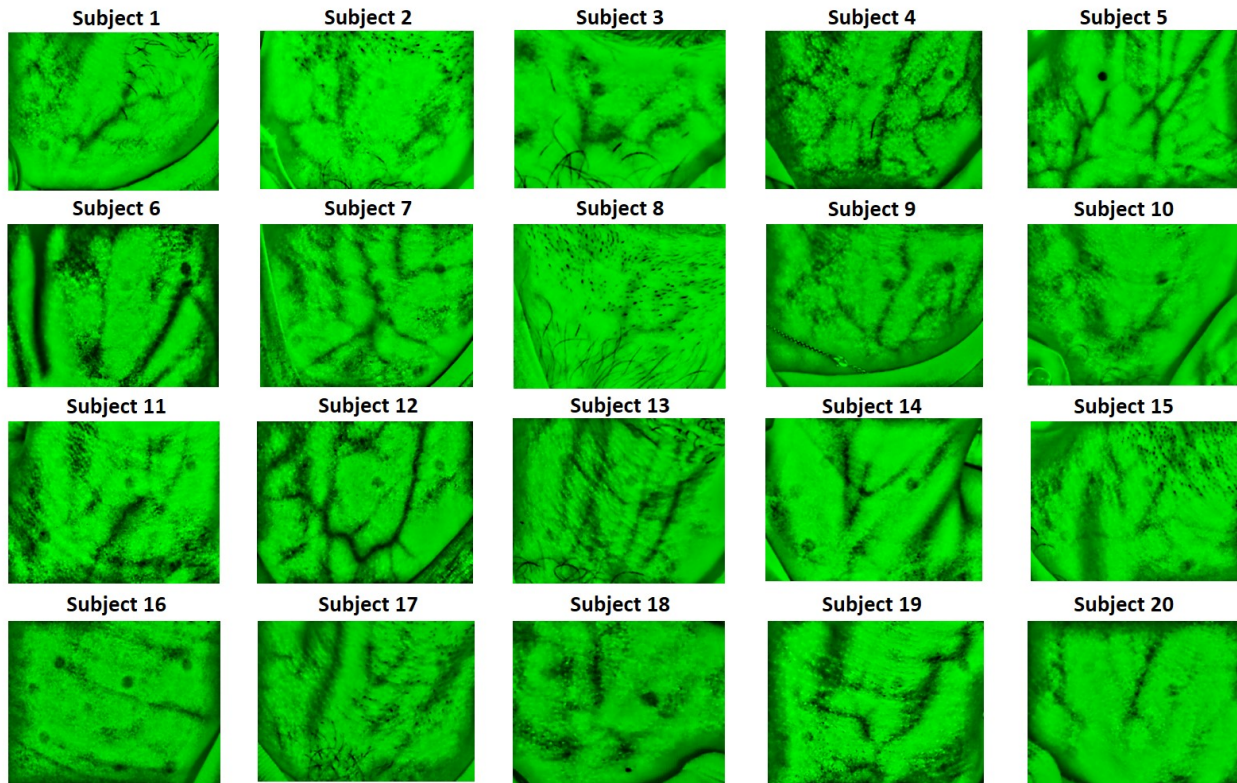


Figure 4.5: *Anterior neck veins vasculature imaging for all participants ($n=20$).*

jugular venous arch can be identified in most of the images. For other subjects (2, 8, 10, 16, 18), the expected venous tree is hardly observable, but instead, less sharp grey lines highlight some thinner isolated veins.

4.3.2 Ultrasound internal jugular vein measurements validation

Figure 4.6 shows the comparison between the ultrasound (US) derived jugular venous pulse and our proposed neck JVP-PPG signals, for three different subjects. The top panel shows the US B-mode image of the carotid artery (CA) and jugular vein (JV). In red, the segment of 3.2 cm in length, indicates along which line the cross-section of the jugular vein was measured at each time instant for a duration of 6.5 seconds. Underneath, the resulting 2D topography graph shows the jugular venous walls and surrounding tissues displacement in the transverse plane over the recording period. The middle panel shows the extracted US jugular venous pulse (US-JVP) after segmenting uniquely the JV cavity area profile. This US-JVP corresponds to the variation in cross-sectional jugular diameter (in cm) over time. When comparing this US reference waveform morphology to our neck JVP-PPG signals, shown in the bottom panel, it can be seen that they match perfectly in terms of shape and timings. In both signals, US-JVP and neck JVP-PPG, the typical (a, c, v) JVP waves can be identified. This validates our proposed neck contact PPG technique for extracting the JVP from the anterior neck.

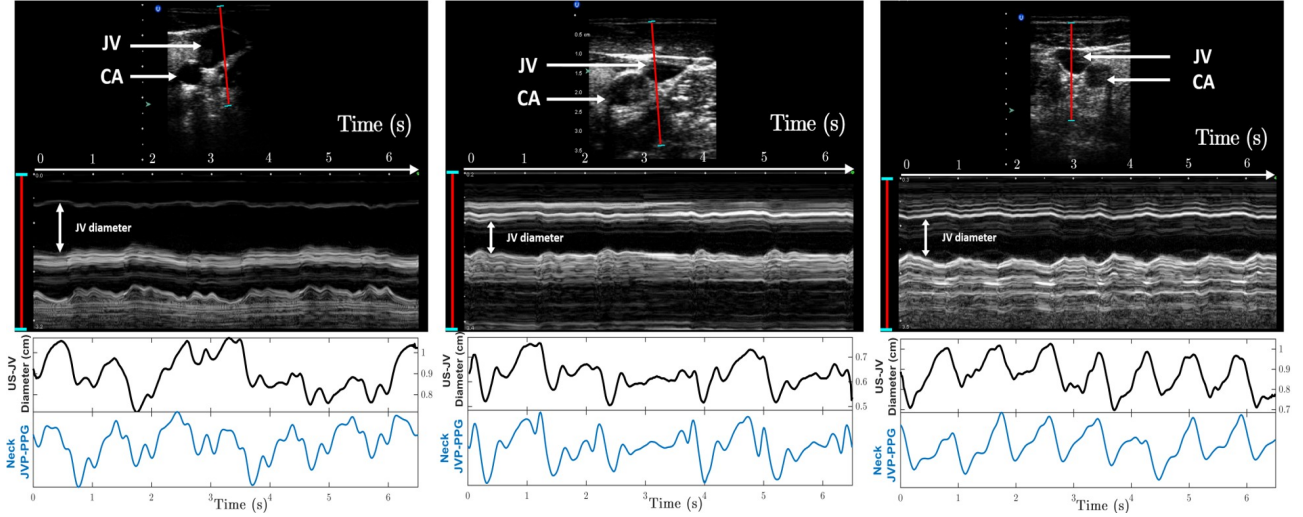


Figure 4.6: Validation of the proposed neck JVP-PPG signals with reference ultrasound (US-JVP) for three different subjects. Top panels show the B-mode image of the carotid artery (CA) and jugular vein (JV) in the transverse plane together with the 2D cross-section profile along the red (3.2 cm) segment over time. The middle panel displays the US-derived JVP signal, i.e. the JV cross-sectional diameter variations over time, after image post-processing of the US 2D topography for the specified time interval. The bottom panel presents the corresponding neck JVP-PPG signal recorded in parallel over the 6.5s period of time, that perfectly follows the reference US-JVP morphology.

4.3.3 Wavelet coherence between neck JVP and finger PPG signals

The raw neck and finger PPG signals measured for all subjects are shown in Figure 4.7. As it can be observed, the two PPG signals are not in phase, but instead the neck JVP pulses seem to lead the arterial finger PPG pulses. This is already consistent with cardiac cycle physiology, since the JVP *a* and *c* waves mark the right atrial and ventricular contractions, that give rise to the arterial PPG pulse observed *a posteriori* at the peripheral parts of the body.

Indeed, an interesting property of the JVP identified by Amelard *et al.* [15] was that it is strongly negatively correlated with the arterial PPG pulse. In order to verify whether our recorded signals were also inverted, the magnitude-squared wavelet coherence was computed, as a measure of the correlation in the time-frequency plane between neck inverted JVP and reference finger PPG signals. The coherence plot of these, is shown for one of the subjects in Figure 4.8. The phase values of the wavelet cross-spectrum are represented by arrows, the orientation of which indicates the relative lag between the neck JVP and finger PPG pulses. Phase arrows are only displayed for coherence values larger than a manually set threshold of 0.8, therefore showing strong correlation. At the cardiac frequency of 1Hz there is a concentration of large wavelet coherence values almost equal to 1, between the neck inverted JVP and the reference finger PPG signals. In addition, along this band, for the whole duration of the recording, it can also be observed that the phase arrows are pointing in the antiphase direction ($\pm 180^\circ$). This confirms the inversion of the proposed neck contact JVP with respect to the arterial pulse, and hence the strong negative correlation.

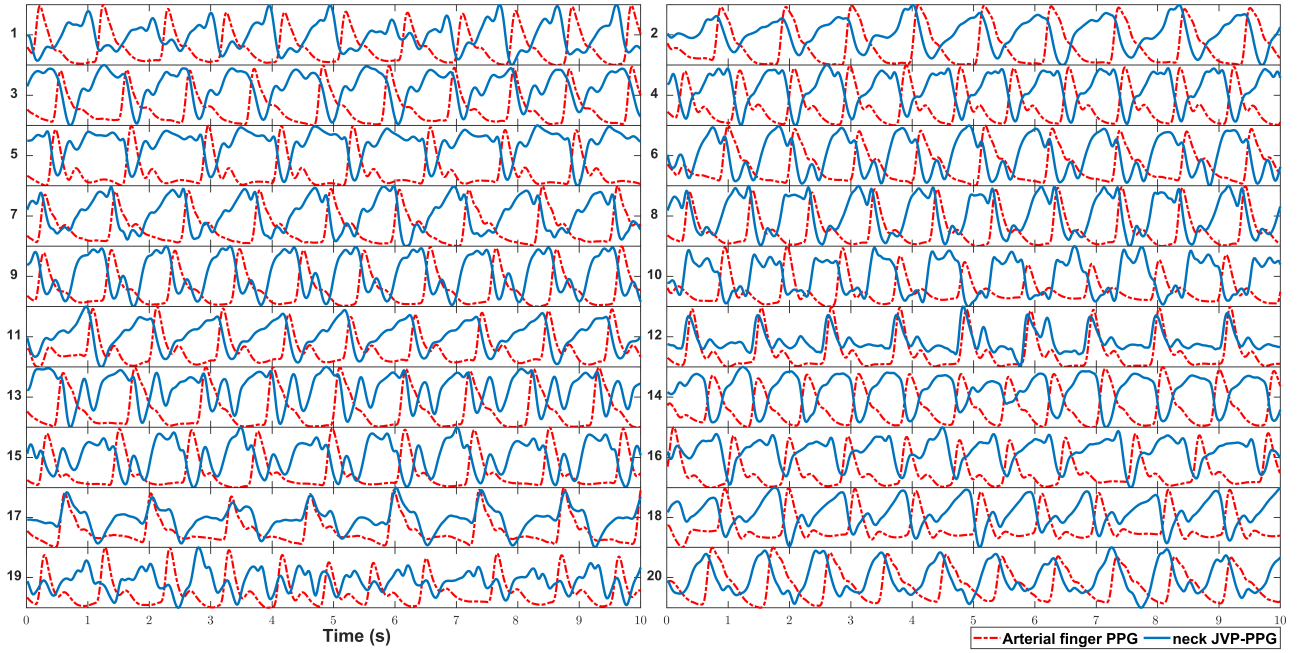


Figure 4.7: Recordings showing neck JVP and finger arterial PPG signals for all subjects ($n=20$). Signals were normalized in amplitude in the range $[0,1]$ for visualization purposes.

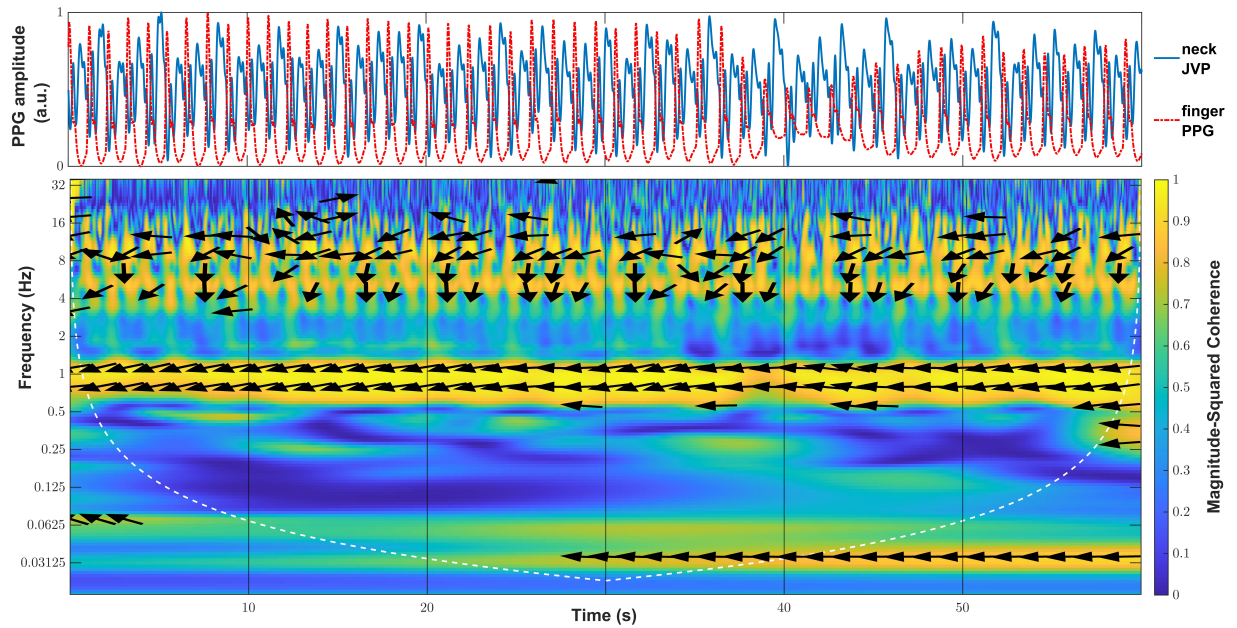


Figure 4.8: Wavelet coherence plot for neck inverted JVP and arterial finger PPG signals. Arrows represent the phase of the wavelet cross-spectrum at magnitude-squared coherence values greater than 0.8.

4.3.4 Neck JVP signals annotation and characteristic waves location

Figure 4.9 shows the neck JVP signals manually annotated (blue dots), together with the ECG (yellow diamonds), and PPG (red dots) reference signals, for 10 subjects. The three characteristic JVP a , c and v waves were identified in all subjects, except for subject 20 for which the a wave was not recognized. The green dots point the onsets of the JVP pulses used to calculate the v - v interval

distances. These were needed to normalize the locations of the different waves and peaks at each cycle, and to calculate the differences between them. Note that for subjects 12 and 17, the neck JVP signals appear as a mix of the expected venous JVP and the arterial PPG pulse. Indeed the latter part is so prominent that it seems to overlap with the finger PPG ground-truth signal. Due to this, it was not possible to accurately define the location of the v wave and, therefore, the v wave was not annotated. The onsets of the JVP pulses could not be defined and, as a consequence, no time differences calculations were considered in these two cases (12&17).

For the rest of the subjects, all the characteristic JVP waves were successfully annotated. To verify the consistency of this annotation across subjects, Fig. 4.10 (a) shows the normalized location of each a , c , v , R and S waves within the $O-O$ interval averaged over 5 JVP cycles for each subject. Standard deviation error bars were also specified. As it can be observed, overall for all subjects, the order of waves in the normalized $v-v$ interval is: v , a , R , c , S . This corresponds to the same cardiac cycle timings recognized in the literature, as presented in Fig. 4.1. This confirms the validity of the signals, but also shows that the JVP waveform is susceptible to change its morphology over time, as the standard deviations appear quite variable depending on the subject. Also for some subjects (1,16 and 18) it can be noticed that the systolic peak of finger PPG exceeds the normalized $O-O$ interval boundary, as the R peak occurs later in time for these subjects.

Figure 4.10(b) presents a summary of the order and time location distributions of each peak for all subjects, in the form of boxplots. These give an idea of the range in which each significant peak is expected to occur within the $v-v$ interval, between the lower and upper adjacent values. Based on these results, the v wave occurrence is in the range of [0.101-0.228] of the normalized $v-v$ interval of a JVP pulse; the a wave in [0.370-0.600], the R peak in [0.418-0.699], the c wave in [0.534-0.810], and the S peak in [0.815-1.074]. As it can be observed, all distributions show similar dispersion, apart from the v wave which is slightly smaller. Also, they all appear symmetric, except for the a wave which is slightly skewed.

4.3.5 Timings between JVP a , c , v waves and ECG and PPG peaks

The relative timings between the main (a, c, v) waves and the R and S peaks, were calculated to compare with the durations extracted by Lam Po Tang *et al.* [17] in the work quantifying the JVP from skin displacements. Even if that study only considered the $v-R$ and $R-c$ time differences, which correspond to the $R-O$ and $c-R$ respectively in this work, the following were also investigated: $S-R$, $S-c$, $c-v$, $c-a$, $a-v$, $R-v$, $R-a$. Average and standard deviations values of the normalized time differences are presented for each subject in Table 4.1. Figure 4.11 summarizes the information for all subjects by means of boxplots.

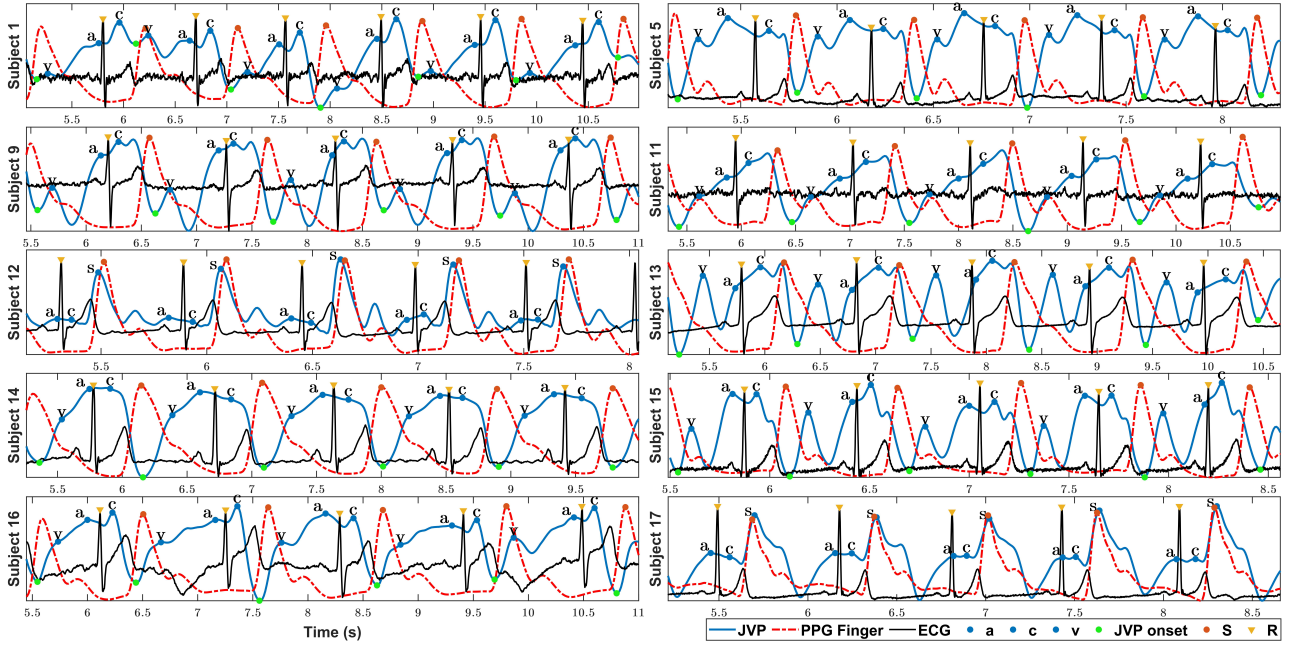


Figure 4.9: Manual *a*, *c*, *v* waves annotation of neck JVP signals with respect to ECG and PPG reference signals for 10 subjects. *S*: finger arterial PPG systolic peak. *R*: QRS complex peak of the ECG. *a*, *c*, *v*: characteristic JVP waves.

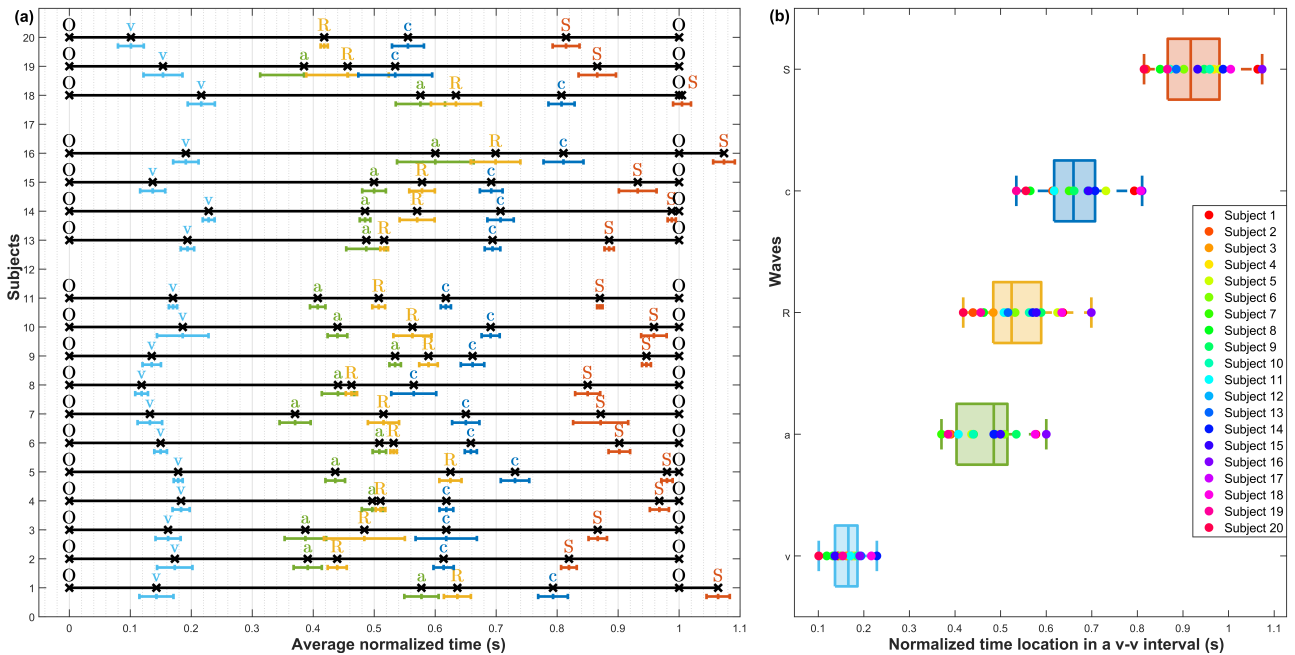


Figure 4.10: Location of the annotated *a*, *c*, *v*, *R* and *S* peaks within a JVP cycle. (a) Average normalized locations of the most characteristic JVP, ECG, PPG waves over a *v-v* cycle for each subject ($n=20$). (b) Boxplots summarizing the normalized time locations for each *v*, *a*, *c*, *R* and *S* waves for all participants ($n=20$).

As it can be observed in Figure 4.11, the largest time differences, normalized by the *v-v* interval duration, correspond to the *R-O* (JVP onset to EGC peak) and *c-v* waves pairs, with median values of 0.52 and 0.50, accordingly. The smallest time difference found is for the *R-a* pair, with a median

Table 4.1: AVERAGE TIME DIFFERENCES BETWEEN CHARACTERISTIC PEAKS AND WAVES FROM NECK JVP, FINGER PPG, AND ECG SIGNALS.

Mean differences (in ms) normalized by the v - v interval (1s) are presented for each subject ($n=20$) together with the standard deviation ($\mu \pm \sigma$). **S** represents the arterial finger PPG systolic peak, **R** is the QRS complex peak of the ECG and **a,c,v** correspond to the JVP waves. **O** indicates the onset of the JVP pulse.

Subj.	1	2	3	4	5	6	7	8	9
S-R	427.6 \pm 18.9	380.1 \pm 10.5	382.7 \pm 54.7	457.3 \pm 10.9	355.1 \pm 10.8	370.2 \pm 15.1	356.1 \pm 21.4	387.3 \pm 14.4	357.4 \pm 8.6
S-c	270.8 \pm 18.2	205.6 \pm 22	248.4 \pm 39.6	349.3 \pm 13	249.2 \pm 16	243.6 \pm 18.9	221.1 \pm 26	285.2 \pm 34.5	285.1 \pm 15.6
c-R	156.9 \pm 5.5	174.5 \pm 19.7	134.3 \pm 17.4	108 \pm 5.2	105.9 \pm 5.7	126.6 \pm 6.8	135 \pm 11.6	102.1 \pm 34	72.3 \pm 13.4
c-v	650.4 \pm 49.5	440.8 \pm 31	456.2 \pm 66.8	435.1 \pm 16.7	552.4 \pm 25.7	508.7 \pm 11.2	518.1 \pm 30.5	446.3 \pm 40.2	526.1 \pm 30.3
c-a	215.8 \pm 13	222.8 \pm 9	231 \pm 38.1	121.5 \pm 14.3	294.8 \pm 36.2	150 \pm 14.1	280 \pm 34.2	124.2 \pm 13.4	127 \pm 15.4
a-v	434.6 \pm 52.7	218 \pm 29.6	225.1 \pm 42.5	313.6 \pm 9.9	257.6 \pm 19.1	358.8 \pm 12.2	238.1 \pm 10.4	322.1 \pm 28.9	399.2 \pm 16.6
R-v	493.5 \pm 48.5	266.3 \pm 37.2	321.9 \pm 81.5	327 \pm 17.6	446.4 \pm 21.1	382.1 \pm 12.5	383.1 \pm 36.4	344.2 \pm 19.1	453.8 \pm 29.8
R-O	636.3 \pm 22.1	439.3 \pm 15.6	483.8 \pm 66.3	510.3 \pm 7.9	625 \pm 18.2	531.8 \pm 5.3	515.2 \pm 25.6	462.7 \pm 9.2	589 \pm 15.2
R-a	58.9 \pm 13.6	48.3 \pm 28.1	96.7 \pm 48.1	13.5 \pm 16.8	188.9 \pm 31.8	23.3 \pm 12.1	145 \pm 39.4	22.1 \pm 24.7	54.7 \pm 16.1

Subj.	10	11	13	14	15	16	18	19	20
S-R	396 \pm 32.6	362.6 \pm 10.9	368.9 \pm 7.9	417.5 \pm 27.1	353.7 \pm 14.9	374.7 \pm 37.3	370.7 \pm 43.4	409.5 \pm 44.5	396.6 \pm 17.3
S-c	267.8 \pm 8.9	252.6 \pm 6.8	191.2 \pm 9.9	280.7 \pm 18.3	240.2 \pm 19.1	263.3 \pm 28.1	197.6 \pm 9.1	331.3 \pm 42	259.3 \pm 21.1
c-R	128.2 \pm 27.7	110 \pm 7.8	177.7 \pm 15.4	136.8 \pm 18.1	113.5 \pm 4.7	111.4 \pm 12.2	173.2 \pm 42.5	78.2 \pm 14	137.4 \pm 24.3
c-V	504.9 \pm 34.3	447.4 \pm 13.5	500.3 \pm 23.6	478.7 \pm 15	555.2 \pm 38.8	619.3 \pm 42.9	590.8 \pm 43.2	380.9 \pm 58.4	454.3 \pm 30.6
c-a	251.2 \pm 23.3	210 \pm 12.2	207 \pm 42	222.2 \pm 27.1	192.1 \pm 37.3	210.2 \pm 36.6	231.4 \pm 44.1	149.6 \pm 19.8	-
a-v	253.7 \pm 55.3	237.4 \pm 11	293.3 \pm 28.3	256.5 \pm 16.2	363 \pm 5	409.1 \pm 74.8	359.4 \pm 49.1	231.3 \pm 63	-
R-v	376.7 \pm 24.2	337.4 \pm 14.9	322.6 \pm 11.4	341.9 \pm 21	441.7 \pm 41.5	507.9 \pm 51	417.6 \pm 51.4	302.7 \pm 59.9	316.9 \pm 23.5
R-O	562.7 \pm 31.3	507.4 \pm 10.5	516.3 \pm 6.4	570.4 \pm 28.7	578.4 \pm 21	699 \pm 40.7	634 \pm 40.8	456.3 \pm 67.7	418 \pm 5.9
R-a	123 \pm 46.3	100 \pm 13	29.3 \pm 36.2	85.4 \pm 30.9	78.6 \pm 39.9	98.8 \pm 26	58.2 \pm 8.3	71.4 \pm 8.6	-

value of 0.071, as ventricular depolarisation (R) directly follows atrial contraction (a) in the cardiac cycle time sequence. Similarly, c - R time variations also show a small median value of 0.13.

