KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Tensor-based ECG analysis in sudden cardiac death

Griet Goovaerts

Supervisors:
Prof. dr. ir. S. Van Huffel
Prof. dr. R. Willems

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

December 2018

Tensor-based ECG analysis in sudden cardiac death

Griet GOOVAERTS

Examination committee:
Prof. dr. ir. P. Verbaeten, chair
Prof. dr. ir. S. Van Huffel, supervisor
Prof. dr. R. Willems, supervisor
Prof. dr. ir. L. De Lathauwer
Prof. dr. ir. J. Suykens
Prof. dr. ir. C. Van Hoof
Prof. dr. X. Hu
(UCSF)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

December 2018

© 2018 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Griet Goovaerts, Kasteelpark Arenberg 10 - box 2446, B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

You often hear that in life, the journey is more important than the destination. I believe the same thing is true for a PhD. Today marks the end of a journey that turned out to be longer and better than I could have ever imagined when I applied for this position almost six years ago. While I am very happy that I reached the end, I will never forget everything that happened along the way and the people that helped me to get here.

I was lucky enough to have two supervisors that complemented each other in many different ways. Sabine, thank you for giving me the opportunity to start as a PhD student in your group, and for giving me enough freedom throughout the years to find my own way to do so. From the start, you were a caring supervisor that made time for us whenever necessary despite your busy schedule. Your vision on biomedical signal processing and your goal to develop algorithms that can have a real impact on the life of patients are truly inspiring. Additionally, the social events you organised at your house were always very nice and helped Biomed to become the close research group it is today.
When I started my PhD, I specifically selected a topic where I could work on real clinical applications. Rik, you ensured that I didn't lose track of this side of my research on the way and broadened my 'engineering view' multiple times. While I very much enjoyed our monthly meetings, I enjoyed the discussions and dinners at conferences even more!

I would also like to thank the other members of my examination committee, Chairman Prof. Pierre Verbaeten, Prof. Lieven De Lathauwer, Prof. Johan Suykens, Prof. Chris Van Hoof and Prof. Xiao Hu for the feedback and comments both during the intermediate presentations and preliminary defense. Lieven, thank you for introducing me to tensor methods. While challenging at times, they certainly added another dimension to my research. Xiao, the four months that I spent in San Francisco were for sure one of the highlights of my PhD. Thank you for welcoming me in your research group, for the nice trips and dinners and for coming all the way to Belgium for my defense. I sure hope

to be back one day. Ran, Del, Jacob, Koa, Kais, Rich and Andrea, thank you for making me feel part of the *squad* from day one!

Throughout the past five years I collaborated with many people, both inside and outside Biomed, the university and Belgium. I want to thank all collaborators for all the joint efforts and for broadening my knowledge on everything from tensor methods to clinical applications. Special thanks to Bert for the nice collaboration from day one, and for extensively answering all my questions at any time of the day. I hope I could convince you that all these complicated methods were worth it after all!

When I was looking for a PhD position, many factors influenced my decision: the topic, the supervisor, the university,... but I never really considered what turned out to be one of the most important things: the research group I would end up in. Without a doubt, I made the perfect choice there: Biomed is the best group of colleagues I could have wished for. A PhD is filled with ups and downs, and in Biomed any success is a reason to celebrate, from pancakes to celebrate an accepted paper to cookies to celebrate the end of another week. Equally important however, there is also always someone around to talk you through the difficult moments, no matter if they are personal or professional. I therefore want to thank everybody who is or was part of Biomed during five amazing years: Thank you all for the *interesting* discussions during lunch and coffee breaks, for all the cake, the parties, the drinks, the sports activities and so much more. I truly appreciated getting to know all of you, and while I could write an additional book about all the things I will remember, there are several people that deserve some extra words:
To everyone in the 'party office', present and past: Thank you for the company, the talks, the chocolates and the pleasant atmosphere in general. Dorien and Margot, thanks for the much-needed moral support the last months and for taking such good care of our avocado plant. With all those cooking lessons, we really have to plan that office dinner very soon! Special thanks to Laure for staying part of the office even after you moved to Gasthuisberg: it honestly feels like you never *really* left. Your unexpected messages and pictures these last months could always make me smile.
Thomas and Rob, we started everything together five years ago, and I am really happy both of you stuck with me until the end. Thomas, it was comforting to know someone was going through the same thing as me last year, from getting those final ADS credits to writing our PhD. Rob, thank you for being such a good friend: for the games, the parties, and Papegaeien in de Reinaert but even more important for being there (literally) whenever I needed to talk.
Alex, you were always up for 'just one more drink'. Let's have more fancy lunches in the future, in Colombia, Belgium or anywhere in between. Jasper, thank you for the many discussions during our coffee breaks and occasional

carpooling, and for sharing your opinion on everything from foodsharing to checkered shirts. Simon V.E., I am sorry that I could never join the football trainings. Maybe I will have more time now? (I doubt it though). Mario, you will forever be the kiwi guy for me, but I will remember your passion for everything in life even longer. Lieven, you were often our much-needed moral compass. Please keep on sending me facts of the day.
During my PhD, I supervised a number of very good master students, three of which became colleagues later on. Ofelie, Jonathan and Simon G., I am still very happy you chose one of my topics. I learned a lot while supervising you and I hope you can say the same thing. Jonathan, I especially enjoyed your continuous enthousiasm and our interesting discussions, from cell biology to crazy projects with fancy figures. I hope we can have many more of those discussions in the future.
Thank you to the biotensors team, most importantly Otto, Nico and Martijn, for guiding me through the tensor world and for the nice collaborations and discussions, both inside and outside the office.
Carolina, you were there for me from before I even started until the very last day. I know you think I was Creepy every once in a while... but maybe that's just who I am. Nevertheless, the times we spent together at both at conferences and in Leuven are some of my happiest PhD memories, and I would gladly buy you many more cocktails and/or purses if it meant we could keep working together forever. Thank you for being a great colleague and an even better friend.

Many thanks also to the rest of the people in STADIUS, in particular Ida, Aldona, John, Elsy, Wim and Maarten for taking care of all our administrative issues and for letting me in my office every time I forgot my badge. You make all our lives at STADIUS a lot easier! I also greatly acknowledge IWT and VLAIO for providing me with a PhD grant for Strategic Basic Research and FWO for the travel grant that made my research stay in San Francisco possible.

Hoewel mijn doctoraat regelmatig de nodige vrije tijd in beslag nam, waren er gelukkig genoeg mensen die ervoor zorgden dat ik op tijd en stond alle werk-gerelateerde dingen even kon vergeten. Een meer dan verdiende dankjewel aan mijn vrienden, vriendinnen en familie voor de interesse in mijn werk, maar vooral voor de ontelbare leuke momenten!

Bedankt aan de *Kempische furies*, Chloë, Evelien, Eveline, Jessica, Liesbeth en Lisanne: ik ken jullie al zo lang dat het lijkt alsof jullie er altijd geweest zijn, en ik weet dat jullie er ook altijd zullen zijn. We maakten de voorbije jaren veel mee samen, en ik kijk al uit naar wat de toekomst nog brengt. Ook veel dank aan Joost, Gilles, Yves en Sam om onze decibels af en toe te tolereren.
Karen, Lotte en Elise, tien jaar nadat we in Analyse 1 naast elkaar kwamen te zitten, sluiten we onze uniefjaren nu bijna echt af. Karen en Lotte, jullie toonden

op zoveel verschillende manieren wat een luxe het is om je beste vriendinnen op wandelafstand te hebben wonen. Elise, Brussel is net iets verder, maar daardoor gelukkig altijd een goed excuus om te brunchen. Of het nu was om mijn hart te luchten of gewoon een avond buiten in de tuin te zitten, ik kon altijd bij jullie terecht. Ook bedankt aan Pieter, Bram, Ann, Christiaan, Lisa, Klaas, Annelies, Thibaut, Michiel, Niels B., Marijn, Maarten, Philippe J., Philippe L., François en Niels V. voor de fijne avonden en Ardennen-weekendjes de voorbije jaren. Of om het (nog 1 keer) met de woorden van Maarten te zeggen: '*Jullie verdienen de Nobelprijs!*'
De busclub-etentjes zijn een vaste afspraak in mijn agenda geworden waar ik telkens weer erg naar uitkijk. Liesbeth, Enid en Heleen, hoewel ik het gevoel heb dat we elke keer dezelfde verhalen oprakelen, ben ik ze toch nog altijd niet beu geraakt. Hopelijk houden we deze traditie nog heel lang vol.
Niels, Klaas en Leen, onze Alma-lunches waren soms moeilijk in te plannen maar daardoor niet minder plezant. Bedankt voor het gezelschap en de ontspannende babbels.

Mama en papa, Heleen en Charlotte, Nicholas en Evert, jullie waren altijd mijn grootste supporters. Dankzij jullie weet ik dat waar ik ook ben of wat er ook gebeurt, ik gelukkig altijd een enorm warme thuis heb om naar terug te komen. Bedankt voor alles, van apero-donderdagen en zussenweekends tot knuffels en ondersteunende woorden wanneer het even wat minder vlot ging. Een weekend thuis zorgde er altijd voor dat ik alles weer helemaal kon relativeren en ontspannen terug naar Leuven kon vertrekken. Zonder jullie stond ik vandaag niet waar ik nu sta!

Thank you all! Bedankt iedereen!

Griet
Leuven, December 2018

Abstract

Sudden cardiac death (SCD) is one of the main causes of death worldwide, accounting for approximately 4.5 million deaths per year. Since it occurs relatively often in younger people, its socio-economic impact is much higher than the impact of other major health issues like cerebrovascular disease. It is therefore important to accurately determine which patients are at risk for developing dangerous arrhythmias in order to implement optimal treatment and prevention strategies. Prediction of sudden cardiac death is however not an evident task, and providing reliable indicators has been a very active area of research for many decades. This research therefore focuses on the development of algorithms to extract potential SCD risk factors from the ECG signal, through a combination of tensor methods and machine learning approaches.

Tensors are multilinear generalizations of vectors and matrices, that can be used to analyse all leads of the ECG channel simultaneously. Since the different spatial leads give a global view of the heart in three dimensions, it makes sense to fully exploit the shared information by combining the information from all leads. The first part of this thesis presents four tensor-based methods to detect and analyse different ECG characteristics. We show that by modifying the tensor decomposition, specific signal characteristics such as changes in heart rate or increased noise levels can be taken into account. This ensures that the developed methods can be optimally used in real-life scenarios, which is confirmed by the good results on different clinical datasets.

The second part of this research is focused on QRS fragmentation (fQRS), a promising risk factor for sudden cardiac death. Detection of fQRS heavily relies on visual inspection, which has been shown to be dependent on rater experience. Therefore, we propose a method to detect and quantify QRS fragmentation using machine learning methods. Quantification of fQRS is a novel approach to examining the biomarker, and we demonstrate that this innovative fQRS score largely correlates to the certainty of QRS fragmentation in a signal. Since the proposed fQRS score is determined objectively, the obtained results can

be easily repeated in different datasets, which promotes the clinical use of this parameter.

Finally, the last part of this thesis investigates to what extent advanced machine learning methods can provide added value in modelling the survival of patients. We show that the combination of the proposed fQRS score with advanced survival models is better capable of predicting the survival time of patients than commonly used statistical models.

Beknopte samenvatting

Plotse hartdood is een van de voornaamste doodsoorzaken wereldwijd, die jaarlijks verantwoordelijk is voor ongeveer 4.5 miljoen sterfgevallen. Aangezien plotse hartdood relatief vaak voorkomt bij jongere mensen, is de socio-economische impact ervan veel groter dan bij andere grote gezondheidsproblemen zoals cerebrovasculaire ziekten. Het is daarom erg belangrijk om op een accurate manier te bepalen welke patiënten risico lopen op het ontwikkelen van gevaarlijke hartritmestoornissen, zodat optimale behandelings- en preventiestrategieën kunnen opgestart worden. Het voorspellen van plotse hartdood is echter geen eenvoudig probleem, en het bepalen van betrouwbare indicatoren is reeds verschillende decennia een zeer actief onderzoeksgebied. Dit onderzoek spitst zich daarom toe op het ontwikkelen van algoritmen om potentiële risicofactoren voor plotse hartdood uit het ECG signaal te extraheren. We maken hierbij gebruik van een combinatie van tensor methoden en machinaal leren.

Tensoren zijn multilineaire veralgemeningen van vectoren en matrices, die gebruikt kunnen worden om alle kanalen van het ECG signaal gelijktijdig te analyseren. Aangezien de verschillende ruimtelijke kanalen een globaal zicht op het hart geven in drie dimensies, ligt het voor de hand om deze gedeelde informatie ten volle uit te buiten door de informatie uit de verschillende kanalen te combineren. Het eerste deel van deze thesis stelt vier tensor-gebaseerde methoden voor om verschillende ECG karakteristieken te detecteren en te analyseren. We tonen aan dat door het aanpassen van de tensorontbinding, we rekening kunnen houden met specifieke signaaleigenschappen zoals veranderingen in hartritme of toenames van het ruisniveau. Dit zorgt ervoor dat de ontwikkelde methoden optimaal gebruikt kunnen worden in levensechte scenario's, wat aangetoond wordt door goede resultaten op diverse klinische datasets.

Het tweede deel van dit onderzoek is gefocust op QRS fragmentatie (fQRS), een veelbelovende risicofactor voor plotse hartdood. Detectie van fQRS maakt voornamelijk gebruik van visuele inspectie, waarvan aangetoond is dat de resultaten afhankelijk zijn van de ervaring van de beoordelaar. We ontwikkelden

daarom een methode om QRS fragmentatie te detecteren en te kwantificeren, gebruik makend van technieken uit het machinaal leren. Kwantificatie van fQRS is een nieuwe manier om deze biomerker te onderzoeken, en we demonstreren dat deze nieuwe fQRS score nauw aansluit bij de zekerheid over de aanwezigheid van fQRS. Aangezien de voorgestelde fQRS score op een objectieve manier bepaald wordt, zijn de verkrijgde resultaten makkelijk repliceerbaar in andere datasets, en kan op deze manier het klinische nut van deze parameter vergroot worden.

Tenslotte wordt in het laatste deel van dit onderzoek onderzocht in welke mate geavanceerde methoden uit het machinaal leren toegevoegde waarde kunnen bieden om de overleving van patiënten te modelleren. Hierbij tonen we aan dat de ontwikkelde fragmentatie score in combinatie met geavanceerde overlevingsmodellen beter in staat zijn om de overlevingstijd van een patiëntengroep te voorspellen dan standaard gebruikte statistische modellen.

List of Abbreviations

AAMI Association for Advancement of Medical Instrumentation.
Acc Accuracy.
AF Atrial Fibrillation.
ANOVA Analysis of Variance.
ApEn Approximate Entropy.
AUC Area Under the Curve.
AV node Atrioventricular node.
AVRR Average RR-interval length.

BPM Beats per minute.

CAD Coronary artery disease.
CI Confidence interval.
CinC Computing in Cardiology.
CoV Coefficient of variance.
Cox PH model Cox proportional hazards model.

DWT Discrete wavelet transform.

ECG Electrocardiogram.
EMD Empirical Mode Decomposition.

FN False negative.
FP False positive.
FPR False positive rate.
fQRS QRS fragmentation.

HR Hazard ratio.
HRV Heart rate variability.
hsd honest significant difference.
HTI HRV triangular index.

ICA Independent Compontent Analysis.
ICD Implantable Cardioverter Defibrillator.

KM Kaplan-Meier.
kNN k-nearest neighbors.

LMLRA Low Multilinear Rank Approximation.
LQTS Long QT syndrome.
LVEF Left ventricular ejection fraction.

MAD Median absolute deviation.

NB Naive Bayes classifier.
NPV Negative predictive value.
NSR Normal sinus rhythm.
NYHA New York Heart Association.

PAC Premature atrial contraction.
PEA Pulseless electrical activity.
PPV Positive predictive value.
PRSA Phase-Rectified Signal Averaging.
PVC Premature ventricular contraction.

RBF Radial Basis Function.
RMS Root Mean Square.
RMSSD Root Mean Square of Successive Differences.
ROC Receiver Operating Characteristics curve.
RWT Redundant wavelet transform.

SA node Sinoatrial node.
SAECG Signal-averaged electrocardiogram.
SCD Sudden cardiac death.
Se Sensitivity.
SNR Signal-to-noise ratio.
Sp Specificity.
SVM Support Vector Machine.
SVR Support vector regression.

TB Treebagger.
TN True negative.
TP True positive.
TWA T wave alternans.

UCLA University of California, Los Angeles.
UCSF University of California, San Francisco.

VF Ventricular fibrillation.
VMD Variational Mode Decomposition.
VT Ventricular tachycardia.

WCPD Weighted canonical polyadic decomposition.
WLS Weigthed least squares.

Contents

Abstract v

Beknopte samenvatting vii

List of Abbreviations x

List of Symbols xi

Contents xi

List of Figures xix

List of Tables xxiii

1 Introduction 1

1.1 Research motivations 1

1.2 The electrocardiogram 4

1.2.1 Physiological origin 4

1.2.2 Abnormal patterns 6

1.2.3 Measurement 7

1.3 Risk prediction of sudden cardiac death 10

1.3.1 ECG-derived risk factors 10

1.3.2 Other risk factors 12
1.4 Research objectives 13
1.5 Chapter overview and personal contributions 13
1.5.1 Part I: Background 15
1.5.2 Part II: Tensor-based methods 15
1.5.3 Part III: QRS fragmentation 17
1.6 Collaborations 17

Part I: Background **21**

2 Tensors and ECG **23**

2.1 Introduction 23
2.2 Motivation 24
2.3 Basic concepts and notations 25
2.4 Tensor operations 27
2.4.1 Tensor-matrix transformations 27
2.4.2 Matrix- and tensor multiplications 28
2.4.3 Tensor decompositions 30
2.5 An overview of tensors in ECG signal processing 34
2.6 Conclusion 35

3 Machine learning and classification **37**

3.1 Introduction 37
3.2 Unsupervised classification methods 38
3.3 Supervised classification techniques 41
3.3.1 Support Vector Machines 41
3.3.2 Other techniques 44
3.4 Performance evaluation 45

3.4.1 Confusion matrix 45
3.4.2 Receiver Operating Characteristics curve 47
3.5 Conclusion 47

Part II: Tensors **49**

4 Unsupervised detection of irregular heartbeats **51**
4.1 Introduction 51
4.2 Data 53
4.2.1 INCART database 54
4.2.2 MIT-BIH Arrhythmia database 54
4.3 Methods 55
4.3.1 Preprocessing 55
4.3.2 Tensorization 56
4.3.3 Tensor decomposition 57
4.3.4 Clustering techniques 57
4.3.5 Evaluation of performance 58
4.4 Results 59
4.4.1 Case study 59
4.4.2 INCART database 62
4.4.3 MIT-BIH database 64
4.5 Discussion 65
4.6 Conclusion 66

5 Automatic detection of T wave alternans **67**
5.1 Introduction 67
5.2 Data 68
5.2.1 Simulation data 69

5.2.2 Physionet dataset 69
5.2.3 University Hospitals Leuven dataset 70
5.3 Methods 70
5.3.1 Preprocessing 70
5.3.2 T wave segmentation 71
5.3.3 Tensor construction 72
5.3.4 Tensor decomposition 72
5.3.5 Tensor decomposition results 74
5.3.6 Detection of T wave alternans 75
5.4 Results 76
5.4.1 Case study 76
5.4.2 Artificial signals 79
5.4.3 Physionet database 81
5.4.4 Clinical dataset from the University Hospitals Leuven . 83
5.5 Discussion and Conclusion 84

6 Analysis of changes in heartbeat morphology prior to in-hospital cardiac arrest 87

6.1 Introduction 87
6.2 Data 89
6.3 Methods 90
6.3.1 Preprocessing 90
6.3.2 Tensorization 91
6.3.3 Tensor decomposition 91
6.3.4 Parameter calculation 94
6.3.5 Analysis of changes in parameter values 95
6.4 Results 99
6.4.1 Changes from baseline 100

6.4.2 Dominant trend analysis 100
6.5 Discussion 100
6.6 Conclusion 103

7 Automatic detection of atrial fibrillation in single and multilead ECG signals 105

7.1 Introduction 106
7.2 Data 106
7.2.1 Physionet/Computing in Cardiology Challenge 2017 . . 107
7.2.2 MIT-BIH AFIB & AFTDB dataset 107
7.3 Methods 109
7.3.1 SVD-based detection in single lead ECG signals 109
7.3.2 MLSVD based detection for single lead ECGs 113
7.3.3 MLSVD-based detection for multilead ECGs 117
7.3.4 Combination of morphological and HRV characteristics 121
7.3.5 Evaluation of results 124
7.4 Results and Discussion 124
7.4.1 AF detection in single lead ECG signals 124
7.4.2 AF detection in multilead ECG signals 131
7.4.3 Detection of AF in single lead ECG signals in combination with multilead ECG signals 135
7.5 Conclusion 137

Part III: QRS Fragmentation 139

8 Detection and Quantification of QRS Fragmentation 141

8.1 Introduction 141
8.2 Data 144
8.3 Methods 145

8.3.1 Preprocessing . . . 145
8.3.2 VMD-based QRS segmentation . . . 147
8.3.3 Feature extraction . . . 151
8.3.4 Classification . . . 154
8.3.5 Performance evaluation metrics . . . 155
8.4 Results . . . 155
8.4.1 QRS segmentation . . . 155
8.4.2 Analysis of feature values . . . 156
8.4.3 Classifier performance . . . 158
8.5 Discussion . . . 161
8.6 Conclusion . . . 163

9 Risk Assessment of All-Cause Mortality using a QRS fragmentation score 165

9.1 Introduction . . . 165
9.2 Survival analysis . . . 167
9.2.1 Kaplan-Meier analysis . . . 168
9.2.2 Cox proportional-hazards model . . . 169
9.2.3 Survival SVM . . . 170
9.3 Data . . . 173
9.4 Methods . . . 174
9.4.1 Optimal cut point determination . . . 174
9.4.2 Comparison of survival models . . . 175
9.5 Results . . . 176
9.5.1 Optimal cut point determination . . . 176
9.5.2 Comparison of survival models . . . 178
9.6 Discussion . . . 179
9.7 Conclusion . . . 181