Focusing on the JVP waves differences, it can be noticed that in increasing order, c - a has the shortest distance (med=0.21), followed by a - v (med= 0.29), and c - v has the largest duration (med=0.50). This is consistent with the literature, which validates the manual annotation of these novel neck JVP-PPG signals and shows the potential they have to be clinically used for biomarkers extraction.

The distance between the PPG systolic peak and the ECG peak (S - R), i.e. the pulse transit time (PTT), showed a median normalized value of 0.38. This was larger than the distance between the same PPG peak and the JVP c wave (S - c) (med=0.25), which is in accordance with the cardiac cycle too, as c wave always occurs after the QRS complex of the ECG.

R - v and R - O refer to the same distance, but the R - v measure proposed in this work is calculated from the v peak to R peak, whereas R - O is the equivalent to v - R measured by Lam Po Tang *et al.* [17]. It was observed that they shared more or less similar ranges in the v - v normalized interval. In this work, the normalized c - R interval ranged from 72.3 \pm 13.4 (subject 9) to 177.7 \pm 15.4 (subject 13); whereas in their work it ranged from 139.6 \pm 6.1 to 248.3 \pm 10.5. For the normalized R - O , values ranged between 418 \pm 5.9 and 699 \pm 40.7; whilst their v - R interval spanned from 360.7 \pm 8.0 and 547.2 \pm 20.4.

Overall, the normalized time differences showed symmetric distributions except for the c - a waves which appeared slightly skewed. The dispersion of the data varied for the various pairs. For S - R , S - c and c - R , the dispersion was very small, with interquartile range (IQR) values of 0.034, 0.40 and 0.29

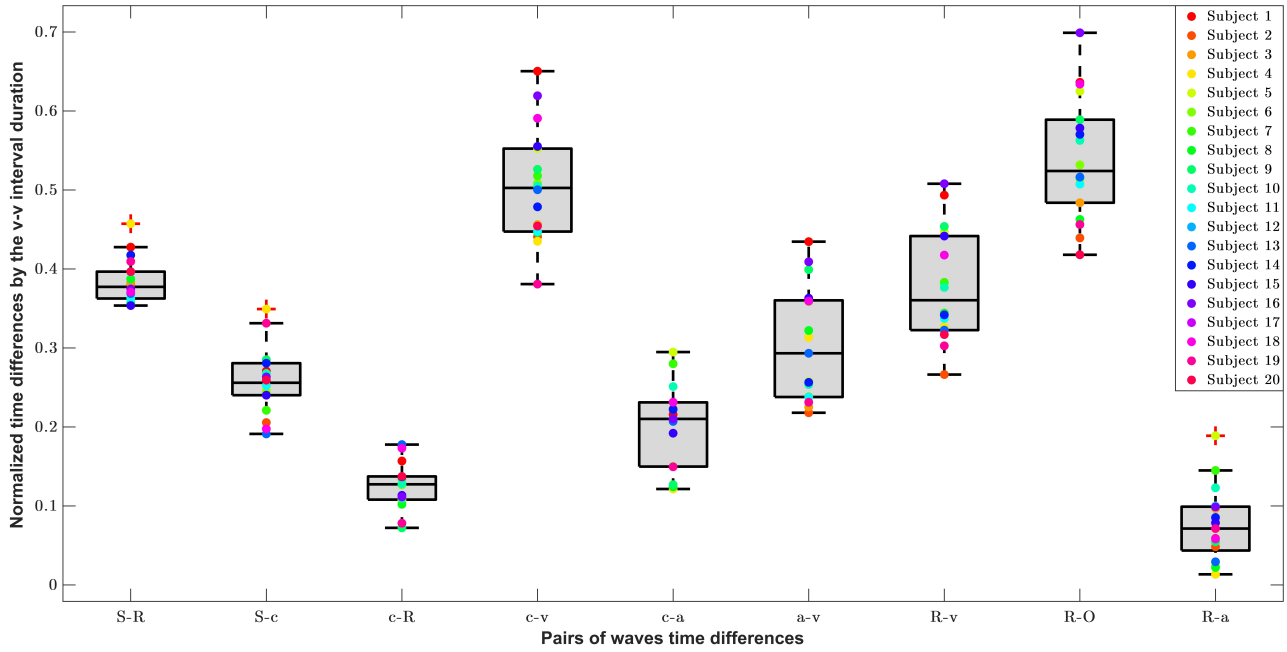


Figure 4.11: Boxplots showing the distributions of average time differences between characteristic peaks and waves from neck JVP, finger PPG and ECG ($n=20$). *S*: arterial finger PPG systolic peak. *R*: QRS complex peak of the ECG. *a,c,v*: characteristic JVP waves. *O*: onset of the JVP pulse.

respectively; for *R-a* (IQR=0.055) and *c-a* (IQR=0.081) was medium; and for the rest larger, with IQR around 0.1. The variance for each subject time difference also varied depending on the pair of waves. The average standard deviation for each *c-R* and *S-c* point was around 0.010; for *S-R*, *R-a*, *c-a*, *c-v*, was around 0.013 and for *a-v*, *r-v* and *R-O* of 0.02.

4.3.6 Average JVP waveforms for all subjects

Figure 4.12 shows the mean neck JVP waveforms extracted for the 20 participants. The pulse shapes in blue, resulted from the averaging of all the individual JVP pulses present in the whole length of the recording (60s), plotted in grey. Despite the fact that the JVP morphology is subject to change from one cardiac cycle to another (as in subjects 1, 10 or 13), in the majority of the cases the individual pulses appeared very similar to each other, facilitating the computation of an accurate average estimate. As a result, the characteristic *a*, *c*, *v* waves were marked, based on the manual annotation performed on each subject's recording with respect to reference ECG and PPG signals.

Different types of average JVP pulse shapes were visualized among different subjects. In all the cases, the *v*, *a* and *c* waves could be clearly identified visually, except for subject 20, for whom the *a* wave did not show enough prominence to be annotated. Indeed, in the majority of the signals (e.g. subjects 1, 6, 7, 8, 9, 10, 13, 15, 18), the *v* wave appears significantly isolated in time, with a prominent *y* descent, compared to the *a* and *c* greater amplitude waves. For half of the participants in the study (subjects 4, 5, 6, 7, 8, 11, 13, 14, 15, 20), a total of 4 distinctive waves were noticed. The

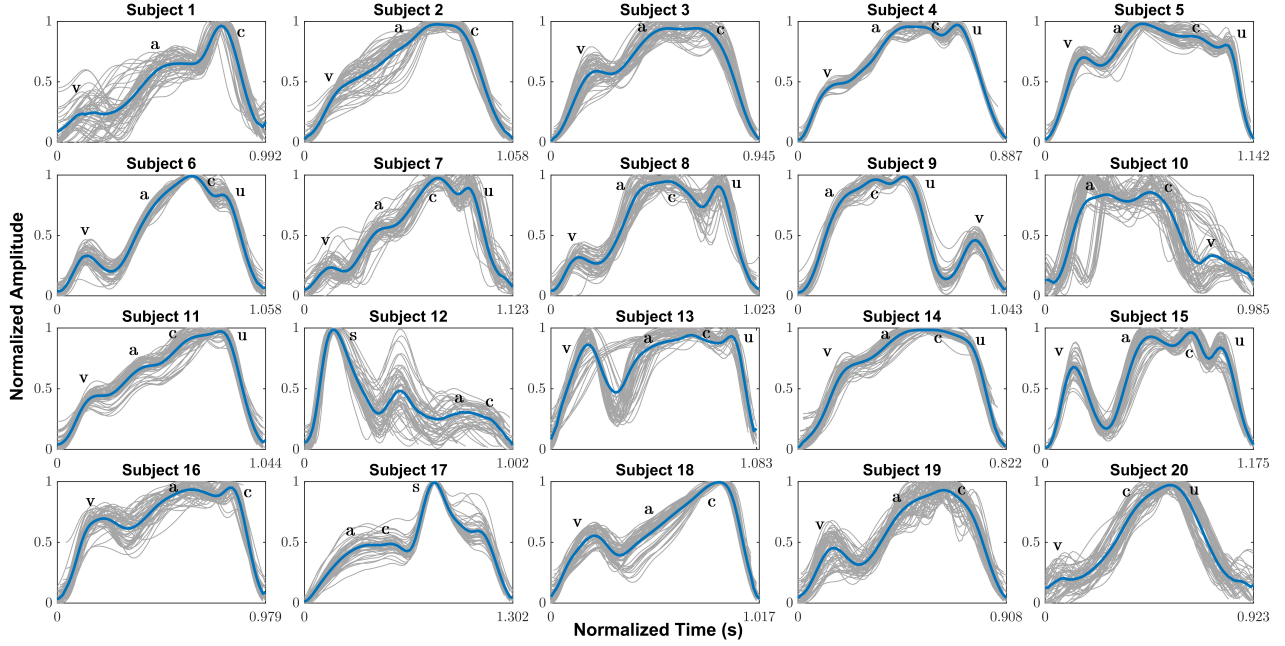


Figure 4.12: Average JVP waveforms manually annotated for all participants ($n=20$). *a, c, v*: characteristic JVP waves as described in the literature. *u*: unknown extra peak specific to the proposed method. *S*: arterial systolic peak overlapping with JVP waveform.

first three were marked as the JVP characteristic peaks and the last one was denoted as “*u*”, standing for unknown. For subjects 12 and 17, the average pulse shape reflects the superposition of the venous JVP *a* and *c* waves, and the arterial PPG pulse with the distinctive *S* peak.

Moreover, some average JVP waveforms show a more triangular shape (e.g. subjects 2, 3, 4, 11, 14, 16), with the *v* wave closer in time to the *a* and *c* waves, collapsing the *y* descent slope which appears very flat or non-existent. This happened independently of the presence of the 4th additional “*u*” wave.

4.4 Discussion

In this chapter, the feasibility of extracting the JVP waveform from the anterior neck, using a reflectance PPG sensor, was demonstrated. Neck JVP-PPG signals were recorded from a total of 20 participants in supine position. Since it is the first time that the JVP is sensed with a contact PPG sensor from the anterior neck, a series of validation metrics were presented to further support the hypothesis and conclusions of this work, as well as to assess the similarities and differences with other methods.

Although it was possible to obtain the JVP signal in the majority of the subjects, JVP-PPG signals were not equally accessible for all participants at once. For some, an exhaustive exploration of the anterior neck area with the PPG sensor was necessary to find the most suitable sensing location, while

in others it was straightforward. A possible reason is the differences in anatomy. Indeed, depending on the subject, veins are likely to be located less superficially in the tissue or occluded by more fat, which would avoid the IR light to penetrate very deep subcutaneously. This was confirmed with the IR camera system that showed how the vein vasculature of the anterior neck appeared visibly different from one subject to another. However, it is worth pointing out that a basic IR imaging low cost system was chosen for the proof of concept in this study as first approach, and might not have revealed the complete venous vasculature precisely.

In other cases, despite relocating the sensor and confirming with the camera that there was a vein underneath, it was still not obvious how to obtain the JVP signals. A clear example was subject 12, for whom vein imaging was one of the most clear ones among all subjects, but all the sensor locations explored resulted in recordings with a mix of the venous JVP and arterial PPG pulses. Indeed the PPG arterial mixed part of the neck JVP-PPG signal was almost perfectly matching in time with the reference PPG pulses by a small difference of 0.05s. A possible hypothesis is that the subject was very athletic, meaning that its stroke volume and cardiac output would be higher. The increase in the force of ventricular contraction allows an increased amount of blood to be ejected, resulting in an arterial pulse of higher amplitude at the neck, being sensed straight away following the *a* and *c* waves of the JVP pulse. Another possibility, however, might have been, that the light is passing through superposed veins and arterioles or at different skin depths, this causing the two signals to be sensed simultaneously, as a mix, in a delayed manner.

On another extreme, in some cases, such as for subject 19, it was not possible to identify the neck arterial PPG pulse anywhere in the neck. It was only possible to extract the JVP (even when exploring the suprasternal notch area that is typically optimal for neck arterial PPG recording, as previously demonstrated [21]). This could be due to the fact that the venous JVP pulse is stronger than the arterial one, even in regions that are supposed to be more irrigated by arterioles. Probably, a reduced cardiac output in this subject was the cause of an increase in CVP.

In this study it was demonstrated, for the first time, that novel neck JVP-PPG signals can be extracted from the anterior neck veins using a contact PPG reflective sensor. Ultrasound reference measurements of the jugular vein in the cross-sectional transverse plane, validated our proposed method.

The computation of the wavelet coherence corroborated the inversion property of the JVP waveform with respect to arterial PPG in concordance with negative correlation observed by Amelard *et al.* [15]. Indeed, this verifies that the proposed neck JVP signal really represents the atrial and ventricular contractions of the right heart (*a-c* waves) that typically occur before the arterial pulse (PPG) is transmitted towards the periphery (finger).

The three characteristic JVP waves *a*, *c*, *v* were identified in the majority of the signals, except

for subject 20 for whom the a wave was not recognized. The order of the peaks in the normalized v - v interval appeared to be v , a , R , c , S , consistent with cardiac physiology. This demonstrates that, the novel contact neck JVP-PPG signals recorded in this study, morphologically present all the conventional features of the well-known JVP and, have the potential to be exploited for CVDs clinical diagnosis. However, the location of each JVP wave was observed to be slightly variable. This suggests that the JVP waveform is predisposed to change its timing morphology from one cardiac cycle to another, as already suggested by Lam Po Tang *et al.* [17]. However, we additionally defined the normality ranges of the different waves within the JVP cycle for the proposed technique, to be able to reliably detect abnormalities in JVP a , c , v timings that appear outside of the established bounds.

When analysing the time differences between the JVP a , c , v waves and the ECG and PPG peaks, it was noticed how the dispersion of the boxplots varied from one pair of waves to another. The reduced dispersion in S - R , S - c , c - R distributions could be explained by the fact that ECG and PPG peaks, are well known and accurately detected, and they represent, together with the c wave, the same physiological process of systolic contraction. So, it could be expected that the time variations within the cardiac cycle between these dependent processes would not vary much among subjects in a healthy population sample. On the opposite, larger variability of time differences between other JVP waves pairs (c - v , a - v , R - v and R - O), could be explained by the dissimilarities in anterior veins vasculature and cardiac system anatomy, unique to each individual.

The comparison with previous work by Lam Po Tang *et al.* [17], of the c - R and R - O time intervals, showed comparable range of values but not exactly the same. The small inexactitude could be due to the smaller cohort of subjects (6 in total) that they used to calculate these differences, compared to the 18 participants we employed. But the close similarity, validates our neck contact JVP-PPG signals as a commensurable method for non-invasive JVP monitoring.

For the first time, average JVP waveforms were computed for each subject. A wider variety of JVP pulse morphologies were observed than in previous works measuring the JVP remotely at the external or internal jugular vein [15, 16, 17]. The different pulse shape classes could be a specific trait of the sensing location of the anterior neck, or simply different morphologies were not revealed in previous studies because averaging of the obtained JVP waveforms was not investigated.

The most distinctive feature of our contact neck JVP-PPG signals, is the presence of an extra 4th wave, of unknown origin, denoted as “ u ” and located after the c wave. The recurrent occurrence of this, in half of the cohort, led us to discard the idea that its origin was the result of some random artifact. After measuring some neck arterial PPG signals from the suprasternal notch and verifying the close superposition with the reference finger PPG, we abandoned the idea that it was the result

of the mixing of the venous JVP and arterial PPG pulses (as in subjects 12 & 17). We therefore hypothesize that the origin of this “*u*” wave is the delayed arrival of a pressure wave reflected through the intricate venous tree. Indeed, as discussed in the Methods (section 4.2.2), the anterior jugular veins can either connect directly to the subclavian vein or indirectly via the external jugular vein. The latter will introduce an extra bifurcation in the direct path between the RA and the measurement site at the anterior neck. Additionally, if we consider other bifurcations in the different veins (thyroid, anterior or internal jugular) that connect and intertwingle at the anterior neck, this will increase the likelihood of reflections sites. Therefore, singularities in different people’s anatomy and the confluence of different venous blood columns might explain the appearance of new waves in the neck contact JVP contour and the diversity in morphology.

These findings set for the first time the key attributes of the novel neck contact JVP-PPG signals and add extra value to the state of the art of JVP recording techniques. Our proposed method is an easy to use, low-cost solution for measuring JVP non-invasively from the anterior neck. It benefits from not requiring expensive equipment or infrastructure, as others solutions do. Unlike previous studies, that extract the JVP from the external and internal jugular veins, this work explores for the first time the original sensing location of the anterior jugular veins. Another advantage is that no expert personnel is required for JVP exploration due to the simplicity of moving the sensor around the anterior neck for signal acquisition. These findings will have a great impact in reducing the risks associated to central venous line catheterization, and therefore, are very promising for the future of CVDs clinical diagnosis. In spite of this, this technique suffers from some minor limitations. It is not always easy to locate the JVP in all subjects equally, due to the variability in veins anatomy, fat content, and pressure in the venous system. Therefore, future research should focus on tackling some of the open questions that remain unanswered. Is there an optimal sensor location and specific conditions to measure this neck contact JVP-PPG waveform in a more standardized manner? What is the origin of the additional unknown “*u*” wave? Is it relevant for JVP monitoring and diagnosis? In future studies, a deeper exploration of the unique intricate venous anatomy should help understand the differences in JVP waveforms morphologies. Ultimately, a clinical validation study, to assess the neck contact PPG modality against invasive catheterization should be carried out to evaluate the viability of implementing our alternative in continuous CVDs monitoring.

4.5 Conclusions

This work proposes, for the first time in literature, the use of reflectance contact PPG on the anterior neck, as a non-invasive, low-cost alternative to sense JVP and obtain physiological parameters of relevance for CVDs. Data acquisition from a total of 20 participants provided a snapshot of the pressure

variations occurring at the right atrium of the heart. This was validated by reference ultrasound measurements of the internal jugular vein. To demonstrate the hypothesis (i.e. that the JVP could be observed through PPG), as well as the feasibility of the new proposed method, the characteristic a , c , v waves of the neck JVP-PPG signals were manually annotated with the help of some reference ECG and finger arterial PPG signals. Calculation of time differences between significant features confirmed the validity of the novel JVP signals, their annotation, and the consistency with previous methods in the literature. Wavelet coherence values proved that neck JVP-PPG signals were also inversely correlated with respect to reference arterial finger PPG. In addition, the extracted average neck JVP waveforms highlighted some singularities of the presented technique. Despite the fact that all distinctive a , c , v waves could be identified, some pulse shapes showed a more triangular contour than typical, as a result of the reduced prominence of the y descent following the v wave. An additional “ u ” wave, of unknown origin, appeared unexpectedly right after the c wave in half of the JVPs in the cohort.

These findings are of great significance for the future design of low-cost, wearable PPG-based sensors to continuously monitor changes in central venous pressure. Patients could easily wear the sensor, with the only condition being having to comfortably lie down during measurements; as opposed to having to resort to either invasive and/or non-invasive but more restrictive and costly methods. Thus, this will aid in both, the efficiency in diagnosis of CVDs, as well as their management; whilst also eliminating some of the risks of invasive alternatives.

References

- [1] I. García-López and E. Rodríguez-Villegas, “Extracting the jugular venous pulse from anterior neck contact photoplethysmography,” *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.
- [2] World Health Organization (WHO), “Cardiovascular diseases (CVDs),” [http://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](http://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)), 2017.
- [3] E. Wilkins, L. Wilson, K. Wickramasinghe, P. Bhatnagar, J. Leal, R. Luengo-Fernandez, R. Burns, M. Rayner, and N. Townsend, “European cardiovascular disease statistics 2017,” European Heart Network, Tech. Rep., 2017.
- [4] E. J. Benjamin, S. S. Virani, C. W. Callaway, A. M. Chamberlain, A. R. Chang, S. Cheng, S. E. Chiuve, M. Cushman, F. N. Dellinger, R. Deo *et al.*, “Heart disease and stroke statistics—2018 update: a report from the American Heart Association,” *Circulation*, vol. 137, no. 12, pp. e67–e492, 2018.
- [5] S. R. McGee, “Inspection of the neck veins,” in *Evidence-based physical diagnosis*. Saunders, 2001, ch. 32.
- [6] N. Garg and N. Garg, “Jugular venous pulse: an appraisal,” *Journal, Indian Academy of Clinical Medicine*, vol. 1, no. 3, p. 261, 2000.
- [7] J. Constant, “Using internal jugular pulsations as a manometer for right atrial pressure measurements,” *Cardiology*, vol. 93, no. 1-2, pp. 26–30, 2000.

-
- [8] N. Ranganathan, V. Sivaciyan, and F. B. Saksena, “Jugular venous pulse,” in *The Art and Science of Cardiac Physical Examination*. Springer, 2006, pp. 67–111.
 - [9] R. N. Smith and J. P. Nolan, “Central venous catheters,” *Bmj*, vol. 347, p. f6570, 2013.
 - [10] S. Tibby and I. Murdoch, “Monitoring cardiac function in intensive care,” *Archives of disease in childhood*, vol. 88, no. 1, pp. 46–52, 2003.
 - [11] R. Mathews and D. L. Brown, “Invasive hemodynamic monitoring in the cardiac intensive care unit,” in *Cardiac intensive care*, 2nd ed. Saunders, 2010, ch. 45, pp. 558–569.
 - [12] M. M. Applefeld, “The jugular venous pressure and pulse contour,” in *Clinical methods: The history, physical, and laboratory examinations*, 3rd ed. Butterworth-Heinemann, 1990, ch. 19, p. 107.
 - [13] F. Sisini, M. Tessari, G. Gadda, G. Di Domenico, A. Taibi, E. Menegatti, M. Gambaccini, and P. Zamboni, “An ultrasonographic technique to assess the jugular venous pulse: a proof of concept,” *Ultrasound in medicine & biology*, vol. 41, no. 5, pp. 1334–1341, 2015.
 - [14] T. T. Dang, C. Huynh, A. Dinh, and K. Tran, “Recognizing area of pulsations on the neck via video camera systems,” in *International Conference on Advanced Technologies for Communications (ATC)*. IEEE, 2015, pp. 139–144.
 - [15] R. Amelard, R. L. Hughson, D. K. Greaves, K. J. Pfisterer, J. Leung, D. A. Clausi, and A. Wong, “Non-contact hemodynamic imaging reveals the jugular venous pulse waveform,” *Scientific reports*, vol. 7, p. 40150, 2017.
 - [16] A. Moço, P. Hamelmann, and G. Haan, “The importance of posture and skin-site selection on remote measurements of neck pulsations: An ultrasonographic study,” in *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 5918–5921.
 - [17] E. J. Lam Po Tang, A. HajiRassouliha, M. P. Nash, P. M. Nielsen, A. J. Taberner, and Y. O. Cakmak, “Non-contact quantification of jugular venous pulse waveforms from skin displacements,” *Scientific reports*, vol. 8, no. 1, p. 17236, 2018.
 - [18] R. Drake, A. W. Vogl, and A. W. Mitchell, *Gray’s anatomy for students*, 3rd ed. Elsevier Health Sciences, 2015.
 - [19] D. Dalip, J. Iwanaga, M. Loukas, R. J. Oskouian, and R. S. Tubbs, “Review of the variations of the superficial veins of the neck,” *Cureus*, vol. 10, no. 6, 2018.
 - [20] K. Chopra, D. Calva, M. Sosin, K. K. Tadisina, A. Banda, C. De La Cruz, M. R. Chaudhry, T. Legesse, C. B. Drachenberg, P. N. Manson *et al.*, “A comprehensive examination of topographic thickness of skin in the human face,” *Aesthetic surgery journal*, vol. 35, no. 8, pp. 1007–1013, 2015.
 - [21] I. García-López, S. A. Imtiaz, and E. Rodriguez-Villegas, “Characterization study of neck photoplethysmography,” in *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 4355–4358.

Chapter 5

Characterization of Artifact Signals in Neck Photoplethysmography

The research presented within this chapter is an edited version of research previously published in:

I. Garcia-Lopez and E. Rodriguez-Villegas, “Characterization of artifact signals in neck photoplethysmography,” IEEE Transactions on Biomedical Engineering, 2020

5.1 Introduction

Arterial photoplethysmography (PPG) signals are commonly measured in clinical settings, using pulse oximetry sensors, for the purpose of determining cardio-respiratory parameters. Peripheral saturation of oxygen (SpO₂%) and heart rate (HR) can be extracted in a reliable manner as soon as clean noise-free PPG signals are obtained. Despite the clinical value of pulse oximetry, the recording of arterial PPG signals is limited by the presence of noise artifacts components, that severally distort the quality.

Artifacts are understood as any source that corrupts a signal of interest. They modify the time and frequency characteristics of the signal with varying severity, and are prone to numerous sources of variability. They are the major cause of inaccurate vital signs measurements and undesirable false alarms in different clinical monitoring contexts [2]. Some artifacts are easy to filter out when they do not directly affect the PPG frequency bandwidth. Others, make it very difficult to recover the clean signal when their broad spectrum overlaps with the fundamental frequency of interest. Artifacts are measurement site specific, since different body locations are subjected to various types/severity of interference, and have dissimilar tissue optical properties and capillarity. The variations in vascular volume across the body imply that the AC and DC components are diversely distorted under the same artifact. Unique artifacts can affect particular body parts as, PPG sensors positioned along the arm

(including finger, wrist and upper arm) demonstrated to be mainly susceptible to arm movements whereas forehead sensors were not affected by these but by head movements instead ([3, 4]). Artifacts are also oximetry mode dependent. Reflectance PPG mode will favor more the uncoupling of the sensor with the skin than transmission PPG due to the weaker attachment, making the former more susceptible to ambient light interference [5]. LED wavelength also impacts the severity of artifacts. For example, green sources of light are less sensitive to noise compared to IR [3], since the optical path in reflectance PPG is shorter for green light. Consequently, fewer artifacts associated with DC components are likely to be sensed at this wavelength.

In standard finger pulse oximetry, PPG artifacts are well studied and controlled. The most critical ones are usually due to motion, and many signal processing techniques have already been proposed in the literature to filter them, as covered in section 2.2. However, the specificity of artifacts to the site of measurement and the type of sensor, do not make these solutions straightforwardly transferable to other body parts.

In order to fully exploit neck PPG for precise biomarkers monitoring, it is necessary to deal with the new and unknown sources of artifacts specific to this reflective site. In this context, this chapter aims to characterize for the first time the noise artifacts of arterial PPG signals sensed at the neck, in order to investigate which features would be most relevant to discriminate them from noise-free PPG signals. To our knowledge, an exhaustive study of arterial PPG artifacts has not been reported in the literature so far, and certainly not focusing on the neck. We believe this step to be crucial for the correct identification of the different sources of interference/noise that could corrupt this neck signal and prevent accurate physiological measurements to be recorded. The outcomes of this work will set the foundations necessary to develop filtering algorithms for the implementation of wearable neck arterial PPG sensors and, therefore, enhance the clinical utility of the recorded neck signals.

5.2 Methods

In this chapter, we propose a general method to characterize the artifacts of the neck PPG measurements, that have not been described before, with the aim of finding which are the most promising features to distinguish each of them from clean PPG signals. The approach consists in: first, identifying the specific artifacts of the new sensing location; second, extracting a set of features capable of differentiating these sources of noise from clean PPG signals; and third, evaluating the discrimination potential of these statistically, to find out the top 10 features that best describe each artifact. This technique could be equally applied to other measurement locations, or even other physiological signals, as long as the relevant artifacts and features are properly selected in each particular case scenario.

5.2.1 Neck artifacts definition

The effects of artifacts in neck PPG signals can vary from critically hindering the most important information to just mildly altering the shape of the signal. As already explained, in the case of PPG, artifacts are measurement site dependent. Therefore, it is important from the outset to have some understanding of which are the most likely ones to affect the PPG signal in the location of interest, i.e. in this case the neck. Following this, artifacts can be grouped into three different categories, depending on the type of source generating them [6]. These are further discussed in subsequent sections.

5.2.1.1 Intrinsic

Intrinsic artifacts are the ones originating from other physiological processes inside the body. In the context of some applications, these signals might not be considered artifacts, because these physiological processes superimposed to the PPG signals contain information which is useful to determine the physiological parameters of interest. However, within the context of this study, which is application agnostic, they will all be treated as artifacts. Examples of these include respiration, eye movements, temperature changes, cardiac signals or muscles contraction. In the particular case of neck PPG, the major intrinsic contributors to consider are: neck muscles tension, trachea sounds vibrations, autonomic system reflexes, tissue heterogeneity, skin temperature changes and fast respiration. The importance of the latter was demonstrated in section 3.4.1, in which it was explained how the fast respiratory fundamental frequency and its harmonics critically interfered with the standard arterial PPG power spectrum, overriding the PPG cardiac information of interest [7].

In terms of autonomic system reflexes, it was initially hypothesized that swallowing, coughing, yawning and snoring would entail some neck and jaw muscles contractions, as well as some vibrations or displacement of the sensor, and this could severely impact the signal quality. Similarly, trachea reverberations as a result of talking might also have the same negative effect.

Other physiological factors that, were thought, might less critically impair the recordings at the neck, were: low perfusion, skin temperature, skin colour and tissue heterogeneity. All these would account for more anatomical localized changes that could erroneously shrink the amplitude of the PPG waveform. For example, cold skin temperature would trigger vasoconstriction of the peripheral capillaries, reducing the perfusion of blood and making it more difficult for actual hemoglobin concentrations to be sensed. Moreover, when an excess of adipose tissue is present in the neck, the properties of light absorption would vary and would not appropriately reflect the real PPG signal amplitude. Thus, in severe cases of obesity in which the fat content of the neck is excessive, the amplitude of the PPG might be reduced. In the context of this work this had no effect on the extraction of the pulsatile signal, but it is worth considering that the number of subjects and the corresponding BMI range was

relatively small. In future larger studies it would be worth investigating this further.

5.2.1.2 Environmental

Optical measurement devices are not affected by electromagnetic, magnetic or capacitive coupling interference, as bioelectric signals (assuming that the interfacing electronics is properly designed). However, PPG recordings are sensitive to natural or artificial ambient light sources. External light interference can be minimized by enclosing the photodetector sensor in a dark isolated enclosure, and correctly placing it in close contact with the skin. If this precaution is taken when using these sensors, the effect of this type of artifact is minimum.

5.2.1.3 External

External artifacts are the ones that originate from external factors associated with real world daily-life activities. The most common ones are motion artifacts caused by subjects moving. These artifacts are unpredictable and can critically corrupt the signal quality. The consequences include changes in the coupling between the sensor and the skin. This results in a distortion of the path-length of the light transmitted through the tissue. Motion artifacts can also lead to an erroneous increase in the blood volume at the recording site. This is called the blood pooling effect. In the particular case of the neck, head and body movements, were expected to be the most corrupting artifacts in this category. In clinical applications, head movements can occur when flexing or rotating the neck and body movements when arms or legs twitch or the user switches posture. In addition, the application of excessive or insufficient contact pressure could also negatively impact blood perfusion and hence, the condition of the signal too.

A summary of the type, incidence frequency and severity of the most relevant artifacts that affect the neck are presented in Table 5.1. The colour gradient represents the importance of the artifact, from high (red) to low (yellow), based on some preliminary laboratory tests and the PPG literature [8]. The most frequent artifacts with the highest severity were estimated to have the highest priority for the neck and were separated, in the upper left corner of the table, through a diagonal border. A total of 10 artifacts, above this border in each category were selected for experimental data acquisition together with the reference PPG, denoted as normal breathing PPG.

Additionally, two other experimental situations that are not considered as artifact generating, since they do not lead to significant corruption of the signal, but are of important clinical interest for respiratory rate monitoring, were also taken into consideration in this study. These are slow breathing and breath-holding periods (simulating apnea events). Different breathing conditions modulate the DC baseline of the PPG signal with various frequencies and amplitudes. It is therefore interesting

Table 5.1: CLASSIFICATION OF NECK ARTIFACTS OCCURRENCE IN SLEEP CONDITIONS: BASED ON SEVERITY AND FREQUENCY OF INCIDENCE.

		<i>Always present</i>	<i>Very likely</i>	<i>Sporadic</i>	<i>Very rare</i>
Intrinsic	<i>Severe</i>	—	Swallowing Talking	Fast Breathing	Yawn Coughing
	<i>Medium</i>	—	—	Snoring	Low perfusion
	<i>Low</i>	Tissue heterogeneity Skin colour	—	Muscle contraction	Skin temperature
Environmental	<i>Severe</i>	—	—	—	—
	<i>Medium</i>	—	—	Ambient light	—
	<i>Low</i>	—	—	—	—
External	<i>Severe</i>	—	Body and Head movements	Movement of sensor probe	—
	<i>Medium</i>	—	—	Excessive/Insufficient sensor contact pressure	—
	<i>Low</i>	—	—	—	—

to study their characteristics together with the rest of the artifacts, to make sure these can be easily differentiated and not confused with any type of interference.

Snoring is also an artifact (or a signal of interest, depending of the application), worth considering in future research involving non-controlled experiments, ideally with patients with obstructive apnea.

5.2.2 Experiment 4 - data acquisition

5.2.2.1 Experimental set up

A total of 19 healthy subjects (12 males and 7 females), with average age of 25 ± 3 years old, height of 1.73 ± 0.09 m and weight of 69 ± 13.6 kg, participated in the experiments (average BMI of 23.02 ± 2.89 kg/m²). The study was approved by the Local Ethics Committee of Imperial College London (ICREC ref.: 18IC4358). Written consent was obtained from all of them. Signals from two PPG sensors, as shown in Fig. 3.1, were collected simultaneously while participants were lying down on a bed. In order to record neck PPG signals, a reflectance pulse oximeter sensor (8000R, Nonin) coupled to a processing module (Xpod, Nonin), was located at the suprasternal space. Also, to have a reference PPG signal, a transmission pulse oximeter (Onyx II 9560, Nonin), was placed at the index finger of the left hand. PPG signals were acquired from both sensors synchronously at a sampling rate of 75Hz.

5.2.2.2 Experimental protocol

13 recording sessions of 140s duration each were carried out per subject. In the first recording, participants were asked to be at rest, breathing at their normal pace. The 12 subsequent recordings aimed to test different artifacts (10) and respiratory states (2), with various intensities and duration,

as listed in Table 5.2. The selected artifacts belonged to either the intrinsic or external categories. Each recording was composed of three periods of artifacts, preceded by 20 seconds of rest. Artifacts could further be categorized depending on their duration as continuous, if they happened for long periods of time, or transient, if they were more time localized. As it can be observed on the table, continuous artifacts are represented with a straight red line, whereas transients are represented by a vertical red arrow. Since certain external artifacts can happen in both forms, in order to re-create more realistic conditions, the alternating artifacts in these cases were tested as follows: 1) the first was continuous and movements were fast, 2) the second, continuous and slow, 3) the third included only transient movements.

Table 5.2: EXPERIMENTAL PROTOCOL: LIST OF ARTIFACTS AND BREATHING STATES TESTED WITH CORRESPONDING RECORDING TIME SEQUENCES (IN SECONDS)

Artifact category	Recording	Experiment time-sequence
INTRINSIC	Breathing Normal	
	Breathing Apnea	
	Breathing Slow	
	Breathing Fast	
	Talking	
	Swallowing	
EXTERNAL	Yawning	
	Coughing	
	Head Up/Down	
	Head Right/Left	
	Head Rotation	
	Body Movements	
	Sensor Rubbing	

5.2.2.3 Recordings annotation

During the experiments, precise time-stamps, pointing the start and end of the artifacts, were marked in real-time and their corresponding type was assigned. After data acquisition, reference finger PPG signals were used to verify the manual segmentation and labelling of neck artifacts. This ensured that features from clean and noise-corrupted data were accurately extracted in posterior steps. The finger pulse oximeter was also essential to guarantee that all the neck artifacts recorded in this study were specific to the novel measurement site of the neck and did not affect the finger PPG signals.

5.2.3 Features extraction

In order to characterize neck PPG artifacts, a statistical comparison between normal, respiratory and artifact PPG segments was carried out. For this, recordings were normalized using the Z-score method, and a total of 41 characteristics from time, correlogram and frequency domains, were extracted. The

window for analysis was selected as the duration of the annotated (artifact, breathing or rest) interval. Features were first calculated individually for successive pulses in the window (time domain), or bins (frequency domain), and were subsequently averaged out over the whole window to output a single value per segment to evaluate statistically. ANOVA test served to indicate which were the most relevant features to significantly discriminate each artifact from normal noise-free PPG signals and hence describe each artifact specific traits. Here we present the list of features explored in this study. Please, refer to Table 5.3 for a detailed description of the calculations.

5.2.3.1 Time domain features extraction

Since time features mainly focus on beat-to-beat characteristics, individual PPG pulses ought to be segmented first. Based on the same approach as used in [7], the lower envelope of the signal was used to obtain the pulse onsets and then, the maximum between onsets identified the systolic peak. The time domain pulse characteristics used were as follows:

Amplitude [F_1], which corresponded to the height difference between the onset of a pulse and the systolic peak.

Width [F_2], which corresponded to the distance between the onset and offset of each pulse, i.e. its duration in time units (seconds).

Peak Height Difference [F_3], i.e. the absolute height difference between peaks of successive pulses in the signal.

Peak Distance [F_4], i.e. the time difference between peaks of successive pulses.

Trough Difference [F_5], i.e. the absolute height difference between onsets of successive pulses.

Rise Time [F_6], which corresponded to the time duration between the onset of the pulse and the systolic peak.

Skewness [F_7], i.e. the degree of symmetry of a pulse.

Kurtosis [F_8], i.e. the degree of sharpness of a pulse.

Change of F_{1-8} features [F_{9-16}], which represented the absolute difference between successive values of a calculated feature for consecutive pulses.

Standard Deviation of F_{1-8} features [F_{17-24}], which represented the standard deviation of feature values for a rest or artifact interval. As demonstrated in [4], the standard deviation value of time domain features captures the variability of successive poor-quality PPG pulses.

Zero-Crossing Rate [F_{25}], as shown in [9], it was expected that the number of times that the PPG signal crosses zero would increase when an artifact occurred, as the PPG signal would have a more chaotic behaviour than clean PPG.

Table 5.3: FEATURES DEFINITIONS AND FORMULAS

Amplitude [F_1]: is the height difference between the onset (O_1) and the systolic peak (S) of a pulse (i). $F_1(i) = S_i - O_{1i}$
Width [F_2]: is the distance between the onset (t_{O_1}) and offset (t_{O_2}) of each pulse (i), i.e. its duration in time units (seconds). $F_2(i) = t_{O_2} - t_{O_1}$
Peak Height Difference [F_3]: is the absolute height difference between peaks (S) of successive pulses (i) in the signal. $F_3(i) = S_{i+1} - S_i $
Peak Distance [F_4]: is the time difference between peaks (t_S) of successive pulses (i). $F_4(i) = t_{S_{i+1}} - t_{S_i}$
Trough Difference [F_5]: is the absolute height difference between onsets (O) of successive pulses (i). $F_5(i) = O_{i+1} - O_i $
Rise Time [F_6]: corresponds to the time duration between the onset ($t_{O_{1i}}$) of the pulse and the systolic peak (t_{S_i}). $F_6(i) = t_{S_i} - t_{O_{1i}}$
Skewness [F_7]: It was calculated as: $skewness = \frac{E(x - \mu)^3}{\sigma^3}$ (5.1) where, $E()$ is the expectation operator, μ is the mean and σ the standard deviation. The skewness of the normal distribution is zero. The more the skewness deviates from zero the less symmetric the pulse shape is. A negative skewness corresponds to a pulse with a longer left tail, whereas a positive skewness represents a longer right tail.
Kurtosis [F_8]: It was calculated as: $kurtosis = \frac{E(x - \mu)^4}{\sigma^4}$ (5.2) The kurtosis of the normal distribution is 3. If the kurtosis is smaller than 3, the pulse will be flatter and less sharp. If the kurtosis is greater than 3, the pulse will have a sharper pointed shape.
Change of F_{1-8} features [F_{9-16}]: The instantaneous change in feature i was calculated as: $\Delta(F_i) = F_i(j) - F_i(j - 1) $ (5.3) where, F_i corresponds to the i th feature [1-8] and j is the pulse index.
Standard Deviation of F_{1-8} features [F_{17-24}]: In a vector A containing N observations of any of the features above, for each of the pulses in an interval, the standard deviation corresponds to: $\sigma = \sqrt{\frac{1}{N - 1} \sum_{j=1}^N A_j - \mu ^2}$ (5.4)
Zero-Crossing Rate [F_{25}]: is the number of times per second that the PPG signal crosses zero.
Correlogram Peaks [F_{26-27}]: are the autocorrelation values of the first and second peaks of the correlogram.
Correlogram Lags [F_{28-29}]: are the lags corresponding to the first and second correlogram peaks.
Shannon Spectral Entropy (0-1.5Hz) [F_{30}]: It is given by: $H = - \sum_{s=1}^N P(s) \log_2 P(s)$ (5.5) Where $P(s)$ is the probability distribution of the power spectrum, given by: $P(s) = \frac{S(s)}{\sum_i S(i)}$ (5.6) With $S(s) = X(s) ^2$ being the power spectrum, where $X(s)$ is the discrete Fourier transform of the signal $x(t)$.
Spectral Kurtosis (0-1.5Hz) [F_{31}]: It is defined as the normalized fourth-order moment of the real part of the short-time Fourier transform, computed as: $K(f) = \frac{\langle X(t, f) ^4 \rangle}{\langle X(t, f) ^2 \rangle^2} - 2$ (5.7) where, $X(t, f)$ is the short-time Fourier transform (STFT) of the signal $x(t)$, and $\langle \cdot \rangle$ is the time average operator.
Relative Power [F_{32-36}]: The relative power was calculated by summing the power contained within the frequency bands (0-1Hz, 0.8-1.2Hz, 1-2Hz, 2-3Hz, 3-4Hz), and dividing it by the total power spanning all the frequencies.
Average Band Power [F_{37-41}]: A total of 5 features were extracted for each of the bands: 0-1Hz, 0.8-1.2Hz, 1-2Hz, 2-3Hz, 3-4Hz. They were calculated by averaging the power contained in the specified frequency bands.

5.2.3.2 Correlogram features

Since autocorrelation features extracted from the correlogram were the most highly ranked in discriminating between noise free and noise corrupted PPG in [9]; these were also explored in this work. Normal PPG signal correlograms were expected to exhibit high peaks equally separated, whereas artifacts correlograms were expected not to show such a periodic pattern, with the peaks having small correlation values. The rationale behind this is the fact that when convoluting a segment with itself at lags proportional to the fundamental frequency, signals are going to be almost superposed, hence resulting in high correlation values. However, for artifacts that are random and chaotic, the convoluted segment is less likely to be overlapped with itself at any lag, and this will result on little correlation. It is important to note that this will only apply to stationary scenarios (e.g. clinical applications) and not to ambulatory situations, in which repetitive motions could still result in high correlation even if introducing interference. Based on this, the following features were extracted:

Correlogram Peaks [F_{26-27}], these are the autocorrelation values of the first and second peaks of the correlogram.

Correlogram Lags [F_{28-29}], these are the lags corresponding to the first and second correlogram peaks.

5.2.3.3 Frequency domain features extraction

All the frequency domain features selected relied on the one-sided modified periodogram estimate of the power spectral density (PSD) of the PPG signal. To obtain it, the spectrogram of each recording was computed, based on the squared magnitude of the Short-Time Fourier Transform (STFT), with a window of 15s and 90% overlap. The output power, typically expressed in decibels logarithmic units (dB/Hz), with time in the x-axis and frequency across the y-axis, was segmented in time to have the separate PSDs corresponding to clean PPG and artifacts-corrupted intervals, as it is shown in Fig. 5.1. The whole subset of the twelve frequency features extracted are described as follows:

Shannon Spectral Entropy (0-1.5Hz) [F_{30}], or the degree of “disorder” of the spectral power distribution. The rationale behind calculating the entropy is that a clean PPG spectrum is expected to be highly ordered, as the frequency information is localized at specific frequencies, typically with the fundamental frequency at $\sim 1\text{Hz}$ and its harmonics at 2, 3, 4Hz etc. This will result in low spectral entropy values. On the contrary, chaotic artifacts will distort the frequency spectrum and will hence show high entropy values, since they carry less meaningful information. This can be actually observed in Fig. 5.1. The frequency band ranging from 0 to 1.5Hz was selected for the calculation, since it is where most of the power of the neck PPG signal is concentrated.

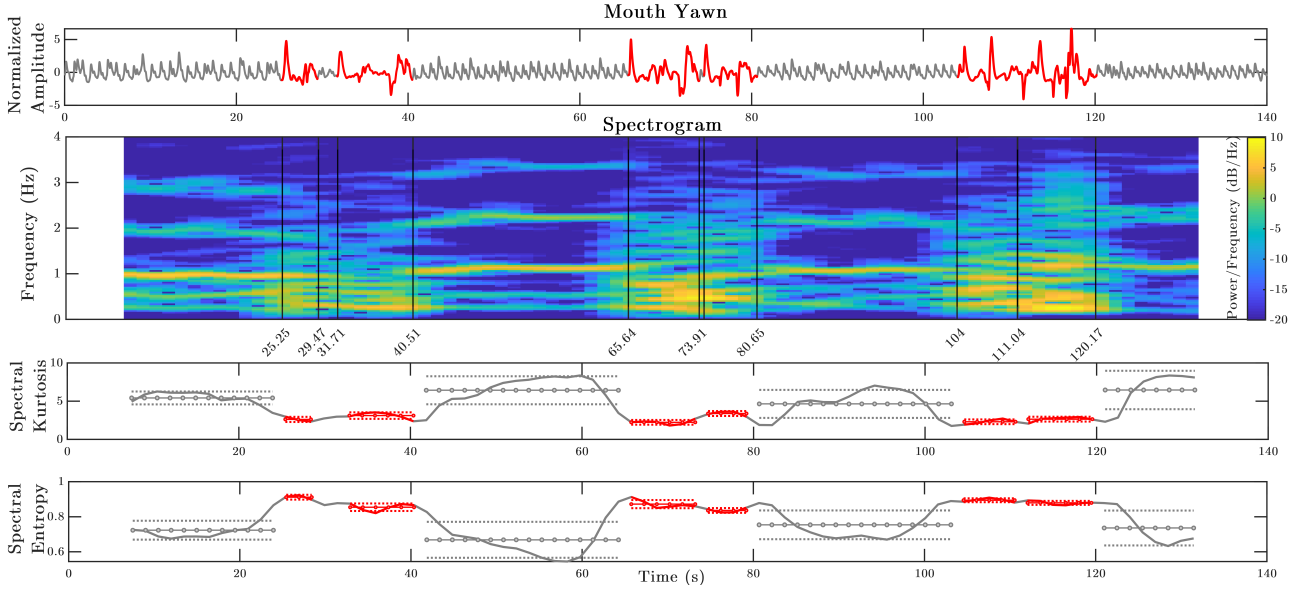


Figure 5.1: *Extraction of some frequency domain features from the spectrogram of a yawn recording.* Top panel shows the PPG recording corrupted by yawn artifacts. Second panel displays the computed spectrogram, from which the Spectral Kurtosis and Entropy features were calculated in the (0-1.5Hz) band.

Spectral Kurtosis (0-1.5Hz) $[F_{31}]$, is a measure of time dispersion of the time-frequency energy distribution, i.e. a measure of the peakedness of the PSD at each specific frequency f . Thus it can indicate the presence and the location of non-stationary transients in the frequency domain [10]. The reason to choose kurtosis as a feature was that in artifacts that are not stationary, the spectrum will be very noisy, broad, and with no distinguishable peaks. Therefore, the spectral kurtosis was expected to be smaller than for clean PPG or stationary breathing artifacts, as shown in Fig. 5.1.

Relative Power $[F_{32-36}]$: A total of 5 features were extracted for each of the following bands: 0-1Hz, 0.8-1.2Hz, 1-2Hz, 2-3Hz, 3-4Hz. The relative power was calculated by summing the power contained within the specified frequency bands, and dividing it by the total power spanning all the frequencies. The 0-1Hz band was chosen to capture power of the respiration frequencies. The 1Hz-centered band, ranging from 0.8Hz-1.2Hz, was selected because it corresponds to the location of the fundamental frequency of the PPG signal. The rest of the bands (1-2Hz, 2-3Hz, 3-4Hz), were chosen to inspect whether there was any abnormality in power density due to artifacts, in between the harmonics of the fundamental PPG frequency (2Hz, 3Hz, 4Hz etc.).

Average Band Power $[F_{37-41}]$: A total of 5 features were extracted for each of the bands: 0-1Hz, 0.8-1.2Hz, 1-2Hz, 2-3Hz, 3-4Hz. They were calculated by averaging the power contained in the specified frequency bands.

Fig. 5.1 shows an example of a recording with the segmented spectrogram and the variation in time for some of the frequency features presented. The features corresponding to normal PPG are displayed in grey, whereas the artifact ones are displayed in red. The dotted straight lines in each

segment show the average value of the corresponding feature for the coloured time period. These correspond to single observations, either for rest or artifacts, that will be input into the statistical tests for evaluation of significance, as explained in the next section. The standard deviation is also annotated with two parallel lines above and below the mean.

5.2.4 Artifacts vs. Normal PPG statistical comparison

5.2.4.1 Statistical tests evaluation

In order to test the statistical significance between pairs of clean-artifacts features, the statistical test selected was a one-way-ANOVA with repeated measures. Standard ANOVA could not be used instead because all features were measured under 13 conditions (1 normal case, 2 respiratory and 10 artifacts), and hence the assumption of independence was not valid. The selected one-way-ANOVA test is an extension of the paired-samples t-test but taking into account within subjects effects. The assumptions made when applying this test are normality of the residuals and sphericity (i.e. variances of the differences between each pair of conditions have to be the same). Probability plots and Lilliefors normality tests were used to assess the normality of the distributions. However, in the context of this work, the second assumption was not always met, so a Greenhouse-Geisser correction was applied to the degrees of freedom. The interpretation of the test is: if the main ANOVA is significant, then there is a difference between at least two conditions. Performing post-hoc multiple pairwise comparisons tests with Tuckey-Kramer correction helped to find out exactly which of those normal-artifacts paired differences were significant.

5.2.4.2 Characterization significance matrix construction

For each feature, p-values for every normal-artifact multiple comparison were obtained. To illustrate, the matrix on the left in Fig. 5.2 shows the results for the amplitude feature $[F_1]$ for all normal and other conditions combinations. Only statistically significant pairs are in color. This corresponds to p-values smaller than the level of significance $\alpha = 0.05$. The smaller the p-value the higher the significance.

The first row of the significance matrix, corresponding to the normal vs. all conditions p-values for the i th feature, was reorganized into the i th column of the characterization significance matrix. This was repeated for all features, as shown on the right panel of Fig. 5.2. The matrix obtained had 41 columns, corresponding to the number of features tested, and 12 rows, accounting for the 10 artifacts and the two additional respiratory states (breathing slow and apnea) against which normal PPG was compared.

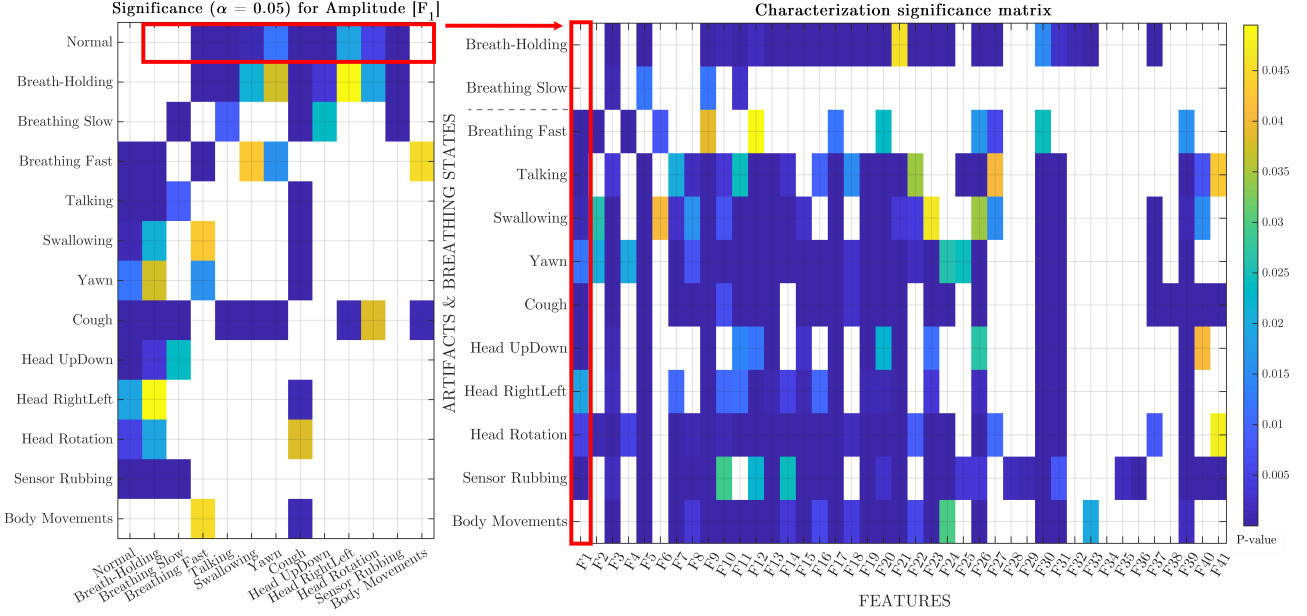


Figure 5.2: Results of pairwise multiple comparisons performed after repeated measures one-way-ANOVA statistical test. The left matrix shows the p-values for all pairs for the Amplitude feature [F_1]. The right matrix displays the statistical pairwise comparisons between Normal vs. Artifacts for all features ($\alpha = 0.05$). F_1 = Amplitude, F_2 = Width, F_3 = PeakHeightDiff, F_4 = PeakDistance, F_5 = TroughDiff, F_6 = RiseTime, F_7 = Skewness, F_8 = Kurtosis, F_9 = ChangeAmplitude, F_{10} = ChangeWidth, F_{11} = ChangePeakHeightDiff, F_{12} = ChangePeakDistance, F_{13} = ChangeTroughDiff, F_{14} = ChangeRiseTime, F_{15} = ChangeSkewness, F_{16} = ChangeKurtosis, F_{17} = StdAmplitude, F_{18} = StdWidth, F_{19} = StdPeakHeightDiff, F_{20} = StdPeakDistance, F_{21} = StdTroughDiff, F_{22} = StdRiseTime, F_{23} = StdSkewness, F_{24} = StdKurtosis, F_{25} = ZeroCrossingRate (ZCR), F_{26} = CorrelogramPeak1, F_{27} = CorrelogramPeak2, F_{28} = CorrelogramLag1, F_{29} = CorrelogramLag2, F_{30} = SpectralEntropy (0-1.5Hz), F_{31} = SpectralKurtosis (0-1.5Hz), F_{32} = RelativePower (0-1Hz), F_{33} = RelativePower (0.8-1.2Hz), F_{34} = RelativePower (1-2Hz), F_{35} = RelativePower (2-3Hz), F_{36} = RelativePower (3-4Hz), F_{37} = AvgPower (0-1Hz), F_{38} = AvgPower (0.8-1.2Hz), F_{39} = AvgPower (1-2Hz), F_{40} = AvgPower (2-3Hz), F_{41} = AvgPower (3-4Hz).