10 Conclusions and Future directions **183**

10.1 Conclusions 183

10.1.1 Algorithms for SCD risk factor extraction 184

10.1.2 Tensors in ECG analysis 185

10.1.3 Risk assessment in ICD patients 186

10.2 Future directions 187

10.2.1 Algorithms for SCD risk factor extraction 187

10.2.2 Tensors in ECG analysis 188

10.2.3 Risk assessment in ICD patients 189

A List of Heart Rate Variability features **191**

A.1 Feature Description 191

A.1.1 Statistical features 192

A.1.2 Geometric and other non-linear characteristics 195

A.2 Linear analysis of features 199

B Construction and Validation of Noise Model **201**

B.1 Construction of SNR model 202

B.2 Validation of SNR model 202

B.2.1 MIT-BIH Noise Stress Database 203

B.2.2 Clinical database 204

Bibliography **205**

Curriculum vitae **229**

List of publications **231**

List of Figures

1.1 Electrical activity in the heart [199]. 5

1.2 Electrode placement and spatial angles of the clinical 12-lead ECG 8

1.3 The impact of SCD in different populations [161, 224]. 11

1.4 Chapter-by-chapter graphical overview of the PhD. 14

2.1 Schematic representations of a vector, matrix and tensor. . . . 26

2.2 The mode-1 vectors and mode-$(2, 3)$ slices of a tensor 26

2.3 The mode-2 unfolding of a tensor. 27

2.4 Visualization of the MLSVD and LMLRA of a third-order tensor $\mathcal{T}$. 31

2.5 Schematic representation of the Canonical Polyadic Decomposition of a third-order tensor. 32

3.1 Illustration of the OPTICS clustering algorithm. 40

3.2 Example of a confusion matrix, with correct labels indicated in green and incorrect labels in red. 46

4.1 Tensorization of the ECG signal. 56

4.2 Excerpt of the ECG signal used in the case study 60

4.3 Multilinear singular values σ_i of the signal used in the case study. 60

4.4 Results of the selected case study signal for CPD ranks varying from one to five. 61

4.5 Results of different clustering algorithms for different CPD ranks. 63

4.6 Influence of number of channels on final result. 64

5.1 Example of a simulated signal from the same patient without noise (left), moderate noise (middle) and heavy noise (right). . 69

5.2 Construction of the T wave tensor: the T waves of all channels are segmented and stacked beat-by-beat in a 3D manner. . . . 73

5.3 Comparison of the tensor decomposition methods used for TWA detection. 75

5.4 Excerpts of ten seconds of all three signals used in the case study, taken from the Physionet TWA database. 77

5.5 Factor vectors obtained by CPD and PARAFAC2 on the signals used as case study. 78

5.6 Results for CPD (red) and PARAFAC2 (black) for four types of artificial signals: Clean signals with varying amount of TWA, clean signals with changing T wave shift, artificial signals with a moderate and high noise level. 80

5.7 Comparison between the reference ranking of the Physionet database and the ranking obtained by CPD and PARAFAC2. . 82

5.8 Estimated levels of T wave alternans using CPD (left) and PARAFAC2 (right) in TWA group vs control group. 83

5.9 ROC curve for classification using CPD (AUC = 0.64) and PARAFAC2 (AUC = 0.88). 84

6.1 Two examples of typical technical artifacts found in the dataset. 93

6.2 Illustration of the factor vectors of the weighted CPD for a typical ECG signal. 94

6.3 Illustration of the different steps in the detection of the dominant trend applied to a heart rate signal. 97

6.4 Parameters derived from dominant trend analysis which showed statistically significant differences between all groups. 101

7.1 Overview of the SVD-based method for single lead signals. . . . 110

7.2 Segmentation, alignment and compression of heartbeats. 111

7.3 Singular values of the heartbeat matrix. 112

7.4 Visualisation of the model tensor after RWT 116

7.5 Integration of morphological and HRV features into one global method 123

7.6 The singular values and results of the cross-validation for the model matrix of the Physionet/CinC challenge dataset 126

7.7 The F_1-scores for different sizes of model set for the SVD-based method. 126

7.8 The two-dimensional projections of the morphological feature vector for the Physionet/CinC Challenge 2017 dataset. 128

7.9 An example of a representative heartbeat that strongly resembles a normal beat. 129

7.10 The multilinear singular values of the model tensor, after tensorization with the RWT. 129

7.11 The multilinear singular values of the model tensor of the MIT-BIH AFIB & AFTDB dataset 131

7.12 The morphological features for the training- and test set of the MIT-BIH AFIB & AFTDB dataset, for the SVD- and MLSVD-based approaches 133

7.13 The ROC curves for the results of the linear SVM on MIT-BIH AFIB & AFTDB test set, for the SVD- and MLSVD-based approach. 134

7.14 ROC curves for the SVMs on the combination of the PhysioNet/CinC Challenge 2017 testset and the MIT-BIH AFIB & AFTDB dataset 136

7.15 The tachogram and representative heartbeat of a signal that was wrongly classified based on the HRV, but correctly classified when adding morphology information. 137

8.1 Examples of different subtypes of fragmented QRS complexes. . 142

8.2 Block diagram of the proposed method for fQRS detection and quantification. 146

8.3 Example of ECG signal with the corresponding modes of the output of Variational Mode Decomposition with $k = 5$ and $\alpha = 100$. 149

8.4 Illustration of the three main steps for QRS segmentation using VMD. 150

8.5 Illustration of the three steps to create the PRSA curve for a normal heartbeat (top) and a heartbeat with fragmentation (bottom). 152

8.6 Box plots showing all feature values grouped by the total score given by five experts. 157

8.7 ROC curves for fQRS detection with all classifiers together with the corresponding AUC values. 159

8.8 Illustrating the fQRS quantification score for the second group (Section 8.3.4) using SVM (Linear, Polynomial and RBF kernels), KNN, NB and TB classifiers. 160

9.1 Kaplan-Meier plots of channels with statistically significant differences ($p < 0.05$) between survival curves. The risk tables represent the number of patients alive at different time instances for fQRS scores higher and lower than θ_{ch}. 177

A.1 Comparison of tachograms of signals with normal sinus rhythm and atrial fibrillation. 192

A.2 The estimated non-parametric probability distributions for all ten HRV-features calculated using the Physionet/CinC Challenge 2017 dataset [83]. 193

A.3 Poincaré plots of the RR- en ΔRR-intervals 196

B.1 Boxplots depicting the results of applying the SNR model on the test set. 203

B.2 Examples of ECG segments from the UCSF database without noise, with moderate noise and with heavy noise. 203

B.3 Histograms of SNR values for each noise class. 204

List of Tables

4.1 List of classes and superclasses of irregular heartbeats as defined by AAMI [1]. 53

4.2 Number of heartbeats of each of the types (normal, supraventricular and ventricular) in both databases used in this Chapter. 55

4.3 Confusion matrix for the detection of abnormal heartbeats and corresponding derivations for each class. 58

4.4 Sensitivy and positive predictive value for each class obtained with different methods on the MIT-BIH database. 65

5.1 Kendall coefficient scores obtained by comparing the rankings from different methods found in literature and the two proposed tensor-based methods with the reference ranking for the Physionet database. 82

6.1 Number of patients with preceding rhythms in the dataset. . . 90

6.2 Results of the analysis of changes of feature values from baseline. 99

7.1 Overview of the Physionet/CinC Challenge 2017 dataset 108

7.2 Overview of the MIT-BIH AFIB & AFTDB dataset 109

7.3 Results for the Physionet/CinC Challenge 2017 dataset 125

7.4 Confusion matrix of the best-scoring SVD-based method, combining morphological and HRV features ($F_1 = 0.770$) 127

7.5 Results for the different multilead methods on the MIT-BIH AFIB & AFTDB dataset 132

7.6 Data profile of the combination of the Physionet/CinC Challenge 2017 and the MIT-BIH AFIB & AFTDB dataset. 135

7.7 Results of the combinatio of the Physionet/CinC Challenge 2017 and the MIT-BIH AFIB & AFTDB dataset. 136

8.1 Frequency of occurrence of different scores, obtained by summing the annotations from all five readers, in the database. 145

8.2 Accuracy results of QRS segmentation of the proposed method and three state-of-the art alternative approaches on the QT database. 156

8.3 Significance results of the post-hoc analysis for the comparison of the VMD- and PRSA-based features using Tukey's hsd test. 158

8.4 Comparison of the results from the proposed method with methods from literature. 159

9.1 Optimal cut points, 95% confidence intervals and results on the test set for each channel. 176

9.2 Results of the different survival models using only clinical parameters or the combination of clinical parameters and fQRS scores. 178

A.1 Covariance table of the (normalised) HRV feature set, showing the Pearson correlation coefficients between the different features. 200

Chapter 1

Introduction

1.1 Research motivations

Sudden cardiac death (SCD) is defined as 'an unexpected natural death from a cardiac cause within a short time period, generally ≤ 1 hour from the onset of symptoms, in a person without any prior condition that would appear fatal' [244]. It is one of the main causes of death worldwide, accounting for approximately 4.5 million deaths per year [179]. Since SCD occurs relatively often in younger people (40% of all cases occurs before the age of 65 [41]), its socio-economic impact is much higher than the impact of other major health issues such as cerebrovascular disease, chronic lower respiratory disease or diabetes [201].

The majority of sudden deaths are attributed to acute cardiac arrhythmias. The three major types of presenting rhythms are ventricular tachyarrhythmia (either ventricular tachycardia (VT) or ventricular fibrillation (VF)), bradyasystole or pulseless electrical activity (PEA) [179]. While these arrhythmia are mostly caused by an underlying heart condition (up to 80% of patients who experience sudden cardiac death have coronary artery disease (CAD) [244]), they are often the first manifestation of a cardiac problem. Cardiac arrests may be reversed by using a defibrillator that delivers an electric shock to restore the normal heart rhythm [79]. However, since most cardiac arrests occur out of hospital in a non-monitored environment and a shock must be administered within minutes after the start of the arrhythmia, the overall survival rate for cardiac arrest is lower than 5% [163].

It is therefore important to accurately determine which patients are at risk for developing dangerous arrhythmias in order to implement optimal

treatment and prevention strategies. When an underlying cardiac condition is diagnosed in time, it can often be managed to prevent deterioration. Patient management involves lifestyle interventions, pharmacotherapy and/or device therapy. Lifestyle interventions can aim at preventing deterioration of both disease and comorbidities [176]. In patients with certain conditions such as hypertrophic cardiomyopathy or long QT syndrome (LQTS), intense physical activity is known to provoke arrhythmias [193, 36]. This specific patient group can thus be restricted from endurance training. For CAD prevention on the other hand, a sedentary lifestyle is known to potentially cause deterioration, and these patients could thus benefit from additional activity [238].
The goal of pharmacotherapy is to control and improve the general heart condition. It can include discontinuation of known pro-arrhythmic drugs or prescription of anti-arrhythmic drugs such as beta-blockers [179].
Finally, an implantable cardioverter-defibrillator (ICD) detects ventricular arrhythmia and ends most of them by delivering an electric shock, similar to an external defibrillator. They were introduced more than 30 years ago and have become indispensable for SCD prevention. An ICD can be implanted as primary or secondary intervention [179]. Primary intervention consists of ICD implantation in patients that did not have a previous cardiac arrest or arrhythmias, but that are known to be at increased risk of SCD. Secondary prevention on the other hand are patients who have experienced previous cardiac arrests. While an ICD manages to terminate most ventricular arrhythmia, its implantation can cause complications such as infections or lead failure. Furthermore, when the device wrongly detects a ventricular arrhythmia, it may administer an inappropriate shock which may result in adverse effects [55]. Therefore, proper selection of patients who would benefit from ICD implantation is an important concern.

Prediction of SCD is however not an evident task, and providing reliable SCD indicators has been a very active area of research for many decades. Screening of patients can consist of a combination of invasive and non-invasive examinations. The non-invasive approach includes imaging techniques such as echocardiography to assess the function of the left ventricle together with analysis of electrical conduction system of the heart measured by the electrocardiogram (ECG).

The ECG is a well-known diagnostic tool and one of the most preferred tests in every day clinical practice [94]. It is widely-used in both hospitals and ambulatory environments because it is easy to measure and contains an immense amount of information about the condition of the heart. Moreover, its associated cost is relatively low compared to most imaging techniques. In recent years, advances in sensor technology and the introduction of wireless technologies have lead to the development of various new ECG technologies, including wearable devices and smartphone set-ups [94]. The rise of these novel technologies has

introduced both opportunities and challenges in the field of ECG monitoring. Improvements in digital filters led to more accurate noise removal methods and increased signal qualities, which allows the detection and analysis of more refined ECG characteristics [6, 144, 172]. Expansion of computing power and storage capacity permit the use of more advanced signal processing techniques and advances in material sciences have lead to the development of sensors that can be worn for many days in a row [171, 182, 206].

Manual analysis of these enormous amounts of data has become a tedious, time-consuming and expensive task. Also, visual interpretation is by definition subjective and can be different for different observers, or even for the same observer at different points in time, causing inter- and intra-rater variability. Furthermore, for real-time applications where an immediate output is needed, visual inspection is not feasible. Therefore, the need for automated ECG processing methods that analyse the ECG signal in a computationally efficient way increases. As digital health gains importance, it is expected that the use of computerized ECG analysis will become an even more important tool that can complement clinicians in their daily practice [136].

In a clinical context, ECG signals are mostly recorded with different leads, where each lead corresponds to the cardiac electrical signal viewed from a different spatial angle. The combination of these leads gives a global view from the heart in three dimensions. It makes thus sense to analyse the signals from all leads simultaneously, in order to fully exploit the information that is shared over all dimensions. This can be done through the use of multilinear tensor methods. Nowadays, automated ECG analysis in clinical practice and SCD risk assessment is mainly limited to algorithms based on '*if-then*' logic: a number of logical rules is defined based on previous knowledge from clinical practice and implemented in an automated way. Machine learning methods however permit to extract much more complicated patterns from data, and to combine different features in a linear and non-linear way. It can thus be expected that the combination of machine learning and tensor methods can provide significant added value to current ECG analysis methods.

This research therefore aims at developing novel signal processing methods to extract potential risk factors from the ECG signal in an automated and reliable way through a combination of tensor methods and machine learning algorithms. The next Sections first give a comprehensive introduction to the physiological origin of the ECG signal, the abnormalities that can be observed and the different measurement set-ups. Afterwards, an overview of the current state of invasive and non-invasive SCD risk stratification is given, In Section 1.4, the principal research goals of this thesis are given, together with an overview of the different Chapters in Section 1.5. Finally, the major collaborations that were set up during the course of this PhD are outlined in Section 1.6.

1.2 The electrocardiogram

Since the ECG is used throughout the rest of the manuscript, it is evident to start with some background information on the main concepts related to both the physiological and technological basis of the signal. The explanations and Figures are mainly based on [4, 44, 199] .

1.2.1 Physiological origin

The heart is the muscle that is responsible for pumping blood throughout the body, providing it with oxygen and nutrients and removing waste products. It consists of four chambers, two atria in the upper part of the heart and two ventricles in the lower part. The heart receives deoxygenated blood in the right atrium, from where it is transported via the right ventricle to the lungs, where oxygen exchange takes place. The oxygen-rich blood is then transported back to the left atrium and left ventricle where it is pumped out through the aorta into the vestibular system. Atria and ventricles need to contract regularly to keep this cycle going and provide a continuous flow of blood through the body. When the cardiac cycle is interrupted, the heart fails to pump effectively and oxygen supply to the tissues is halted. If the blood flow is not restored within minutes, it leads to brain damage, tissue degeneration and ultimately death.
The synchronized contraction and relaxation of the cardiac muscle cells generates an electrical potential difference, which can be measured by placing electrodes on the body surface. The resulting signal is known as the ECG. Each part of the cardiac cycle corresponds to a particular wave or segment in the ECG signal:

1. The sinoatrial node (SA node) contains a group of pacemaking cells that have the ability to spontaneously depolarize. They determine the heart rate and autonomously generate an action potential. The spread of this electrical impulse corresponds with the iso-electrical line preceding the P wave.

2. As the electrical signal propagates through the atria, it causes depolarisation of the muscle cells. This causes the atria to contract, resulting in the P wave in the ECG signal. The electrical activity spreads through the atria via specialized internodal pathways from the sinoatrial node to the atrioventricular node (AV node).

3. The AV node slows down the signal to avoid that atria and ventricles contract simultaneously, which would hinder efficient blood flow between the chambers. This leads to the iso-electrical PQ segment.

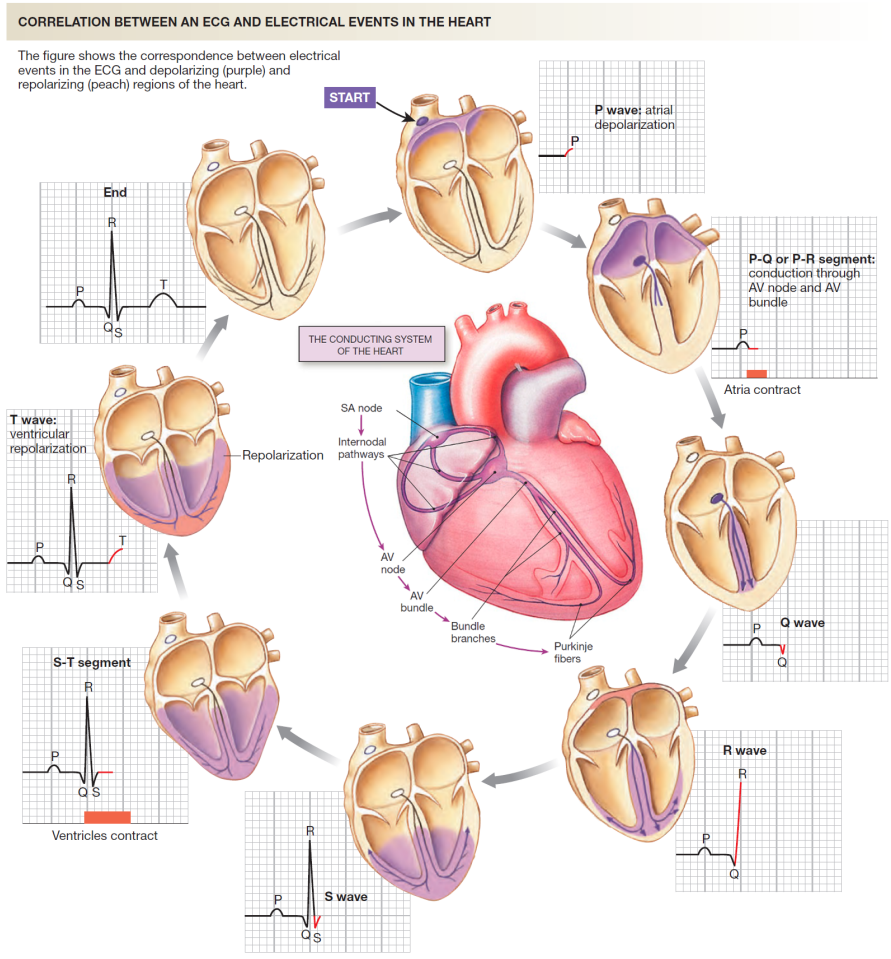


Figure 1.1: Illustrations of the different steps in the cardiac cycle that give rise to the different waves and segments in the electrocardiogram. Figure taken from [199].

4. The electrical signal is then passed to the bundles of His, the bundle branches and Purkinje fibers. This starts depolarisation of the cells in the ventricle and thus leads to ventricular contraction, visible in the ECG as the QRS complex. Conduction in the Purkinje network happens very rapidly (4 m/s), so that all contractile cells in the ventricle contract almost simultaneously. Repolarisation of the atrial muscle cells happens simultaneously, but is masked by the QRS complex.

5. After depolarisation, the muscle cells reach a plateau in the action potential, during which no electrical activity takes place. This corresponds to the iso-electrical ST-segment.

6. Finally, repolarisation and relaxation of the ventricles causes the T wave in the ECG signal.

An illustration of the different steps is shown in Figure 1.1, taken from [199]. In normal cases, the cardiac cycle as described above repeats itself in a very regular way, leading to a stable heart rhythm which is referred to as *normal sinus rhythm* (NSR).

From Figure 1.1, it is clear that the ECG contains information about the different electrical events in the heart. If there is an abnormality or disturbance in any of the stages of the cardiac cycle, this is often also visible in the ECG. It is therefore a valuable diagnostic tool to detect and analyse abnormalities in the propagation pattern.

1.2.2 Abnormal patterns

Clinicians use the ECG signal to assess the condition of the electrical conduction system of the heart. Changes in the cardiac behaviour will be reflected in the ECG signal, and abnormalities in different stages of the cardiac cycle will affect different waves and segments of the ECG. For example if the issue is related to ventricular repolarisation, this will mainly be visible in the T wave and can be diagnosed as such. In the remainder of this Section, we follow the structure described in Clifford, Azuaje, and McSharry [44], which makes the distinction between four major types of abnormal patterns that can be detected from the ECG signal.

The first type are abnormalities in the heart rate. As explained in the previous Section, in normal conditions the heart rate is regular and determined by the pacemaker cells in the SA node. When the SA node fires more quickly or slowly than usual, this is referred to as respectively sinus tachycardia or bradycardia. Both types can be normal physiological responses to for example stress or fear,

but might also be signs of underlying issues. The heart rate can be assessed by calculating the time differences between consecutive QRS complexes (also called RR-intervals). A time series of RR-intervals can be collected in a tachogram and used to examine changes in heart rate and heart rate variability (HRV).

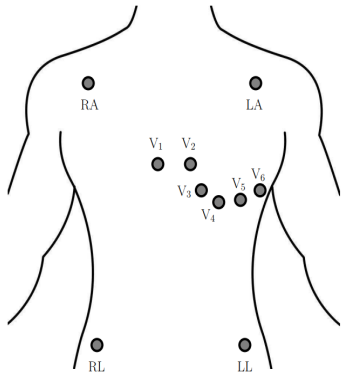
Apart from the SA node, the heart also contains additional regions in the AV node, atria and ventricles that can generate an electrical impulse. In some cases, for example when the rate of the other pacemakers exceeds the rate of the SA node, these regions take over the role of pacemaker, this is known as *ectopic depolarization* leading to *ectopic beats*. Depending on the origin of the electrical impulse, they are known as premature atrial contractions (PAC) or premature ventricular contractions (PVC). Other types of abnormal heartbeats exist, such as for example escape beats that arise when there has been an excessively long pause in the SA node.

The ECG can also reveal metabolic abnormalities such as *ischemia*, which can occur when part of the heart is not receiving enough blood flow and which might ultimately lead to myocardial cell death. It is often caused by coronary artery disease and mainly changes the appearance of the T wave and ST segment in the ECG. While typical ischemia patterns exist, they are only seen in a minority and most ischemic events are characterized by non-specific ECG changes. Other metabolic abnormalities which can be detected in the ECG signal are electrolyte abnormalities such as hyper- and hypokalemia and calcium disturbances.