In the characterization significance matrix, it is possible to see differences in the discrimination capability between features. In fact, the *Change in Amplitude* feature [F_9] has the potential to discriminate normal clean PPG from any condition, whereas the *Relative Power (1-2Hz)* feature [F_{34}] is not significant for any normal-artifact pair. Another interesting thing to notice is that each feature will have a higher or lower p-value (i.e. different discrimination potential), depending on the source of variability. For instance, the *Standard Deviation Skewness* [F_{23}], is not significant for Breathing Fast, Slow and Talking, and neither has a very large p-value (≈ 0.045) for the Swallowing artifact, meaning little discriminatory ability. But, it shows very small p-values (≈ 0.01) for the rest of the artifacts, meaning it would be a useful feature to differentiate these from normal PPG. This demonstrates the need of ranking the features for each category independently, in order to determine what are the most important traits that differentiate each artifact or breathing state from normal PPG.

5.2.4.3 Ranking of features for each artifact

The p-values of the characterization significance matrix were sorted in ascending order for each artifact independently. Hence, the first ranked feature corresponded to the one with the lowest p-value, and

thus, with the greatest discriminatory potential with respect to normal PPG. However, in order to keep the most representative features describing each category, a feature selection technique was proposed and explained in the next section.

5.2.4.4 mRMR feature selection

For each category, the best 10 features were selected, using the minimum Redundancy Maximum Relevance (mRMR) approach. But instead of applying this procedure as usual based on the mutual information [11], the ranked statistical tests p-values, were exploited for the maximum relevance criterion, and the correlation values, for the minimum redundancy criterion. The idea was to find a subset of features that maximized the discriminability between normal PPG and each artifact or respiratory condition, and reduced the correlation among selected features at the same time. This was accomplished through an iterative process in which one new feature was added at a time to the final ranking. For that, the value of the operator $\Omega_i(P, R)$ was calculated for each of the remaining features, and the feature that maximized it was added, so that:

$$\max[\Omega_i(P, R)], \quad \Omega_i = P_{ij}/R_i \quad (5.8)$$

where, P_{ij} corresponds to the p-value of the i th feature differentiating between normal PPG and the j th artifact; and R_i refers to average correlation value between the i th feature and all the features already selected in the final ranking.

5.3 Results

5.3.1 Features ranking based on statistical results

Before selecting the top 10 more relevant and less redundant characteristics for each artifact independently, features were ranked from most to least significant based on the p-values obtained from the characterization significance matrix. The results are shown in the heatmap in Fig. 5.3. Note in grey the features that were not significant (p-value>0.05).

As it can be observed, there were a few features that concentrated the most important rankings (in dark blue) for most of the artifacts. These were: *Peak Height Difference* [F_3], *Trough Difference* [F_5], *Change in Amplitude* [F_9], *Standard deviation of Amplitude* [F_{17}], *Spectral Entropy* [F_{30}], *Spectral Kurtosis* [F_{31}], and *Average Power (1-2Hz)* [F_{39}]. This suggests that these characteristics are the ones with higher capability to differentiate between normal neck PPG and artifacts. On the opposite, some of the extracted features, did not show overall statistical significance for almost any artifact. As seen in

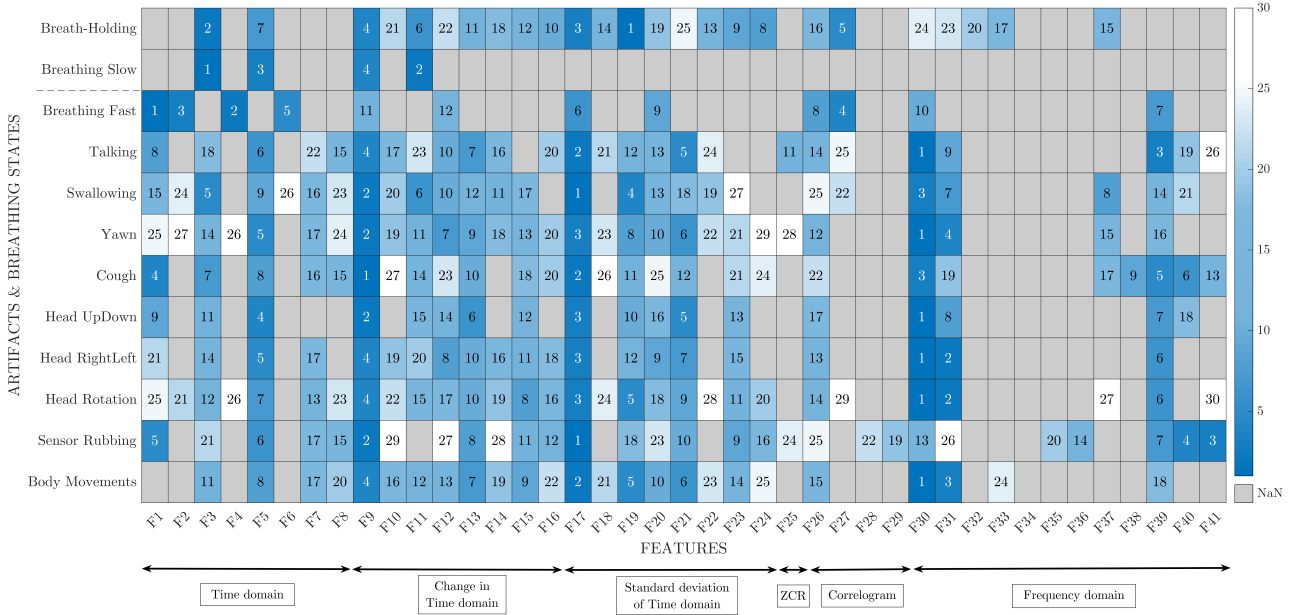


Figure 5.3: Ranking of features for each independent artifact and breathing state to discriminate against normal PPG, based on the p-values obtained from the characterization matrix. The colour scale indicates the ranking numbers ordering. The darkest blue colours represent the most discriminative features with the lowest ranking values (i.e. the lowest p-values) and the lightest colours display the least significant features with the largest ranking numbers (i.e. higher p-values). In grey, NaN values indicate the features that did not show significance ($p > 0.05$). F_1 = Amplitude, F_2 = Width, F_3 = PeakHeight-Diff, F_4 = PeakDistance, F_5 = TroughDiff, F_6 = RiseTime, F_7 = Skewness, F_8 = Kurtosis, F_9 = ChangeAmplitude, F_{10} = ChangeWidth, F_{11} = ChangePeakHeightDiff, F_{12} = ChangePeakDistance, F_{13} = ChangeTroughDiff, F_{14} = ChangeRiseTime, F_{15} = ChangeSkewness, F_{16} = ChangeKurtosis, F_{17} = StdAmplitude, F_{18} = StdWidth, F_{19} = StdPeakHeightDiff, F_{20} = StdPeakDistance, F_{21} = StdTroughDiff, F_{22} = StdRiseTime, F_{23} = StdSkewness, F_{24} = StdKurtosis, F_{25} = ZeroCrossingRate (ZCR), F_{26} = CorrelogramPeak1, F_{27} = CorrelogramPeak2, F_{28} = CorrelogramLag1, F_{29} = CorrelogramLag2, F_{30} = SpectralEntropy (0-1.5Hz), F_{31} = SpectralKurtosis (0-1.5Hz), F_{32} = RelativePower (0-1Hz), F_{33} = RelativePower (0.8-1.2Hz), F_{34} = RelativePower (1-2Hz), F_{35} = RelativePower (2-3Hz), F_{36} = RelativePower (3-4Hz), F_{37} = AvgPower (0-1Hz), F_{38} = AvgPower (0.8-1.2Hz), F_{39} = AvgPower (1-2Hz), F_{40} = AvgPower (2-3Hz), F_{41} = AvgPower (3-4Hz).

the concentrated grey areas, some examples are *RelativePower* for all bands [F_{32-36}], *AvgPower* (0.8-1.2Hz) [F_{38}], *RiseTime* [F_6] or *Correlogram Lags1&2* [F_{28-29}]. It is also worth pointing out that there were only four statistically significant features capable of differentiating Breathing Slow from normal PPG. Therefore, these small differences imply that this respiratory state is not very dissimilar from normal PPG, and therefore it will be similarly differentiable from the rest of the artifacts too.

5.3.2 Artifacts characterization

In order to provide a summary of the main features characterizing each artifact and respiratory state, the 10 most highly ranked features selected with the mRMR approach are shown in Table 5.4. The first column in the table indicates the duration type: continuous or transient. The second column graphically shows a recording example of each artifact or breathing state. The third column lists the top ranked features resulting from the mRMR selection. The last column corresponds to the difference of each feature distribution mean, i.e. $Sign = \mu_{artifact} - \mu_{normal}$. On the basis of this, it

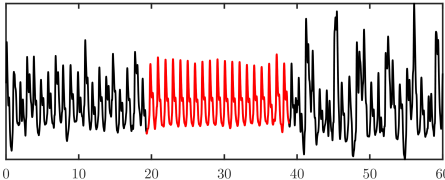
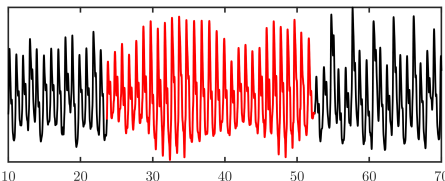
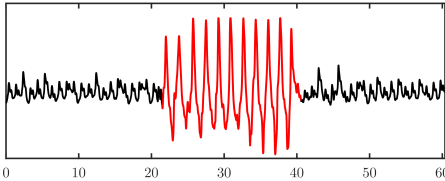
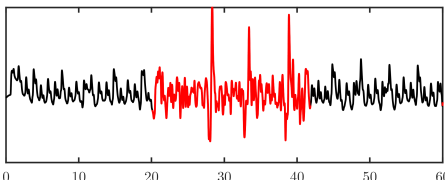
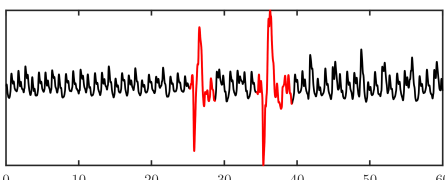
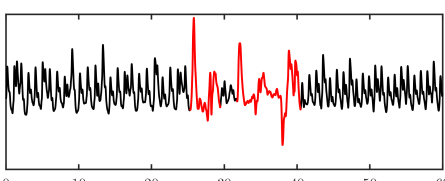
can be assessed whether the listed features for each artifact or respiratory category are greater ($>$) or smaller ($<$) than normal PPG.

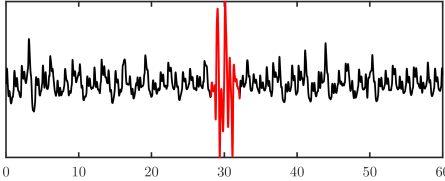
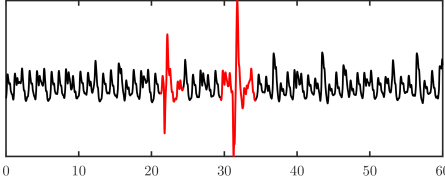
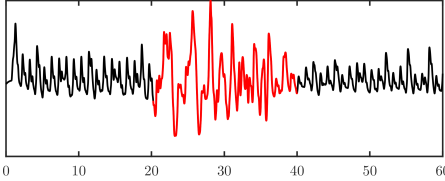
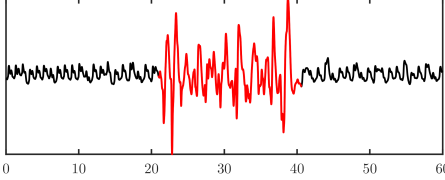
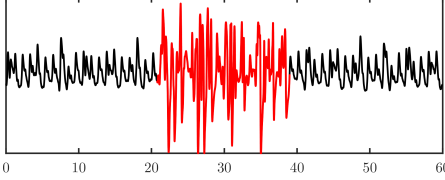
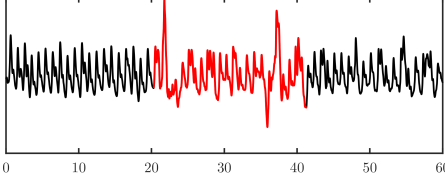
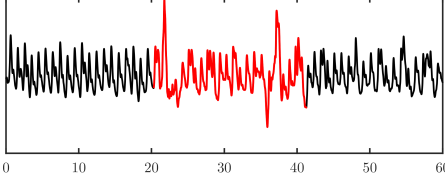
For **Breathing Fast**, the PPG signal is greatly distorted. The *Amplitude* and *StdAmplitude* are considerably larger ($>$) than normal, as movements of the thoracic cage displace the sensor in a larger range and with higher frequency. Observed fluctuations do no longer correspond to the cardiac pulse, being actually more separated between them ($> PeakDistance, StdPeakDistance$), and show greater *Width* and *RiseTime*. However, the fast respiration frequency is periodic and consistent, this resulting on the *CorrelogramPeaks1&2* features having high correlation values. In terms of frequency features, the *SpectralEntropy (0-1.5Hz)* is smaller ($<$) than in the normal case. This is due to the fact that the peak of the fast respiratory frequency at $\sim 0.5\text{Hz}$ concentrates around $\sim 60\%$ of the total power, and so the degree of disorder is smaller than for normal PPG, where the power is split in between the two respiration peaks ($\sim 0.3\text{Hz}$ and $\sim 0.6\text{Hz}$) and the fundamental PPG one ($\sim 1\text{Hz}$). The *AvgPower (1-2Hz)* is greater ($>$) than for normal PPG, since, in the power spectrum, the second harmonic of the breathing fundamental respiratory frequency is located at $\sim 1.2\text{Hz}$ and its relative power is around 30% .

Swallowing, **Yawn** and **Cough** share many characteristics related to their short instantaneous duration. The three of them exhibit a larger *SpectralEntropy (0-1.5Hz)*, smaller *SpectralKurtosis (0-1.5Hz)* and larger *AvgPower (1-2Hz)* than normal PPG. This suggests that these artifacts are non-stationary and that their power spectrum is not centered at any frequency in particular, but broadly overlaps with the whole PPG bandwidth of interest with high disorder. The corrupted frequency spectrum will have abnormally high power content in between PPG principal frequency peaks. In addition, the **Cough** artifact is characterized by containing greater *AvgPower* than the reference PPG spectrum for all the frequency bands tested: $0-1\text{Hz}$, $0.8-1.2\text{Hz}$, $1-2\text{Hz}$, $2-3\text{Hz}$, $3-4\text{Hz}$, since it is of higher frequency. **Cough** artifacts also give rise to sharper, non-symmetric beats of high amplitude variations as demonstrated by the features *Kurtosis*, *Skewness* and *ChangeAmplitude*. The most distinctive trait of **Swallowing** transient artifacts is their sudden drop and slightly delayed increase in amplitude, that is captured by greater *StdAmplitude*, *StdPeakHeightDiff*, *ChangePeakDistance*, *ChangeAmplitude*, *PeakHeightDiff* and *ChangePeakHeightDiff* features. **Yawn** artifacts are characterized by having an unsteady wider separation between peaks than normal PPG ($> ChangePeakDistance, StdPeakDistance$).

Motion artifacts including **Head Up/Down**, **Right/Left**, **Rotation** and **Body**, also have a big predominance of frequency characteristics; specially a larger *SpectralEntropy (0-1.5Hz)*, smaller *SpectralKurtosis (0-1.5Hz)* and larger *AvgPower (1-2Hz)*. This demonstrates the high degree of disorder of the power spectrum, the poor frequency peaks localization and the larger concentration in power

Table 5.4: CHARACTERIZATION OF ARTIFACTS AND BREATHING STATES OF INTEREST

BREATHING STATES		Ranked Features	Sign
Continuous	Breath-Holding (Apnea)		<ol style="list-style-type: none"> 1. StdPeakHeightDiff < 2. StdKurtosis < 3. CorrelogramPeak2 > 4. ChangeAmplitude < 5. StdWidth < 6. StdAmplitude < 7. PeakHeightDiff < 8. StdRiseTime < 9. ChangePeakHeightDiff < 10. CorrelogramPeak1 >
Continuous	Breathing Slow		<ol style="list-style-type: none"> 1. PeakHeightDiff < 2. ChangePeakHeightDiff < 3. ChangeAmplitude < 4. TroughDiff <
ARTIFACTS		Ranked Features	Sign
Continuous	Breathing Fast		<ol style="list-style-type: none"> 1. Amplitude > 2. StdPeakDistance > 3. PeakDistance > 4. Width > 5. AvgPower (1-2Hz) > 6. CorrelogramPeak2 > 7. RiseTime > 8. CorrelogramPeak1 > 9. StdAmplitude > 10. SpectralEntropy (0-1.5Hz) <
Continuous	Talking		<ol style="list-style-type: none"> 1. SpectralEntropy (0-1.5Hz) > 2. ZeroCrossingRate > 3. StdTroughDiff > 4. AvgPower (1-2Hz) > 5. ChangeAmplitude > 6. CorrelogramPeak1 < 7. StdAmplitude > 8. TroughDiff > 9. SpectralKurtosis (0-1.5Hz) < 10. Amplitude >
Transient	Swallowing		<ol style="list-style-type: none"> 1. StdAmplitude > 2. SpectralEntropy (0-1.5Hz) > 3. StdPeakHeightDiff > 4. ChangePeakDistance > 5. ChangeAmplitude > 6. PeakHeightDiff > 7. SpectralKurtosis (0-1.5Hz) < 8. ChangePeakHeightDiff > 9. AvgPower (1-2Hz) > 10. Skewness <
Transient	Yawn		<ol style="list-style-type: none"> 1. SpectralEntropy (0-1.5Hz) > 2. TroughDiff > 3. ChangePeakDistance > 4. StdPeakHeightDiff > 5. SpectralKurtosis (0-1.5Hz) < 6. StdTroughDiff > 7. CorrelogramPeak1 < 8. AvgPower (1-2Hz) > 9. StdPeakDistance > 10. ChangeAmplitude >

ARTIFACTS		Ranked Features	Sign
<i>Transient</i>	Cough		1. ChangeAmplitude >
			2. AvgPower (2-3Hz) >
<i>Cont. & Trans.</i>	Head Up/Down		3. Skewness <
			4. CorrelogramPeak1 <
			5. AvgPower (0.8-1.2Hz) >
			6. SpectralEntropy (0-1.5Hz) >
			7. Kurtosis >
			8. AvgPower (1-2Hz) >
			9. AvgPower (3-4Hz) >
			10. SpectralKurtosis (0-1.5Hz) <
			1. SpectralEntropy (0-1.5Hz) >
			2. AvgPower (2-3Hz) >
<i>Cont. & Trans.</i>	Head Right/Left		3. TroughDiff >
			4. StdAmplitude >
			5. SpectralKurtosis (0-1.5Hz) <
			6. AvgPower (1-2Hz) >
			7. ChangeAmplitude >
			8. StdTroughDiff >
			9. ChangeTroughDiff >
			10. Amplitude >
			1. SpectralEntropy (0-1.5Hz) >
			2. AvgPower (1-2Hz) >
<i>Cont. & Trans.</i>	Head Rotation		3. SpectralKurtosis (0-1.5Hz) <
			4. StdAmplitude >
			5. ChangePeakDistance >
			6. ChangeAmplitude >
			7. CorrelogramPeak1 <
			8. TroughDiff >
			9. StdPeakDistance >
			10. ChangeSkewness >
			1. SpectralEntropy (0-1.5Hz) >
			2. ChangePeakDistance >
<i>Cont. & Trans.</i>	Sensor Rubbing		3. SpectralKurtosis (0-1.5Hz) <
			4. StdAmplitude >
			5. AvgPower (1-2Hz) >
			6. CorrelogramPeak1 <
			7. Skewness <
			8. Width >
			9. ChangeSkewness >
			10. StdPeakHeightDiff >
			1. StdAmplitude >
			2. ZeroCrossingRate >
<i>Cont. & Trans.</i>	Body Movements		3. StdSkewness >
			4. Amplitude >
			5. StdPeakDistance >
			6. ChangeAmplitude >
			7. CorrelogramLag2 <
			8. TroughDiff >
			9. ChangeSkewness >
			10. ChangeTroughDiff >
			1. SpectralEntropy (0-1.5Hz) >
			2. ChangeAmplitude >
<i>Cont. & Trans.</i>	Body Movements		3. SpectralKurtosis (0-1.5Hz) <
			4. StdAmplitude >
			5. Skewness <
			6. StdTroughDiff >
			7. StdPeakHeightDiff >
			8. CorrelogramPeak1 <
			9. StdPeakDistance >
			10. ChangeTroughDiff >
			1. SpectralEntropy (0-1.5Hz) >
			2. ChangeAmplitude >

in between normal peaks of the PPG spectrum. **Head Up/Down** and **Body movements** mainly introduce changes in “vertical” axis features such as higher *TroughDiff*, *StdAmplitude*, *ChangeAmplitude*, *StdTroughDiff*, *ChangeTroughDiff* and *Amplitude*. **Head Right/Left** and **Rotation** are characterized by a mix of “vertical” and “horizontal” traits, demonstrating their irregular and random behaviour in amplitude and time. They also show lower and almost null correlation values for the *CorrelogramPeak1*, meaning that these corrupted signals are very variable and not replicable in time. This is reinforced by the larger *ChangeSkewness* that suggests a higher alteration rate of successive beats symmetry, as the signal loses its standard shape. Similarly, for **Body Movements**, the *Skewness* appears to be smaller than for normal PPG meaning that these artifact’s beats are less symmetric and, in particular, more left skewed.

Talking is a continuous artifact that mainly manifests with changes in “vertical” time domain features. As it can be observed in the corresponding recording figure on Table 5.4, there are sudden increases in *Amplitude* and *TroughDiff* characteristics. The frequency spectrum is therefore severely corrupted ($>SpectralEntropy$ (0-1.5Hz), *AvgPower* (1-2Hz)). The most distinctive characteristic is the *ZeroCrossingRate*, that is greater than for normal PPG, as it is probably modulated by the frequency of talking.

Last but not least, **Sensor Rubbing** is an artifact that can occur in a continuous or transient manner, depending on how the sensor is rubbed and the sources of rubbing. This artifact can be characterized by overall irregular larger amplitude ($>StdAmplitude$, *Amplitude*, *ChangeAmplitude*, *TroughDiff*, *ChangeTroughDiff*) and asymmetry of beats (*ChangeSkewness*, *StdSkewness*). It also presents a higher *ZeroCrossingRate* than normal, but this could be biased by the intentionally fast movement of the sensor during the experimental procedure.

5.3.3 Characterization of additional respiratory states of interest

Characteristics of the two extra respiratory stable states of interest, Breathing Slow and Breath-Holding (Apnea), were also compared against normal PPG. Results were likely presented in Table 5.4.

For **Breathing-Holding (Apnea)**, it can be observed that the PPG signal is more stable, with no baseline wander modulating the overall envelope. This is reflected by the “vertical” characteristics: *StdPeakHeightDiff*, *StdKurtosis*, *ChangeAmplitude*, *StdAmplitude*, *PeakHeightDiff* and *ChangePeakHeightDiff*. These all have a lower ($<$) sign, showing that the instantaneous and overall variation in amplitude among pulses is very small, and the peak height difference is reduced compared to normal. Even the standard deviation of “peakedness” (i.e. *StdKurtosis*) of successive pulses varies significantly less than normal. PPG pulses for this respiratory state seem all invariable and uniform, with the standard deviations in *Width* and *RiseTime* also being lower than normal. The clear simi-

larity between pulses is exposed by the *CorrelogramPeak1* and *CorrelogramPeak2* features, that show higher correlation values than for normal PPG. Finally, there are no significant features from the frequency domain in the ranked features, reinforcing the idea of a perfectly periodic PPG signal, that does not corrupt the PPG cardiac frequency bandwidth of interest. Indeed, the relative power from 0 to 0.8Hz is less than 20% whereas for normal PPG is $\sim 40\%$, and from 0.8 to 1.2Hz is of $\sim 60\%$ versus $\sim 40\%$ for normal.

For **Breathing Slow**, the situation is analogous: pulses look very stable and uniform. Similar “vertical” features, as for **Breath-Holding (Apnea)**, are: *PeakHeightDiff*, *ChangePeakHeightDiff*, *ChangeAmplitude* and *TroughDiff*. These are smaller ($<$) than for normal PPG, showing that the signal is less variable. Even if the PPG signal baseline is modulated by the slow respiration frequency, with around 30% of the total power in the 0-1Hz band, it fluctuates less than for normal breathing PPG.

These findings suggest that these respiratory states, which are likely to occur during sleep, do not severely disrupt the PPG signal, but on the contrary, they seem to be more stable and closely related to normal PPG than to the rest of the artifacts. Indeed, it is worth pointing out that, for the features ranked in these respiratory cases, the signs are in general opposite to the signs found for the same corresponding features in the rest of the artifacts. In other words, they have more stable characteristics than normal PPG, while artifacts have more variable and abnormally larger ones. Therefore, the breathing states of interest, **Breath-Holding (Apnea)** and **Breathing Slow**, have the potential to be clearly distinguishable from interference, by applying simple thresholds or more complex classification models on the identified top features.

5.4 Discussion

In this chapter, artifacts specific to arterial PPG measured on a new measurement site, the neck, were characterized using signals recorded from 19 participants. A total of 41 features from the time, correlogram and frequency domains were extracted and input into statistical tests for evaluation. Several features showed significance in differentiating normal PPG from the rest of the artifacts. But since, depending on the source of the interference, features demonstrated distinct degrees of significance, a mRMR approach was applied to sort the most important characteristics for each artifact independently. This consisted in ranking features based on increasing p-values and minimizing the presence of highly redundant information relying on their correlation. A table, listing all the artifacts with the top 10 ranked features, presented an exhaustive characterization of neck PPG artifacts.

In summary, even though it was shown that each artifact can be characterized by a particular

subset of ranked features, there are some similarities appearing among them. First, there is a group of key features that showed a higher capability to discriminate normal PPG from all the artifacts. These include: *Peak Height Difference*, *Trough Difference*, *Change in Amplitude*, *Standard deviation of Amplitude*, *Spectral Entropy*, *Spectral Kurtosis* and *Average Power (1-2Hz)*. On the contrary, some others, like *RelativePower* for all bands, *AvgPower (0.8-1.2Hz)*, *RiseTime* or *CorrelogramLags1&2*; did not show great statistical significance. This leads to the conclusion that the latter are not a suitable choice for differentiating normal neck PPG from the rest of the artifacts. Second, overall, all artifacts considered in this study were characterized by being very chaotic, unsteady and irregular. They were distinguished by having larger and highly variable amplitude related characteristics, as well as noisier and disordered frequency spectrum than normal PPG. **Breathing Fast**, was the only one that showed fewer irregular features, even if still abnormally large. In fact, the cardiac pulses information was notably distorted with great amplitude and separation between “beats”, but maintaining the particular fast respiratory rate, as the predominance of high values of autocorrelation features (*CorrelogramPeak1&2*) reflected.

These findings set for the first time the characteristics of neck PPG artifacts, presenting the features with the greatest potential to differentiate each of them. The recurrent identified patterns can be used to develop techniques to detect and classify corrupted neck PPG segments. For this purpose, thresholds could be established, or classification models could be trained, by exploiting the most suitable features, as identified in this study, for each specific artifact category. Moreover, the periodic respiratory modulation of neck PPG signals [7] could be similarly exploited as a discrimination tool for breathing states. The distinctive fundamental frequencies of each breathing type observed in the spectrum, contrast with the highly disordered and unlocalized frequency spectra of artifacts. These results will be of great importance when trying to develop artifacts detection methods for novel neck PPG signals denoising, since the sources of interference they are susceptible to, are now well known and analyzed. In future research of neck PPG, engineers devising their signal processing pipelines will directly have access to the best tuned features for their end-application. This will, on its turn, facilitate a more accurate extraction of clinical biomarkers of interest.

In addition, it was also verified that eliminating artifacts would not eliminate extreme cases of signals of interest such as **Breathing Slow** and **Apneas**. These, did not impair the PPG as artifacts did, but instead showed to preserve the signal quality comparably or better than Breathing Normal. A possible reason explaining this is that a slower, more controlled or absent respiration frequency, makes the sensor more stable. Therefore, the signals under these circumstances are less affected by the DC respiration baseline modulation. This suggests that even though they alter some “normal” PPG traits, they are signals of interest that enhance the clinical utility of the neck.

In finger pulse oximetry, artifacts altering the PPG signals are mainly due to motion, as hands and arms are in constant movement. In comparison, neck PPG artifacts sources are more diverse. **Head movements** are the source of strong artifacts, but the neck position makes them much more stable in terms of frequency and amplitude. Other intrinsic artifacts resulting from autonomic reflexes (**Swallowing, Yawning, Coughing** or **Talking**) are unique to the neck location, but, in relative terms they are shorter and generally less frequent. In addition, respiratory movements noticeably modulate the DC baseline of PPG signals, but this can be information of great interest, in some clinical contexts, such as in apnea detection for SUDEP prevention. It is for this reason, that the neck could potentially be a very interesting alternative to finger PPG within the context of continuous respiratory monitoring and/or diagnosis, mostly in uncontrolled environments where artifacts caused by hand movements are the strongest.

This characterization of neck artifacts might also be very valuable to extract additional information for numerous applications for which the final objective is to identify the occurrence of particular events. For example, based on the ranked features it would be possible to detect whether a person is swallowing its medication, its respiration frequency has been altered, or the person is suffering apnea events. Finally, this study has only focused on recording neck signals in supine position, but to increase the generality, other positions could be similarly explored. It could be likewise interesting to record and analyze combinations of artifacts and respiratory states happening simultaneously at the same time.

5.5 Conclusion

This work experimentally extracts and analyzes, for the first time in literature, the 10 most important neck PPG sources of artifacts and two other complementary respiratory states of interest. These artifacts of intrinsic and external origin include respiratory, autonomic reflexes, sensors, and motion types. An exhaustive signal processing characterization study of these is presented in this chapter. This characterization, exploits time, correlogram and frequency domain characteristics. Statistical features comparisons between normal PPG and artifacts, revealed the group of significant features with the greatest capability of discrimination between the two groups. A mRMR ranking approach, selected the 10 most characteristic features for each condition independently, among the total 41 extracted. These findings are crucial for the complete characterization of sensed neck PPG, as they set the foundations for mathematically understanding the sources of undesired signal interference in this novel pulse oximetry measurement site. This is of great importance for the future development of artifact elimination methods, with the objective to clean up the PPG signals and to increase the accuracy of future algorithms aiming to extract physiological parameters for a variety of clinical contexts. In

addition, this characterization could also support algorithms aiming to extract information in which the specific artifact is a signal of interest itself, such as detection of swallowing or apnea events.

References

- [1] I. Garcia-Lopez and E. Rodriguez-Villegas, “Characterization of artifact signals in neck photoplethysmography,” *IEEE Transactions on Biomedical Engineering*, 2020.
- [2] M. T. Petterson, V. L. Begnoche, and J. M. Graybeal, “The effect of motion on pulse oximetry and its clinical significance,” *Anesthesia & Analgesia*, vol. 105, no. 6, pp. S78–S84, 2007.
- [3] Y. Maeda, M. Sekine, and T. Tamura, “Relationship between measurement site and motion artifacts in wearable reflected photoplethysmography,” *Journal of Medical Systems*, vol. 35, no. 5, pp. 969–976, 2011.
- [4] J. W. Chong, D. K. Dao, S. Salehizadeh, D. D. McManus, C. E. Darling, K. H. Chon, and Y. Mendelson, “Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection–reduction approach. Part I: motion and noise artifact detection,” vol. 42, no. 11, pp. 2238–2250, 2014.
- [5] M. Engin, A. Demirel, E. Z. Engin, and M. Fedakar, “Recent developments and trends in biomedical sensors,” *Measurement*, vol. 37, no. 2, pp. 173–188, 2005.
- [6] K. T. Sweeney, T. E. Ward, and S. F. McLoone, “Artifact removal in physiological signals: Practices and possibilities,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 488–500, 2012.
- [7] I. García-López, S. A. Imtiaz, and E. Rodriguez-Villegas, “Characterization study of neck photoplethysmography,” 2018, pp. 4355–4358.
- [8] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, no. 3, p. R1, 2007.
- [9] N. Pradhan, S. Rajan, A. Adler, and C. Redpath, “Classification of the quality of wristband-based photoplethysmography signals,” in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2017, pp. 269–274.
- [10] R. Dwyer, “Detection of non-gaussian signals by frequency domain kurtosis estimation,” in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 8, 1983, pp. 607–610.
- [11] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

Chapter 6

Artifacts Classification and Apnea Events Detection in Neck Photoplethysmography Signals

6.1 Introduction

Finger PPG pulse oximetry signals have been previously used in the literature for apnea detection in conjunction with other monitoring sensors (e.g ECG, EEG, respiration, sound) or on their own ([1, 2, 3]). Among those exclusively using PPG sensors, most of the efforts have focused on first, detecting oxygen desaturations from the surrogate $\text{SpO}_2\%$ signal [4]; and then extracting relevant apneic characteristics ([5, 6]). Some of the most typical features include: time series statistics of the SpO_2 signal, the oxygen desaturation index quantifying the severity of the drop in oxygen levels by 2%, 3%, 4% (ODI2, ODI3, ODI4), and the desaturation area under these thresholds. Other studies, directly employed the PPG signal to extract time and frequency domain features, such as the PPG amplitude, beat-to-beat characteristics, or the low (0.04-0.15Hz) and high (0.15-0.5Hz) frequency powers ([7, 8]). Papini *et al.* [9] included both pulse rate variability (PRV) and respiratory activity derived features from the PPG signal. Lázaro *et al.* [10] focused on detecting decreases in amplitude of the PPG signal (DAP) that were previously shown to be correlated with apnea ([11]). However, these still depend on the detection of the delayed DAP segment of the signal occurring after the apnea. Present PPG apnea detection methods could therefore be effective in clinical scenarios, where recordings are post-processed offline. However, they show limited utility in more real-time applications. For example, in Sudden Unexpected Death in Epilepsy (SUDEP), the prompt detection of apneic events could be a matter of life or death. Neck PPG signals could offer a solution to the current limitations, as apneic respiratory arrests can be instantaneously recognized by monitoring time and frequency features ([12, 13]).

The acquisition of neck PPG signals is however limited by the presence of artifacts that superimpose to the signal of interest. Hence, the occurrence of head movements, coughing or swallowing could lead to unreliable and inaccurate SpO₂ and HR readings; which in certain situations could put the patient's life at risk, and in others could lead to discontinuous adoption due to false alarms. In order to improve the accuracy on the quantification of these physiological parameters, artifacts removal and signal reconstruction methods have been extensively developed and reported in the literature. Some include time and frequency filtering approaches like discrete wavelet transforms ([14, 15]), Fourier series analysis [16] or source separation techniques (e.g. independent component analysis [17] or singular value decomposition [18]). These approaches are, however, prone to the introduction delays and/or distortion in the noise-free PPG segments. Adaptive filtering strategies have also been widely explored ([19, 20]), using additional sensors (accelerometers) to provide a noise reference estimate. Other approaches have focused, instead, on detecting and removing artifact-corrupted PPG sections, prior to the estimation of the physiological parameters of interest ([21, 22, 23]). Following this approach, several machine learning algorithms have been proposed in the literature to discriminate artifacts from clean PPG. Some examples of classification models include: decision lists ([24, 25, 26, 27]), decision trees [28], naïve Bayes classifiers [29], support vector machines (SVM) ([21, 30, 31, 32]), multi-layered perceptrons [33], personalized neural networks (NN) [34], and 1-D CNNs [35].

In the specific case of neck PPG, we have previously defined and characterized the most common neck PPG artifacts [12]. However, there is no evidence of any previous research devising algorithms for neck PPG artifacts classification. Since artifacts removal is crucial for neck PPG to work in real life conditions, the first goal of this chapter was to design a high performance classifier capable of discriminating artifacts from clean PPG signals. In addition, given that neck PPG signals have a big potential to instantaneously detect apneic events, the second objective of this work was to develop, for the first time in literature, an apnea classification model utilizing neck PPG.

6.2 Methods

6.2.1 Experimental protocol

In the previous chapter, a set of artifacts, including fast breathing, talking, head and body movements, swallowing, coughing, yawning and sensor rubbing; as well as two additional respiratory states of interest (slow breathing and breath-holding apnea); were recorded in Experiment 4 [12]. Two PPG sensors were used for data acquisition in supine position: a reflectance pulse oximeter (8000R, Nonin) placed at the suprasternal notch of the neck and a transmission one (Onyx II 9560, Nonin) placed on the index finger for reference purposes.

This dataset was used in this work for both artifacts classification and apnea detection. It consisted of 13 recordings per subject, of 140s duration each. During the first control recording, participants were instructed to breath at their normal respiratory pace. Then, to test other respiratory states, they were asked to modulate their respiratory frequency at three different moments in the recording for 20-30s. In one recording at a slower pace, and in another recording by holding their breaths to simulate apneic events. Ultimately, the last 10 recordings introduced the different neck PPG artifacts in alternating periods of 20s with spontaneous breathing in between. During data acquisition, the onsets and offsets of artifacts were marked in real-time. After the experiments, the annotation was verified by comparing with reference finger PPG signals. Each recording was independently normalized.

6.2.2 Features extraction

6.2.2.1 Windows segmentation and labelling

In order to obtain relevant features for further classification, recordings were segmented in small data fragments. The extracted features were averaged within a defined time window that was repeatedly shifted by 2s along the whole recording. Each average feature corresponded to an independent observation to be inputted into the classification model. In this manner, every new upcoming bit of data was evaluated, simulating real-time processing conditions. Various window lengths ($W = 4, 5, 6, 7, 8$ and 10s) were explored to assess which one maximized the accuracy of classification.

The labelling of each window, as *artifact / clean PPG* for the artifacts classification model, or as *apnea / normal PPG* for the apnea detection model was defined based on a percentage (%) threshold Thd of window corruption. In other words, if let's say $Thd = X\%$ or more of the evaluated PPG segment's total length contained an artifact (or apnea) signal, then the window was assigned to the positive class. Otherwise, if the percentage of corruption was less than $Thd = X\%$, the window was labelled as the negative class: clean PPG (or normal PPG respectively). Several thresholds of corruption ($Thd = 20\%, 30\%, 40\%, 50\%$) were tested as well to explore how the different labelling criteria affected the sensitivity and specificity of the algorithms.

6.2.2.2 Features

Most of the features proposed in Chapter 5, were also considered in this work, since they demonstrated strong statistical significance in the differentiation between normal clean PPG from artifacts or breathing states [12]. New additional features derived from the envelope of the PPG signal were additionally included, to increase the classification performance. The 51 features considered in this chapter for both classification models are presented below. Further details on these features can be found in the previous chapter.

- *Time domain morphological features:*

Amplitude $[F_1]$ vertical distance between the onset of a PPG pulse and the systolic peak.

Width $[F_2]$ time duration between the onset and offset of a PPG pulse in time units (seconds).

Peak Height Difference $[F_3]$ relative amplitude between successive pulses peaks.

Peak Distance $[F_4]$ horizontal distance between successive pulses peaks (in seconds).

Trough Difference $[F_5]$ relative amplitude difference between onsets of successive pulses.

Rise Time $[F_6]$ time period between the onset of a PPG pulse and its systolic peak.

Skewness $[F_7]$ degree of symmetry of a PPG pulse.

Kurtosis $[F_8]$ degree of sharpness of a PPG pulse.

Change of F_{1-8} features $[F_{9-16}]$ instantaneous difference of feature's values for consecutive pulses.

Standard Deviation of F_{1-8} features $[F_{17-24}]$ feature's standard deviation over the whole window length.

Zero-Crossing Rate $[F_{25}]$ number of times per second that the PPG signal crosses zero.

- *Correlogram features:*

Correlogram Peaks $[F_{26-27}]$ autocorrelation values of the first and second peaks of the correlogram.

Correlogram Lags $[F_{28-29}]$ lags of the first and second correlogram peaks.

- *Frequency domain features:*

The one-sided modified periodogram estimate of the power spectral density (PSD) was used to calculate the frequency features. For that, the spectrogram was derived using the squared magnitude of the Short-Time Fourier Transform (STFT) with a window of 10s and 90% overlap. The output power (dB/Hz) was then sliced in time to obtain each window PSD.

Shannon Spectral Entropy (0-1.5Hz & 1-4Hz) $[F_{30,31}]$ degree of “disorder” of the power spectrum's probability distribution.

Spectral Kurtosis (0-1.5Hz & 1-4Hz) $[F_{32,33}]$ peakedness of the PSD at each specific frequency.

It is calculated as the normalized fourth-order moment of the real part of the short-time Fourier transform.

Relative Power $[F_{34-36}]$ calculated by adding the power contained within specific frequency bands (0-0.8Hz, 0.8-1.3Hz, 1.3-1.8Hz) and dividing it by the total power spanning all frequencies.

Average Band Power $[F_{37-41}]$ power of the signal was averaged within the five frequency bands: 0-0.8Hz, 0.8-1.3Hz, 1.3-1.8Hz, 2.2-2.8Hz, 3.2-3.8Hz.

- *Envelope features:*

The upper envelope of the PPG signal was extracted using spline interpolation over local maxima separated by at least 50 samples ($>0.667s$). A total of 10 features were extracted from this envelope signal.

Envelope standard deviation $[F_{42}]$ variance in the envelope signal within the window.

Envelope maximum $[F_{43}]$ maximum value of the envelope signal within the specific window.

Envelope minimum $[F_{44}]$ minimum value of the envelope signal within the specific window.

Envelope range $[F_{45}]$ difference between the maximum and minimum values of the envelope signal within the current window.

Envelope approximate Entropy $[F_{46}]$ regularity statistic that measures the unpredictability of repetitive patterns. In other words, a PPG envelope signal including repetitive fluctuations, such as spontaneous breathing, would show small approximate entropy values, whereas a less predictable signal (e.g. artifact) would be characterized by larger ones. It was computed using the *approximateEntropy()* function in MATLAB 2020 [36].

Envelope area $[F_{47}]$ area under the envelope absolute signal, computed by numerical integration via the trapezoidal method.

Envelope Average Power $[F_{48-51}]$ power of the envelope signal was averaged within the following frequency bands: 0-0.15Hz, 0.2-0.5Hz, 0.5-1Hz and 1-1.5Hz.

6.2.3 Classification pipeline

In this study, two classification algorithms were developed: an artifacts classifier and an apnea classifier. According to our previous findings [12], on one hand, neck PPG artifacts, with similar noisy characteristics, could be clearly distinguished from normal PPG. On the other hand, normal, slow breathing and apnea PPG signals shared common stable clean PPG features. As a consequence, for the artifacts classifier, all the artifact types were grouped together under the *artifacts* positive class; while the negative *clean PPG* class encompassed: the normal, apnea and slow breathing PPG signals.

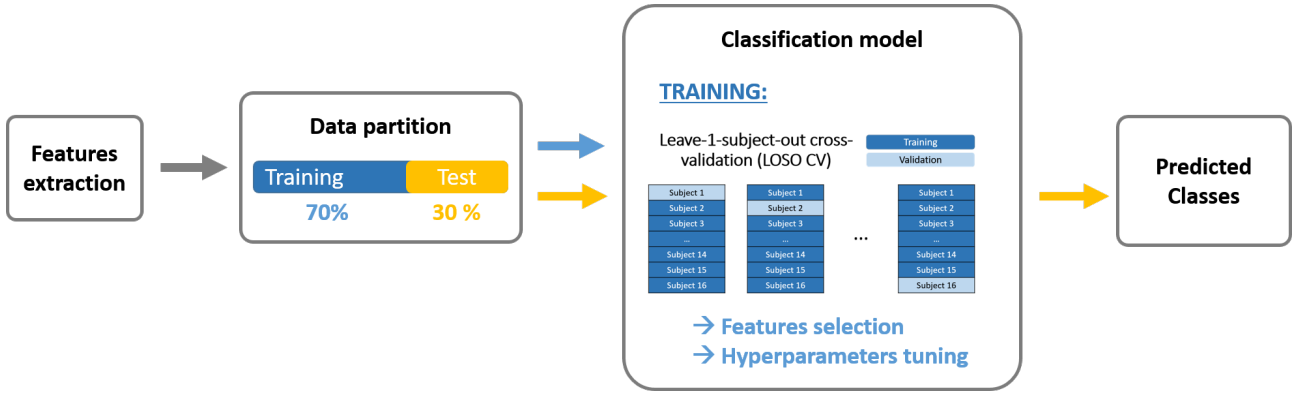


Figure 6.1: Classification pipeline. The predicted output classes for artifacts classification are: Artifacts/clean PPG; for apnea classification: Apnea/Normal PPG. This process was repeated 30 times with different randomization in the data partition stage, for each window length and threshold of corruption (%) combination.

In order to detect apneic events among the clean PPG signals category, an apnea classifier was also engineered. The positive class consisted of the *apnea* PPG signals. And the *normal PPG* negative class comprised the normal and slow breathing categories. The number of artifacts and breathing states were evenly sampled at random in order to perform balanced binary classification.

Figure 6.1 shows an overview of the classification pipeline for both classifiers. This process was repeated 30 times with different randomization in the data partition stage, for each combination of window length and threshold of corruption (%). Each stage is further detailed in the subsections below.

6.2.3.1 Data partition

Since there was window overlapping, a random partition of data could no longer be used, as the condition of independence between training and test data would be violated.

As it can be observed in Figure 6.2, two types of data partitions were used for classification. For artifacts classification, a Leave-30%-of-Subjects-Out approach was implemented. As Fig. 6.2(a) shows, for every seed, 70% of the subjects were selected at random for training (with all the recordings), and the other 30% were left for testing. This ensured that the classifier was tested against completely unseen data, which is one of the most strict validation strategies. All subjects were assigned to the test set evenly, at least 7 times each. This avoided any subject-bias.

For apnea classification, due to a limited number of breath-holding recordings, an alternative Leave-1/3-of-Recording-Out per subject partition was adopted instead. An illustration of three intercalated breath-holding events that simulate apneic events can be observed in Fig. 6.2(b). Apnea recordings were thus divided in three even segments for each subject. The same number of normal and apnea PPG windows were included in each of them and no overlapping windows (in the border) were allocated to

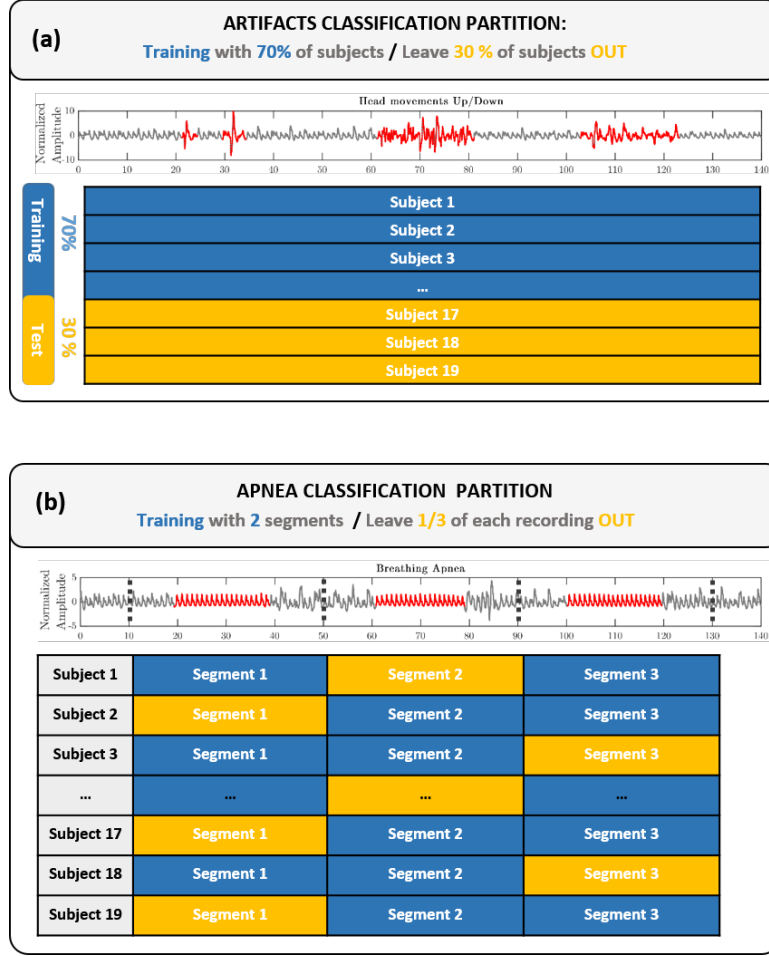


Figure 6.2: Data partitions for artifacts classification and apnea detection models. (a) Leave-30%-of-Subjects-Out approach for artifacts classification. (b) Leave-1/3-of-Recording-Out per subject for apnea detection.

either of the neighbouring segments. This prevented overfitting and guaranteed independence of the training and test sets. For each random seed repetition, one of the three segments was selected for the test set and the remaining two were used for training.

This data partition step was repeated 30 times for both classifiers, with different randomization of the training and test sets, to verify that the proposed algorithms showed a good generalization performance.

6.2.3.2 Training

A SVM classifier with a radial basis function (RBF) kernel was chosen for the artifacts and apnea classification. The objective of the SVM classification problem was to find the weights vector \vec{w} and bias term b defining the optimal hyperplane, that maximizes the margin between classes and minimizes the loss term such that:

$$\min_{w,b,\xi} \frac{1}{2} \vec{w}^T \vec{w} + C \sum_{i=1}^n \xi_i \quad (6.1)$$

subjected to the condition:

$$y_i(\vec{w}^T \phi(\vec{x}_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (6.2)$$

where, $\phi(\vec{x}_i)$ are the training vectors in the input space, y_i the classes labels $[-1,1]$ and ξ_i the slack variables. C corresponds to the regularization parameter that controls the trade-off between maximizing the margin ($C \rightarrow 0$) and minimizing the penalty term ($C \rightarrow \infty$). The function $K(\vec{x}_i, \vec{x}_j)$ maps the training vectors into a higher dimensional space in order to gain linear separation. The RBF gaussian kernel used was defined such that:

$$K(\vec{x}_i, \vec{x}_j) = \phi(\vec{x}_i)^T \phi(\vec{x}_j) = \exp(-\gamma \|\vec{x}_i - \vec{x}_j\|) \quad (6.3)$$

where, $\gamma = \frac{1}{2\sigma^2}$ is the inverse of the radius of influence of the samples selected by the model as support vectors.

During training, the best features and hyperparameters, that optimized the model's performance, were selected using the Leave-One-Subject-Out Cross-Validation (LOSO-CV) strategy. Similarly to k-fold cross-validation, the training data was repeatedly split, by selecting one subject at a time for testing, and the rest of the subjects for training. This approach avoids overfitting and prevents subject bias during feature selection and hyperparameters optimization.

6.2.3.3 Features selection

The features selection step was included within the LOSO-CV and was performed only on the training subjects. It consisted of two stages. First, the total 51 features were ranked using chi-square tests. These evaluated whether the features were independent of the classes labels, and then ranked the features based upon the output p-values. A small p-value revealed that the corresponding feature was dependent on the response variable, and therefore, was an important feature to consider for classification.

The top 30 features ranked with the Chi-square tests were fed into a forward sequential feature selection algorithm. In a wrapper fashion, the subsequent ranked features were sequentially added to the top 30 candidate set until the addition of further features did not decrease the average misclassification error by more than a relative tolerance of 1e-6.

6.2.3.4 Hyperparameters optimization

In order to boost the SVM training performance, the soft-margin misclassification cost (C) and the RBF kernel gamma (γ) hyperparameters were optimized by grid search. For the different classifiers, all the combinations of C and γ , listed as follows, were evaluated using LOSO-CV.

Artifacts classification:	$C = 0.5, 1, 4, 6, 8, 16, 32, 64, 80, 128$ $\gamma = 2^{-15}, 2^{-13}, 2^{-11}, \dots, 2^{-1}, 2^1, 2^3$
Apnea classification:	$C = 0.125, 0.75, 1, 2, 3, 4, 5, 6, 8, 32$ $\gamma = 2^{-15}, 2^{-13}, 2^{-11}, \dots, 2^{-1}, 2^1, 2^3$

The hyperparameters that maximized the cross-validation training accuracy were selected for the artifacts classifier, and those showing the highest F1-score were chosen for the apnea classifier.

6.2.3.5 Performance metrics and model selection

Once the most optimal hyperparameters and features were selected through LOSO-CV, the final SVM model was trained with the whole training data partition. Subsequently, it was evaluated on the independent test set (in yellow in Fig. 6.1), to output the predicted classes.

In order to assess the classification performance of both classifiers, the following metrics (in %) were calculated as the average over the 30 randomization repetitions: accuracy (ACC), sensitivity (SE), specificity (SP), precision, and F1-score (F1).

The best artifacts classification model was chosen based on the combination of window length and threshold of corruption (%) (W/Thd) that maximized the accuracy metric. In apnea classification, the harmonic mean of precision and recall, i.e. the F1-score, was used instead to select the best W/Thd model. Indeed, the F1 metric is more relevant in this case, as the Type I (false positives) and Type II (false negatives) errors are crucial for safety in critical apnea detection applications.

6.2.4 Statistical evaluation of the classification results

In order to assess whether the different windows and corruption thresholds (%) had an effect on the classification performance of both classifiers, a two-way ANOVA statistical test was carried out for each performance metric. The normality and homoscedasticity assumptions were verified using Lilliefors and Levene's tests. This confirmed the homogeneity of variance among different sample groups and the Gaussianity of the distributions. Post-hoc multiple comparisons, based on the Tukey's honest significant difference criterion, were subsequently performed in order to investigate which pairs of means were significantly distinct, for the different windows and corruption thresholds (%) evaluated.