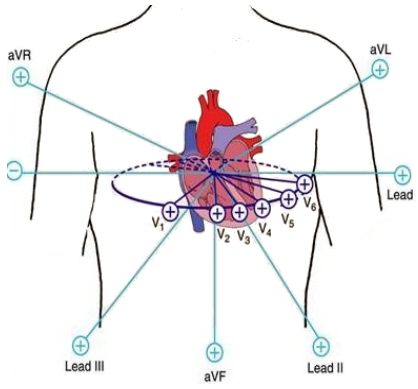
Finally, certain abnormalities of the geometry of the heart can also be assessed with the ECG. This includes pathologies where part of the heart enlarges or part of the heart undergoes cell death and scarring. While imaging techniques can give a more comprehensive view of the location and extent of these geometrical defects, examination of the ECG signal has become a convenient if imperfect screening test for structural abnormalities, since they can change the trajectory and/or magnitude of the electrical impulse. ECG analysis can then be used as a first screening tool, after which imaging techniques can give a more comprehensive view of the irregularities.

1.2.3 Measurement

The ECG signal can be measured *in-subject*, *on-subject* and *off-subject*, by placing electrodes in the chest, on the body surface and in close proximity to the subject respectively. The most common measurement set-up is however the on-subject approach where multiple electrodes are placed on the chest and/or limbs. Any pair of electrodes can be used to measure the potential difference between the two corresponding electrode locations, and is called an ECG lead or channel. The polarity of deflections in an ECG lead depends on



(a) Standard placement of electrodes for a 12-lead ECG recording. Figure taken from [229].



(b) The different spatial angles recorded with a 12-lead ECG. Figure adapted from [184].

Figure 1.2: The 12-lead ECG is recorded with ten electrodes placed on the chest and limbs, and gives a comprehensive three-dimensional view of the cardiac electrical behaviour.

the direction of the electric wavefront: when depolarisation propagates towards the positive electrode, the voltage is seen as positive and corresponds with an upward deflection in the ECG. Vice versa, propagation in the opposite direction is visible as a downward deflection.

In clinical practice, a 12-lead ECG configuration is a standard measurement set-up. It is measured with ten electrodes on standardized places: four electrodes are placed on the limbs and six on the chest. The electrodes are labelled according to their location on the body surface:

- LA = Left Arm Electrode
- RA = Right Arm Electrode
- LL = Left Leg Electrode
- RL = Right Leg Electrode
- Vx = Chest electrodes V1 to V6

The placement of all electrodes for a 12-lead ECG recording is illustrated in Figure 1.2a. Note that the limb leads are not placed *on* the limbs, but on a

location on the chest *near* the limb. Electrically, placing an electrode at any location on the limb is the same. However, if an electrode is placed further away from the heart, this adds impedance of tissue resistance. Therefore placement on the chest might be preferred.

The 12 leads derived from these electrodes can be separated into limb leads, augmented leads and chest leads. The limb leads are derived from three pairs of limb electrodes: lead I $= LA - RA$, lead II $= LL - RA$ and lead III $= LL - LA$. Note that the RL electrode is not used to obtain an ECG signal, but as reference electrode to reduce the common mode interference. The augmented limb leads aVL, aVR and aVF combine the signals from the limb leads to give additional views. They take the difference between one limb electrode and a virtual electrode consisting of the average of the two remaining electrodes. Finally, the precordial or chest leads V1–V6 are derived from the six chest electrodes. The negative electrode is a virtual electrode called Wilson's Central Terminal, which is the average of the signals from electrodes LA, RA and LL and corresponds to the electrical centre of the heart.

Each ECG lead shows the electrical activity from one spatial angle, as can be seen in Figure 1.2b. Together, the leads completely characterize the electrical activity of the heart and give a comprehensive three-dimensional view. The chest leads record the different angles in the horizontal plane while the limb leads and augmented leads provide information about the vertical plane. Clinically, the 12 leads are further divided into different regions, depending on which part of the heart can be monitored with that lead. This way, a distinction is made between the inferior region (leads II, III, aVF), lateral region (leads I, aVL, V6), anterior region (leads V1–V5) [223]. The different regions allow localization of cardiac effects. If there is for example ST elevation in leads V3 and V4 this points to an anterior myocardial infection.

Twelve lead ECG signals are mainly used in a clinical context and are often short-term measurements of ten seconds. They are mostly high-quality signals since the patient can be asked to stay still for the duration of the recording. Long-term ambulatory measurements are done with Holter monitors, which are portable ECG recorders. They are required to detect events that only occur occasionally and thus cannot be captured during a physical examination [120]. The patient has to wear the recording device during normal activities for several days, and measurements are done with two or three chest leads. Nowadays several wearable ECG recording systems exist that allow patients to record their ECG signals with for example a smartphone. While these recording devices are mostly used for single-lead ECG (typically lead I), they are very well suited for intermittent monitoring, where patients can take an ECG measurement whenever they experience an abnormal sensation.

In short, the optimal type of measurement depends on the required recording length and the desired number of leads. Depending on the application, clinical ECGs, ambulatory recordings or wearable devices might be preferred.

1.3 Risk prediction of sudden cardiac death

The goal of SCD risk prediction is to identify patients at high risk of developing dangerous arrhythmic events, in order to adapt their treatment strategies so such events can be avoided. The ideal risk stratification method detects all patients who will experience sudden cardiac death, but excludes everyone who will not develop arrhythmia or who will die from other causes. Current risk stratification approaches are not ideal for two main reasons. First, in many cases, a population is divided into a low- and high-risk group. This dichotomization however ignores the fact that risk is mostly a fluctuating continuum and hereby reduces the amount of information available. Second, most strategies focus on the patient group with the highest *relative* SCD risk while the *total* number of deaths from this population often only accounts for a small proportion of the number of deaths overall. This is referred to as the *Myerburg paradox* [161]. In SCD risk prediction for example, current guidelines are mainly focused on patients with a left ventricle ejection fraction (LVEF) $\leq 35\%$ because the incidence in this population is up to 30 times higher than average [179]. The total number of events in the general population however greatly exceeds the number of events in this population. This is illustrated in Figure 1.3, which shows the incidence and total number of events in several risk groups.

To predict the risk of sudden cardiac death in an individual, different risk factors have been identified that can be derived from the ECG and other clinical examinations. The next sections give a short overview of some of the most important risk factors. For a more comprehensive overview, we refer the reader to [224]. Although many studies have been conducted to identify and detect good prognostic risk factors, current strategies are far from ideal. This is illustrated by the fact that currently only 20–30% of patients who receive an ICD in primary prevention get appropriate shocks, and that the overall number of SCDs is still very high, meaning that a substantial part of the population that is at risk for SCD is not identified as such.

1.3.1 ECG-derived risk factors

SCD is the result of a combination of several factors, many of which can be detected using the ECG.

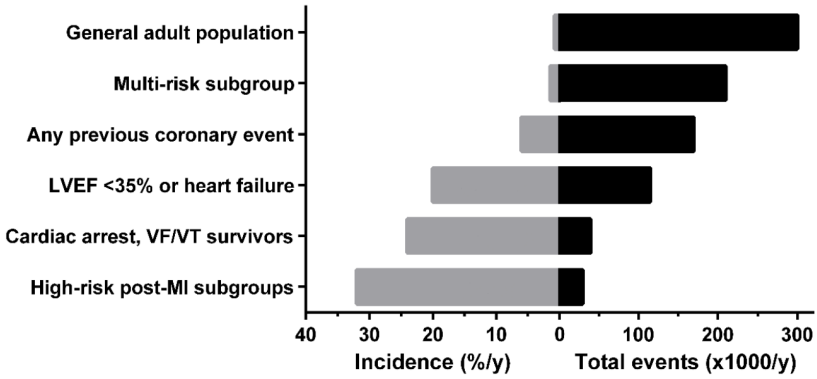


Figure 1.3: The impact of SCD in different populations, shown by comparing the incidence of SCD in each group in grey and the total number of SCD events per year in black. Figure adapted from [161, 224].

The first factor are disturbances in autonomic tone, which are caused by imbalances between the sympathetic and parasympathetic nervous system. Together they form the autonomic nervous system which regulates the physiology of the heart (including the heart rate, rhythm and contractility). They can be monitored with different HRV-measures in the time- and frequency-domain that measure the variability and predictability of the RR-interval of normal beats [191]. Some examples can be found in Appendix A, which lists a number of HRV-features that can be used for the detection of atrial fibrillation. Additionally, heart rate turbulence parameters have been defined that describe the behaviour of the RR-interval after spontaneous PVCs [15]. Finally, the acceleration and deceleration capacity of the heart rate over a long time can be quantified with Phase-Rectified Signal Averaging (PRSA) [16]. The deceleration capacity has been linked to SCD and total mortality in post-myocardial infarction patients [14].

The QRS duration corresponds with the duration of ventricular activation. An increased QRS duration ($\geq$ 120 ms) indicates a delayed conduction of the electrical signal in the ventricles. It has been shown that patients with an abnormal QRS duration benefit more from a SCD implant [12, 160]. It can be measured directly from the 12-lead ECG signal or from the *signal-averaged ECG* (SAECG). The SAECG is calculated by taking the average of multiple heartbeats in time. This is a form of synchronized averaging which is known to reduce noise. In the ECG signal, it can furthermore reveal so-called *late potentials*. Another ECG interval which is a known SCD risk factor is the QTc

interval, corresponding to the duration of ventricular repolarization normalized for heart rate. In large population studies, prolonged QTc is an independent predictor of SCD [32]. It is however also influenced by non-cardiac factors such as diabetes and obesity and is therefore perhaps rather a global risk factor [96]. Furthermore variability of the QT interval duration has also been suggested as marker of repolarization instability which is also linked to arrhythmia [95].

T wave alternans (TWA) is an abnormal ECG pattern where the amplitude of the T wave shows a beat-to-beat change in a characteristic ABABAB-pattern [200]. It can be detected in healthy hearts at high heart rates, but if it also arises at lower heart rates ($\leq$ 110 beats per minute) it is a sign of electrically unstable tissue and associated with increased mortality [40, 107].
Another abnormal pattern is QRS fragmentation (fQRS), where the QRS complex exhibits additional deflections or notches [53]. It is caused by myocardial scarring and its presence in specific cardiac regions is linked to ICD shocks and mortality in certain patient groups [223].

Finally, most ventricular arrhythmia are initiated by premature ventricular beats. The frequency and complexity of PVCs in a ECG recording has therefore also been used as risk factor, which has shown promising results in ischemic heart disease [13]. It should however be noted that PVCs are also common in the general population.

1.3.2 Other risk factors

As mentioned earlier, the LVEF is one of the most widely used prognostic risk factors. It is usually measured with an echocardiography. A LVEF $\leq 35\%$ is generally used as cut-off value, but has its limitations as mentioned before and illustrated in Figure 1.3. Advanced imaging techniques such as cardiac MRI can provide extra information about tissue characterization such as the extent of tissue injuries after myocardial infarction.

The New York Heart Association (NYHA) classification is a risk scale that combines several symptoms of congestive heart failure, which is associated with many factors attributing to ventricular arrhythmias [179]. While some studies have shown that the NYHA scale is a very strong independent predictor [237], it is also a subjective scale which suffers greatly from inter-observer variability, restricting its practical use.

Finally, invasive risk stratification is done as well by performing electrophysiological studies with programmed electrical stimulation, but their clinical value is debatable [179].

1.4 Research objectives

As discussed in Section 1.1, the development of automated signal processing methods is an important challenge in ECG analysis. Many of the risk factors described earlier show conflicting results in different studies, making efficient risk stratification additionally difficult. This can partially be explained by the inter-observer variability related to visual analysis. The principal goal of this research is therefore the development of objective and reliable algorithms for the extraction of potential SCD risk factors from the ECG signal. Since ultimately these algorithms would be used in clinical practice, they should be easy to interpret and compute from the ECG signal.

Technically, a substantial part of this research is focused on exploring the use of tensors in ECG analysis. Tensors are multilinear generalizations of vectors and matrices that have been extensively used in many domains, but are a rather novel concept in cardiac applications. A second objective of this thesis is therefore to evaluate the potential of tensors in ECG processing, and to apply tensor decomposition methods in a manner that takes into account the specific characteristics of the signal. Many tensor methods exist as they rapidly gain popularity, and the goal of this thesis does not consist in the development of new mathematical techniques, but rather in applying existing methods in an innovative way.

The first two research goals target the extraction of SCD risk factors from the ECG signal. The final objective consists of using these features for risk assessment in ICD patients. While many statistical risk models exist, we will investigate whether the machine learning methods used throughout the research can also provide added value here compared to state-of-the-art methods.

1.5 Chapter overview and personal contributions

The manuscript is divided in three main parts. The first part comprises two background chapters that give the reader an introduction in respectively tensor methods and machine learning methods necessary for the remainder of the book. The next chapters are each focused on a different ECG application, and present methods to detect and examine different ECG abnormalities. The first four chapters specifically use tensor-based methods to analyse the ECG signal, the final two chapters tackle the problem of detection and quantification of QRS fragmentation. In Figure 1.4, the structure of this dissertation is presented graphically, showing the connections between the different chapters.

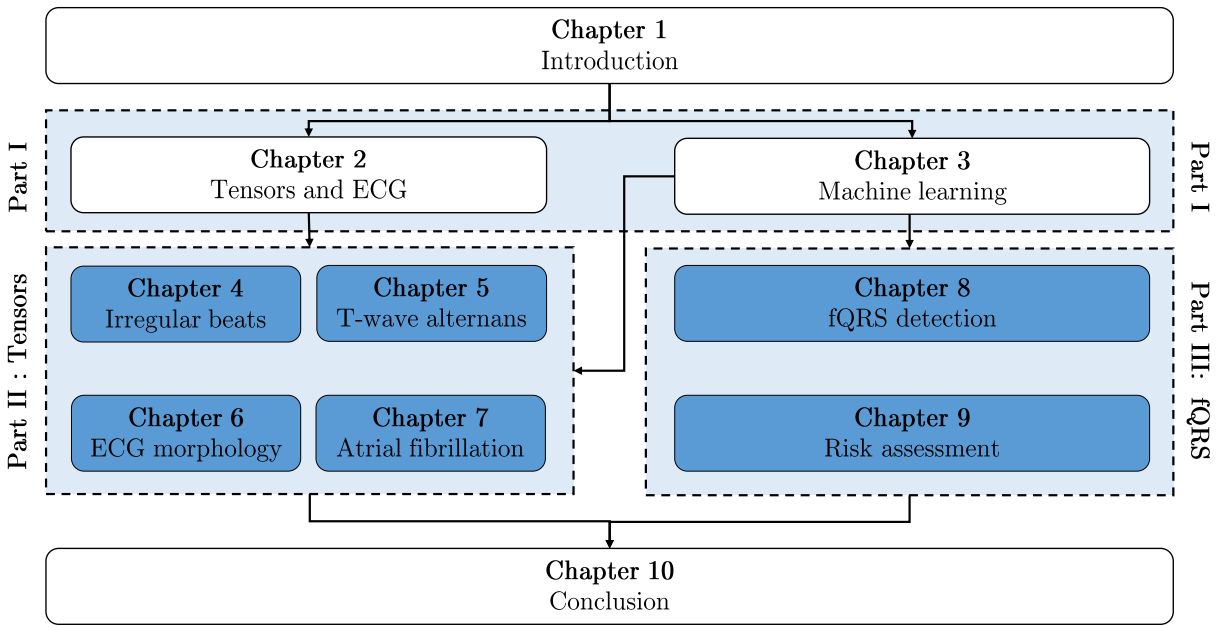


Figure 1.4: Chapter-by-chapter graphical overview of the PhD and the division in three different parts. The chapters with personal contributions are marked in blue.

The chapters with the main personal contributions of this thesis are highlighted in blue.

1.5.1 Part I: Background

The first part of this thesis consists of two Chapters meant to give some background information related to the most important concepts connected with tensors and machine learning.

Chapter 2

In this chapter, the use of tensors in ECG processing is introduced and motivated to the reader. Tensors are multilinear generalizations of vectors and matrices, and have been successfully applied in different domains. The main concepts and methods that are used in Part II are explained, together with an overview of the use of tensors in cardiac applications until now.

Chapter 3

The majority of the studies in this PhD apply different classification methods to allocate signals or patients to a specific group. Classification is part of the large field of machine learning, which is presented in this Chapter. It includes the most important techniques for supervised and unsupervised classification, as well as performance measures to quantify the results.

1.5.2 Part II: Tensor-based methods

The second part of the manuscript contains four Chapters that each present a method to detect and analyse different ECG characteristics. All algorithms proposed in this part rely on tensor-based approaches to decompose the ECG signal.

Chapter 4

Chapter 4 presents a first application where tensors are used for ECG processing and a first personal contribution. Here, the goal is to detect irregular heartbeats such as PVCs and PACs in an unsupervised manner. The method was first published as a conference paper [88], and has been extended to include a full

validation on two publicly available datasets. Additionally, this Chapter contains an analysis of the most important tensor decomposition parameters, which will be used throughout the rest of the dissertation.

Chapter 5

The second personal contribution of this work consists of an algorithm for automatic detection of T wave alternans, presented in Chapter 5. TWA is a pattern that is associated with increased heart rates, and that becomes problematic when detected at normal rates. The method uses the same approach as described in Chapter 4, but is modified to deal with time changes within the signal, as can happen when the heart rate changes. Preliminary results were presented at two conferences [90, 91], and the final method was published as [87].

Chapter 6

A method to analyse the changes in ECG morphology right before in-hospital cardiac arrested is discussed in Chapter 6. This study was conducted during a research visit in the University of California, San Francisco. We analysed long-term signals of patients in the intensive care unit who suffered a cardiac arrest. The main difficulty here is signal quality, as the signals contain significant amounts of noise. The proposed method incorporates information about the quality in the tensor decomposition to perform a more robust analysis, making it well-suited for use with real-life signals. Results of a preliminary study on a smaller dataset were published as a conference paper [86], and a paper describing the full analysis is currently in review.

Chapter 7

The final tensor-based algorithm detects atrial fibrillation in single and multilead ECG signals. Contrary to the previous Chapters, it uses the multilinear singular value decomposition to decompose the ECG tensor, and classifies signals with a combination of morphological features and heart rate variability features. Part of the results have been presented at the 2018 IEEE Asilomar conference.

1.5.3 Part III: QRS fragmentation

QRS fragmentation is one of the promising SCD risk factors described in Section 1.3. Detection however heavily relies on visual inspection, which has been shown to be dependent on rater experience [225]. The final part of the research consists of methods that use machine learning to detect and quantify fQRS.

Chapter 8

This Chapter describes a method to detect and quantify the presence of QRS fragmentation in an ECG signal. It uses Phase-Rectified Signal Averaging and Variational Mode Decomposition to segment and characterize the QRS complex. The method is an original contribution of this work, which is accepted for publication in IEEE Journal of Biomedical and Health Informatics. Preliminary results were published as [89].

Chapter 9

The method described in the previous Chapter is used here to examine whether the proposed fQRS score can be used as risk factor to predict mortality in a population of ICD patients. Standard statistical techniques are compared with more advanced machine learning methods to verify whether they have added value. The method described in the first part of this chapter was presented at Computing in Cardiology 2018, where it received a nomination as finalist for the Rosanna Degani Young Investigator Award.

Chapter 10

The final chapter summarizes the thesis by listing the main contributions and suggesting additional directions for future research. Two appendices contain respectively a description and analysis of the heart rate variability features used for AF detection and the method to construct and validate the noise model used in Chapter 6.

1.6 Collaborations

This PhD research was conducted in the Biomed research group for biomedical signal processing, which is part of the STADIUS Center for Dynamical

Systems, Signal Processing and Data Analytics in the Department of Electrical Engineering of KU Leuven, under the supervision of professor Sabine Van Huffel. I collaborated with several people both within and outside the Biomed group, leading to different (conference) publications.

The first collaboration, with Thomas De Cooman, Carolina Varon, Devy Widjaja and Tim Willemen, was done in the framework of the 2014 Computing in Cardiology Challenge, where we reached the first place in phase II of the competition and the third place in phase III. Thomas De Cooman was the main responsible who did most of the analysis, but I helped with development and finetuning of the method. The method was first presented at Computing in Cardiology 2014 [57], and an extended version was later published in Physiological Measurement [56].

Next, I assisted Alexander Suárez León with his work on tensor-based detection on irregular heartbeats, which can be seen as a follow-up to the method presented in Chapter 4. I also presented this as a poster at Computing in Cardiology 2015 [205]. A next extension of this method was presented at the 2017 Computing in Cardiology [204].

During my PhD I was the main supervisor of four master thesis students, whose research also lead to several publications. The master thesis of Ofelie De Wel was the start of the work on tensor-based irregular heartbeat detection. Discussions during her thesis lead to the idea of the method proposed in Chapter 4, which I developed and presented at Computing in Cardiology 2015 [88]. Jonathan Moeyersons worked on T wave end detection methods, which was published in the proceedings of BIOSIGNALS 2017 [150]. Together with Lieven Billiet, I supervised the master thesis of Adriaan Lambrechts who examined a clinical QT prolongation dataset collected by Eline Vandael from the department of Clinical Pharmacology in the context of her PhD research [219]. The results of this study are currently being written into two separate publications. Finally, I supervised the thesis of Simon Geirnaert together with Sibasankar Padhy, who worked on tensor-based detection of atrial fibrillation. This lead to the method described in Chapter 7 which was presented at the IEEE Asilomar 2018 conference. The main concept of this method was developed by myself in the context of the 2017 Computing in Cardiology challenge, Simon Geirnaert worked out the full method and optimized the final classification.

Within the context of the research included in this thesis, I mainly collaborated with Carolina Varon. She helped to start up the work on T wave alternans detection based on a preliminary student project she supervised and provided the code for PRSA used in Chapter 8. Part III, which includes all work related to QRS fragmentation, was done in collaboration with her and Sibasankar Padhy.

Apart from collaborations within the BIOMED group, several collaborations with other groups were also set up. The collaboration with the tensor group of Lieven De Lathauwer was essential for the development of the tensor-based methods in Part II of this thesis. Together with Martijn Boussé, I worked on the development of a supervised approach for irregular heartbeat detection which forms the basis for the AF detection method in Chapter 7 and which was presented at the 2017 IEEE EMBS conference [26]. Furthermore, he provided the Matlab code for the weighted tensor decompositions used in Chapter 6 and provided valuable comments to the manuscript of the resulting paper. The initial discussions that led to the method described in Chapter 7 were done with him and Otto Debals. Michiel Vandecappelle developed the updating method shortly described in Chapter 6. Furthermore, Nico Vervliet and the rest of the group developed the Matlab toolbox Tensorlab [231] which was used throughout the thesis.

Throughout my PhD, I contributed to the ERC Advanced Grant BIOTENSORS (no. 339804), which aimed at the development of tensor methods for biomedical signal processing applications. Within the project, I mainly worked on work package 7, 'tensor-based blind source separation in functional monitoring'. The methods described in Chapters 4-7 all fit within the BIOTENSORS framework. Additionally, in 2014 I obtained an IWT research grant for Strategic Basic Research (duration January 2015-December 2018). The research stay in UCSF was supported by an FWO travel grant.