6.3 Results

6.3.1 Classification results

Figure 6.3 shows the average results for both artifacts and apnea classification algorithms, across all windows and thresholds of corruption (%). The bar graphs represent the mean performance over the

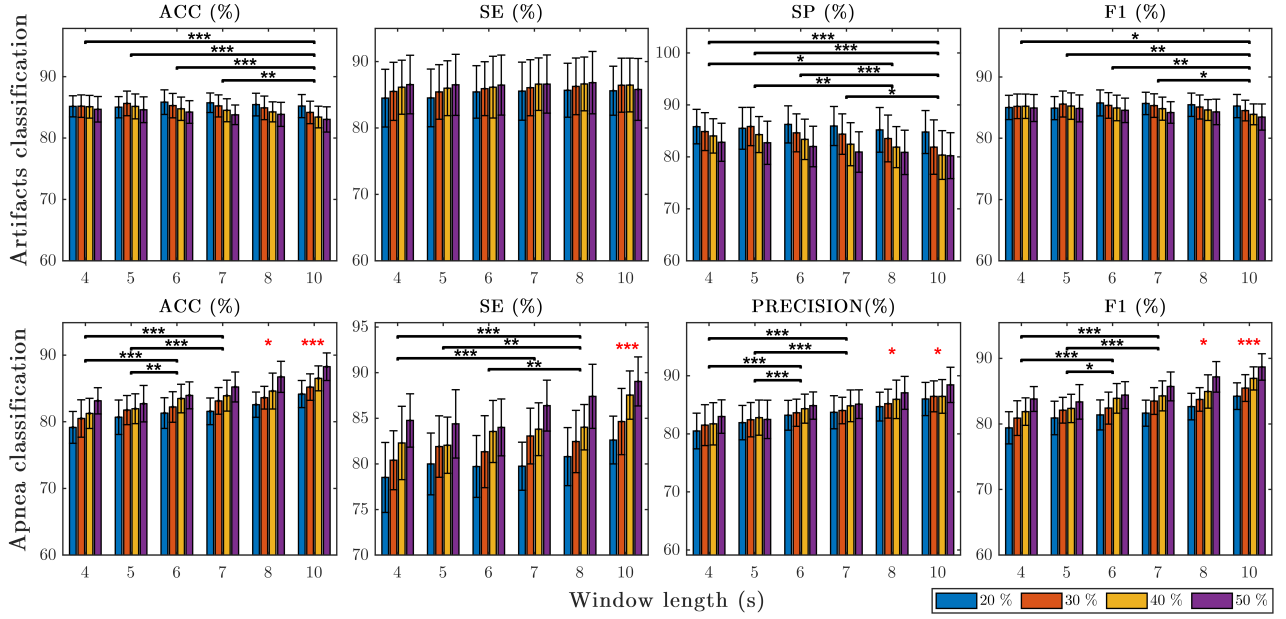


Figure 6.3: Average classification results for the proposed artifacts and apnea classification algorithms, over the 30 randomization experiments. Bar graphs show the average performance metrics across the different windows and corruption thresholds (%) tested. The error bars represent the extent of the standard deviation above and below the mean. Different thresholds of corruption ($Thd = 20\%$, 30% , 40% , 50%) are specified as separate coloured bars for each window length ($W = 4, 5, 6, 7, 8, 10s$). The statistical results of the multiple pairwise comparisons testing for the window effect are displayed with a horizontal line and a black asterisk symbol indicating the alpha significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The red asterisks on top of some window groups indicate that all the multiple comparisons were statistically significant for that specific window.

30 repetitions and the error bars, the corresponding standard deviations. Overall, both classifiers demonstrated good performance with average metrics' values larger than 80% for the majority of the W/Thd models. A more exhaustive analysis is presented in the following subsections.

6.3.1.1 Artifacts classification

The results presented in the upper panels of Figure 6.3, show a good performance of around 86% for the various windows and thresholds. The ACC , SE and $F1$ mean values oscillate in a short range of 2-3% for the different W/Thd combinations. However, the SP mean values expand across a larger range of 6%, probably due to a threshold effect. On average, the standard deviations for ACC and $F1$ are very small (1.8%), whereas for SE and SP are slightly higher (4%). But still, these values remain acceptable considering that a Leave-30%-of-Subjects-Out validation approach was used, which is one of the most strict ones.

Table 6.1 presents the average performance results for the best (W/Thd) artifacts classification model. The window and corruption threshold (%) combination that maximized the accuracy of artifacts classification was $W = 6s - Thd = 20\%$, with a value of $85.84 \pm 2.00\%$. The $F1$ -score ($85.77 \pm 2.12\%$), SP ($86.26 \pm 3.57\%$), and $precision$ ($86.29 \pm 2.92\%$) values of this W/Thd model were also the largest compared to all other parameters pairs.

Table 6.1: AVERAGE PERFORMANCE RESULTS ($\mu \pm \sigma$, N=30) FOR THE BEST ARTIFACTS AND APNEA CLASSIFICATION MODELS

Best W/Thd model	Artifacts*	Apnea**
	classification	classification
	W=6s - Thd=20%	W=10s - Thd=50%
ACC	85.84 \pm 2.00	88.25 \pm 2.07
SE	85.43 \pm 3.95	89.03 \pm 2.69
SP	86.26 \pm 3.57	87.42 \pm 3.63
Precision	86.29 \pm 2.92	88.42 \pm 3.04
F1	85.77 \pm 2.12	88.68 \pm 2.01

* The artifacts classifier discriminates between noise-corrupted PPG segments and clean data (normal breathing, slow breathing and apnea PPG fragments)

**The apnea classifier distinguishes apnea events from the rest of clean PPG data (normal breathing, slow breathing)

6.3.1.2 Apnea classification

In the lower panels of Figure 6.3 are exposed the average classification results for the apnea classification algorithms. Although the various metrics demonstrated a good performance of around 83-84% in average for all the W/Thd , there was a clear ascending trend that reasonably increased the range of mean values. The difference between extreme values could span from an 8% in *precision* and up to a 10.5% in *SE*. This suggested that the windows and thresholds parameters might have had an effect. The standard deviations, pictured as error bars, occur in general very small ($< 3.2\%$) for all metrics.

The best apnea classification model (W/Thd) and the corresponding performance metrics are listed in Table 6.1. The maximum F1 score of $88.68 \pm 2.01\%$ was obtained for the apnea classification model with a window of $W = 10s$ and a threshold of corruption of $Thd = 50\%$. This W/Thd combination also maximized the *ACC* ($88.25 \pm 2.07\%$), *SE* ($89.03 \pm 2.69\%$), *SP* ($87.42 \pm 3.63\%$) and *precision* ($88.42 \pm 3.04\%$), compared to the other W/Thd pairs.

Figure 6.4 shows the predicted classes output of the best artifacts and apnea classification models. Some of the most characteristic features that were inputted in the classifiers are also displayed such as the Peak Height Difference, the Envelope's maximum value and the Spectral Entropy ($< 1.5Hz$).

6.3.2 Statistical tests results:

6.3.2.1 Two-way ANOVA

Overall, the resulting ANOVA tables for both classifiers, showed that the window length and the threshold of corruption (%) affected the classification performance metrics significantly ($p < 0.05$). Some exceptions to this were the window length effect for the sensitivity of artifacts classification ($p = 0.707$) and the threshold effect (%) for the specificity of apnea classification ($p = 0.065$). No statistical evidence of an interaction effect between the two factors was shown for any metric ($p > 0.05$).

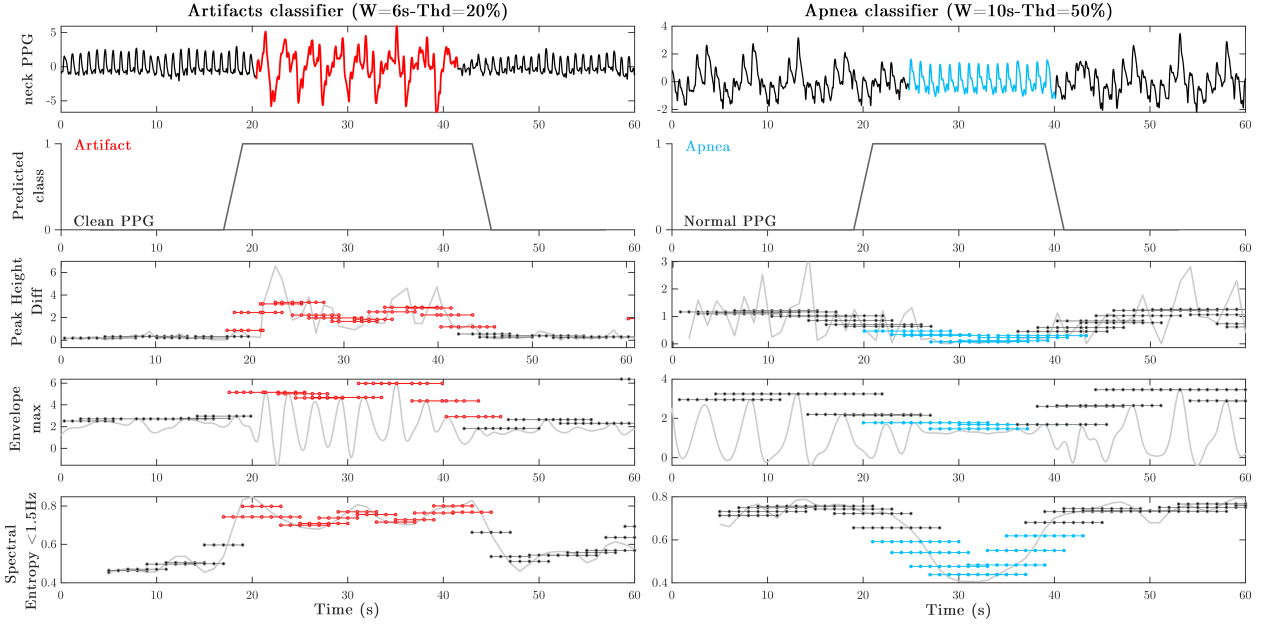


Figure 6.4: Classification decision results of the best models for one head movement artifact and an apnea event. The variation of some of the features used are displayed in the lower panels: Peak Height Difference, the maximum of the envelope and the Spectral Entropy ($< 1.5\text{Hz}$). True artifacts and apnea windows are labeled in red and blue respectively.

The results of the post-hoc multiple comparisons for the W and Thd effects are described in the next subsections.

6.3.2.2 Window length effect

In Figure 6.3, the statistically significant pairwise differences among window lengths groups ($W = 4, 5, 6, 7, 8, 10\text{s}$) are shown in the form of horizontal lines with an asterisk symbol representing the p-values ranges (* $0.01 < p < 0.05$, ** $0.001 < p < 0.01$ and *** $p < 0.001$). For the sake of visualization, a unique red asterisk symbol was used when any group was statistically significant with all the others simultaneously. The largest p-value was chosen for the asterisk representation.

As it can be observed, in **artifacts classification**, the window $W = 10\text{s}$ shows the greatest significance. Indeed, for the average ACC , SP and $F1$ metrics, $W = 10\text{s}$ is the only group that is statistically different from all the rest of the windows (except from $W = 8\text{s}$). For SP , besides $W = 10\text{s}$, the average specificity values of $W = 8\text{s}$ are also statistically distinct from the $W = 4$ and 5s ones. This could be explained by a slight decrease in performance, from $W = 5 - 6\text{s}$, with increasing window length of ACC (-1.1%), SP (-2.7%) and $F1$ (-0.87%). No significant pairwise comparisons appear among window groups for SE , since, according to the ANOVA findings, the window length did not have an effect ($p > 0.05$). Actually, no dissimilarity in the average SE values is noticeable among window groups, being all roughly equal to 86% in average. The fact that the standard deviations of $\sim 4\%$ are some of the largest compared to other performance metrics, might also explain the non-significance.

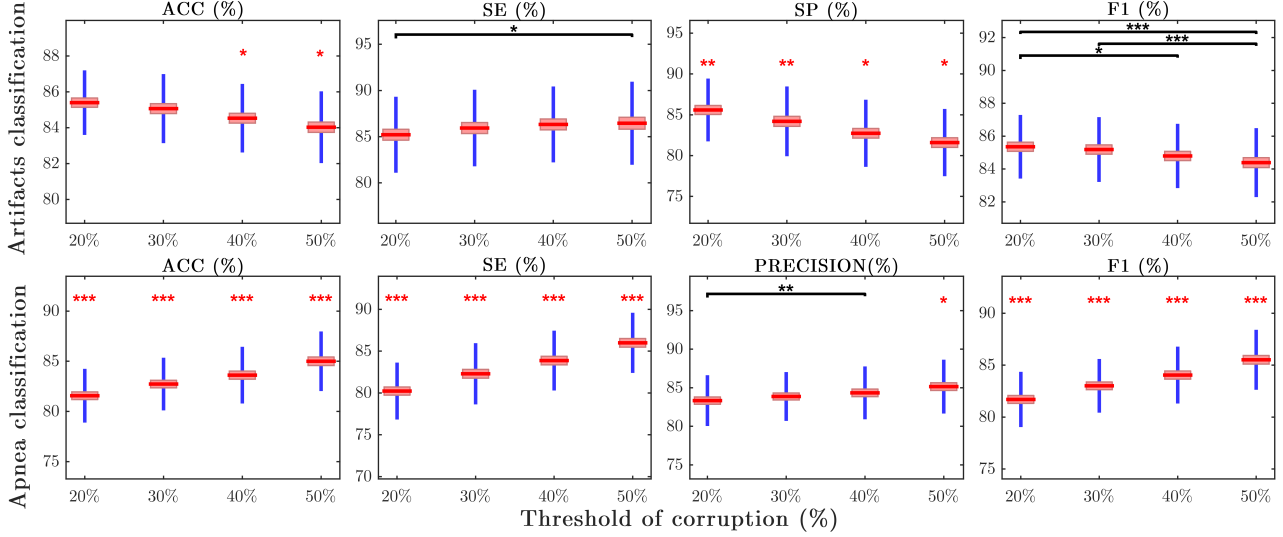


Figure 6.5: Means plots of the classification performance metrics across different thresholds of corruption (%). The means, with the corresponding 95% confidence intervals, are represented in red. Standard deviations above and below the mean are shown in blue. The statistical results of the multiple pairwise comparisons testing for the threshold effect are displayed with a horizontal line and a black asterisk for different alpha significance levels: * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$. The red asterisks on top of some threshold groups indicate that all the multiple comparisons were statistically significant for that specific window.

In **apnea classification**, the lower panels of Fig. 6.3 show that the overall performance increases with longer window lengths. A rise of $\sim 5\%$ in the window means can be noticed throughout from $W = 4s$ to $W = 10s$. This is corroborated with the average results of $W = 10s$ and $W = 8s$ being statistically distinct from the shorter windows' lengths groups. In addition, the pairwise differences between $W = 6, 7s$ and $W = 4, 5s$ are statistically significant for the *ACC*, *precision* and *F1* values. In the case of *SE*, the average values of the $W = 8s$ window are also statistically distinct from the $W = 4, 5, 6s$ lengths ($p < 0.01$), as well as $W = 7s$ is different from $W = 4s$ ($p < 0.001$).

6.3.2.3 Threshold of corruption (%) effect

Figure 6.5 shows the means plots of the classification performance metrics across different thresholds of corruption (%), for both artifacts and apnea classifiers. The statistical pairwise differences between various thresholds values ($Thd = 20\%, 30\%, 40\%, 50\%$) were displayed with asterisks as in Fig. 6.3.

In **artifacts classification**, it can be observed that the mean *ACC*, *SP* and *F1* decrease with increasing percentage of corruption threshold (%), whereas the opposite happens for *SE*. The drop in average *ACC* and *F1* performance from $Thd = 20\%$ to $Thd = 50\%$ is very subtle ($1 - 2\%$), whereas for *SP* it is a bit more meaningful with a 4% reduction. Indeed, the mean specificity values for all the *Thd* groups are statistically distinct from one another, with a p-value of $p < 0.01$ for $Thd = 20, 30\%$ and $p < 0.05$ for the $Thd = 40, 50\%$ groups. For the other performance metrics (*ACC*, *SE* and *F1*), due to the small changes in mean differences among groups, only the most extreme thresholds appear

to be statistically different. In fact, the pair $Thd = 20\%-Thd = 50\%$ accumulates the largest number of statistically significant differences overall, followed by $Thd = 20\%-Thd = 40\%$.

In **apnea classification**, the performance metrics' average values increased significantly with the threshold of corruption (%). The increment for ACC , SE and $F1$, was of around 5% from $Thd = 20\%$ to $Thd = 50\%$. The mean values of all the thresholds groups for these metrics were statistically different from one another ($p < 0.001$). The mean *precision* value for $Thd = 50\%$ was also statistically significant ($p < 0.05$) with respect to the rest of the threshold groups. However, the gain in *precision* from $Thd = 20\%$ to $Thd = 50\%$ was only of 2%.

6.3.3 Features selection results

Figure 6.6 displays the features selection frequency of occurrence (%) over the 30 randomization experiments, for the best artifacts and apnea classification models. The features were ranked in decreasing order. The most relevant features for each classification task were likely to be selected 100% of the times, while the the most irrelevant ones were never chosen for the final model in any of the 30 repetitions (0%).

For **artifacts classification**, in the upper panel of Fig. 6.6(a), a total of 26 features were selected 100% of the times, out of the 30 repetitions. Some examples are the *Amplitude*, *PeakHeightDiff*, *TroughDiff* and the corresponding *Changes* and *Standard deviations* of these. In the frequency domain, the *AvgPower* and *Spectral Entropy* features for all the specified bands were also some of the most important. In addition the *Envelope* characteristics were likewise predominantly selected. An additional set of 7 features that were chosen more than 50% of the times, showed good discriminative potential. But, the 18 lowest ranked features, appeared less than (30%) of the times in the final classification model.

In the lower panel of Fig. 6.6(b), it can be observed that a set of 24 features were selected in all randomization experiments (100%) for **apnea classification**. These mainly included time domain vertical characteristics of the signal (e.g. *PeakHeightDiff*, *ThroughDiff*), as well as the *Changes* and *StandardDeviations* of these features. All the *Envelope* characteristics (except *approxEntropy*) and the *Correlogram peaks*, were also part of the most highly selected features. In the frequency domain, the *AvgPower* (0-0.8Hz), *RelPower* (0.8-1.3Hz), *Spectral Entropy* (< 1.5Hz) and *Spectral Kurtosis* (< 1.5Hz) were also some of the most important features to consider for apnea detection. Besides the top (100%) features, another extra 9 were also significantly chosen more than 50% of the times. Among the rest of the 18 features selected in less than half of the 30 repetitions, the *Pulse Width*, *PeakDistance*, *SpectralEntropy* (1-4Hz) and *RelPower* (1.3-1.8Hz) were never picked for apnea classification.

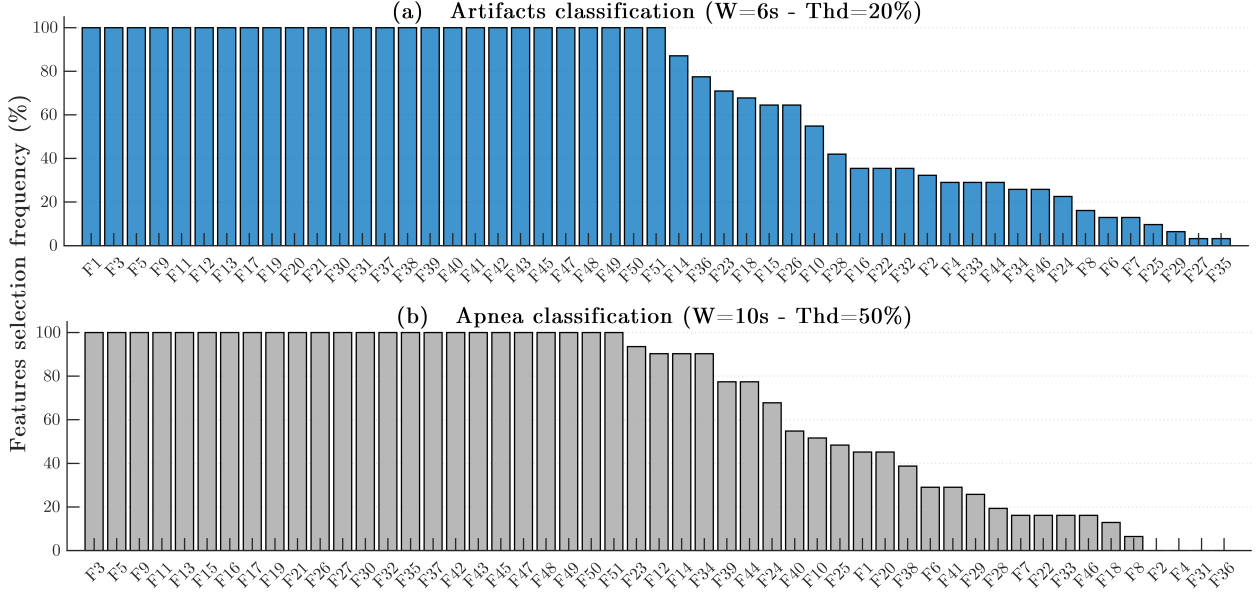


Figure 6.6: Features selection ranked by frequency of occurrence over the 30 randomization experiments for the best artifacts and apnea classification models. F_1 = Amplitude, F_2 = Width, F_3 = PeakHeightDiff, F_4 = PeakDistance, F_5 = TroughDiff, F_6 = RiseTime, F_7 = Skewness, F_8 = Kurtosis, F_9 = ChangeAmplitude, F_{10} = ChangeWidth, F_{11} = ChangePeakHeightDiff, F_{12} = ChangePeakDistance, F_{13} = ChangeTroughDiff, F_{14} = ChangeRiseTime, F_{15} = ChangeSkewness, F_{16} = ChangeKurtosis, F_{17} = StdAmplitude, F_{18} = StdWidth, F_{19} = StdPeakHeightDiff, F_{20} = StdPeakDistance, F_{21} = StdTroughDiff, F_{22} = StdRiseTime, F_{23} = StdSkewness, F_{24} = StdKurtosis, F_{25} = ZeroCrossingRate (ZCR), F_{26} = CorrelogramPeak1, F_{27} = CorrelogramPeak2, F_{28} = CorrelogramLag1, F_{29} = CorrelogramLag2, F_{30} = SpectralEntropy (0-1.5Hz), F_{31} = SpectralEntropy (1-4Hz), F_{32} = SpectralKurtosis (0-1.5Hz), F_{33} = SpectralKurtosis (1-4Hz), F_{34} = RelativePower (0-0.8Hz), F_{35} = RelativePower (0.8-1.3Hz), F_{36} = RelativePower (1.3-1.8Hz), F_{37} = AvgPower (0-0.8Hz), F_{38} = AvgPower (0.8-1.3Hz), F_{39} = AvgPower (1.3-1.8Hz), F_{40} = AvgPower (2.2-2.8Hz), F_{41} = AvgPower (3.2-3.8Hz), F_{42} = EnvelopeStd, F_{43} = EnvelopeMax, F_{44} = EnvelopeMin, F_{45} = EnvelopeRange, F_{46} = EnvelopeApproxEntropy, F_{47} = EnvelopeArea, F_{48} = EnvelopeAvgPower (0-0.15Hz), F_{49} = EnvelopeAvgPower (0.2-0.5Hz), F_{50} = EnvelopeAvgPower (0-0.5Hz), F_{51} = EnvelopeAvgPower (0.5-1Hz),

6.4 Discussion

In this chapter, two automated algorithms were developed to classify noise artifacts and detect apneic events from novel neck PPG signals. A total of 51 features from the time, correlogram and frequency domains were extracted to fit both classifiers. These included morphological, statistical, and envelope characteristics of the PPG signal, as well as PSD-derived features. A SVM classifier with a RBF kernel was trained for different windows ($W = 4, 5, 6, 7, 8$ and 10s) and thresholds of corruption ($Thd = 20\%, 30\%, 40\%, 50\%$). A LOSO-CV strategy was implemented to protect against overfitting and subject bias, during features selection and hyperparameters optimization. The classifiers were tested in unseen data, to predict whether each PPG window belonged to the *artifacts/clean PPG* classes; and whether within the clean PPG category, it was an *apnea/normal PPG* segment. This process was repeated 30 times with different randomizations of the data in order to evaluate the generalization capability of the models. Overall, the results demonstrated a good average performance for both classifiers ($\sim 86\%$). The standard deviations for the different (W/Thd) models were also small enough ($\sim 2\%$) to suggest that the algorithms were very stable and could generalize well accross data.

This increases the confidence that the results obtained could be reliably replicated in the future, with a similar range of values, no matter the data partition. Specially, for the artifacts' algorithm that is tested in a totally independent set of subjects (Leave-30%-of-Subjects-Out partition), the low variance indicates that the method is robust. However, some substantial differences in the performance metrics were observed among several (W/Thd) models.

The analysis of the features selected for the best (W/Thd) classification models indicated that overall, there was a recurrent set of features for each classifier, with a high chance ($\sim 100\%$) of being chosen. This suggested that features like *PeakHeightDiff* and *TrougtDiff*, as well as the corresponding *Changes* and *Standard deviations* of these, had a higher discriminative potential. The final set of features, also included *Envelope*, *AvgPower* and *Spectral Entropy* characteristics for specific frequency bands. However, around 18 features out of the total 51, were not selected many times (or even none), implying that they were not very informative for classification. The presented ranking of features offers, at hand, the most promising set of features for neck PPG artifacts classification and apnea detection. This analysis would be relevant for future studies aiming at processing neck PPG signals and improving the current classification results. It could likewise be a good starting point for additional features engineering for other related neck PPG applications.

For the **artifacts classification** results, the best W/Thd model, with the largest average accuracy ($85.84 \pm 2\%$), was $W = 6s - Thd = 20\%$. This model also maximized all the other performance metrics, except for SE which did not show statistical significance. Even though there is a decrease in performance from $W = 6s$ with increasing window length, the $W = 6s$ window group only appeared to be statistically distinct from $W = 10s$ in terms of ACC , SP and $F1$; and from $W = 8s$ in terms of SP . Therefore, it cannot be straightforwardly concluded that in general, $W = 6s$ is the most optimal window length for neck artifacts classification. But, since $W = 4, 5, 6$ and $7s$ are statistically equally valid, and $W = 6s$ slightly improves the overall performance, it would still be preferable to pick $W = 6s$ as the most suitable window for future algorithms. Indeed, other PPG studies have also found appropriate window lengths in a similar range for their proposed artifacts classifiers ([21, 33, 34]).

In terms of threshold of corruption (%), the classification performance decreased with larger Thd values. Specially, the average SP for the optimal $Thd = 20\%$ was statistically larger than the rest of groups, hence increasing the ACC and $F1$ too. This suggests that, in future works, a smaller threshold of corruption for window labelling, would considerably benefit the performance of the algorithm. However, if in turn, SE is deemed more important, a model with larger $Thd > 20\%$ would be recommended instead.

Comparing these results with other artifacts classification studies in the literature, leads to the conclusion that our algorithm performed well. Indeed, as it can be observed in Table 6.2, our model

Table 6.2: COMPARISON OF ARTIFACTS CLASSIFICATION RESULTS IN THE LITERATURE WITH OUR BEST ($W=6s$ - $Thd=20\%$) MODEL

	ACC (%)	SE (%)	SP (%)
Our method	85.8 ± 1.65	83.8 ± 4.1	87.43 ± 3.7
Couceiro [30]	87.5 ± 0.6	78.4 ± 1.2	94.4 ± 0.6
Chong [21]	93.9	94.3	92.4
Sukor [28]	83 ± 11	89 ± 10	77 ± 19
Tabei [34]	98.07 ± 2.02	92.6 ± 6.54	99.78 ± 0.93
Cherif [37]	83 ± 8	84 ± 16	83 ± 12
Fischer [24]	98.3	99.6	90.5

showed similar ACC , SE and SP than the SVM classifier proposed by Couceiro [30], the decision tree by Sukor [28] or the adaptive thresholding approach by Cherif [37]. However, some algorithms exploiting fine tuned decision lists (Fischer [24]), personalized neural networks (Tabei [34]), or a linear SVM with major voting (Chong [21]), outperformed our results. But, these are just for reference and are not straightforwardly comparable because each classification problem is distinct. The measurement sites in other works are different and consequently are susceptible to different types of artifacts. Different works also implement different validation strategies.