Different clinical partners were involved in this thesis, who helped to understand the obtained results from a clinical point of view and provided the majority of the ECG signals. This research was possible due to a strong collaboration with the department of Cardiology of UZ Leuven. Prof. dr. Rik Willems, co-supervisor of this PhD, and dr. Bert Vandenberk were involved in the clinical interpretation and analysis of most of the results of this thesis. Additionally, the extensive fQRS database described in part III was essential for the presented results and was annotated by Bert Vandenberk, Tomas Robyns, Mathias Claeys, Frederik Helsen and Sofie Van Soest, all associated with the department of Cardiology.

An international clinical collaboration was set-up with prof. Xiao Hu from the department of Physiological Nursing from the University of California San Francisco during a three month research visit in San Francisco. He collected and provided the large dataset of intensive care patients together with Duc Do and Robin Ma. This resulted in the work presented in Chapter 6, of which preliminary results were presented at Computing in Cardiology 2017 [86].

I would like to acknowledge and thank all collaborators, because the joint efforts allowed me to gain different insights and perspectives which greatly influenced

the work presented in the next Chapters. Furthermore the availability of clinical ECG signals was essential to ensure that the developed methods could deal with authentic ECG and noise characteristics, improving the practical usability in the future.

Part I

Background

Chapter 2

Tensors and ECG

As explained in the introduction, this research proposes several tensor-based ECG processing methods. While tensor methods have been successfully applied in many fields from telecommunication to psychometrics, they are a rather novel concept in the analysis of cardiac signals. This Chapter provides a background to concepts and methods related to tensors, which will be used throughout the rest of the thesis. First a motivation for the use of tensor methods in ECG processing is given, after which the main concepts, methods and calculations are introduced. Finally, a short literature overview of past applications of tensor techniques in cardiac applications is given. The notations and definitions given in this Chapter will be followed throughout the rest of the manuscript.

2.1 Introduction

In many applications, data and signals are measured and presented in a natural two-dimensional structure: a speech signal can be captured by a set of microphones, a digital image consists of horizontal and vertical pixels and a clinical ECG signal is mostly measured with multiple spatial leads. Digital signals like the previous examples are generally stored and analysed in matrix structures.

It is however easy to imagine different situations where more than two parameters describe these signals: speech signals from different speakers can be recorded with the same microphone set, a colour image contains different RGB-values for each pixel, and one might wish to compare the different heartbeats in

an ECG signal for example. While matrices could technically be used to analyse these multidimensional signals by constructing multiple matrices for each additional parameter value and concatenating them in one big matrix; there is no reason why the original multidimensionality of the data should not be preserved and exploited maximally. This way the information that is shared over all dimensions can be analysed simultaneously. This can be done in a straightforward way through the use of *tensors*. Mathematically speaking, tensors are higher-order generalizations of vectors and matrices. They have been successfully applied in many different fields, including but not limited to chemometrics, psychometrics and telecommunication applications [119, 42, 197]. In biomedical signal processing, they have gained popularity in neuroscience applications such as epilepsy monitoring [18, 63, 3], brain tissue segmentation [158, 23] and EEG processing [243, 126, 49].

This chapter introduces the use of tensors in ECG signal processing, a concept which will be further expanded on in the next Chapters. Since the field of tensors and tensor decomposition methods is increasing rapidly, the concepts and methods explained in this Chapter are restricted to the methods that are relevant for the rest of the thesis. For an extensive summary of the remaining, we refer the reader to [119, 42, 197].
The next Sections first give a motivation for the use of tensors in ECG processing (Section 2.2). Then, the most important concepts and tensor operations are described in respectively Sections 2.3 and 2.4. Finally, Section 2.5 reviews the use of tensor methods in different ECG applications in literature.

2.2 Motivation

ECG signals can be measured with one or multiple leads. Especially in clinical contexts, multilead recordings are considered the golden standard. Furthermore, long-term measurements are done with Holter monitors which typically have a minimum of two channels. A big advantage of multilead measurements is that each lead records the cardiac electrical activity from a different spatial angle, and different leads thus complement each other. Together, they give a comprehensive three-dimensional view of the electrical activity of the heart.

Most methods dealing with multilead ECG signals process different ECG leads separately and later combine the results from all channels in a final step (also referred to as *late integration*) [123, 101, 233]. This however fails to exploit the correlations between the different leads and removes any structural link between the different spatial angles. The use of tensors allows simultaneous analysis of the complete ECG signal, making use of all the information present

in the signal. For applications which require assessment of the global behaviour of the heart, this forms a clear advantage.

While matrix-based methods have been used successfully in many applications, they often put constraints on the signals in order to reach unique solutions. Principal Component Analysis (PCA) for example requires the different components to be orthogonal to each other, an assumption which is not necessarily met in real-life conditions. Tensor techniques are however unique under mild conditions (see also Section 2.4.3).

Finally, when dealing with biomedical signal processing problems, interpretability of the final outcome is an important concept to keep in mind, especially when the results need to be communicated with clinicians and patients. Nowadays, deep learning methods such as artificial neural networks are rapidly gaining popularity. While they have shown to be very powerful, they are in their core *black box* methods. This means that it is not always straightforward to interpret how the final result is calculated from the input values. Tensor methods on the other hand can lead to interpretable components which can be physiologically interpreted.

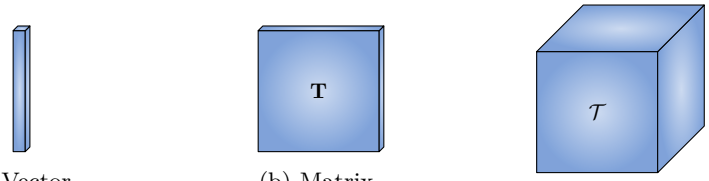
2.3 Basic concepts and notations

Generally speaking, N-dimensional data are represented by a Nth order tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$. The number of modes of a tensor is equal to the *order*, e.g. a tensor with three modes is a third-order tensor. Within this framework, it is also possible to define *scalars* ($N = 0$), *vectors* ($N = 1$) and *matrices* ($N = 2$). Figure 2.1 shows an example of these structures for orders up to three. From here on, we refer to scalars with lower-case letters (t), to vectors with bold lower-case letters ($\mathbf{t}$) and to matrices with bold upper-case letters ($\mathbf{T}$). Finally, tensors are written as calligraphic letters ($\mathcal{T}$). Since, in the next Chapters, only third-order tensors are used, the rest of the Chapter is mainly focused on these structures. All Equations can however be easily adapted to higher-order tensors.

A matrix $\mathbf{T}$ consists of row- and column vectors. In a similar way, mode-n vectors or fibers are defined for tensors:

Definition 2.1 (mode-n vector)**.** A mode-n vector of a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is the vector that is the result of fixing all indices except the n^{th} index.

In the matrix case, the columns form the mode-1 vectors and the rows the



(a) Vector. (b) Matrix. (c) Third-order tensor.

Figure 2.1: Schematic representations of a vector ($N = 1$), matrix ($N = 2$) and tensor ($N = 3$).



(a) Mode-1 vectors or fibers. (b) Mode-$(2, 3)$ slices.

Figure 2.2: The mode-1 vertical vectors and mode-$(2, 3)$ horizontal slices of a $4 \times 4 \times 4$-tensor. The other mode-n vectors and mode-(m, n) slices can be constructed in a similar way.

mode-2 vectors. Figure 2.2a shows the mode-1 vectors (fibers) of a third-order tensor.

A slice is defined in the same way:

Definition 2.2 (mode-(m, n) slice)**.** A mode-(m, n) slice of a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is the matrix that is the result of fixing all indices except the m^{th} and n^{th} index.

In a third-order tensor, the different types of slices are typically named after their corresponding directions (similar to row and column vectors in matrices): this way horizontal (mode-(2,3)), lateral (mode-(1,3)) and frontal (mode-(1,2)) slices are defined.

Finally, the Frobenius-norm of a tensor is defined as:

Definition 2.3 (Frobenius-norm)**.** The Frobenius-norm $||\mathcal{T}||_F \in \mathbb{R}$ of a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is calculated as:

$$||\mathcal{T}||_F = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \ldots \sum_{i_N=1}^{I_N} t_{i_1,i_2,\ldots,i_N}^2}$$

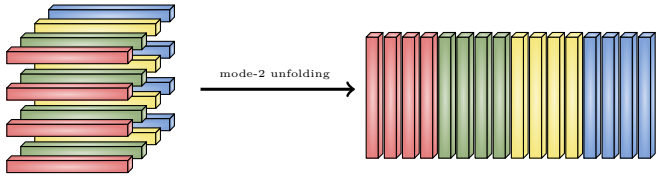


Figure 2.3: The mode-2 unfolding of a $4 \times 4 \times 4$ - tensor. The mode-2 vectors make up the columns of the resulting 4×16 - matrix.

It is essentially the multilinear generalization of the L_2-norm commonly used in vectors.

2.4 Tensor operations

The different tensor operations can be grouped in three main classes: transformations, multiplications and decompositions. The main operations of each class are summarized in the next Subsections.

2.4.1 Tensor-matrix transformations

In many applications, tensors are converted to matrices and vice versa. The transformation from a tensor to a matrix is referred to as tensor unfolding or flattening, a process where the elements of the tensor are reformatted in a lower-order structure [119].

One way of transforming tensors to matrices is the case of mode-n unfolding or matricization, which places the mode-n fibers of a tensor $\mathcal{T}$ as column-vectors in a matrix:

Definition 2.4 (mode-n unfolding)**.** The mode-n unfolding of a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is written as $\mathbf{T}_{(n)}$. It maps the tensor elements $(i_1, i_2, \cdots, i_N)$ to matrix elements (i_n, j), with:

$$j = 1 + \sum_{\substack{k=1 \\ k \neq n}}^{N} (i_k - 1) J_k, \text{ met } J_k = \prod_{\substack{m=1 \\ m \neq n}}^{k-1} I_m.$$

The mode-2 unfolding of a third-order tensor is presented in Figure 2.3. The mode-1 and mode-3 unfoldings can be visualized in a similar way.

Similarly to matricization, tensors can also be vectorized, e.g. transformed into a vector. In that case, the mode-n fibers of the tensor are rearranged in a vector.

When data or signals are naturally collected in matrix-format, they have to be transformed into tensors in order to be able to apply tensor methods. This is done through tensorization, where a vector or matrix is mapped onto a tensor by creating additional modes. Many different tensorization methods exist, and the choice of method is dependent on both the data and the application. An overview of the most well-known deterministic and statistical techniques can be found in [64].
In the next Chapters, tensorization will be mainly done through segmentation, e.g. by dividing the signals in equal-length segments and stacking these in the frontal slices of a third-order tensor. The main advantage of ECG signals is that they contain 'natural' segments in the form of heartbeats or individual ECG waves. Tensorizing them in such manner gives rise to third-order tensors with modes *time* $\times$ *channels* $\times$ *heartbeats*, where each mode-2 vector contains the temporal profile of one heartbeat in one channel.

2.4.2 Matrix- and tensor multiplications

Although this Chapter is focused on tensors, several matrix multiplications play an important part in tensor calculations. The most relevant operations are the Kronecker product, Kathri-Rao product and Hadamard product.

Definition 2.5 (Kronecker product)**.** The Kronecker product $\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{I_1 I_2 \times J_1 J_2}$ between two matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times J_1}$ and $\mathbf{B} \in \mathbb{R}^{I_2 \times J_2}$ is defined as:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1J_1}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2J_1}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I_1 1}\mathbf{B} & a_{I_1 2}\mathbf{B} & \cdots & a_{I_1 J_1}\mathbf{B} \end{bmatrix}$$

The Kathri-Rao product is the columnwise Kronecker product, and it is defined by:

Definition 2.6 (Kathri-Rao product)**.** Given two matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2}$ and $\mathbf{B} \in \mathbb{R}^{I_3 \times I_2}$, then the Kathri-Rao product $\mathbf{A} \odot \mathbf{B} \in \mathbb{R}^{I_1 I_3 \times I_2}$ can be calculated as:

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} \mathbf{a_1} \otimes \mathbf{b_1} & \mathbf{a_2} \otimes \mathbf{b_2} & \ldots & \mathbf{a_n} \otimes \mathbf{b_n} \end{bmatrix}$$

Note that the Kathri-Rao product is only defined for matrices with the same number of columns.
Finally, the Hadamard product is the element-wise multiplication of two matrices of the same size:

Definition 2.7 (Hadamard product)**.** Given two matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2}$ and $\mathbf{B} \in \mathbb{R}^{I_1 \times I_2}$, the Hadamard product $\mathbf{A} * \mathbf{B}$ can be calculated as:

$$\mathbf{A} * \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1I_1}b_{1I_1} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2I_1}b_{2I_1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I_11}b_{I_11} & a_{I_12}b_{I_12} & \cdots & a_{I_1I_2}b_{I_1I_2} \end{bmatrix}$$

An important tensor-matrix multiplication is the mode-n product, which can be seen as a generalisation of the matrix-matrix multiplication:

Definition 2.8 (mode-n product)**.** The mode-n product $\mathcal{T} \cdot_n \mathbf{U} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ between tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$ has elements:

$$(\mathcal{T} \cdot_n \mathbf{U})_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} t_{i_1 i_2 \ldots i_N} u_{j i_n},$$

for all indices.

The mode-n product multiplies each mode-n vector with $\mathbf{U}$ [119]. It can also be explained in terms of the mode-n unfolding:

$$\mathcal{S} = \mathcal{T} \cdot_n \mathbf{U} \Leftrightarrow \mathbf{S}_{(n)} = \mathbf{U}\mathbf{T}_{(n)}. \tag{2.1}$$

In a similar way, the multiplication between two matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times J}$ and $\mathbf{B} \in \mathbb{R}^{J \times I_2}$ can be written as a mode-2 product:

$$\mathbf{A}\mathbf{B} = \mathbf{A} \cdot_2 \mathbf{B}^{\mathrm{T}}. \tag{2.2}$$

The inner product between two tensors is defined as:

Definition 2.9 (Inner product)**.** The inner product $\langle \mathcal{A}, \mathcal{B} \rangle \in \mathbb{R}$ between tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_N}$ is defined as:

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} a_{i_1 i_2 \ldots i_N} b_{i_1 i_2 \ldots i_N}.$$

Tensors with an inner product equal to zero are orthogonal.

2.4.3 Tensor decompositions

Tensor decomposition methods decompose a tensor in the sum of a number of low-rank components. Different approaches exist, of which the Canonical Polyadic Decomposition (CPD) [99], Multilinear Singular Value Decomposition (MLSVD) [61] and Block Term Decomposition (BTD) [58] are some well-known methods. The choice of decomposition method depends on the application, and more specifically on the characteristics of the data available and the desired output. The next Chapters make use of the MLSVD and CPD to decompose tensors, which will be explained next.

Multilinear Singular Value Decomposition

The multilinear singular value decomposition or higher-order singular value decomposition is the multilinear generalization of the singular value decomposition (SVD) for matrices [61]. As a recap, the SVD of a matrix $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2}$ can be written as follows:

Definition 2.10 (Singular Value Decomposition)**.** The singular value decomposition of a matrix $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2}$ is equal to:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}} = \mathbf{\Sigma} \cdot_1 \mathbf{U} \cdot_2 \mathbf{V},$$

with:

- $\mathbf{U} \in \mathbb{R}^{I_1 \times I_1}$ and $\mathbf{V} \in \mathbb{R}^{I_2 \times I_2}$ orthogonal matrices,
- $\mathbf{\Sigma} \in \mathbb{R}^{I_1 \times I_2}$ a diagonal matrix with non-negative, non-increasing real numbers on the diagonal

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(I_1, I_2)} \geq 0.$$

The diagonal entries σ_i are known as the singular values of $\mathbf{A}$, the columns of $\mathbf{U}$ and $\mathbf{V}$ are respectively called the left and right singular vectors. The rank of the matrix is defined as the number of non-zero singular values. The SVD of a matrix $\mathbf{A}$ is often used when one wants to compute a low-rank approximation of $\mathbf{A}$, since it can be proven (in the *Eckart-Young theorem* [74]) that the optimal rank-r approximation of $\mathbf{A}$ is calculated by removing the singular vectors corresponding to the smallest singular values of $\mathbf{A}$ until r singular vectors remain.

The MLSVD of a third-order tensor $\mathcal{T}$ is now defined as follows:

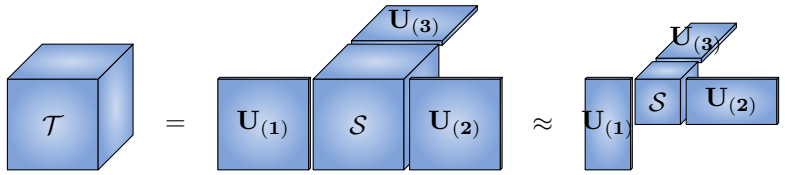


Figure 2.4: Visualization of the MLSVD and LMLRA of a third-order tensor $\mathcal{T}$.

Definition 2.11 (Multilinear Singular Value Decomposition)**.** The MLSVD of a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is equal to:

$$\mathcal{T} = \mathcal{S} \cdot_1 \mathbf{U}^{(1)} \cdot_2 \mathbf{U}^{(2)} \cdot_3 \mathbf{U}^{(3)}$$

with:

- $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ orthogonal factor matrices
- $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ a core tensor with following properties:
  - *All-orthogonality* : Two subtensors $\mathcal{S}_{i_n=\alpha}$ and $\mathcal{S}_{i_n=\alpha}$ are orthogonal for all values of n, α and β (with $\alpha \neq \beta$).
  - *Ordening* : for all values of n holds:

$$||\mathcal{S}_{i_n=1}||_{\mathrm{F}} \geq ||\mathcal{S}_{i_n=2}||_{\mathrm{F}} \geq \cdots \geq ||\mathcal{S}_{i_n=I_n}||_{\mathrm{F}} \geq 0,$$

The Frobenius-norms $||\mathcal{S}_{i_n=i}||_{\mathrm{F}} = \sigma_i^{(n)}$ are the multilinear singular values of $\mathcal{T}$ and the column vectors of $\mathbf{U}^{(n)}, \forall n \in \{1, 2, \ldots, N\}$ are the multilinear singular vectors. A visualisation of the MLSVD for a third-order tensor is shown in Figure 2.4.

The matrix rank can also be generalised to tensors: the multilinear rank of a tensor is defined as:

Definition 2.12 (Multilinear rank)**.** The multilinear rank $(R_1, R_2, \ldots, R_N)$ of tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as the n-tuple of the mode-n ranks R_n, with:

$$R_n = \mathrm{rank}_n(\mathcal{T}) = \mathrm{rank}(\mathbf{T}_{(n)}).$$

The n-rank is thus the dimension of the vector space of the mode-n vectors, or the number of linearly independent mode-n vectors.

The MLSVD can also be used to calculate a low multilinear rank approximation (LMLRA) of a tensor $\mathcal{T}$, but contrary to the matrix case it is not necessarily

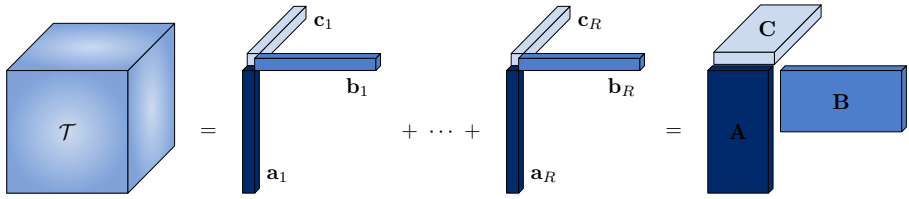


Figure 2.5: Schematic representation of the Canonical Polyadic Decomposition of a third-order tensor. The individual factor vectors $\mathbf{a}_i$, $\mathbf{b}_i$ and $\mathbf{c}_i$ for $i = 1, ..., r$ can be combined into factor matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$.

the optimal approximation. Truncation can be done in a similar way as the matrix case, e.g. by only keeping the mode-n singular vectors corresponding to the highest mode-n singular values. While the LMLRA calculated in this way is not perfect, it is usually a good approximation since the multilinear singular values are ordered (see Definition 2.11), meaning that most 'energy' is concentrated in the vectors corresponding to the first singular values (the part that is kept in the LMLRA). It is therefore considered a suitable solution for most applications, which can be refined by iterative algorithms if necessary [108].

Canonical Polyadic Decomposition

A Polyadic Decomposition expresses a tensor as a sum of rank-one components. If the number of components is the minimal number of components necessary to exactly decompose the tensor, this decomposition is called the Canonical Polyadic Decomposition. For a third-order tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ it is expressed as:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r \tag{2.3}$$

with R the total number of components or rank of the decomposition and $\mathbf{a}_r \in \mathbb{R}^{I_1}$, $\mathbf{b}_r \in \mathbb{R}^{I_2}$ and $\mathbf{c}_r \in \mathbb{R}^{I_3}$ for $r = 1, \ldots, R$ the factor vectors. The factor vectors can be combined into factor matrices for each mode, i.e.

$$\mathbf{A} = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \ldots & \mathbf{a_r} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{b_1} & \mathbf{b_2} & \ldots & \mathbf{b_r} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{c_1} & \mathbf{c_2} & \ldots & \mathbf{c_r} \end{bmatrix}$$

Using this notation, we can write Equation 2.3 more compactly as:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \tag{2.4}$$

A schematical presentation of CPD for a third-order tensor is shown in Figure 2.5.

In Equations 2.3 and 2.4, R stands for the rank of the tensor decomposition or the number of rank-1 components. For a canonical decomposition, R is equal to the rank of the tensor. Note that the tensor rank is different from the multilinear rank (Definition 2.12). While the definition of tensor rank is straightforward, its determination is a non-trivial problem. Different approaches for automatic rank estimation exist such as the core consistency diagnostic [31] or Rankest contained in Tensorlab [231]. Both approaches gradually increase the tensor rank and compare the results of the decomposition with the original tensor. They however lead to an overestimation of the tensor rank for noisy signals [48], and are therefore less suitable for biomedical signals which inherently contain noise. In practice, visual inspection of the obtained components for different ranks is very informative for evaluating the results and determining suitable rank values. Another starting point are the multilinear singular values, which can provide a lower-bound for determining the rank [42].

One of the main advantages of tensor decompositions compared to matrix-based methods is that the resulting components are unique under mild conditions (up to permutation and scaling). A general framework for uniqueness properties of third-order tensors is presented in [70, 69]. This means that no additional constraints have to be imposed on the data or factor matrices which is required for matrix decompositions such as Principal Component Analysis (orthogonality) or Independent Component Analysis (independence).

In the literature, many algorithms for CPD calculation can be found. Throughout the thesis, Tensorlab [231] is used for virtually all tensor calculations. Tensorlab is a Matlab-based toolbox which contains numerically optimized methods for tensor calculations. CPD is computed with an optimization based algorithm which starts from a (random) initialization that is iteratively updated until convergence. It is possible that the algorithm obtains different outcomes for different initializations, since it may converge to a local minimum instead of the optimal solution. It can therefore be useful to evaluate the stability of the results by evaluating the method over a considerable number of initializations.

While many other tensor decomposition methods exist, CPD can be considered one of the most accessible approaches. Generally, it leads to components which can be easily visualised and interpreted. Furthermore, different computation

methods have been developed that deal with practical issues such as missing data [232] or noisy signals [25]. In Chapters 4–7, these methods will be applied to identify different ECG features, highlighting the power of these tensor decompositions in cardiac applications.

2.5 An overview of tensors in ECG signal processing

Multilead ECG signals are described by two parameters, namely time and space. Compared to other biomedical signals they are relatively noisefree and contain an inherent structure. They therefore form prime candidates for use with tensor methods. However, despite these favorable characteristics, tensors methods are still only used in a limited number of ECG applications.

Tensors were first used in ECG processing to separate the fetal and maternal electrocardiogram measured in pregnant women. This is essentially a blind source separation problem where multiple electrodes are placed on the chest and stomach of the expecting mother. The signal measured by these electrodes contains a mixture of the heart signal of both the mother and the baby. Blind source separation techniques such as Independent Component Analysis (ICA) aim to separate both signals as accurately as possible. They have been applied numerous times since 1995 [60, 62, 241, 38] and have been shown to outperform classical methods such as adaptive filters [241]. While ICA is applied directly to matrices, it can be seen as a tensor decomposition technique since in its core it relies on diagonalization of third- or fourth-order tensors [59].

A large number of methods were developed in the context of the PhysioNet/-Computing in Cardiology (CinC) Challenge, an annual competition that aims at tackling ECG processing problems that are unsolved or not well-solved. The 2013 challenge dealt with the problem of fetal ECG processing. For this purpose, more than 25 teams developed techniques, of which seven teams used tensor decomposition techniques. While a number of methods rely on the aforementioned ICA [183, 20, 228, 133], different tensor-based methods were also used. Niknazar, Rivet, and Jutten [165] reshape the measured signals in a three-way tensor which is then decomposed with CPD, separating the maternal and fetal ECG signals. A similar approach is taken by Akhbari et al. [7], but here a weighted version of CPD is used in order to improve the robustness of the method. While the CPD approaches did not lead to very good results, the ICA-based methods by Behar, Oster, and Clifford [20] and Varanini et al. [228] were the best-scoring methods overall. Other tensor decomposition methods used for FECG separation are PARAFAC2 [2], a variant on CPD which allows variations in one mode, and the periodic Tucker Decomposition [5].

Automatic detection of various cardiac abnormalities is an important problem, as discussed in Chapter 1. Irregular heartbeat detection refers to the task of distinguishing normal sinus beats from abnormal beats. Li et al. [128] and Huang and Zhang [104] both proposed tensor-based schemes, where spectral information is used to construct tensors which are then decomposed with higher-order variations of respectively Principal Component Analysis and Linear Discriminant Analysis. Another abnormal heart rhythm is atrial fibrillation (AF). It is one of the most prevalent arrythmia affecting an estimated one percent of the general population. Block term decomposition has been applied to analyse the atrial activity [185, 242] and automatically extract atrial sources [166] in multilead ECG signals of patients with AF. Padhy et al. developed a multiscale approach combining wavelet decompositions and MLSVD for detection of T wave alternans [170] and classification of different types of myocardial infarction [169]. Finally, some other examples of applications within ECG processing which have been tackled with tensors are ECG denoising [121] and compression [168, 52].

Note that several of the applications mentioned here are also discussed in the remainder of the manuscript. Chapter 4 presents an unsupervised tensor-based approach to irregular heartbeat classification and Chapter 5 deals with automatic detection of T Wave Alternans. Finally the problem of AF detection is tackled in Chapter 7.

2.6 Conclusion

Tensors are multilinear generalizations of vectors and matrices. This Chapter introduces the use of tensors in ECG processing and gives the background necessary for the next Chapters. The tensor decomposition methods described in Section 2.4.3 will be used in Chapters 4–7 for automatic detection and analysis of various cardiac problems.

Chapter 3

Machine learning and classification

After the previous chapter introduced tensors and tensor methods, this chapter gives a background to the different machine learning and classification techniques used in this thesis. The different ECG applications handled all require the classification of an ECG signal in different classes. Several classification methods can be used to tackle this problem. They can be divided into two main groups, explained in the different sections of this chapter: unsupervised techniques and supervised techniques. A final section explains how the results obtained by these classification methods can be evaluated.

3.1 Introduction

Machine learning was first defined in 1959 by Arthur Samuel as 'a field of study that gives computers the ability to learn without being explicitly programmed' [189]. Nowadays it has become a large, interdisciplinary field, spanning computer science, statistics, information theory and even philosophy. It involves many learning problems, including classification, regression, ranking and dimensionality reduction.

In many biomedical applications, including the ECG applications tackled in the next Chapters, the ultimate goal is classification: a patient or signal has to be assigned to a certain group, and this decision has to be made using a number of predefined features or characteristics. Take for example a doctor

that diagnoses a patient: he effectively determines whether a patient is ill based on patient characteristics such as body temperature or blood results. In the next chapters, classification will for instance be used to determine whether a heartbeat is normal or irregular or to classify whether a QRS complex contains fragmentation.

Within the large field of classification methods, we can distinguish two main types: *supervised* and *unsupervised* classification methods. Supervised classification uses a set of training points where the true class labels are known. The algorithms use this training set to learn the optimal way of mapping the features onto the class label. Unsupervised approaches do not require training data, the algorithm instead learns the underlying structure of the data. A special class of machine learning methods deals with the problem of *survival analysis*, which is mainly characterized by the presence of censored data where the class label is not (yet) known. Several of the machine learning methods can be adjusted to manage this specific type of data. Since this specific type of problem is only tackled in Chapter 9, the key concepts of survival analysis will be explained then.

The next sections describe the unsupervised and supervised classification methods used in this work. Finally, the metrics that are standard used to evaluate the performance of classification methods are described in section 3.4.

3.2 Unsupervised classification methods

Unsupervised classification methods do not rely on a training set of labeled data to learn the underlying data structure and perform classification. Information is grouped according to similarities or differences between the different data points. One of the most common groups of unsupervised classification methods are clustering methods. Note that in many applications, clustering is not necessarily used for classification, but in an exploratory way to gain insights about the available data.

Clustering refers to the process of defining coherent groups of data points without requiring prior knowledge about the labels. A cluster is defined so all data points within a cluster are similar to each other and different from data points in other clusters [113]. Many different groups of clustering techniques exist. Here, we consider four types of clustering techniques, which each have their specific advantages and drawbacks: partitioning techniques, hierarchical techniques, density-based techniques and spectral techniques [75]. The different groups each contain many different methods. The next subsections explain one

standard method for each group, which will be used in Chapter 4 to detect irregular heartbeats.

For all methods, the number of clusters k is considered to be fixed and known prior to the clustering. When k is not defined beforehand, a number of methods to estimate the optimal value exist [187, 149]. In some applications, domain knowledge can also be used to provide an estimate for k.

Partitioning techniques: k-medoids

The goal of partitioning-based clustering techniques is to divide the data in k partitions that are iteratively updated in order maximize the similarity within a cluster [75]. The most well-known partitioning clustering method is k-means clustering, which detects a cluster by minimizing the squared error between the cluster center (calculated as the empirical mean of a cluster) and the points within the cluster [109]. While still very popular, it is known to be sensitive to outliers. A more robust variation is k-medoids clustering [113]. The approach is very similar to k-means clustering but while k-means clustering calculates the distance between each cluster point and the centroid (e.g. the mean of the cluster), k-medoids clustering uses a member of the cluster as cluster center, referred to as the medoid. Initialization of the method can be done by either selecting k random data points as medoids or by selecting the k farthest points (e.g. the k points with the largest distance between them). The algorithm then iterates over all data points and assigns each point to the closest medoid. In a next step, the medoids are then updated by calculating the new center of the clusters, which can cause some points to move between clusters (since their distance to the new medoids might be smaller than the distance to their current medoid). The last steps are repeated until the method converges and no more data points change clusters.

Hierarchical techniques: agglomerative clustering

Agglomerative hierarchical clustering is a bottom-up approach which initially represents each data point as a cluster of its own. The clusters are then iteratively merged by combining clusters that are similar to each other, until the desired cluster structure is obtained [109]. Merging can be done by different criteria, such as single linkage, complete linkage or average linkage. In complete linkage the distance between two clusters is defined as the maximum distance between any data point in the first cluster and any data point in the second cluster. Single linkage and average linkage use respectively the minimum and average distance. In each step, the two clusters with the smallest linkage distance are

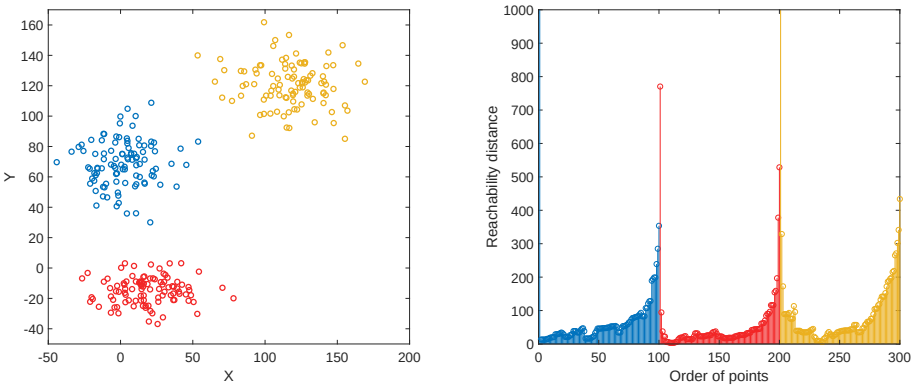


Figure 3.1: Illustration of the OPTICS algorithm: The left panel shows a simulated dataset with three obvious clusters in red, yellow and blue. The OPTICS reachability plot is shown in the right panel, the three clusters show up as valley of points with similar reachability distances.

combined into one cluster. Agglomerative clustering is a bottom-up approach: each data point forms its own cluster in the beginning and by moving up in the hierarchy, clusters are joined. An alternative is divisive clustering, which is a top-down approach: all points start in one global cluster, which is split recursively moving down the hierarchy.

Density-based techniques: OPTICS

Density-based clustering techniques define clusters as areas of high density that are separated by areas of low density. One of the most popular density-based algorithms is OPTICS (*Ordering Points To Identify The Clustering Structure*) [10], which gained popularity since it requires minimal parameter optimization. OPTICS orders the points in the dataset such that the points which are closest to each other become neighbours in a so-called *reachability-plot*. Clusters show up in the reachability plot as valleys of points with similar reachability distances and can be extracted as such.

To illustrate the concept, we created a simulated dataset of 300 points with three distinct clusters, shown in the left panel of Figure 3.1. The corresponding reachability plot is shown on the right panel. It shows the reachability distance from each point to its neighbour. Points within a cluster have similar reachability distances, while points between clusters have higher distances, creating three valleys corresponding to the initial clusters. The valleys can be extracted from

the plot by detecting points with significantly higher reachability distances which form the boundaries between clusters.

Spectral techniques

Spectral techniques do not make assumptions on the shape of clusters, but instead treat it as a graph partitioning problem, where the data points that need to be clustered are considered as nodes in a graph. An extensive introduction can be found in Von Luxburg [235]. The first step of most spectral clustering methods is to construct a *similarity matrix* of the data, e.g. a symmetric matrix $\mathbf{A}$ with $\mathbf{A}_{i,j} \geq 0$ a similarity measure between points i and j. Next, the *degree matrix* is constructed, which contains information about the degree of each point, or the connections it has in the graph. The *Laplacian matrix* is then defined as the difference between the similarity matrix and the degree matrix. Spectral clustering techniques use the first k eigenvectors of a Laplacian matrix of $\mathbf{A}$ as inputs for k-means clustering [235]. Different spectral clustering algorithms make use of different types of graphs and different similarity measures [235].

3.3 Supervised classification techniques

Supervised classification techniques use a training set of labelled data to effectively learn a function $f(\mathbf{x})$ that represents the optimal way of mapping the provided features $\mathbf{x} \in \mathbb{R}^D$ to the correct class label y, with D the length of the feature vector or the dimension of the feature space. The direct consequence is that a database of signals with known features and class labels has to be available. Ultimately, the goal is to apply $f(\mathbf{x})$ to new, unseen examples from a test set $\mathbf{x}_{\text{test}}$ and still obtain a reliable prediction of the new class label y_{test}: the method should be capable of generalizing.
As in unsupervised classification, many supervised techniques have been developed. In the next Chapters, we mainly use Support Vector Machines to perform classification in a supervised manner. Some alternatives are however briefly discussed in Section 3.3.2.

3.3.1 Support Vector Machines

Support Vector Machines (SVM) are a class of supervised methods that can be used for both classification and regression. It has been shown that they are superior to other popular classifiers in many types of problems [148]. They are also known to have good generalization properties. The basic idea of an SVM

is to find a hyperplane which perfectly separates two classes [207, 227]. This is done by maximizing the margin between the classes, and more specifically by maximization of the distance between the hyperplane and the closest data points of both classes (also called the *support vectors*). If all points closest to the hyperplane satisfy $|w^T x_i + b| = 1$, then the margin between the classes is defined as $2/||w||_2$ and can be maximized as such. When perfect separation cannot be achieved, missclassifications are allowed by introducing a soft margin constant (the *slack variable*) that penalizes points that lie on the wrong side of the margin.
SVMs were originally developed for linearly separable data [226], and the methods were later extended to include non-linear classification [227]. This is done with a so-called *feature map* $\phi(\cdot) : \mathbb{R}^D \rightarrow \mathbb{R}^{D_h}$. The input data are mapped to a higher-dimensional feature space which is used to train a linear SVM.

The integration of all previous elements (e.g. the margin maximization, introduction of slack variables and mapping of input points) leads to the following convex optimisation problem [207]:

$$\begin{aligned} \min_{\mathbf{w},b,\xi} J_P(\mathbf{w},\xi) =& \frac{1}{2}\mathbf{w}^T\mathbf{w} + c\sum_{i=1}^{N}\xi_i \\ s.t. \qquad & y_i[\mathbf{w}^T\phi(\mathbf{x_i}) + b] \geq 1 - \xi_i, \\ & \xi_i \geq 0, \qquad \forall i = 1, \ldots, N, \end{aligned} \tag{3.1}$$

with ξ_i the slack variables, c a regularization constraint and N the total number of points in the (training) dataset. The objective function in 3.1 consists of two parts: a regularization term and minimization of the number of missclassifications. The hyperparameter c provides a balance between these two parts: when c is too high, this will lead to *overfitting*, and vice versa: a value of c that is too low causes *underfitting*.

The dual formulation of this optimization becomes the following quadratic programming problem:

$$\begin{aligned} \max_{\alpha} J_D(\alpha) &= -\frac{1}{2}\sum_{i,j=1}^{N} y_i y_j K(x_i, x_j)\alpha_i\alpha_j + \sum_{i=1}^{N}\alpha_i \\ s.t. \qquad & \sum_{i=1}^{N}\alpha_i y_i = 0, \qquad 0 \leq \alpha_i \leq c, \qquad i = 1, \ldots, N, \end{aligned} \tag{3.2}$$

This optimisation problem has the following properties:

- When solving 3.2, many of the Lagrange multipliers α_i are equal to zero. Points corresponding to $\alpha_i \neq 0$ are the support vectors. Only the distances between those points and the hyperplane are maximized. The final formulation of the SVM is then:

$$y(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^{N}\alpha_i y_i K(\mathbf{x}, \mathbf{x_i}) + b\right) \tag{3.3}$$

  This decision function can be expressed only in terms of the support vectors.

- Neither the dual problem 3.2 nor the classification function 3.3 contain the feature map $\boldsymbol{\phi}(\cdot)$. Only the inner products of the feature maps are required. The *kernel trick* replaces these inner products with a positive definite kernel function :

$$K(\mathbf{x}, \mathbf{z}) = \boldsymbol{\phi}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\phi}(\mathbf{z}). \tag{3.4}$$

  The main advantage of the kernel trick is that it avoids the (cost-intensive) calculation of the high-dimensional feature space.

  Different kernel function can be used. Three examples of popular kernels are:

  - The RBF (*Radial Basis Function*) kernel: $K(\mathbf{x}, \mathbf{z}) = e^{-\frac{||\mathbf{x}-\mathbf{z}||_2^2}{\sigma^2}}$, with σ^2 a new hyperparameter that defines the non-linearity of the kernel.
  - The linear kernel : $K(\mathbf{x}, \mathbf{z}) = \mathbf{x}^{\mathrm{T}}\mathbf{z}$.
  - The polynomial kernel: $K(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^{\mathrm{T}}\mathbf{z} + \tau)^d$, with d the degree of the polynomial and $\tau \geq 0$.

  Recently, the clinical kernel was proposed [51], which is a kernel that was specifically designed for clinical data. It has separate expressions for continuous, ordinal, categorical and binary data.

Since both the primal and dual problem are convex optimization problems, a global optimum can be found. This is a big advantage compared to other supervised techniques such as neural nets, for which the optimization problem is not convex so it can result in local optima.

SVMs are defined for binary classification problems, where the goal is to distinguish between two classes. For multiclass problems, different binary SVMs can be combined in different configurations:

- *One-versus-all*: in this paradigm, the k classes give rise to k binary classifiers: one for each class versus the collection of the remaining classes. Potential problems are the creation of unbalanced classes and contradictions between assigned labels. A possible solution is to convert the scores to posterior probabilities, which can be done with for example Platt scaling [178].
- *One-versus-one*: this method trains a binary classifier for each combination of classes, resulting in $\frac{k(k-1)}{2}$ SVMs, for k classes. This method focuses on details that are different between classes. The final label belongs to the class with the largest number of votes.

3.3.2 Other techniques

K-nearest neighbors

In k-nearest neighbors (kNN), a new datapoint $\mathbf{x}_{\text{test}}$ is compared with all points from the training set. The k points closest to $\mathbf{x}_{\text{test}}$ are selected and it is assigned the class label of the most frequent class among its k nearest neighbours. k is usually small compared to the size of the training set. It is mostly used since the results are easy to interpret and do not require a lot of calculation time.

Naive Bayes Classifier

Naive Bayes classifiers (NB) are examples of probabilistic classifiers, based on the application of Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{3.5}$$

which describes the conditional probability of B given the prior probabilities of A and B and the conditional probability of A. The prior probabilities $P(A)$ and $P(B)$ and the conditional probability ($P(B|A)$) are all learned from the

training set. It assumes independence among the features, e.g. the value of a particular feature should be independent of the values of all other features.

Decision trees

Decision trees predict the output value y by answering a series of questions about the input features $\mathbf{x}$. The different features are repeatedly split into subsets, which is repeated until all subsets are pure, e.g. they all share the same class label y. The main advantage of decision trees is the interpretable nature of the algorithms: the different subquestions can be visualised into a tree-like structure that is very similar to the human way of thinking. Different decision trees can be combined into a random forest, which is an ensemble learning method that corrects for overfitting by balancing out the decisions from different trees.

3.4 Performance evaluation

Classification methods can be evaluated by comparing the results of the algorithm with the true labels of the data points, assuming these labels are known. Supervised methods have to be tested on a separate test set, preferably only using signals of patients not included in the training set. When the true labels of the test signals are not known, which can be the case when evaluating unsupervised methods, internal validation indices can be used. They measure the similarity within a cluster and compare it with the differences between clusters. A widely-used example is the Silhouette index [130].

When the true labels are known, which is the case for all datasets used in this work, different performance metrics can be used. The next Sections discuss the metrics used in the next Chapters, making a distinction between metrics derived from the confusion matrix or Receiver Operating Characteristics (ROC) curve.

3.4.1 Confusion matrix

A confusion matrix is a table which represents the number of instances in each class versus the number of instances in each predicted class. It is an easy way to visualize the performance of a classification method and various metrics have been defined to quantify its contents. Figure 3.2 shows an example of a confusion matrix for a binary classification problem. It reports the number of

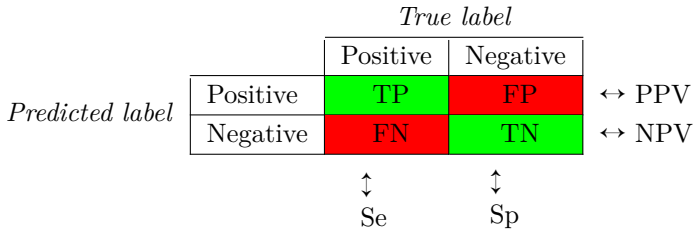


Figure 3.2: Example of a confusion matrix. Correct labels are indicated in green, incorrect labels in red. Based on the confusion matrix, different performance metrics can be evaluated, such as sensitivity, specificity, positive and negative predictive value.

true positive (TP), true negative (TN), false positive (FP) and false negative (FN) detections. With these numbers, following performance metrics can be calculated:

- Sensitivity *(Se)* : $\frac{TP}{TP+FN}$
- Specificity *(Sp)* : $\frac{TN}{TN+FP}$
- Positive predictive value *(PPV)* : $\frac{TP}{TP+FP}$
- Negative predictive value *(NPV)* : $\frac{TN}{TN+FN}$
- False Positive Rate *(FPR)*: $\frac{FP}{FP+TN}$
- Accuracy *(Acc)*: $\frac{TP+TN}{TP+TN+FP+FN}$

The sensitivity can also be called *recall*, the positive predictive value is also known as the *precision*. Note that the accuracy, which represents the fraction of correct predictions, is not a reliable metric for imbalanced problems. Imagine a dataset where 95% of the entries belong to the positive class. A classifier which classifies all observations as positive has an accuracy of 95% in this dataset, while not giving any useful output. Therefore, in unbalanced problems, we opt to replace the accuracy with the F_1-score:

$$F_1 = 2.\frac{\text{Precision.Recall}}{\text{Precision} + \text{Recall}}$$

It is the harmonic mean of the precision and recall, or of the sensitivity and PPV. The F_1-score has different variations such as the F_2- or $F_{0.5}$-score exist, but they are not considered here since they are less commonly used in the applications handled in this research.

3.4.2 Receiver Operating Characteristics curve

The Receiver Operating Characteristics (ROC) curve expresses the diagnostic ability of a classifier by plotting the true positive rate versus the false positive rate when the decision threshold is varied. It is a straightforward way to compare the performance of different models and select the optimal one. Furthermore it can also be used to define a good decision threshold, since changing the threshold results in a different prediction and corresponding performance. The ideal threshold can be determined by selecting the point in the upper left corner of the curve.

The ROC curve can be quantified by calculating the Area Under the Curve (AUC). A perfect classification method has an AUC value of 1, while a method that makes random guesses has an AUC of 0.5. AUC values $\leq$ 0.5 are theoretically possible but are in practice indicative of wrong labeling or an incorrect training strategy.