The findings of this artifacts classification model, are of great importance for denoising and conditioning novel neck PPG signals, and hence, enabling the possibility of exploiting this novel PPG measurement site for physiological monitoring. The removal of PPG-corrupted sections, would significantly improve the accuracy of HR and SpO_2 readings of neck pulse oximeter sensors. Ameliorating the quality of neck PPG signals, would similarly facilitate the accurate derivation of other biomarkers of interest.

In **apnea classification**, the average performance increased with the window length and the threshold of corruption (%) by a considerable amount ($> 5\%$), reaching its maximum at $W = 10s - Thd = 50\%$. In addition, both the $W = 10s$ window and the $Thd = 50\%$ threshold effects were shown to be statistically significant with respect to the other windows' and thresholds' groups for all the performance metrics. Therefore, it can be inferred that the $W = 10s - Thd = 50\%$ parameter's combination is the most suitable for detecting apnea events with neck PPG, as it maximizes not only the F1-score ($88.68 \pm 2.01\%$), but all the other performance metrics too ($ACC = 88.25\%$, $SE = 89.03\%$, $SP = 87.42\%$, $precision = 88.42\%$).

Since $W = 10s$ and $Thd = 50\%$ are the largest values in the ranges explored, in future studies the grid search bounds of the window length and threshold (%) parameters could be even expanded to investigate whether the performance could potentially improve. However, even though the choice of longer windows could benefit the detection, the reason behind proposing neck PPG signals as an alternative to common approaches, was to reduce the latency of apnea detection. So, increasing the

Table 6.3: COMPARISON OF APNEA CLASSIFICATION RESULTS IN THE LITERATURE WITH OUR BEST ($W=10s$ - $THD=50\%$) MODEL

	Signals used	ACC (%)	SE (%)	SP (%)	Precision (%)
Our method	<i>neck</i> PPG	88.25 \pm 2.07	89.03 \pm 2.69	87.42 \pm 3.63	88.42 \pm 3.04
Knorr-Chung [7]	PPG	75.4	91.6	84.7	85.9
Lázaro [10]	PPG	70.37	81.82	68.57	-
Papini [9]	PPG	86	39	94	51
Jung [4]	SpO ₂	91	83	89	-
Deviaene [6]	SpO ₂	82.8	64.3	88.6	64.2
Deviaene [8]	PPG+SpO ₂	83.4	73.7	86.6	64.8

window length to 30s or 1min segments, would limit the utility of the proposed method for real time applications. To illustrate, in the context of SUDEP, a longer window processing duration could increment the risk of mortality, as short apneic events might not be that promptly detected.

Reviewing other apnea detection approaches in the literature, the proposed RBF SVM model exploiting time and frequency characteristics directly derived from the PPG signal, outperformed both the studies exclusively extracting PPG features and the ones relying on the surrogate SpO₂ time series. As observed in Table 6.3, the *SE* and *precision* values of the SpO₂-based algorithms proposed by Deviaene *et al.* are poor ([6, 8]). In these approaches, features extraction focused on the signal segment corresponding to the oxygen desaturation, which is delayed from the actual respiratory apnea onset by 20-40s. This lag could be critical for real-time applications. The same issue applied to the work by Jung *et al.* [4]. Even though they claimed to accomplish real-time apnea detection by locating the original apneic event in the preceding 25 seconds prior to the onset of the desaturation; they first had to detect the lagged response of the SpO₂. Other SpO₂-based algorithms in the literature, which performed epoch-based classification with windows length of 1min or larger ([38, 39]), were likewise not suitable for real-time implementations.

Among the PPG works, the linear discriminant classifier proposed by Lázaro *et al.* [10], evaluating pulse rate variability (PRV) features from 4 windows preceding, following and spanning the delayed decreases in amplitude (DAP) events; also suffered from the same limitation. Papini *et al.* [9] achieved the highest specificity (*SP*) by inputting PPG-derived PRV and respiratory features into a deep learning model, but the *SE* and *precision* were insufficient for robust online monitoring. The results obtained by Knorr-Chung *et al.* [7], with an ANN trained on PPG time and frequency characteristics, were good but the classification model was not implemented in an epoch-by-epoch online manner. Instead, the most representative PPG fragments showing normal breathing and apneic patterns, were manually segmented for classification.

This work, in contrast, is a significant advancement in the field, since it demonstrates, for the first

time in literature, that it is possible to robustly detect apnea events from neck PPG signals in an instantaneous manner. The proposed method has the advantage to be simple and has the potential to be used for near real-time applications, hence lifting the long waiting times of offline processing. As a consequence, in the future, the large number of cumbersome PSG sensors could be reduced to a unique, wearable neck PPG system. This could potentially have a great impact in SUDEP monitoring, by supporting airflow measurements in the decision of apnea classification.

Overall, the methods in this work present useful recommendations for future designers of neck PPG processing algorithms, in terms of suggested features, window lengths, labelling thresholds and classification models. This is important for future adoption of the neck as a PPG site. Indeed, the proposed artifacts classification algorithm presents the first proof-of-concept classifier for neck PPG artifacts removal. However, once the corrupted PPG fragments are identified, a decision on how to process them should be taken. This study was devised with the idea that corrupted fragments could just be discarded, to improve the accuracy of HR and SpO₂ parameters estimation. It does not tackle, however, the reconstruction of detected artifact signals. This should be explored in future work, specially when artifacts are expected to be predominant. Another limitation of this study is that the proposed classification models were trained using experimental artifacts or breath-holding events. These need to be tested in real sleep scenarios to validate their performance. Also, a wider number of participants, including patients suspected to have apneas, should be recruited. Indeed, the majority of studies developing apnea detection algorithms in the literature, make use of polysomnography databases, with real apneas of different kinds (obstructive, central, mixed). The accuracy of the current apnea algorithm, would probably decrease when tested against this variety of respiratory events.

Future work, should then focus on combining complementary respiratory signals measuring air-flow [2], to support the classification decision and improve the performance. Tracheal sounds, for example, can be easily sensed from the multipurpose site of the neck. Further improvements and validation of this proof-of-concept would ideally lead to the implementation of these classifiers in the WADD wearable apnea monitoring system for SUDEP prevention.

6.5 Conclusion

In order to fully exploit the novel PPG measurement site of the neck, specifically to support real-time apnea detection applications, corrupted PPG segments need to be first recognized for removal. Two automatic algorithms were designed in this work to achieve these. The first classifier demonstrated good performance in distinguishing neck PPG-corrupted segments from clean PPG data; and the second, showed a promising capability of promptly detecting apneic events, in a near real-time manner,

both uniquely exploiting PPG time and frequency features. The preliminary results of this study, provide useful tools to facilitate neck PPG signals processing, that could encourage the future usage of the neck as a PPG measurement site.

References

- [1] F. Mendonca, S. S. Mostafa, A. G. Ravelo-García, F. Morgado-Dias, and T. Penzel, “A review of obstructive sleep apnea detection approaches,” *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 825–837, 2018.
- [2] M. Uddin, C. Chow, and S. Su, “Classification methods to detect sleep apnea in adults based on respiratory and oximetry signals: a systematic review,” *Physiological measurement*, vol. 39, no. 3, p. 03TR01, 2018.
- [3] V. Monasterio, F. Burgess, and G. D. Clifford, “Robust classification of neonatal apnoea-related desaturations,” *Physiological measurement*, vol. 33, no. 9, p. 1503, 2012.
- [4] D. W. Jung, S. H. Hwang, J. G. Cho, B. H. Choi, H. J. Baek, Y. J. Lee, D.-U. Jeong, K. S. Park *et al.*, “Real-time automatic apneic event detection using nocturnal pulse oximetry,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 706–712, 2017.
- [5] P. I. Terrill, “A review of approaches for analysing obstructive sleep apnoea-related patterns in pulse oximetry data,” *Respirology*, vol. 25, no. 5, pp. 475–485, 2020.
- [6] M. Deviaene, D. Testelmans, B. Buyse, P. Borzée, S. Van Huffel, and C. Varon, “Automatic screening of sleep apnea patients based on the SpO₂ signal,” *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 607–617, 2018.
- [7] B. R. Knorr-Chung, S. P. McGrath, and G. T. Blike, “Identifying airway obstructions using photoplethysmography (PPG),” *Journal of clinical monitoring and computing*, vol. 22, no. 2, pp. 95–101, 2008.
- [8] M. Deviaene, J. Lázaro, D. Huysmans, D. Testelmans, B. Buyse, S. Van Huffel, and C. Varon, “Sleep apnea detection using pulse photoplethysmography,” in *2018 Computing in Cardiology Conference (CinC)*, vol. 45. IEEE, 2018, pp. 1–4.
- [9] G. B. Papini, P. Fonseca, M. M. van Gilst, J. W. Bergmans, R. Vullings, and S. Overeem, “Wearable monitoring of sleep-disordered breathing: estimation of the apnea–hypopnea index using wrist-worn reflective photoplethysmography,” *Scientific reports*, vol. 10, no. 1, pp. 1–15, 2020.
- [10] J. Lázaro, E. Gil, J. M. Vergara, and P. Laguna, “Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children,” *IEEE journal of biomedical and health informatics*, vol. 18, no. 1, pp. 240–246, 2013.
- [11] E. Gil, J. M. Vergara, and P. Laguna, “Detection of decreases in the amplitude fluctuation of pulse photoplethysmography signal as indication of obstructive sleep apnea syndrome in children,” *Biomedical Signal Processing and Control*, vol. 3, no. 3, pp. 267–277, 2008.
- [12] I. Garcia-Lopez and E. Rodriguez-Villegas, “Characterization of artifact signals in neck photoplethysmography,” *IEEE Transactions on Biomedical Engineering*, 2020.

-
- [13] I. García-López, S. A. Imtiaz, and E. Rodriguez-Villegas, "Characterization study of neck photoplethysmography," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* IEEE, 2018, pp. 4355–4358.
 - [14] G. Joseph, A. Joseph, G. Titus, R. M. Thomas, and D. Jose, "Photoplethysmogram (ppg) signal analysis and wavelet de-noising," in *Annu. Int. Conf. IEEE on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD)*, 2014, pp. 1–5.
 - [15] A. K. Bhoi, S. Sarkar, P. Mishra, and G. Savita, "Pre-processing of ppg signal with performance based methods," *International Journal of Computer Application*, vol. 4, no. 2, pp. 251–256, 2012.
 - [16] K. A. Reddy, B. George, and V. J. Kumar, "Use of fourier series analysis for motion artifact reduction and data compression of photoplethysmographic signals," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 5, pp. 1706–1711, 2009.
 - [17] B. S. Kim and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 566–568, 2006.
 - [18] J. Lee, M. Kim, H.-K. Park, and I. Y. Kim, "Motion artifact reduction in wearable photoplethysmography based on multi-channel sensors with multiple wavelengths," *Sensors*, vol. 20, no. 5, p. 1493, 2020.
 - [19] H. H. Asada, H.-H. Jiang, and P. Gibbs, "Active noise cancellation using mems accelerometers for motion-tolerant wearable bio-sensors," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 1, 2004, pp. 2157–2160.
 - [20] H. Han, M.-J. Kim, and J. Kim, "Development of real-time motion artifact reduction algorithm for a wearable photoplethysmography," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 1538–1541.
 - [21] J. W. Chong, D. K. Dao, S. Salehizadeh, D. D. McManus, C. E. Darling, K. H. Chon, and Y. Mendelson, "Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection–reduction approach. Part I: motion and noise artifact detection," *Ann. Biomed. Eng.*, vol. 42, no. 11, pp. 2238–2250, 2014.
 - [22] S. Salehizadeh, D. K. Dao, J. W. Chong, D. McManus, C. Darling, Y. Mendelson, and K. H. Chon, "Photoplethysmograph signal reconstruction based on a novel motion artifact detection–reduction approach. Part II: Motion and noise artifact removal," *Ann. Biomed. Eng.*, vol. 42, no. 11, pp. 2251–2263, 2014.
 - [23] R. Krishnan, B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1867–1876, 2010.
 - [24] 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*, vol. 21, no. 2, pp. 372–381, 2017.
 - [25] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, "Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring," *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 832–838, 2014.
 - [26] R. Krishnan, B. Natarajan, and S. Warren, "Analysis and detection of motion artifact in photoplethysmographic data using higher order statistics," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE Int. Conf. on.* IEEE, 2008, pp. 613–616.

- [27] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in photoplethysmogram," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 4972–4975.
- [28] J. A. Sukor, S. Redmond, and N. Lovell, "Signal quality measures for pulse oximetry through waveform morphology analysis," *Physiological measurement*, vol. 32, no. 3, p. 369, 2011.
- [29] N. Pradhan, S. Rajan, A. Adler, and C. Redpath, "Classification of the quality of wristband-based photoplethysmography signals," in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2017, pp. 269–274.
- [30] R. Couceiro, P. Carvalho, R. P. Paiva, J. Henriques, and J. Muehlsteff, "Detection of motion artifact patterns in photoplethysmographic signals based on time and period domain analysis," *Physiological measurement*, vol. 35, no. 12, p. 2369, 2014.
- [31] F. Tabei, R. Zaman, K. H. Foysal, R. Kumar, Y. Kim, and J. W. Chong, "A novel diversity method for smartphone camera-based heart rhythm signals in the presence of motion and noise artifacts," *PloS one*, vol. 14, no. 6, p. e0218248, 2019.
- [32] T. Pereira, K. Gadhomi, M. Ma, X. Liu, R. Xiao, R. A. Colorado, K. J. Keenan, K. Meisel, and X. Hu, "A supervised approach to robust photoplethysmography quality assessment," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 3, pp. 649–657, 2019.
- [33] Q. Li and G. Clifford, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiological measurement*, vol. 33, no. 9, p. 1491, 2012.
- [34] F. Tabei, R. Kumar, T. N. Phan, D. D. McManus, and J. W. Chong, "A novel personalized motion and noise artifact (mna) detection method for smartphone photoplethysmograph (ppg) signals," *IEEE Access*, vol. 6, pp. 60 498–60 512, 2018.
- [35] C.-H. Goh, L. K. Tan, N. H. Lovell, S.-C. Ng, M. P. Tan, and E. Lim, "Robust ppg motion artifact detection using a 1-d convolution neural network," *Computer Methods and Programs in Biomedicine*, p. 105596, 2020.
- [36] MathWorks, "approximateEntropy: Measure of regularity of nonlinear time series," <https://uk.mathworks.com/help/predmaint/ref/approximateentropy.html>, 2020.
- [37] S. Cherif, D. Pastor, Q.-T. Nguyen, and E. L'Her, "Detection of artifacts on photoplethysmography signals using random distortion testing," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2016, pp. 6214–6217.
- [38] S. S. Mostafa, F. Mendonça, A. G. Ravelo-Garcia, G. G. Juliá-Serdá, and F. Morgado-Dias, "Multi-objective hyperparameter optimization of convolutional neural network for obstructive sleep apnea detection," *IEEE Access*, vol. 8, pp. 129 586–129 599, 2020.
- [39] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, "An oximetry based wireless device for sleep apnea detection," *Sensors*, vol. 20, no. 3, p. 888, 2020.

Chapter 7

Conclusions

In this thesis, neck photoplethysmography, has been thoroughly studied for the first time in literature. Neck PPG signals have been recorded under different respiratory and artifact conditions, to provide a full characterization of this novel measurement site. Extensive processing of the key features of the proposed signals, has enabled the design of two artifacts classification and apnea detection algorithms. The neck has been proven to be a very promising pulse oximetry measurement location, for future application in cardiorespiratory monitoring, and specially within the context of sudden unexpected death (SUDEP) prevention.

7.1 Contributions

The original contributions of this thesis are summarized as follows.

Neck PPG, extensively studied, for the first time.

Although the finger is the gold standard measurement site in pulse oximetry, the neck, can be very attractive for sensing several physiological signals, besides oximetry, with the same monitoring device. Specifically for respiratory diseases, the body region of the neck is particularly suited for sensing airflow from tracheal sounds. In combination with oxygen saturation, airflow measurement is of great relevance for apnea detection. Nevertheless, neck PPG signals have not received much attention in the literature so far, since, at first glance, the neck region could appear sub-optimal for sensor attachment and patient's comfort. The very limited works have only proven the feasibility of sensing it in a reduced number of subjects, and very controlled conditions.

In this thesis, the neck PPG has been explored in depth for the first time in literature, in terms of pulse characteristics, spectral content at different breathing rhythms and artifact conditions. It is

also the first time that this novel measurement site has been thoroughly compared to standard finger PPG, to assess its advantages and limitations.

In *Chapter 3*, the principal differences between neck and finger PPG were established in terms of morphological pulse shape characteristics. For instance, the neck PPG waveform showed a higher diastolic peak, probably explained by a reduced number of bifurcations in the arterial tree transmission path, compared to when the pressure wave travels to the periphery (finger). This and other pulse contour findings defined the normal traits of neck PPG pulses.

When inspecting the spectral differences between neck and finger PPG signals under different breathing conditions, the respiratory frequency component was found more prominent in the novel site. This singular property could facilitate the extraction of the respiratory rate in different breathing situations. The suitability of neck PPG for apnea detection is hence two-fold: not only $\text{SpO}_2\%$ desaturations could indicate hypoxia, but also the absence of respiration could be simply detected by examining the power spectral content of neck PPG.

Besides exploring the time and frequency characteristics, neck PPG was shown to be a promising pulse oximetry measurement site, as precise HR and SpO_2 (%) were estimated compared to ground truth values. However, for this proof-of-concept study, the extraction of these two vital parameters was not tested under artifact conditions. This demonstrated the necessity to further study the unknown artifacts affecting the novel site of the neck, in order to further develop filtering algorithms and ensure reliable readings.

Tools for neck PPG signal processing.

Another important contribution of this thesis was to provide tools to facilitate the use and the processing of neck PPG signals, including the characterization of the most common artifacts, the identification of the most discriminative features, and the development of validated classification models, with the most optimal hyperparameters, to reliably detect artifacts and apnea events. Having all these at hand, could boost the future adoption of this PPG measurement site in cardiorespiratory monitoring.

In *Chapter 5*, the 10 most common neck PPG artifacts, and two other breathing states of interest, were defined and statistically compared to normal PPG. An exhaustive characterization study presented a detailed description of the most discriminative features differentiating each artifact from clean PPG. Indeed, artifacts severely distorted the PPG signal with abnormally larger and highly variable amplitude related characteristics, as well as higher spectral entropy frequency features, among others. The breathing states of interest, Breathing Slow and Apnea, shared stable characteristics with clean PPG, implying they could be easily differentiated from artifacts too. These results are of relevant importance for denoising neck PPG signals, since the sources of artifacts susceptible to corrupt them, are now well known and detectable. These findings, could provide valuable tools to future engineers

devising their signal processing pipelines to detect neck artifacts, as they will have direct access to the best tuned features for their specific application.

Chapter 6, expanded this features characterization analysis into the development of two SVM-RBF classification models. By additionally exploring respiratory envelope-related characteristics, an artifacts classifier and an apnea classifier were trained. These learned to differentiate clean PPG signals from artifacts, and normal PPG from apnea PPG events, respectively. Features were automatically chosen during training using a sequential feature selection approach. The proposed artifacts classification model achieved good performance in comparison to other algorithms in the literature. This could now help remove the corrupted fragments of the PPG signal, and thus ensure reliable physiological parameters extraction. The proposed apnea detection model also outperformed other works. These findings show, for the first time, that neck PPG signals can be reliably exploited for apnea detection. The determination of the best window length (W) and threshold of corruption (Thd) parameters will serve as guidance to future researchers developing neck PPG classification algorithms, in relation to what detection window to choose in different scenarios. The recurrent optimal subset of features selected by each classifier could similarly suggest a good starting point for feature engineering in future research.

Last but not least, all the experiments carried out in this research suggested that the suprasternal notch of the neck is a valid location to extract arterial PPG. However, suggestions on alternative positioning of the sensor on top of the anterior veins, also opened a window into the venous system.

Beyond pulse oximetry...

Apart from being a suitable pulse oximetry alternative, neck PPG signals could also have a huge potential for other cardiorespiratory applications.

In *Chapter 4*, I demonstrated, for the first time, the feasibility of extracting the jugular venous pulse (JVP) in a novel way: exploiting contact neck PPG. The JVP is an important physiological signal that provides unique clinical information about right atrial and central venous pressure (CVP) abnormalities. It is a key biomarker in cardiovascular diseases (CVD) diagnosis, as several cardiac dysfunctions affecting the right side of the heart (e.g. tricuspid stenosis), cannot be diagnosed with common ECG, PPG, or blood pressure signals. However, the gold standard method to measure JVP, invasive central venous catheterization, entails an enormous amount of risks. This is why, demonstrating that it is possible to extract the JVP from the anterior jugular veins with contact photoplethysmography, can potentially overcome the shortcomings of the current state-of-the-art. This was shown by identifying the conventional a, c, v waves of the JVP, in the recorded neck JVP-PPG signals, as established by cardiac-cycle physiology and validated by B-mode ultrasound imaging. The findings of this chapter thus confirmed that neck JVP-PPG is a very promising, safe and cost-

effective tool, that could be an important breakthrough for the future of CVD diagnosis, since it eliminates all the risks associated with invasive catheterization. This work is of great significance to enable continuous non-invasive point-of-care monitoring of CVP.

This is only one example of a CVD application for which neck PPG signals have a great potential. However, further study of the neck *arterial* PPG waveform in a wider population, could also reveal new insights into the head vasculature function. The analysis of contour features of the PPG pulse shape can provide a large amount of information, not only in relation to the pulsatile function of the heart, but also regarding the arteries distensibility properties. For example, by measuring variations in the diastolic peak height and occurrence time, arterial stiffening could be detected. Arterial neck PPG could therefore be used as a new non-invasive diagnostic tool to identify carotid artery blocks, atherosclerosis, or to monitor the aging process of the upper vascular system.

The neck has also been proven to be a propitious measurement site for respiratory diseases monitoring, since the respiratory frequency component was shown to be more prominent than in finger. This was exploited in *Chapter 6*, in the development of the apnea classification algorithm, by engineering envelope-specific features. Indeed, the findings of this chapter show, for the first time in literature that, not only, apneic events can be detected using neck PPG signals, but also, that this can be achieved in a near real-time manner. In contrast to other works that detect delayed SpO₂ desaturations (20-40s), or implement long windows of detection (>30s), our method, using a 10s window shifted every 2s, is a significant advancement in the field. The proposed instantaneous apneic detection, can be essential in critical care monitoring applications such as in SUDEP prevention, since latency of other approaches could result in fatal consequences. These findings could then help investigate the mechanisms and risk factors of SUDEP, and ultimately reduce epileptic patients' mortality.

Overall, the findings of this thesis support the idea that, neck PPG is, in a wide number of healthcare contexts, a feasible pulse oximetry alternative. However, it is important to stress, that neck PPG would only be beneficial for stationary medical applications, such as sleep or in-patient monitoring. For other PPG ambulatory activities like continuous tracking of heart rate during daily activities, the aesthetics of the patient could prevent adoption. In these cases wrist-watch PPG devices could be more appealing to users, compared to a wearable device stucked to the middle of their neck, no matter how small this could be.

7.2 Future work

7.2.1 Areas of improvement of this PhD

This exploratory work was a first proof-of-concept on how to record, analyze and exploit novel neck PPG signals for different cardiorespiratory applications. Consequently, the experimental procedures were quite restricted to a small number of subjects, commercial sensors available and very controlled recording conditions. All these aspects should be improved in the future in order to have a better understanding on how neck PPG signals could be implemented in real scenarios.

For that, I would suggested to future researchers:

- To use a custom neck wearable PPG system, accordingly calibrated to this measurement site. In this research a forehead-based reflectance PPG sensor was used. However, this was not calibrated for the specific measurement site of the neck, being hence useless for $\text{SpO}_2\%$ estimation. Also, even though the LED irradiance was valid for neck PPG sensing, potential differences in tissue characteristics, such as a higher fat content, could make this solution no longer optimal for signals obtention. In these situations where a higher concentration of light absorbers is present, it would be desirable to increment the signal-to-noise-ratio (SNR) by adding, for example, more light sources to the sensor design.
- To recruit a larger number of participants, across different age groups, with an equal ratio of males and females, a wider BMI range and more diverse skin tones. A thorough investigation on how these factors might impact the recording of neck PPG signals, could further validate the results of this thesis or open new research questions to ensure future adoption of the neck site.
- To further explore the neck region in order to try to identify the most optimal recording location to either uniquely obtain the arterial PPG or the venous JVP-PPG. This might be one of the biggest challenges, since the obtention of neck PPG has been shown to be highly dependent on participants' anatomy. However, a pre-screening of the anterior veins with an IR vein camera could provide specific sensor placement recommendations for each independent user.
- To record more breath-holding apnea events per subject to have a larger and more representative number of samples of this class, for future algorithms development. This could benefit the robustness and sensitivity of classifiers, and avoid any possible bias coming from class imbalance issues.
- To record neck PPG signals in real sleep conditions, to evaluate whether the features and classifiers engineered in this thesis, can be straightforwardly applied in a future detection system, or need to be further fine-tuned.

7.2.2 Towards a neck PPG wearable apnea detection device

The findings of this PhD work, have paved the way for the development of a wearable neck-based PPG apnea monitoring system. In this section, the future directions to extend the results achieved in this thesis are described.

7.2.2.1 Artifacts algorithm validation

Initially, the design of the complete filtering strategy, to obtain reliable neck PPG signals, should be of primary focus. The idea is to only filter the PPG sections detected as corrupted to avoid distorting the rest of the clean PPG signal. For that, once the proposed artifacts classification algorithm automatically detects corrupted neck PPG fragments, a decision should be made on how to process them. It should be decided whether the artifacts are worth to be reconstructed, or whether they can be simply discarded.

The performance of this algorithm in accurately detecting and removing artifacts, should then be validated in a wider cohort of participants and real sleep conditions. In this stage, the developed algorithm should be implemented in real-time and tested in a series of sleep monitoring experiments.

As a result, artifacts-free PPG signals will enable a reliable estimation of blood oxygen saturation levels ($\text{SpO}_2\%$), as well as other vital physiological parameters such like heart rate.

7.2.2.2 Apnea algorithm validation in patients

Once the artifacts removal algorithm is validated, it has to be combined into a single classification pipeline with the apnea detection algorithm. However, the proof-of-concept apnea classifier presented in this thesis, was trained on breath-holding events simulated by healthy participants. Testing the apnea detection performance in patients prone to have apneas becomes necessary before integration into a system.

The recording of neck PPG signals in real apneic conditions, could also provide new insights into apnea desaturation characteristics at this measurement site, for the fine-tuning of the classification model. For example, the SpO_2 signal could be additionally exploited to design new features based on the delayed desaturations. The combination of these, with PPG signal features, would increase the pool of apnea biomarkers. This could ultimately reinforce the decision of the classifier.

7.2.2.3 Low power implementation and WADD sensor fusion

Once both models are tested in real sleep conditions, a low power version of the combined algorithm, could be integrated with the wearable apnea detection device (WADD). Indeed, a fusion of the existing apnea detector based on acoustic signals, with the novel PPG parameters, is desired to improve

the performance of the WADD system. However, an updated wearable prototype would need to be designed to accomodate both, the microphone for tracheal sounds monitoring, and the PPG technology for oximetry sensing.