3.5 Conclusion

This Chapter provided an overview of the different machine learning techniques used in this research. Both unsupervised and supervised classification methods were introduced, together with metrics to evaluate the final performance. In the next Chapters, the various techniques will be used to classify ECG signals in different classes. A special class of machine learning techniques, not mentioned in this Chapter, are survival analysis methods. They will be introduced in Chapter 9.

Part II

Tensors

Chapter 4

Unsupervised detection of irregular heartbeats

This Chapter presents a first application of tensors in ECG processing. The tensor methods presented in Chapter 2 are applied for unsupervised detection of irregular heartbeats in multilead ECG signals. The method more specifically uses Canonical Polyadic Decomposition to decompose the ECG signal and applies clustering methods to the resulting factor vectors to detect abnormal beats. This Chapter also includes an analysis of the influence of different parameters on the results. More specifically, the effect of the choice of rank, initialization and number of channels are discussed. The method has been applied on two publicly available databases in order to compare the results with state-of-the-art methods. A first version of the method and preliminary results were previously published as Goovaerts, G., De Wel, O., Vandenberk, B., Willems, R., and Van Huffel, S. (2015) Detection of irregular heartbeats using tensors. In Proceedings of the 43rd Computing in Cardiology Conference (CinC2015) (pp. 573-576). Nice, France [88].

4.1 Introduction

Cardiac arrhythmias affect millions of people and are the main cause of many cases of sudden cardiac death. A large number of people could benefit from a better and more reliable detection of cardiac dysfunction. An arrhythmia is defined as any disturbance of the normal sinus rhythm. This can be a perturbation or abnormality in the rate, the regularity, the site of origin or

the conduction of the electrical impulses of the heart. In general, they can be divided in two main groups: the first group, morphological arrhythmia, contains arrhythmia that consist of only one heartbeat; the second group consists of arrhythmia formed by a group of heartbeats and is called rhythmic arrhythmia [135]. The focus of this Chapter lies in automatic detection of the first class, e.g. morphological arrhythmia.

Various types of morphological arrhythima exist, each differing in the origin and/or electrical pathway of the abnormal beat. The Association for the Advancement of Medical Instrumentation (AAMI) has defined a standard for how heartbeats should be annotated and classified and how different methods should be evaluated and compared. This standard will be followed throughout the Chapter. The complete standard is specified in ANSI/AAMI EC57:1998/(R)2008 [1]. It defines 15 recommended arrhythmia classes that can be further divided in three superclasses: Normal beats, atrial or supraventricular beats and ventricular beats. Normal beats refer to beats that originate during normal sinus rhythm. Atrial and ventricular beats respectively arise in the atria or ventricles. Table 4.1 shows how the division of the 15 recommended classes is done. The annotation types as used on Physionet [85] are mentioned in the rightmost column.

Classification of heartbeats in different classes is important for two main reasons. Firstly, heart rate variability studies, which are useful in many clinical and non-clinical scenarios require series of *normal* RR-intervals. The presence of ectopic beats or other abnormalities introduces abrupt changes in the RR time series which can seriously disturb the values of HRV indices. Secondly, most ventricular arrhythmia, which are potentially lethal, are initiated by premature ventricular beats. Their frequency and complexity have been shown to be predictive for predicting SCD in certain patient groups and have been used as stratification tools in large clinical trials [13].

The irregular heartbeat detection problem is generally composed of three stages: preprocessing, feature extraction and classification. Various features, such as RR interval features [211, 66, 132], ECG morphology features [66, 129] and features drawn from the wavelet transform [180, 181] have been proposed to characterize the ECG. An overview of the most well-known methods, including results on standard databases, can be found in [135]. In these traditional heartbeat classification methods the ECG is typically represented as a vector. However, by representing the ECG as a tensor more structural information can be preserved. The application of tensor methods for heartbeat classification has been done before [104, 128]. However, these methods all require a training stage and more complicated optimization mechanisms. This chapter proposes an unsupervised classification method that is based on tensor decomposition methods which does not require a training stage. Different unsupervised clustering techniques are

Table 4.1: List of classes and superclasses of irregular heartbeats as defined by AAMI [1]. Only the superclasses used here are mentioned; an extra (super)class of paced beats exists, but was omitted since paced recordings are not included in this study. The rightmost column indicates the corresponding annotations as used in Physionet [85].

Heartbeat class	**Type of heartbeat**	**Annotation type**
Normal	Normal Beat	N
	Left Bundle Branch Block	L
	Right Bundle Branch Block	R
Atrial	Atrial Premature Beat (APB)	A
	Aberrated APB	a
	Nodal Premature Beat	J
	Premature Supraventricular Beat	S
	Atrial Escape Beat	e
	Junctional Escape Beat	j
Ventricular	Ventricular Premature Beat	V
	Ventricular Escape Beat	E
	Fusion Ventricular and Normal	F

compared to assess their performance, and the influence of different parameters in the tensor decomposition on the final result is examined likewise.

The next Section describes the datasets used to develop the algorithm. Section 4.3 presents the proposed tensor-based method. In Section 4.4 and Section 4.5, the results of the different analyses are respectively described and discussed, followed by some concluding remarks in Section 4.6.

4.2 Data

The methods described in this Chapter have been applied on two publicly available datasets, the INCART database and MIT-BIH database [153]. Both datasets can be found on Physionet [85]. The following Sections give a brief overview of the main characteristics of both databases.

4.2.1 INCART database

The first database used is the St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia (INCART) Database. This database was measured in the St. Petersburg Institute of Cardiological Technics and contains 75 non-pacemaker recordings of 30 minutes at a sampling frequency of 257Hz. All signals were routine clinical 12 lead ECG signals. The signals were collected from 17 men and 15 women between 18 and 80 years old who were undergoing tests for coronary artery disease. Preference was given to recordings of patients with ECGs consistent with ischemia, coronary artery disease, conduction abnormalities and arrhythmias. Heartbeat annotations, computed by an automatic algorithm and manually corrected, are also provided. The full dataset contains 153 651 normal beats and 21 964 abnormal beats (1959 supraventricular and 20 005 ventricular beats, see Table 4.2)). Since this database is not often used in the literature, it will be mainly used here to define the optimal parameters for the tensor decomposition and to compare different clustering techniques.

4.2.2 MIT-BIH Arrhythmia database

The MIT-BIH Arrhythmia Database was collected in 1980 during a collaboration between Boston's Beth Israel Hospital and the MIT university [153]. It contains 48 half-hour recordings of two-channel ambulatory ECG recordings, obtained from 47 subjects experiencing various types of cardiac arrythmia. The first lead is always lead II, the second a variant of a precordial lead, mostly V1 and in some occasions V2, V4 or V5. Recordings were selected to ensure that most types of clinically significant arrythmias were included in the database. The sampling frequency is 360Hz. All heartbeats were annotated by two or more independent cardiologists, and disagreements were solved to obtain consensus in all recordings. For this study, paced recordings were removed from the dataset. The number of heartbeats in the different classes is summarized in Table 4.2.

Since the database contains a large variety of arrhythmia and has been publicly available for over three decades, it has been widely used to evaluate and compare irregular heartbeat classification methods. A framework to standardize the evaluation of the results has been proposed in Luz et al. [135] and will be used throughout the rest of the chapter.

Table 4.2: Number of heartbeats of each of the types (normal, supraventricular and ventricular) in both databases used in this Chapter.

Type	INCART	MIT-BIH
Normal	153 651	90 089
Supraventricular	1959	2779
Ventricular	20 005	7007
Total	175 615	99 875

4.3 Methods

The tensor-based irregular heartbeat detection method consists of four steps. 1) In the preprocessing stage, the noise is removed from the data before it is tensorized. After 2) tensorization, 3) the tensor is decomposed so 4) the irregular heartbeats can be detected.

4.3.1 Preprocessing

In the preprocessing stage, both baseline wander, powerline interference and high frequency noise were removed from the ECG signal. The same preprocessing approach as described in [66] was used in order to be able to effectively compare the results from our method with the results of the different methods compared in their study. First, baseline wander removal was done by applying two consecutive median filters to each channel. The first median filter has a length of 200 ms and removes the QRS complex and P wave, the second filter has a width of 600 ms to remove the influence of the T wave. The resulting signal represents the baseline of the signal and was finally subtracted from the original ECG signal. Powerline interference and high frequency muscle noise were removed with a Butterworth low-pass filter with order six and cut-off frequency 35 Hz.

Because the ECG signal has to be segmented into distinct heart cycles prior to tensorization, the locations of the R peaks need to be known. Since R peak detection was not the focus of this study, the annotations provided with the databases were used here. In general, standard methods such the Pan-Tompkins algorithm [173] or the wavelet-based ECG delineator developed by Martínez et al. [143] can be used to detect the R-peaks with high accuracy.

4.3.2 Tensorization

As explained in Section 2.4.1, the second-order multilead ECG signal has to be transformed into a higher-order tensor prior to applying tensor decomposition methods. This is referred to as tensorization. It can be done in many different ways, depending on signal characteristics and the application at hand [64]. Since the main interest here are the differences between consecutive heartbeats, a *heartbeats* mode was created by segmenting the signal in different heartbeats and stacking these heartbeats in a 3D manner. This way a tensor was constructed where each frontal slice contained one heartbeat over all channels. Consider a (preprocessed) multilead ECG signal with C leads and N samples that consists of M heartbeats. This 2D signal $\mathbf{X} \in \mathbb{R}^{C \times N}$ was transformed in a 3D tensor $\mathcal{X} \in \mathbb{R}^{C \times J \times M}$, with J the length of the segmentation window used to cut out the heartbeats. The dimensions of $\mathcal{X}$ are thus *channels* $\times$ *time* $\times$ *heartbeats*. While the temporal mode of the original signal contains the total length of the ECG signal, the time mode of the tensor only has the duration of one heartbeat.

The segmentation was done by taking a fixed-length interval around the R peak. The interval starts 200 ms before the R peak and has a length of 700 ms. The different segments were normalized before stacking them in the tensor by subtracting the mean and dividing by the standard deviation. If the heart rate differs greatly among the signal (which can be the case in long-term recordings or during stress tests) it may be necessary to apply resampling since it is important that the different waves of the ECG signal are largely aligned. Note that by segmenting the heartbeats around the R peaks, the QRS complexes of the different beats in a signal were by definition synchronized. Figure 4.1 illustrates the tensorization process.

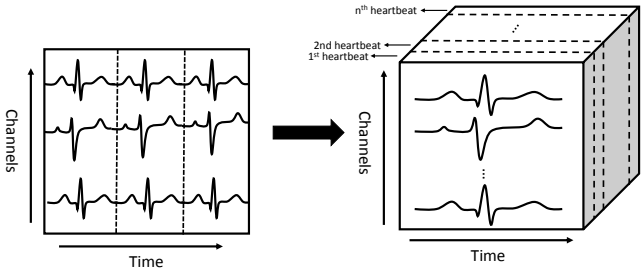


Figure 4.1: Tensorization of the ECG signal: The multilead ECG signal is mapped onto a third-order tensor by segmenting it in individual heartbeats and stacking the heartbeats in the frontal slices of the tensor.

4.3.3 Tensor decomposition

Many tensor decomposition methods exist, all leading to factors with different ranks [119]. Here, Canonical Polyadic Decomposition (CPD), explained in Section 2.4.3, is used to decompose the tensor $\mathcal{X}$ in a sum of R rank 1-tensors:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a_r} \otimes \mathbf{b_r} \otimes \mathbf{c_r} \tag{4.1}$$

As discussed earlier, determination of R, the rank of the decomposition, is a non-trivial problem. We therefore varied the value of R in order to determine the optimal rank.

The results are R rank-one tensors each consisting of three loading vectors $\mathbf{a_r}, \mathbf{b_r}$ and $\mathbf{c_r}$ for $r = 1, \ldots, R$ which match the three dimensions of the original tensor. They can be combined into three loading matrices $\mathbf{A} \in \mathbb{R}^{C \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$ and $\mathbf{C} \in \mathbb{R}^{M \times R}$. The first loading matrix ($\mathbf{A}$), corresponding to the *channels* mode, is associated with the differences in heartbeat morphologies over different channels. The second loading matrix $\mathbf{B}$ (*time*) shows the temporal profile of the components. $\mathbf{C}$ corresponds to the *heartbeats* mode and gives an indication of variations between different heartbeats in the tensor. The columns of this matrix are further used for irregular heartbeat detection.

4.3.4 Clustering techniques

The irregular heartbeats can be distinguished from normal heartbeats by examining the third loading matrix $\mathbf{C}$. When a heartbeat significantly differs from other heartbeats in the signal, this is visible in $\mathbf{C}$. The values corresponding to an irregular heartbeat are significantly different from the value of a normal beat. Detection of the irregular heartbeats can be done in an unsupervised way by applying clustering techniques, discussed earlier in Section 3.2.

We made two main assumptions about the data:

1. The data naturally consist of two main clusters, one representing normal heartbeats and the other representing abnormal heartbeats. This is a reasonable assumption as long as the abnormal heartbeats within a signal all belong to the same class, e.g. if a signal contains either ventricular or supraventricular beats.

2. The number of normal heartbeats is larger than the number of abnormal heartbeats, and the largest cluster will thus be assigned to the normal

Table 4.3: Confusion matrix (right) for the detection of abnormal heartbeats and corresponding derivations for each class (left). The heartbeats of the three classes (N, S, V) are classified as normal (n) or abnormal (a). Correct classifications are marked in green, incorrect labels in red.

$TP_n = Nn$	$TP_s = Sa$	$TP_v = Va$
$TN_n = Sa + Va$	$TN_s = Nn$	$TN_v = Nn$
$FP_n = Sn + Vn$	$FP_s = Na$	$FP_v = Na$
$FN_n = Na$	$FN_s = Sn$	$FN_v = Vn$

Label	Result: n	Result: a
N	Nn	Na
S	Sn	Sa
V	Vn	Va

class. This assumption is met in all signals in the database, since normal heartbeats by definition make up the largest part of the ECG signal.

All clustering methods have their distinct advantages and disadvantages. To select the optimal method, all four methods described in Section 3.2 (e.g. k-medoids clustering, agglomerative clustering, OPTICS and spectral clustering) were therefore applied to the dataset and compared.

4.3.5 Evaluation of performance

All clustering methods group the heartbeats in two classes: a normal class and an abnormal class. By definition, the heartbeats in the largest class were labeled normal. This means that the three original classes (normal, supraventricular and ventricular) were mapped on two predicted classes (normal and abnormal). The corresponding confusion matrix is shown on the right side of Table 4.3.

As mentioned in the Introduction, the AAMI has defined a standard for both annotating irregular heartbeats (summarized in Table 4.1) and evaluating the performance of different methods. The recommended evaluation measures are the Sensitivity (Se), Positive Predictivity Value (PPV), False Positive Rate (FPR) and Overall Accuracy (Acc), also discussed in Chapter 3. Since heartbeat classification is an extremely imbalanced problem, with the number of normal heartbeats multiple times larger than the other classes, the overall accuracy is strongly influenced by the results of the normal heartbeats. We therefore opted to replace it with the F_1-score.

Table 4.3 shows the confusion matrix, with the original labels N, S and V and the resulting labels n and a. The number of TP, TN, FP and FN for each class can be calculated with the equations on the left hand side of the Table. This way, all evaluation measures were calculated for each class individually,

for example:

$$F1_i = \frac{2TP_i}{2TP_i + FN_i + FP_i} \tag{4.2}$$

with i is the class label, which can be either N, S or V. $F1_N$, $F1_S$ and $F1_V$ are thus respectively the F_1-scores for the normal, supraventricular and ventricular classes. The other measures were calculated in a similar way using the Equations in Section 3.4. Aggregate scores for the overall database were calculated by averaging the individual metrics for each class, such as for example:

$$F_1 = \frac{F1_N + F1_S + F1_V}{3} \tag{4.3}$$

4.4 Results

This section is divided into three parts. First the detailed results of one case study are described. Next, the sensitivity of the performance to the choice of rank, initialization and number of channels is evaluated on the INCART database. Finally, the results for the full MIT-BIH dataset are evaluated and compared with standard methods.

4.4.1 Case study

As a case study, a segment containing the first 100 heartbeats of the first signal (*I01*) of the INCART database described in Section 4.2.1 was selected. Figure 4.2 shows a four second sample of this signal. This particular segment contains two irregular heartbeats which are highlighted in red. The full signal contains five irregular heartbeats in total.

The signal was preprocessed and tensorized as described in Sections 4.3.1 and 4.3.2. As discussed in Section 2.4.3, there are different ways to estimate an appropriate choice for CPD rank. Figure 4.3 shows the multilinear singular values σ_i for the three modes. The first mode, the spatial mode, has singular values that gradually decrease until $r = 5$. The other modes have one very dominant singular value, after which the values decrease again gradually. Another possibility for rank estimation is the method *rankest*, available in Tensorlab, which calculates the ratio between the Frobenius norm of the original tensor and the CPD result. Using this method, an estimate of 242 was obtained, which might be optimal to exactly reconstruct the original tensor but is of little practical use.

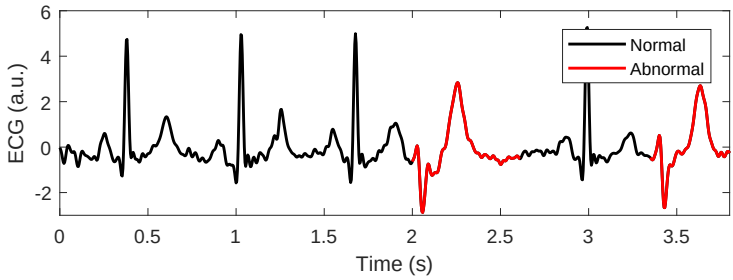


Figure 4.2: Excerpt of four seconds of the ECG signal used in the case study, with the irregular heartbeats indicated in red.

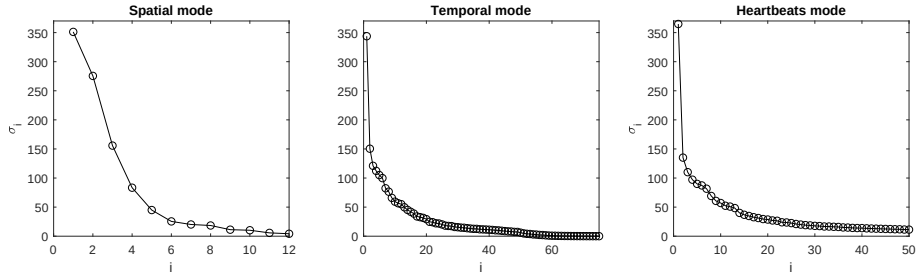


Figure 4.3: Multilinear singular values σ_i of the signal used in the case study.

To explore the effect of changing the CPD rank, the value of R in Equation 4.1 was varied between 1 and 5. The resulting factor vectors are shown in Figure 5.4. The different rows show the different ranks, and the factor vectors for each mode are plotted in the different columns. For $R = 1$, shown in the top row, the feature vectors have a straightforward physiological interpretation, especially for the temporal and heartbeats mode. The temporal factor vector is easily recognizable as a normal (regular) heartbeat, which is expected since the majority of heartbeats are normal. Furthermore the morphology of a normal heartbeat is similar over all channels, making it suitable to capture in a rank-one component. The morphology changes over the different channels are visualized in the spatial factor vector. From this vector we can for example conclude that the polarity of the signal greatly varies over the different channels: heartbeats in the first channel have an opposite polarity than heartbeats in the fourth and seventh channels. This also has a physiological interpretation, since the polarity of an ECG wave is related to the direction of the electrical wavefront, which changes depending on the considered channel. Finally, the factor vector of the heartbeats dimension is the most important for irregular heartbeat detection since it shows the changes in the ECG over different heartbeats. The values

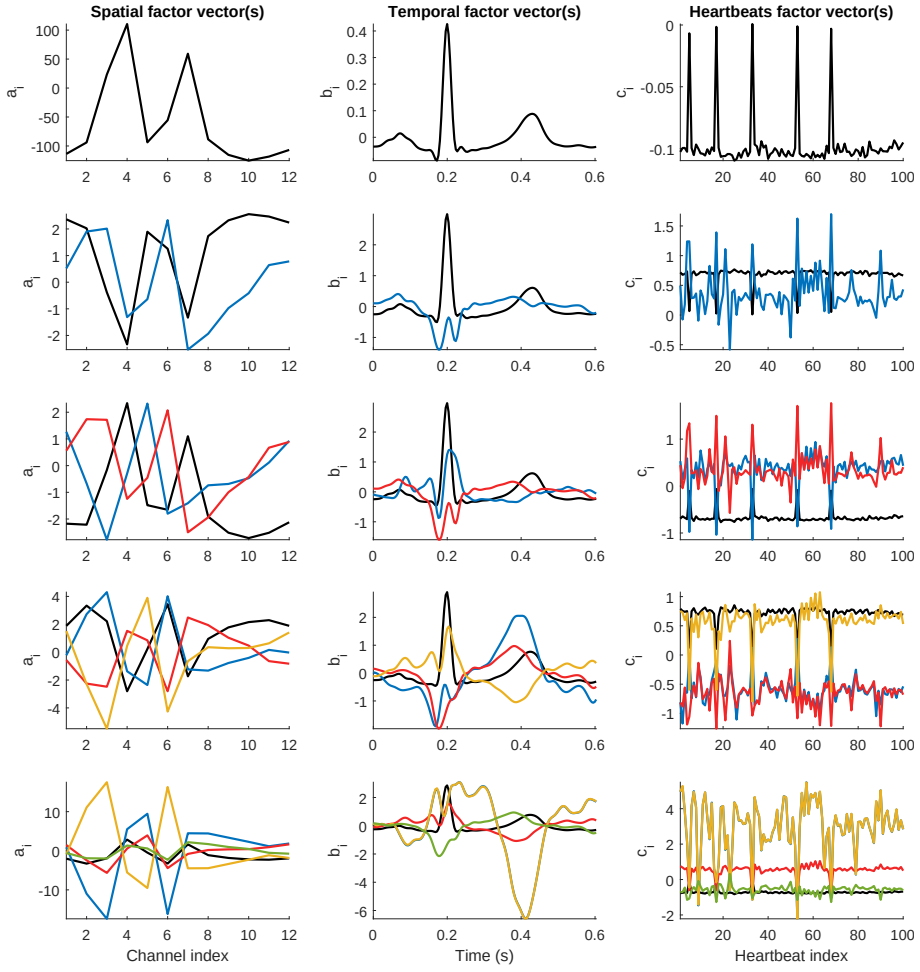


Figure 4.4: Results of the selected case study signal for CPD ranks varying from one to five. The factor vectors for all modes are shown in the different columns.

of this loading vector for normal heartbeats varies around -0.1. The abnormal heartbeats are easily distinguishable by their higher values.

Increasing the CPD rank adds extra terms to the tensor decomposition, visualized in the additional rows in respectively blue, red, yellow and green. It is immediately clear that the first factor (shown in black) resulting from a CPD of rank 1 is also present in the decompositions with higher ranks. While this might seem straightforward, Kolda and Bader [119] show examples where the best rank-one approximation of a tensor is not a factor in the best rank-two approximation and so on. Comparison of the temporal vectors of the CPD with rank 3 and 4 shows that this is indeed the case when the rank is further increased. Another remarkable observation is that for a CPD with rank 5, two components (plotted in yellow and blue) have near-identical temporal and heartbeats profiles and opposite values in the spatial dimension. This is referred to as *degeneracy*, where two components are almost equal up to a sign change and is indicative of a wrongly-set rank [202].
Based on Figure 5.4, it seems that a decomposition with rank $R = 1$ is suitable to distinguish regular from irregular heartbeats.

4.4.2 INCART database

As discussed earlier, the INCART database is a publicly available multilead database. It is however not used very extensively in literature, and was therefore mainly used here to analyse the effect of different parameter values on the final result. More specifically, the influence of the CPD rank, clustering method and number of available channels are discussed next. To evaluate the stability of the obtained results, each experiment was run 100 times with different random initializations. The minimum and maximum values over all runs are shown in each Figure by shaded areas.

The main parameter to optimize in the tensor decomposition is the rank R. To examine the effect of the choice of R on the result, the value was again varied between 1 and 5 as in the previous Section. Figure 4.5 shows the results for all clustering methods considered. $F1_i$ scores for each class are shown together with the overall F_1 score (in yellow). Minimal and maximal values over the 100 runs are indicated with the shaded areas, the average value is indicated with a circle.

For all methods, a decomposition of rank 1 led to the best results. K-medoids clustering and spectral clustering resulted in better F_1-scores than the other alternatives. The results for the classes of normal and ventricular beats were significantly larger than the class of supraventricular beats. The overall F_1-score was largely influenced by the subpar results for the latter class. This could

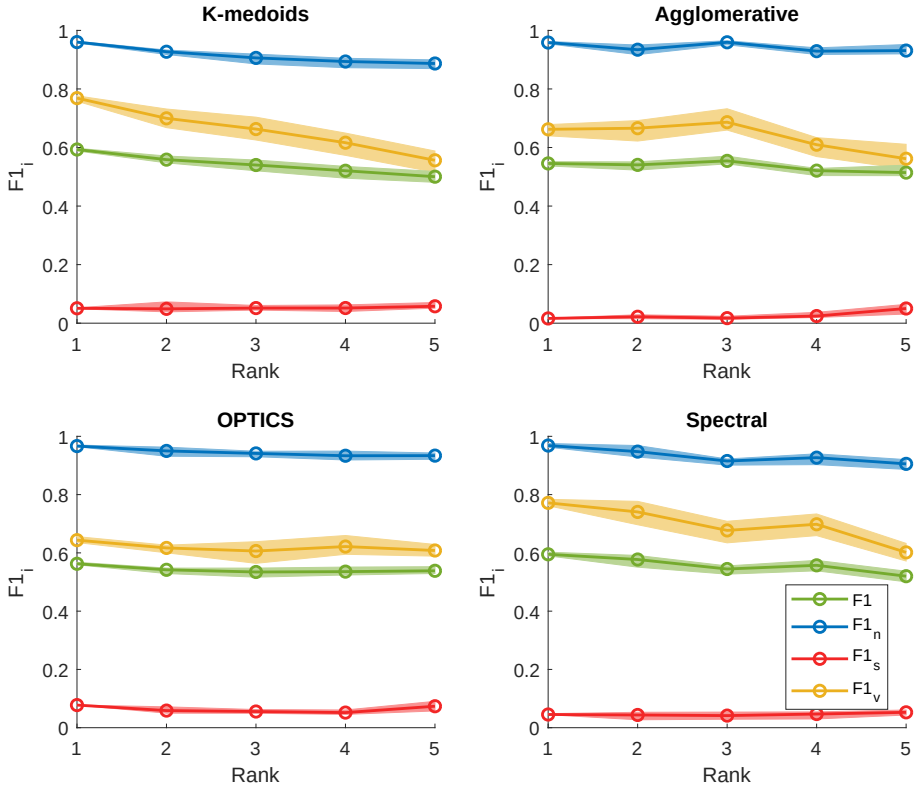


Figure 4.5: Results of different clustering algorithms for different CPD ranks. Each method was run 100 times, and the minimal and maximal results are shown together with the median.

be expected, since the morphology of supraventricular beats is in many cases similar to the morphology of normal beats. In those cases, the classes of normal and supraventricular beats will be overlapping and cannot be separated by unsupervised techniques.

A rank-1 decomposition lead to very stable results, with no fluctuations in the results of different trials. Increasing the rank added variation in the results, which was confirmed by visual inspection of the components that showed different solutions in each iteration. Since the tensor decomposition was initialized randomly, convergence to a local optimum could thus not be avoided in these cases.

Different measurement setups record ECG signals with a different number of

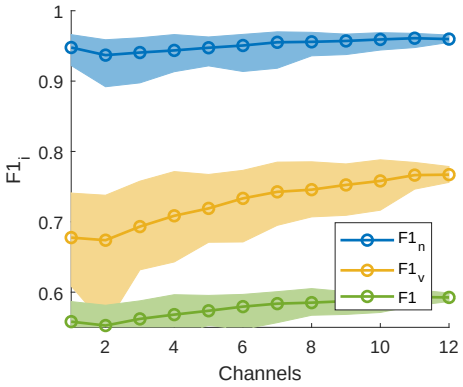


Figure 4.6: Results of changing the number of channels in the tensor on the final result. All experiments were repeated 100 times, and the shaded areas show minimum and maximum values over all runs together with the mean value.

channels. In clinical practice, signals with two, three, six or twelve channels are all encountered frequently. While the number of channels for a dataset is fixed, the effect of the number of channels was examined by randomly selecting n channels. Since the total number of channels here is twelve, n could be varied between one and twelve. The CPD rank was fixed to one. The results can be seen on Figure 4.6. The results for the supraventricular class were omitted for clarity. The trend was however similar to the other classes. For all classes, increasing the number of channels used improved the results. Optimal results were obtained using all twelve ECG leads.

4.4.3 MIT-BIH database

Results on the INCART database indicated that for irregular heartbeat detection, the optimal choice of rank was 1 and k-medoids or spectral clustering lead to better results than the other clustering methods. The number of channels is fixed and equal to 2 for the MIT-BIH database. The results of the proposed tensor-based techniques were calculated using these settings and compared with results from standard methods from literature. For fair comparison, only results from automatic methods which do not require manually annotated beats are reported.

Table 4.4 presents the results obtained using different methods. The methods proposed in this Chapter are in the bottom two rows. From the Table we can conclude that the tensor-based methods perform equally well or better than

	Normal		Atrial		Ventricular	
	Se (%)	PPV (%)	Se (%)	PPV (%)	Se (%)	PPV (%)
de Chazal [66]	87	**99**	76	39	87	43
Mar [141]	90	**99**	**83**	34	87	61
Llamedo [131]	93	**99**	77	39	82	70
$\text{CPD}_{\text{kmedoids}}$	95	94	19	37	**90**	75
$\text{CPD}_{\text{spectral}}$	**96**	93	19	**43**	73	**79**

Table 4.4: Sensitivy and positive predictive value for each class obtained with different methods on the MIT-BIH database. The best results for each column are indicated in bold.

the reference methods for the classes of normal and ventricular beats. The sensitivity results for atrial beats are clearly less than the other methods. The positive predictive value is however the largest between all methods considered.

4.5 Discussion

This Chapter proposed a novel method for unsupervised detection of irregular heartbeats. The method was tested on two publicly available datasets. From the results obtained on both the INCART and MIT-BIH database we can conclude that the tensor-based approach produces satisfiable results that are comparable to or better than state-of-the art methods for the classes of normal and ventricular beats. For atrial beats however, the results are considerably worse. Since the developed method only extracts morphology-based feature and the morphology of normal and atrial beats is very similar, both classes are clustered together. A possible solution would to add a second set of RR-interval features that characterize the irregular heart rhythm that often accompanies episodes of atrial activity. This would however strongly increase the dimension of the feature space, which might cause the performance of the clustering methods to decrease.

Another consequence is that the cluster formed by the normal heartbeats would spread out more and possibly split in multiple subclusters: while normal beats have a very coherent morphology they can have very diverse heart rates. Unsupervised classification with a fixed number of clusters would thus not lead to good results. Several alternatives exist: a first possibility is to automatically determine the number of clusters in the data set (such as for example with the

methods described in [187, 149]). The different clusters can then be annotated by an expert. This is referred to as semi-supervised classification. Another approach is to use the combination of tensor-based morphological features and heart rate variability features as input for a supervised classification method such as support vector machines.

The MIT-BIH dataset contains signals measured in a two channel set-up. The results from Figure 4.6 indicate that the results would improve further if more channels were available. The lack of additional publicly available multilead validation datasets can be considered a drawback of irregular heartbeat detections in general.

An advantage of the proposed tensor-based approach compared to the methods in Table 4.4 is that the results are easy to visualize and interpret: all factor vectors from the tensor decomposition have a physiological interpretation. The first- and second-order factor vectors could be further analysed to get more global information about the ECG signal. Another large benefit of the proposed method compared to the other methods mentioned in Table 4.4 is that only one parameter, the CPD rank, had to be optimized in order to get the required results.

The proposed framework is very general and can be used to detect different ECG abnormalities with minor modifications. Depending on the part of the ECG signal used to construct the tensor $\mathcal{T}$, different ECG waves can be analysed. This will be shown in the next Chapters, which analyse for example variations in the T wave (Ch. 5) or general ECG morphology (Ch. 6).

4.6 Conclusion

This Chapter presented a first application of tensor decompositions in ECG signal processing in the form of a method to detect irregular heartbeats. The method leads to very good results when only taking normal and ventricular beats into account, but fails to distinguish normal and atrial beats. However, since most ventricular arrhythmia are initiated by PVCs, a type of ventricular beats, this class is more important in the context of SCD risk estimation. The frequency of ventricular beats has been used as a risk factor in earlier studies [13], and can be calculated in a straightforward way with the proposed algorithm.

Chapter 5

Automatic detection of T wave alternans

The previous Chapter showed how tensor decompositions can be used for automatic determination of irregular heartbeats. In this Chapter, a similar approach is used to detect T Wave Alternans, a promising risk factor for sudden cardiac death. A practical issue when dealing with long-term ECG signals are the changes in timing between different ECG waves. In this Chapter, this problem is tackled by using an alternative method to decompose the tensor, PARAFAC2, which allows variations in the factor vectors in one mode. This chapter was published as: Goovaerts, G., Vandenberk, B., Willems, R., & Van Huffel, S. (2017). Automatic detection of T wave alternans using tensor decompositions in multilead ECG signals. Physiological measurement, vol. 38, no. 8, pp. 1513–1528 [87]. Preliminary results were also published as [90, 91].

5.1 Introduction

T wave alternans (TWA) is an example of a promising feature for risk stratification of sudden cardiac death that can be derived from the ECG. Studies have shown that the presence of TWA is correlated with the risk of arrhythmia in patients with ischemic cardiomyopathy after previous myocardial infarctions and long QT syndrome [107, 195, 40]. T wave alternans is defined as a periodic beat-to-beat variation in the amplitude of the T wave. Large and small T waves alternate in an ABABAB-pattern. Since the T wave represents the repolarization of the ventricles, it is linked with the repolarization characteristics

of the heart. When the variation between subsequent T waves is large enough, TWA can be detected by visually inspecting the ECG signal. In cases where the difference is only a few microvolts or when dealing with long ECG signals this is however not feasible. Therefore (semi-)automatic methods are used that can reliably detect TWA in the ECG signal.

Martínez and Olmos [142] give an overview of commonly used algorithms for (semi-)automatic T Wave Alternans detection. Most methods start with a preprocessing stage followed by the actual TWA detection. Examples of commonly used methods are the correlation method [34], modified moving average method [162] or spectral method [200]. Most methods decompose the ECG signal to a beat-to-beat time series that contains the T wave characteristics. The actual detection of T wave alternans is then done on this decomposed signal.

The objective of this Chapter is to present a novel TWA detection method based on tensor decompositions. The method is fully automated and both detects and quantifies the amount of T Wave Alternans in a multilead ECG signal. The main structure of the method is similar to the irregular heartbeat detection method from Chapter 4: The signal is segmented and reorganized in a tensor. By decomposing the tensor and analyzing the resulting factor vectors, TWA can be detected and the amount of TWA can be estimated. Preliminary results were presented in [90] and [91]. This chapter includes a full validation on both artificial and clinical signals. Furthermore the use of the more general PARAFAC2 tensor decomposition method makes application in realistic situations possible.

In the next section the datasets on which the method has been evaluated are described. Section describes the different steps necessary to detect T wave alternans, including a description of the tensor decomposition method used here. The results are summarized in Section , followed by discussion and conclusion.

5.2 Data

The method has been evaluated on three different datasets: a simulation dataset, the TWA challenge database available on Physionet [85] and a dataset from the University Hospitals Leuven.

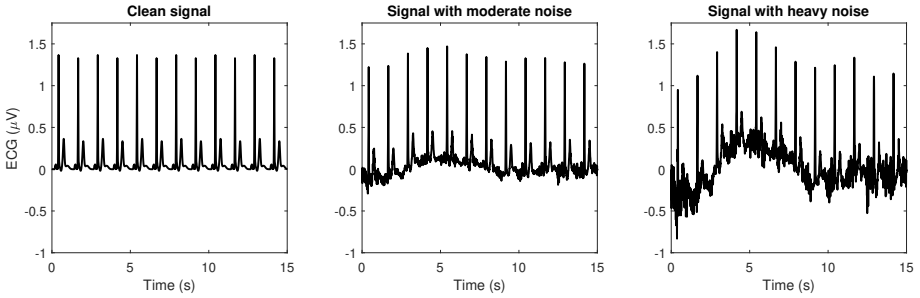


Figure 5.1: Example of a simulated signal from the same patient without noise (left), moderate noise (middle) and heavy noise (right).

5.2.1 Simulation data

Artificial ECG and noise signals were generated using the OSET toolbox [188]. The artificial signals consist of the 12 standard ECG channels and are 30 seconds long. The sampling frequency is 250 Hz. The signals were simulated by modeling a cardiac dipole and linearly projecting it onto the body surface using the Dower transform [71]. T wave alternans was simulated by introducing a constant offset to each odd T wave as demonstrated in [47]. The level of TWA was varied between 5 μV and 50 μV. After all signals were generated, the differences between subsequent T waves were manually checked to ensure the TWA level was consistent with the simulated levels. Fifty different patients were simulated by calculating fifty Dower transformation matrices using least-sqaures fitting based on signals of the PTB database available on Physionet [27, 85]. In order to work with realistic ECG signals and to be able to test the robustness to noise, realistic ECG noise was added to the signals using the noise generator provided with [188]. The noise consists of a mixture of equal amounts of baseline wander, muscle artifacts and electrode movements. A separate noise signal was generated for each synthetic signal and superposed onto the signal. The SNR value of the signals was varied to construct two sets with respectively moderate noise (SNR = 15dB) and high noise (SNR = 5dB). Figure 5.1 illustrates the noise level by showing a signal from one patient in all three conditions (without noise, moderate noise, heavy noise).

5.2.2 Physionet dataset

The TWA database available on Physionet [85] was assembled for the Computing in Cardiology Challenge of 2008 [152]. It contains 100 ECG signals sampled at 500 Hz. Most signals are 12-lead ECG channels, some signals only contain two or

three channels. The duration of the signals is (approximately) two minutes. The records were collected from different databases, and the final dataset contains both real-life and artificial ECG signals. The signals are ranked according to the level of T wave alternans present in the ECG. The exact amount of TWA is not known.

5.2.3 University Hospitals Leuven dataset

The third dataset was obtained in the University Hospitals Leuven (Belgium), and includes 33 Holter ECG signals from nine subjects. Selection of subjects was done based on the results of a clinical TWA test (Cambridge Analytic Spectral Method). Only subjects with multiple consistent positive or negative tests were included in the dataset. During the test, subjects were asked to cycle on a stationary bike. By increasing the workload, their heart rate is increased from resting rate to 110 beats per minute. The dataset included 13 signals from four subjects with positive clinical TWA tests (referred to as TWA group) and 20 signals from five subjects with negative clinical results (referred to as control group). All ECG signals are recorded with a sampling frequency of 256 Hz. Most records have three ECG channels (30/33 records), some only include two channels (3/33 records). The signals are all between 14 and 23 minutes long. The study was approved by the Ethical Committee of the University Hospitals Leuven.

5.3 Methods

The complete TWA detection method consists of five steps. After 1) preprocessing and 2) T wave segmentation, 3) a tensor containing all T waves is constructed. 4) This tensor is then decomposed and 5) the decomposition result is analyzed to detect and quantify T wave alternans.

5.3.1 Preprocessing

The preprocessing stage is necessary to remove noise from the signal so that it does not affect the results of the TWA detection. When applying noise removal methods, it is essential that the shape and amplitude of the T wave are not altered since this could introduce errors in the final result. We considered baseline drift and high frequency noise from for example muscle artifacts as major noise sources in ECG signals. To remove baseline wander a method

that removes the baseline by reducing quadratic variation was used [78]. The quadratic variation [x] of a signal $\mathbf{x} \in \mathbb{R}^n$ is defined as:

$$[x] = \sum_{k=1}^{n-1} (x_k - x_{k+1})^2 \tag{5.1}$$

It can be seen as a measure for the stability in a signal, which increases with increasing noise levels. Quadratic variation reduction calculates an estimate of the baseline by solving a constrained convex optimization problem, with the quadratic variation of the solution one of the constraints. The method has been shown to outperform other state-of-the-art methods [78]. After that, the signal was filtered with a low-pass filter with a cut-off frequency of 60 Hz to remove high frequency noise. Since the T wave is typically contained in a frequency band ranging from 2 to 35 Hz [33], this should not induce significant changes in morphology of the T wave on a macroscopic scale. Finally the signals were normalized to facilitate comparison between different signals.

5.3.2 T wave segmentation

In order to be able to detect transient T wave alternans (i.e. TWA that is only present in certain segments of the ECG), a moving window of 50 heartbeats was used to analyze the signals. T wave segmentation is preferred over T wave delineation since the latter is sensitive to noise. T waves were segmented by selecting a window with length J after each detected R peak. Since the length of the QT interval changes with a changing heart rate, the start point st and length of the windows J are dependent on the mean RR interval length over the current analysis window of 50 heartbeats ($\overline{RR}$). For st we made a distinction between three possible $\overline{RR}$ ranges:

1. $\overline{RR} \leq 0.6s$: st = R peak + 100ms
2. $0.6s < \overline{RR} < 1.1s$: st = R peak + 150ms
3. $\overline{RR} \geq 1.1s$: st = R peak + 250ms

The window length J was $0.4\sqrt{\overline{RR}}$. The interval boundaries have been determined experimentally to ensure accurate T wave segmentation and largely correspond to other values found in literature [34].

5.3.3 Tensor construction

In a similar way as Chapter 4, the ECG signal was transformed into a tensor before applying tensor decompositions. Since the objective is T Wave Alternans detection, the tensor should be constructed in such way that the changes in T waves over different heartbeats were maximally emphasized. Therefore we constructed a *T wave tensor* $\mathcal{T}$: a third-order tensor that consists of the T waves of all channels and heartbeats. The T waves of heartbeat one formed the first frontal slice of the tensor, the T waves of the second heartbeat formed the second slice, etc. This way, one tensor was constructed for each ECG signal. Each tensor has three modes, *channels* x *time* x *heartbeats*. If the ECG signal contains C channels and M heartbeats, the resulting tensor will have dimensions $(C \times J \times M)$. Since we used a sliding window of 50 heartbeats to analyze the signals, M was here fixed to 50. Note that both the second and third dimensions contain temporal information: the second dimension contains the signal of individual T waves while the third dimension shows a more long-term view of the different heartbeats. Figure 5.2 illustrates the construction of the tensor.

There is an important difference between the tensorization in the previous Chapter and here: in Chapter 4, the individual heartbeats were segmented by selecting a window *around* the R peaks. When these heartbeats are stacked one after the other, the different heartbeats are automatically synchronized at the R peak location (which is always at the same time location in the segmentation window). Here, segmentation is done by taking a fixed-size window *after* the R peak that contains the T wave. When for example the heart rate changes within a signal, it is possible that the timing of the T wave with respect to the R peak changes, and that the T waves of different beats are thus not strictly aligned in the third mode.

5.3.4 Tensor decomposition

In the next step, the T-wave tensor $\mathcal{T} \in \mathbb{R}^{C \times J \times M}$ was decomposed to obtain the factor vectors necessary to perform TWA detection. Two different decomposition methods were compared: CPD and PARAFAC2.

Canonical Polyadic Decomposition (CPD)

Similar to the approach explained in the previous Chapter, CPD was the first method used to break down the tensor in different factors. If $\mathcal{T}$ is a third-order

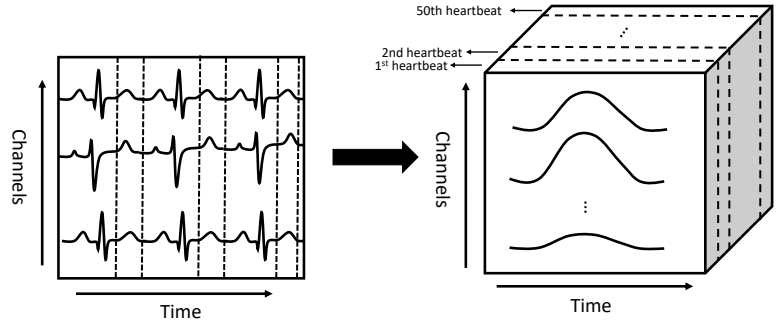


Figure 5.2: Construction of the T wave tensor: the T waves of all channels are segmented and stacked beat-by-beat in a 3D manner. The result is a tensor $\mathcal{T} \in \mathbb{R}^{C \times J \times 50}$ with modes *channels* x *time* x *heartbeats*.

tensor, CPD can be expressed as:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a_r} \otimes \mathbf{b_r} \otimes \mathbf{c_r} \tag{5.2}$$

The number of rank-1 terms R is the rank of the decomposition, $\mathbf{a_r}, \mathbf{b_r}$ and $\mathbf{c_r}$ for $r = 1, ..., R$ are the factor vectors for each order. Building on the results of Chapter 3, and since here the interest lies in the main variation between subsequent T waves, only one component was extracted and therefore R was restricted to one. The result of Equation 5.2 were then three factor vectors $\mathbf{a}_1 \in \mathbb{R}^N$, $\mathbf{b}_1 \in \mathbb{R}^J$ and $\mathbf{c}_1 \in \mathbb{R}^{50}$ that correspond to the three modes of the original tensor, *channels*, *time* and *heartbeats*.