A combined decision of the PPG-based apnea features and airflow acoustic information, could significantly enhance the detection of apneas. Additionally, the final wearable apnea monitoring system could be coupled with an accelerometer and a vibrator/alarm, to enable positional therapy for SUDEP prevention. As a consequence, the improved system could have a meaningful impact in reducing mortality due to SUDEP.

7.2.2.4 Clinical trials

After the integration of the new WADD-PPG system, a pilot study with a polysomnography system could be run to assess the improved detection.

Ultimately, clinical trials would be required for complete validation, prior to bringing a ready-to-use wearable apnea monitoring device to the hands of epileptic patients.

Appendix A

Ethics approval documents

The experiments performed in this research work were approved by the Local Ethics Committee of Imperial College London (ICREC ref.: 18IC4358).

Relevant documents of the Ethics Application are attached as follows:

A.1 Joint Research Compliance Office approval letter

A.2 Experimental Protocol

A.3 Participant Information Sheet

A.4 Informed Consent Form

A.1 Joint Research Compliance Office Approval Letter

Imperial College
London

Imperial College Research Ethics Committee
Imperial College London
Room 221
Medical School Building
St Marys Campus
London
W2 1PG
Tel: +44 (0)207 594 9484
researchethicscommittee@imperial.ac.uk

Esther Rodrigues Villegas
914, Level 9
Department of Electrical and Electronic Engineering
South Kensington Campus
Imperial College, SW7 2AZ

20th February 2018

Dear Professor Rodriguez Villegas

Study Title: Non-invasive photoplethysmography signals at the neck of healthy volunteers

ICREC reference: 18IC4358

The above study was approved by your Head of Department on 20/12/17 and by the Joint Research Compliance Office on 20th February 2018.

Under the Imperial College Research Ethics Committee process, a study that has been reviewed by the Joint Research Compliance Office and Head of Division/Department (or Principal), where no significant ethical issues have been identified in the protocol or ethics application, can be approved without requiring it to go to full committee.

Documents

The documents reviewed were:

- ICREC Application form
- Protocol (v3 19/02/118)
- Recruitment Email (v3 19/02/118)
- Informed Consent Form (v3 19/02/118)
- Participant Information Sheet (v3 19/02/118)
- Sponsorship and Insurance Request
- Questionnaire

Yours sincerely,



Ruth Nicholson,
Research Governance Manager,
Imperial College London

A.2 Experimental Protocol

Non-invasive photoplethysmography signals collection at the neck of healthy volunteers

WPPG

Version 3 – 19/02/2018

MAIN SPONSOR: Imperial College London

FUNDERS: European Research Council

STUDY COORDINATION CENTRE: NA

NRES reference: NA

Protocol authorised by:		
Name & Role	Date	Signature

Study Management Group

Chief Investigator: Professor Esther Rodriguez Villegas

Co-investigators: Irene Garcia

Statistician: NA. The main aim of this study is to record physiological signals that have never been recorded before, for the purpose of carrying out new research on the identification of respiratory biomarkers. Since these signals are new it is not possible to make any assumption for statistical calculations.

Study Management: NA

Clinical Queries

Clinical queries should be directed to Prof. Esther Rodriguez-Villegas who will direct the query to the appropriate person.

Sponsor

Imperial College London is the main research Sponsor for this study. For further information regarding the sponsorship conditions, please contact the Head of Regulatory Compliance at:

Joint Research Compliance Office
5L10C, 5th Floor Lab Block
Charing Cross Hospital Fulham Palace Road
London W6 8RF
Tel: 0203 311 0204
Fax: 0203 311 0203

Funder

This protocol describes the project **'Non-invasive photoplethysmography signals collection at the neck of healthy volunteers'** study and provides information about procedures for entering participants. Every care was taken in its drafting, but corrections or amendments may be necessary. These will be circulated to investigators in the study. Problems relating to this study should be referred, in the first instance, to the Chief Investigator.

This study will adhere to the principles outlined in the NHS Research Governance Framework for Health and Social Care (2nd edition). It will be conducted in compliance with the protocol, the Data Protection Act and other regulatory requirements as appropriate.

TABLE OF CONTENTS

1. INTRODUCTION.....	5
1.1 BACKGROUND.....	5
1.2 RATIONALE FOR CURRENT STUDY.....	5
2. STUDY OBJECTIVES AND DESIGN	6
2.1 METHODOLOGY	6
2.2 STUDY OUTCOME MEASURES.....	7
3. PARTICIPANT ENTRY	8
3.1 PRE-REGISTRATION EVALUATIONS.....	8
3.2 INCLUSION CRITERIA	8
3.3 EXCLUSION CRITERIA.....	8
3.4 WITHDRAWAL CRITERIA	8
4. ADVERSE EVENTS	8
4.1 DEFINITIONS.....	8
4.2 REPORTING PROCEDURES	9
5. ASSESSMENT AND FOLLOW-UP	9
6. DATA ANALYSIS AND HANDLING.....	9
7. REGULATORY ISSUES.....	9
7.1 ETHICS APPROVAL	9
7.2 CONSENT.....	9
7.4 INDEMNITY	10
7.5 SPONSOR	10
7.6 FUNDING.....	10
8. STUDY MANAGEMENT	10
9. REFERENCES.....	11

GLOSSARY OF ABBREVIATIONS

PPG	Photoplethysmography
SUDEP	Sudden Unexpected Death in Epilepsy
WADD	Wearable Apnoea Detection Device

KEYWORDS

Photoplethysmography, Artefacts, Apnoea, Oxygen desaturation, Signal processing.

STUDY SUMMARY

TITLE	Non-invasive photoplethysmography signals collection at the neck of healthy volunteers
DESIGN	<ul style="list-style-type: none"> ▪ Simultaneous collection of physiological signals, including PPG, with wearable sensors in conventionally used locations like the finger and wrist, and atypical ones like the neck. ▪ Participants will be asked to carry out breathing exercises in conventional real-life body positions- like lying down in a supine position to simulate sleeping posture; including inducing different artefacts (also representative of real life scenarios, such as movements that could arise during sleep).
AIMS	To investigate the characteristics of physiological signals in locations of the body not-conventionally used, since those locations, although sub-optimal for the specific signal, could facilitate multimodal recordings from just one wearable device (i.e. small, battery powered, non-invasive device that can be worn). For example, the neck would not be the location of choice for PPG but is optimum to detect breathing sounds.
OUTCOME MEASURES	<ol style="list-style-type: none"> 1. Identify the differences between finger and neck PPG pulse waveforms. 2. Identify what artefacts (also called noise and interference) affect the PPG signal in this part of the body and characterize them to be able to design a filtering technique to obtain a clean signal. 3. Obtain PPG recordings at the novel site of the neck for analysis and further development of signal processing algorithms to obtain physiological biomarkers (potentially in combination with the information provided by other signals).
POPULATION	Around 60 healthy volunteers
ELIGIBILITY	<ul style="list-style-type: none"> • Age: in between 18 and 70 years old • Fluent in English or Spanish • No pacemakers
DURATION	Approximately 3 years.

1. INTRODUCTION

1.1 BACKGROUND

Sudden unexpected death in epilepsy (SUDEP) is one of the most common causes of mortality among epileptic patients, with up to 9 deaths per 1000 patients each year [1]. The incidence in people with epilepsy is also more than 20 times that in the general population. It is difficult to describe such a phenomenon, but a widely adopted definition of SUDEP is: "a sudden, unexpected, witnessed or not witnessed, non-traumatic and non-drowning death in patients with epilepsy, with or without evidence of a seizure and excluding documented status epilepticus, in which post-mortem examination does not reveal a toxicological or anatomical cause of death" [2].

The exact mechanisms leading to SUDEP in epilepsy are still unclear but there is evidence that seizure-related oxygen desaturation apnoea events might play a crucial role [3, 4]. Indeed, sleep apnoeas disorders disturb the oxygen supply to vital organs such as the brain or the heart, by inducing obstruction of the upper airways or cessation of thoracic respiratory movements [5]. Hypoxemia occurs in 1/3 of generalized and non-convulsive seizures [6].

In this framework, a new Wearable Apnoea Detection Device (WADD) sensing acoustic signals of breath placed externally at the neck, was developed. This new device was tested in a pilot study in comparison to a standard polysomnography system [8]. A sensitivity of 88.6% and a specificity of 99.6% in detecting apnoeas even in the presence of artefacts, showed a promising applicability for SUDEP prevention. Nevertheless, the reduced sensitivity of the WADD to hypoapnoea detection could be critical for epileptic patients. This means that during hypoapnoea, when airways are partially obstructed, and a reduced amount of air intake entails a decrease of oxygen levels [9]; this lethal ventilatory dysfunction would not be so accurately detected by the WADD.

Therefore, to avoid neglecting life-threatening hypoapnoea events, it would be necessary to also be able to automatically determine oxygen saturation values from photoplethysmography (PPG) signals obtained in the neck; since the alternative of sensing different signals from various locations in the body would not be realistic considering the usability requirements imposed by the medical devices standards both in the EU and in other countries.

Epilepsy is just one example of a condition that would benefit from having more than one health-related signal sensed using just one wearable sensor from the same location in the body (this could be the neck or a different location)

1.2 RATIONALE FOR CURRENT STUDY

Photoplethysmography (PPG) is a non-invasive way of detecting changes in blood volume, using light, shining at the surface of the skin. Together with computer algorithms this can be used to determine how efficiently some organs, such as the lungs and heart, work. For example, from the PPG signal it is possible to have an indication of the amount of oxygen in the blood. This is called pulse oximetry. PPG is a widely used technique in both, hospitals and in non-medical products. As an illustration, an example of a product that uses PPG is the Apple watch.

The best possible measurement location for PPG is the finger. However, the finger is not the most convenient location for all health-related applications. It is for example not suitable for prevention of Sudden Unexpected Death in Epilepsy (SUDEP). Hence having alternative measuring sites, such as the neck, would be in some cases desirable. One of the reasons for

this is that the finger is not the best body location to measure additional signals, from which other health related parameters, which are important in the context of other diseases, could be detected. Places like the neck however would be far better to allow a larger number of health-related parameters to be measured simultaneously using just one wearable (i.e. a little, battery powered, non-intrusive, non-invasive device worn in the body). The neck however, has not been explored as a PPG measurement site, but it is a priori known that it is a tricky location from the engineering point of view since the signal we are looking for will be weaker than in the finger and will be also “spoilt” due to interference from other signals (which we call “artefacts”). It is for these reasons, that collection of body signals in healthy volunteers from not the typically used locations (such as the neck, but also others) becomes necessary; so that we can carry out research on what the characteristics of these signals are and how we can clean them from the interference of other non-desired signals.

Therefore, the goal of this study is to record new signals from wearable sensors and their inherent artefacts, which will allow us to carry out research leading to novel signal processing algorithms that will enable accurate multimodal recording (more than one signal measured) from just one wearable sensor.

2. STUDY OBJECTIVES AND DESIGN

The study will consist on simultaneously collecting physiological signals (ie. Signals generated by the body), including PPG (see explanation in Section 1.2), from the neck, and signals from other conventionally used non-invasive physiological monitoring locations (such as the finger) of healthy participants under different conditions. For example, a reflectance pulse oximeter sensor will be placed in contact with the anterior neck, and a transmission pulse oximeter sensor will be placed in the index finger as a gold standard.

2.1 METHODOLOGY

Volunteer participants will be welcomed by a research team member in the Department of Electrical and Electronic Engineering at Imperial College London. The experiments will be explained to them, and subsequently, they will be given the information sheet and offered some privacy to read it. When themselves express they are ready, all their questions will be answered. It will only be after that, if they express their desire to continue that the consent form will be given to them. As part of the admission process participants will be explicitly told that they can withdraw at any point in the experiment, and also that they can choose at any point to be themselves, instead of the investigators who put the wearable sensors on.

After the wearable sensors are set up, the signal collection protocol will commence. Throughout, participants will be asked to be in different positions -including lying horizontally with the face and torso facing up to simulate sleeping posture- and to breath under different breathing conditions (such as slow and fast, deep or shallow). Some typical test scenarios would be:

- Participant, at rest, for finger (or alternative) and neck PPG signals comparison in controlled conditions.
- Physiological signals acquired in different breathing and posture conditions.

- Physiological signals obtained forcing typical real-life signal interferences (Interferences are other signals which are not desired but end up superimposed to the signal of interest. They are also called artefacts). For example, signals will be obtained when the subject is moving, in order to carry out research on interference elimination (i.e. signal cleaning) techniques.

The experiments will be carried out in a laboratory-based environment at Imperial College London in the Department of Electrical and Electronic Engineering, where all the infrastructure and equipment necessary are available. The co-investigator will be in charge of preparing the experiment setting, ensuring all sensors are correctly calibrated and instruct the participants how to proceed.

A total of around 60 healthy participants in between >18 and <70 years old, are intended to be recruited randomly to have a variety of genders and group ages. However, at the beginning of the study, we will focus in the age group of 20-30 years old, as the best quality signals are expected in young adults.

Experiments will last about 1 hour in total, including introduction and setting up. It might be necessary to come back to perform experiments up to three times. This will depend on whether the signals acquired in the different sessions are good enough to carry out the research.

Participants might also be asked to fill questionnaires to investigate the usability of existing sensors.

All data will be saved pseudo-anonymized in password-secured computers at Imperial College and will be fully anonymized after the study finishes.

The entire length of the study is estimated to be 3 years as the refinement of the algorithm will be carried out progressively and iteratively as new signals are acquired.

2.2 STUDY OUTCOME MEASURES

The different endpoints of the study are:

- **Primary endpoints:**
 1. Identify the differences between finger and neck PPG pulse waveforms through extracted features comparison.
 2. Identify what artefacts affect the PPG signal in this part of the body and characterize them to be able to design an appropriate filtering technique to obtain a clean signal.
 3. Obtain PPG recordings at the novel site of the neck for analysis and further development of signal processing algorithms.
- **Secondary endpoints:**
 1. Potential study of factors affecting the PPG signal characteristics of the neck: such as patient gender, age, height, weight, lifestyle...
 2. Potential study of correlations of PPG signals characteristics with other physiological signals recorded with typical polysomnography sensors.
 3. Anonymized database of physiological events recorded with wearables both from the neck and other body locations (such as hand and chest) that can be used for future respiratory and cardiac research.

4. Estimation of best location in the body for multimodal (i.e more than one signal) recording of physiological signals.
5. Investigate the usability of wearable sensors and user interfaces.

3. PARTICIPANT ENTRY

3.1 PRE-REGISTRATION EVALUATIONS

There is any pre-screening procedure that participants have to undergo before entering the study. The only requirement is that they fill a questionnaire to provide some basic information regarding their health state.

3.2 INCLUSION CRITERIA

Healthy subjects of all between 18 and 70 years old could be recruited for this study.

3.3 EXCLUSION CRITERIA

- Subjects with pacemakers

3.4 WITHDRAWAL CRITERIA

If the recordings taken at the neck do not show the minimum quality expected to identify acceptable PPG pulses, the subject might be asked to not continue with the experiment. The participant's data would remain confidential as for the other participants and might be used in the future for analysis purposes if deemed necessary, as the participant signed in the informed consent form (clauses 4. & 5.).

4. ADVERSE EVENTS

4.1 DEFINITIONS

Adverse Event (AE): any untoward medical occurrence in a study subject.

Serious Adverse Event (SAE): any untoward and unexpected medical occurrence or effect that:

- **Results in death**
- **Is life-threatening** – *refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe*
- **Requires hospitalisation, or prolongation of existing inpatients' hospitalisation**
- **Results in persistent or significant disability or incapacity**
- **Is a congenital anomaly or birth defect.**

Medical judgement should be exercised in deciding whether an AE is serious in other situations. Important AEs that are not immediately life-threatening or do not result in death or hospitalisation but may jeopardise the subject or may require intervention to prevent one of the other outcomes listed in the definition above, should also be considered serious.

4.2 REPORTING PROCEDURES

All adverse events should be reported. Depending on the nature of the event the reporting procedures below should be followed. Any questions concerning adverse event reporting should be directed to the Chief Investigator in the first instance: e.rodriquez@imperial.ac.uk

5. ASSESSMENT AND FOLLOW-UP

There will not be routine assessments in this study as the recruited participants are healthy subjects and will not be undergoing any medical procedure. Moreover, data acquisition does not cause any harm that will require health assessments nor follow up.

However, if during the analysis and processing of the PPG signals, the investigators find some recording that does not show sufficient quality, they might ask the participant to come for another recording session to repeat it.

6. DATA ANALYSIS AND HANDLING

Research data will be processed and analysed on secure Imperial College university computers. Any data on laptop computers will be anonymised and encrypted. Named investigators and research assistants who are under the direct supervision of the investigators, will be responsible of the analysis and signal processing, without personal information, for the development of algorithms.

Data and all appropriate documentation will be stored for a minimum of 10 years after the completion of the study. All this will be explained in the information sheet that is provided to subjects prior to seeking their consent to participate.

7. REGULATORY ISSUES

7.1 ETHICS APPROVAL

The Chief Investigator has obtained approval from the Head of Department and Joint Research and Compliance office. The study will be conducted in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th World Medical Assembly, Helsinki 1964 and later revisions.

7.2 CONSENT

Consent to enter the study must be sought from each participant only after a full explanation has been given, an information leaflet offered, and time allowed for consideration. Signed participant consent should be obtained. The right of the participant to refuse to participate without giving reasons must be respected. After the participant has entered the study the investigator remains free to record alternative signals to that specified in the protocol at any stage if he/she feels it is in the participant's best interest, but the reasons for doing so should be recorded. In these cases, the participants remain within the study for the purposes of follow-up and data analysis. All participants are free to withdraw at any time from the protocol treatment without giving reasons and without prejudicing further treatment.

7.3 CONFIDENTIALITY

The Chief Investigator will preserve the confidentiality of participants taking part in the study and is registered under the Data Protection Act (1998).

Research data will be processed and analysed on secure password locked Imperial College university computers. Any data on laptop computers will be anonymised. Investigators and research assistants, will be responsible of the analysis and signal processing, without personal information, for the development of algorithms.

Any paperwork will be kept in a locked filing cabinet only accessible by the investigators.

Data will be stored in accordance with Imperial College London's data policy, all data will be stored for 10 years after the end of the study.

7.4 INDEMNITY

Imperial College London holds negligent harm and non-negligent harm insurance policies which apply to this study.

7.5 SPONSOR

Imperial College London will act as the main Sponsor for this study.

7.6 FUNDING

The European Research Council is funding this study.

8. STUDY MANAGEMENT

The day-to-day management of the study will be co-ordinated by the Co-investigator.

9. REFERENCES

- [1] S. Duncan and M. J. Brodie, "Sudden unexpected death in epilepsy," *Epilepsy & Behavior*, vol. 21, no. 4, pp. 344–351, 2011.
- [2] L. Nashef, "Sudden unexpected death in epilepsy: terminology and definitions," *Epilepsia*, vol. 38, no. s11, pp. S6–S8, 1997.
- [3] L. Nashef, F. Walker, P. Allen, J. Sander, S. Shorvon, and D. Fish, "Apnoea and bradycardia during epileptic seizures: relation to sudden death in epilepsy.," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 60, no. 3, pp. 297–300, 1996.
- [4] A. S. Blum, J. R. Ives, A. L. Goldberger, I. C. Al-Aweel, K. Krishnamurthy, F. W. Drislane, and D. L. Schomer, "Oxygen desaturations triggered by partial seizures: implications for cardiopulmonary instability in epilepsy," *Epilepsia*, vol. 41, no. 5, pp. 536–541, 2000.
- [5] C. Guilleminault, A. Tilkian, and W. C. Dement, "The sleep apnea syndromes," *Annual review of medicine*, vol. 27, no. 1, pp. 465–484, 1976.
- [6] Bateman LM et al. Ictal hypoxemia in localization-related epilepsy: analysis of incidence, severity and risk factors. *Brain* 2008; 131:3239-45.
- [7] W. W. Flemons, N. J. Douglas, S. T. Kuna, D. O. Rodenstein, and J. Wheatley, "Access to diagnosis and treatment of patients with suspected sleep apnea," *American journal of respiratory and critical care medicine*, vol. 169, no. 6, pp. 668–672, 2004.
- [8] E. Rodriguez-Villegas, G. Chen, J. Radcliffe, and J. Duncan, "A pilot study of a wearable apnoea detection device," *BMJ open*, vol. 4, no. 10, p. e005299, 2014.
- [9] G. Mbata and J. Chukwuka, "Obstructive sleep apnea hypopnea syndrome," *Annals of medical and health sciences research*, vol. 2, no. 1, pp. 74–77, 2012.
- [10] P. Corbishley and E. Rodríguez-Villegas, "Breathing detection: towards a miniaturized, wearable, batteryoperated monitoring system," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 196–204, 2008.
- [11] A. Romem, A. Romem, D. Koldobskiy, and S. M. Scharf, "Diagnosis of obstructive sleep apnea using pulse oximeter derived photoplethysmographic signals," *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, vol. 10, no. 3, p. 285, 2014.

A.3 Participant Information Sheet

Participant Information Sheet V3 (19/02/2018)

Title of Project Non-invasive photoplethysmography signals collection at the neck of healthy volunteers

Name of PI Professor Esther Rodriguez Villegas

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish.

This participant information sheet tells you the purpose of this study and what will happen to you if you take part. It also gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether you wish to take part. Thank you for reading this.

- **What is the main purpose of the study?**

Photoplethysmography (PPG) is a non-invasive way of detecting changes in blood volume, using light, shining at the surface of the skin. Together with computer algorithms this can be used to determine how efficiently some organs, such as the lungs and heart, work. For example, from the PPG signal it is possible to have an indication of the amount of oxygen in the blood. This is called pulse oximetry. PPG is a widely used technique in both, hospitals and in non-medical products. As an illustration, an example of a product that uses PPG this is the Apple watch.

The best possible measurement location for PPG is the finger. However, the finger is not the most convenient location for all health-related applications. It is for example not suitable for prevention of Sudden Unexpected Death in Epilepsy (SUDEP). Hence having alternative measuring sites, such as the neck, would be in some cases desirable. One of the reasons for this is that the finger is not the best body location to measure additional signals, from which other health related parameters, which are important in the context of other diseases, could be detected. Places like the neck however would be far better to allow a larger number of health-related parameters to be measured simultaneously using just one wearable (i.e. a little, battery powered, non-intrusive, non-invasive device worn in the body). The neck however, has not been explored as a PPG measurement site, but it is a priori known that it is a tricky location from the engineering point of view since the signal we are looking for will be weaker than in the finger and will be also “spoilt” due to interference from other signals (which we call “artefacts”). It is for these reasons, that collection of body signals in healthy volunteers from not the typically used locations (such as the neck) becomes necessary; so that we can carry out research on what the characteristics of these signals are and how we can clean them from the interference of other non-desired signals.

Therefore, the goal of this study is to record new signals from wearable sensors and their inherent artefacts, which will allow us to carry out research leading to novel signal processing algorithms that will enable accurate multimodal recording (more than one signal measured) from just one wearable sensor.

- **Why have I been chosen?**

You were chosen to enter this study as we are looking for 60 healthy participants in between 18 and 70 years old.

- **Do I have to take part?**

It is up to you to decide whether or not to take part. If you do decide to take part, you will be asked to sign a consent form after reading this information sheet. If you decide to take part, you are still free to withdraw at any time and without giving a reason.

If you decide to take part in the study, you will be asked first to answer a questionnaire with some basic information regarding your health state. Then you will be scheduled for an initial recording session lasting from 30 minutes to 1 hour. Up to 2 other recording sessions of similar or slightly longer duration could be scheduled in the following months after your data has been processed. This will depend on whether we consider that the signals that were obtained in previous sessions were good enough to be used for our research.

All recording sessions will take place at the Circuits and Systems Laboratory in Level 9 of the Department of Electrical and Electronic Engineering at Imperial College London, in a controlled environment under the supervision of the investigators.

- **How are the signal recordings taking place and what do I have to do?**

The study will consist on simultaneously collecting respiratory, cardiac and movement signals with non-invasive wearable sensors under different conditions (such as sitting down, standing, lying down or walking), and performing some simple breathing exercises (like breathing deep or shallow, slow or fast). Figure 1 shows an example of a typical set up.

The experiments will be carried out in our laboratory, where all the infrastructure and equipment necessary are available. The co-investigator will be in charge of preparing the experiment setting, ensuring all sensors are correctly calibrated and give you the instructions necessary on how to proceed. You will be given the option of either placing the sensors yourself or letting one of the investigators do it for you.



Figure 1: Typical study set up: a. Transmission pulse oximeter sensor placed in the index finger. b. Reflectance pulse oximeter sensor placed in contact with the anterior neck.

- **What are the possible risks and benefits of taking part?**

The non-invasive recordings have no major risks. All the sensors used in the experiments comply with the safety requirements of the EU standards for medical devices.

We cannot promise the study will help you directly but the information we get might help improve the performance of wearable systems for healthcare monitoring in general, and more specifically for prevention of Sudden Unexpected Death in Epilepsy (SUDEP).

- **What if something goes wrong?**

Imperial College holds Public Liability (“negligent harm”) and Clinical Trial (“non-negligent harm”) insurance policies which apply to this trial. If you can demonstrate that you experienced serious and enduring harm as a result of your participation in this trial, you may be eligible to claim compensation without having to prove that Imperial College is at fault. If the injury resulted from any procedure which is not part of the trial, Imperial College will not be required to compensate you in this way. Your legal rights to claim compensation for injury where you can prove negligence are not affected. Please contact the Principal Investigator Professor Esther Rodriguez Villegas if you would like further information about the insurance arrangements which apply to the trial.

- **Will my taking part in this study be kept confidential?**

To ensure your data’s protection and confidentiality a number of best-practice measures will be taken to ensure this is maintained. All data are saved in password-secured computers at Imperial College and will be anonymized for subsequent review, to comply with the Data Protection Act. Data will be held for 10 years post study end.

- **What will happen to the results of the research study?**

The results will be anonymized and can be published in scientific journals and presented at scientific conferences.

- **Who is organising and funding the research?**

Imperial College London.

- **Who has reviewed the study?**

This study was reviewed by the Head of the Department and the Joint Research Compliance Office.

Next step

Please, confirm if you would like to take part in these experiments, sign the inform consent and let us know your availability. If you still have any questions before deciding, please do not hesitate to contact us, and finally, we would like to remind you that you may withdraw from the study at any time and all your data will be kept confidential and anonymised.

A.4 Informed Consent Form



Study Protocol Number:
Participant Identification Number for this trial:

INFORMED CONSENT FORM

Full Title of Project: *Non-invasive photoplethysmography signals collection at the neck of healthy volunteers*

Name of Principal Investigator: *Esther Rodriguez Villegas*

1. I confirm that I have read and understand the subject information sheet for the above study and have had the opportunity to ask questions which have been answered fully.
2. I understand that my participation is voluntary, and I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.
3. The compensation arrangements have been discussed with me.
4. I agree to take part in the above study, my anonymised data to be used for further analysis and the results be shown in scientific papers and conferences.
5. I agree for my signals to be used for future research projects after all my personal information (i.e. information which could identify me) has been removed.
6. I give permission to the researcher to take audio recordings during the experiment and use them for signal processing analysis.

Please initial box

☐
☐
☐
☐
☐
☐

Name of Patient/Participant

Signature

Date

Name of Person taking consent
(if different from Principal Investigator)

Signature

Date

Principal Investigator

Signature

Date

1 copy for patient/participant; 1 copy for Principal Investigator; 1 copy to be kept with hospital notes

Version 3- 19/02/18