CPD results in one loading vector for each tensor mode. If the tensor is low-rank, the CPD model is accurate and these three vectors will capture the main variations present in the tensor. When the ECG signal is relatively noisefree and the T waves are perfectly aligned in the third mode, this assumption is met and CPD will lead to correct results, similar to the results presented in the previous Chapter. When there is however a variation in for example the timing of subsequent T waves (as can be the case with a changing heart rate), the tensor cannot be decomposed in rank-1 components and CPD will lead to inaccurate results. In those cases it is better to use a more general tensor decomposition method such as PARAFAC2.

PARAFAC2

PARAFAC2 is a tensor decomposition method that is a variation of the standard CPD [30]. The main difference is that PARAFAC2 is less restrictive than CPD in the sense that it allows variations in the factor vectors of one mode. This is especially useful when one of the factors contains a time shift. In this application, such time shift arises if the T waves are not exactly aligned in time, for example when the heart rate differs and the QT interval length changes. The difference is shown in Figure 5.3, which shows schematic representations of CPD (Figure 5.3a) and PARAFAC2 (Figure 5.3b). If $\mathbf{X}_m \in \mathbb{R}^{C \times J}$ is one frontal slice of a tensor with the T waves of one heartbeat in all channels, this slice can be modeled as:

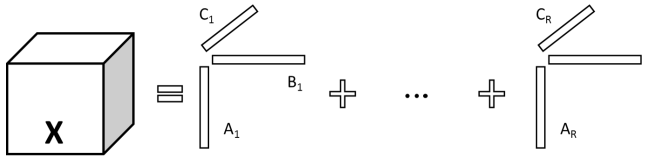
$$\mathbf{X}_m = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_{r,m} \otimes \mathbf{c}_r + \epsilon \tag{5.3}$$

The subscript m implies that there is an individual loading vector $\mathbf{b}_{r,m}$ for each frontal slice of the tensor, effectively making $\mathbf{B}_r$ (the collection of loading vectors for each heartbeat m) a loading *matrix* with dimensions $J \times M$. Each row of this loading matrix corresponds to the temporal profile of a T wave in one heartbeat, taking into account the possible time shifts in T waves due to heart rate changes. M is thus the total number of heartbeats contained in one tensor, and J the length of the T wave segmentation window.

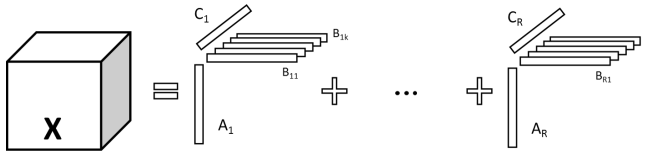
Equation 5.3 will have a unique solution when the cross-product $\mathbf{b}_{r,m}^T \mathbf{b}_{r,m}$ remains constant for all m. This is the case if the different rows of $\mathbf{B}$ are only shifted in time. Under normal circumstances, e.g. when the ECG signal does not contain an excessive amount of noise, this assumption is met and PARAFAC2 accurately decomposes the ECG signal.

5.3.5 Tensor decomposition results

The result of the tensor decomposition is a collection of three factor vectors (or two factor vectors and one matrix in the case of PARAFAC2), which give information about the structure of the signal in the three tensor modes. All three components give valuable information about the T wave characteristics in time and space, but only one was used for the actual detection of T wave alternans. The first mode-1 factor vector corresponds to the spatial dimension. This vector shows the variation of the T waves over the different channels. The mode-2 factor vector (or matrix for PARAFAC2) shows the time course of the T waves. If there is a time shift present this will manifest itself in the different rows of the PARAFAC2 factor matrix. In the case of CPD this vector simply shows the average T wave in the signal. The mode-3 factor vector expresses the



(a) Canonical Polyadic Decomposition



(b) PARAFAC2

Figure 5.3: Comparison of the tensor decomposition methods: The result of CPD are three factor vectors, while PARAFAC2 yields tw factor vectors and one factor matrix since it allows variations in the second mode.

differences in the T wave in subsequent heartbeats. Only this vector was used in the next step for further analysis.

5.3.6 Detection of T wave alternans

The actual detection of T Wave Alternans was done using the third factor vector as explained in the previous paragraph. If a signal contains TWA, the typical ABABAB-pattern that is present in the T wave amplitude will be captured by the tensor decomposition as an alternating time series in the third dimension. To detect TWA the number of consecutive turning points of $\mathbf{c}_r$, the loading vector of the heartbeats dimension, were first detected. Turning points were defined here as local minima or maxima, e.g. points where the derivative of the factor vector changes sign. Since noise also causes (random) variations in the T wave amplitude, a threshold on the minimum number of consecutive turning points was introduced to avoid false positive detections. The threshold was set to 10, since a preliminary study showed that this value leads to good results with real-life signals [91]. The TWA magnitude was then determined by summing the differences between all turning points.

Let TP be the vector of turning point indices (taking into account the minimal number of 10 consecutive turning points) with length L, and $\mathbf{c}_r$ the third loading vector in the result of the tensor decomposition. The total TWA magnitude in

a window can then be calculated as:

$$TWA = \sum_{i=2}^{L} |\mathbf{c}_r(TP(i)) - \mathbf{c}_r(TP(i-1))| \tag{5.4}$$

5.4 Results

This section presents the results of TWA detection on different types of data. In all cases a comparison between CPD and PARAFAC2 is done in order to assess differences between both methods. First the detailed results for three case studies (two signals containing TWA, one with and one without T wave shift, and one signal without TWA) are qualitatively described. Next the results for the three datasets considered are discussed. All experiments were performed using MATLAB R2015b. Tensorlab is used for the tensor construction and CPD decomposition [231]. PARAFAC2 is implemented in the N-way toolbox [9].

5.4.1 Case study

We examined three sample signals taken from the Physionet database. The first signal (*twa11*) did not present TWA, the second signal (*twa99*) contained a significant amount of T wave alternans and the third signal (*twa01*) contained both TWA and shifted T waves caused by changes in heart rate. This way we could evaluate both the effect of the presence of TWA and a shift in T wave position for both methods. All signals consised of 12 channels and were cropped to 50 heartbeats. Figure 5.4 shows a ten second excerpt of the signals. None of the signals contained obvious noise, meaning the results are not influenced by signal quality. The complete signals (30 minutes, 12 channels) can be found on https://physionet.org/pn3/twadb/. After tensorization and tensor decomposition of all three signals (using both CPD and PARAFAC2), the resulting loading vectors could be examined. The results are shown in Figure 5.5.

Figures 5.5a, 5.5d and 5.5g show the spatial vector for the signal without TWA, with TWA and with TWA and T wave shift respectively which show the sign and magnitude of the T wave in different channels. From Figure 5.5a we can for example conclude that T waves in channel 9 had a larger amplitude than those in channel 12. The temporal vectors, shown in Figure 5.5b, 5.5e and 5.5h correspond to the shape of the T wave in time. Here the difference between CPD and PARAFAC2 is obvious, especially in Figure 5.5h. The result for PARAFAC2

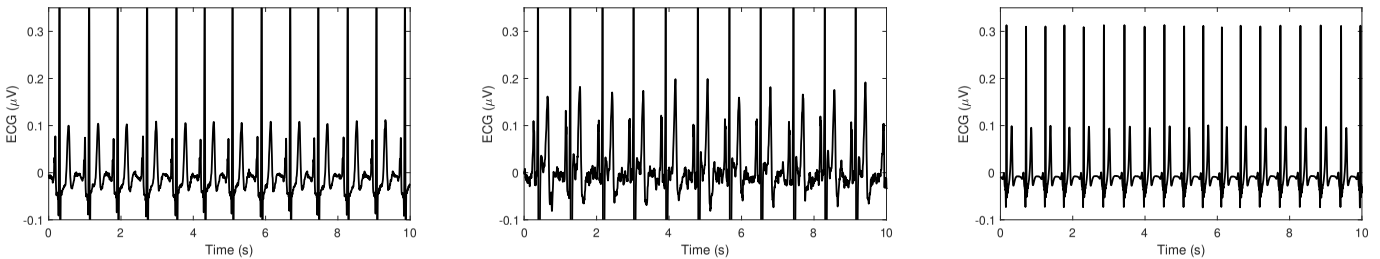
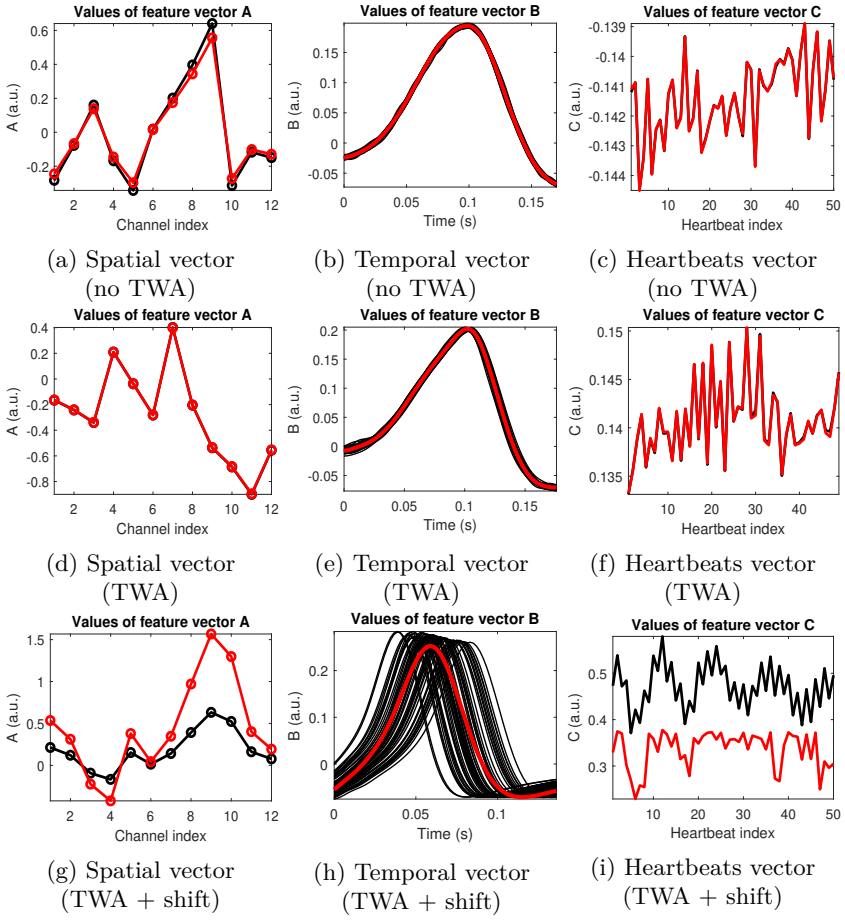


Figure 5.4: Excerpts of ten seconds of all three signals used in the case study, taken from the Physionet TWA database. The left plot shows a sample of the signal without TWA (*twa11*), the middle panel represents the sample with TWA (*twa99*) and the right plot shows the signal with TWA and T wave shift (*twa01*).



(a) Spatial vector (no TWA)
(b) Temporal vector (no TWA)
(c) Heartbeats vector (no TWA)
(d) Spatial vector (TWA)
(e) Temporal vector (TWA)
(f) Heartbeats vector (TWA)
(g) Spatial vector (TWA + shift)
(h) Temporal vector (TWA + shift)
(i) Heartbeats vector (TWA + shift)

Figure 5.5: Vectors **a**, **b** and **c** for signal without TWA (top row), a signal with TWA (middle row) and a signal with TWA and shifted T waves (bottom row). The results for CPD are plotted in red, the results for PARAFAC2 in black. The three different columns show respectively the spatial, temporal and heartbeats feature vectors.

is a matrix, and each row of the matrix (corresponding to the temporal profile for each heartbeat) is plotted on top of each other. The presence of a large T wave shift in the third signal is evident. While CPD succeeded in capturing the mean shape of the T wave, the difference in the timing of the subsequent waves was discarded. This information is contained in the feature matrix obtained by PARAFAC2. The other signals did not contain a large shift, as is demonstrated by the small variation in the PARAFAC2 results of Figure 5.5b and 5.5e. Hence, the results of CPD and PARAFAC2 were much more similar.

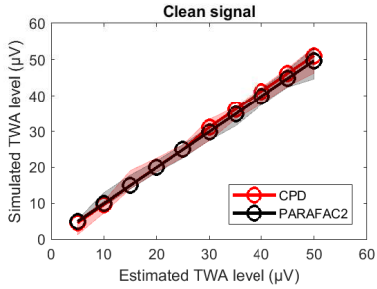
T wave alternans can be detected by analyzing the third loading vector. Figures 5.5c, 5.5f and 5.5i show this loading vector for all analyzed signals. On Figure 5.5c we see that, although there is a certain variation in the T wave magnitude over different heartbeats, no clear pattern is visible. The typical ABABAB-pattern is however clearly visible on Figure 5.5f and 5.5i. Note that in Figure 5.5f, there is only T wave alternans present in the first part of the signal. The total TWA magnitude could then be calculated as defined in Equation 5.4. Note that the results for PARAFAC2 and CPD are almost identical in Figure 5.5c and 5.5f, while the results in Figure 5.5i differ significantly. This is caused by the large T wave shift (see Figure 5.5i).

5.4.2 Artificial signals

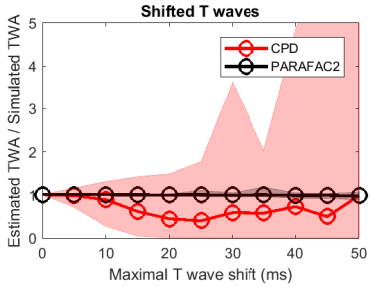
We first analyzed the results of both proposed methods on a clean signal with varying levels of T wave alternans. Then we investigated the influence of a temporal shift in the T waves. Finally the robustness of the methods was tested by changing the noise levels. All results are summarized in Figure 5.6. All results shown are averaged over the 50 simulated patients, with the variance among all patients indicated with the shaded areas.

The results for a signal without noise are presented in Figure 5.6a. The simulated TWA level was varied between 5 μV and 50 μV. There is a clear correlation between the simulated and estimated TWA magnitude, both for CPD and PARAFAC2. There is very little variance between results for all patients: the mean differences in estimated TWA levels over all patients was 3 μV and does not change significantly with the simulated TWA level.

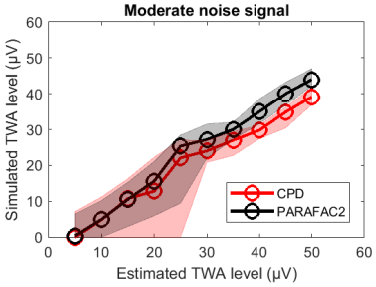
An artificial T wave shift was then introduced by moving the T wave segmentation window a random number of samples (Figure 5.6b). The maximal shift was varied between 0 and 50 ms and the TWA level was fixed to 25 μV. The results are presented as the ratio between the estimated TWA and simulated TWA, which should be 1 for a good detection. The results for PARAFAC2 stay approximately constant (between 0.98 and 1.01), even when the T wave shift is maximal. CPD did not handle the increasing T wave shift well: for a



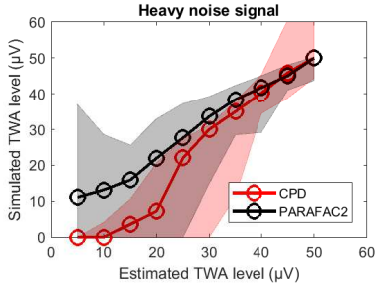
(a) Simulated vs estimated TWA levels for artificial signals without noise.



(b) Estimated TWA levels for different values of T wave shift.



(c) Simulated vs estimated TWA levels for artificial signals with moderate noise.



(d) Simulated vs estimated TWA levels for artificial signals with high noise.

Figure 5.6: Results for CPD (red) and PARAFAC2 (black) for four types of artificial signals: Clean signals with varying amount of TWA (5.6a), clean signals (TWA = 25 μV) with changing T wave shift (5.6b), artificial signals with a moderate (5.6c) and high (5.6d) noise level. The circles represent the mean values over all 50 simulated patients, the shaded areas the variations in the results.

very small shift (between 0 and 10 ms) the result was correct, but for larger shifts the results deteriorated quickly. For a maximal T wave shift of 50 ms the average CPD result seems good, but the large variance among different patients indicate that the obtained results varied greatly and were thus not reliable.

When introducing noise, the correlation between the simulated and estimated TWA levels was maintained, although the estimated CPD levels were distinctly smaller (5.6c). For a simulated TWA level of 50 μV, the average result obtained with CPD was 39 μV, while the result of PARAFAC2 was more accurate (43 μV). The results show a larger spread for lower TWA levels, which could be expected since the noise is more likely to mask the TWA in these cases. Both methods showed less precise results under high noise level (see Figure 5.6d): While the estimated TWA magnitudes on average increased with increasing simulated TWA, there were clearly more variations when compared with Figure 5.6a and 5.6c. Here PARAFAC2 seems to underperform compared to CPD, especially for low TWA levels.

5.4.3 Physionet database

To our knowledge, the Physionet database is the only publicly accessible labeled TWA database. This makes it ideal for use as a benchmarking tool to make comparison with existing methods possible. This can be done by ranking the results from the TWA detection in order of magnitude, and comparing this ranking with the reference ranking by calculating the Kendall rank correlation coefficient between the two. A Kendall coefficient that is greater than 0.436 is considered significant ($p > 0.99$). Both an order using CPD and PARAFAC2 were computed. The Kendall coefficients obtained were both significant, the CPD coefficient being 0.79 and the PARAFAC2 coefficient 0.87. Table 5.1 compares the tensor-based scores with some methods found in literature and the best result of the original TWA challenge. A coefficient of 0.87 and 0.79, achieved using respectively PARAFAC2 and CPD, would have been the fourth and sixth highest rank obtained during the challenge, and are both higher than the values for some of the most widely used methods found in literature.

Figure 5.7 shows a comparison between the reference ranking of the signals and the ranking obtained by respectively CPD (Figure 5.7 - left) and PARAFAC2 (Figure 5.7 - right). There is a clear correlation for both methods, which is reflected in the high Kendall coefficient. The reference ranking and PARAFAC2 ranking correspond largely, and for most signals there is only a small difference between the reference score and the PARAFAC2 score. A remarkable exception is signal *twa63*, which has an intermediate amount of TWA according to the reference ranking (position 52) but where no TWA is detected using the tensor-

Method	**Kendall coefficient**
CPD	0.7851
PARAFAC2	0.87
Periodic Component Analysis [151]	0.766
Modified Moving Average Method [162]	0.73
Spectral Method [200]	0.416
Best challenge [198]	0.911

Table 5.1: Kendall coefficient scores obtained by comparing the rankings from different methods found in literature and the two proposed tensor-based methods with the reference ranking for the Physionet database.

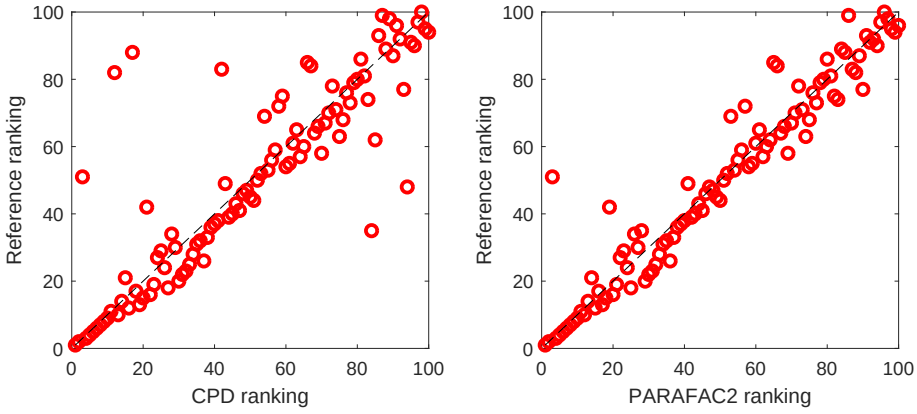


Figure 5.7: Comparison between the reference ranking of the Physionet database and the ranking obtained by CPD and PARAFAC2.

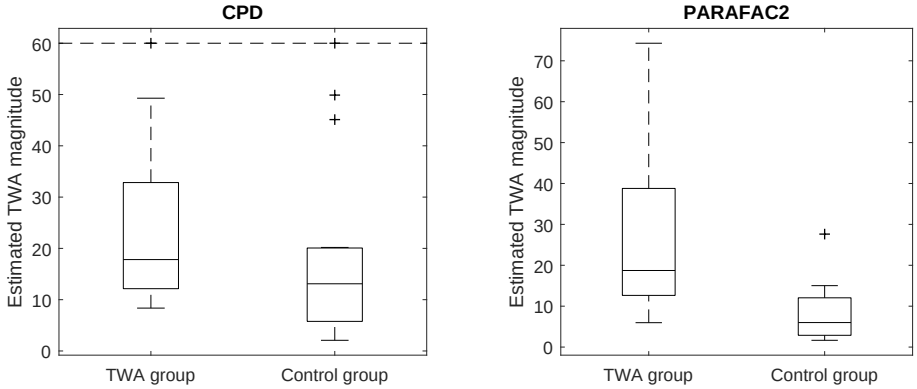


Figure 5.8: Estimated levels of T wave alternans using CPD (left) and PARAFAC2 (right) in TWA group vs control group. The CPD results contain several outliers which are represented in red. Their value was capped at 60 μV (indicated with a dotted line) to increase clarity of the plotted results.

based approach (position 3 with both CPD and PARAFAC2). Further inspection of this signal showed that this signal is taken from the Sudden Cardiac Death Holter Database [92], a database containing signals from patients with sustained ventricular tachyarrhythmia. In this signal in particular, the timing of the T waves was considerably later than in other signals, leading to a suboptimal segmentation and tensor construction. This, combined with the presence of the ventricular tachyarrhythmia, most likely causes the incorrect result in this case.

5.4.4 Clinical dataset from the University Hospitals Leuven

When analyzing the clinical database, T wave alternans levels were calculated in windows of 50 heartbeats as explained in Section 3.6. In order to present results for the complete signals, the TWA levels in all windows were combined by calculating the 90th percentile of all values of windows containing TWA, resulting in one TWA value per signal. The detected levels of T wave alternans in the University Hospitals Database are shown in Figure 6.4 for both tensor decomposition methods.

The results are split according to the patient group (TWA vs control). When comparing between groups, both PARAFAC2 and CPD showed higher TWA amplitudes for the TWA group than for the control group, but the differences were only significant with PARAFAC2 ($p = 0.008$). Additionally we also analysed the number of windows which contain TWA in both groups. The

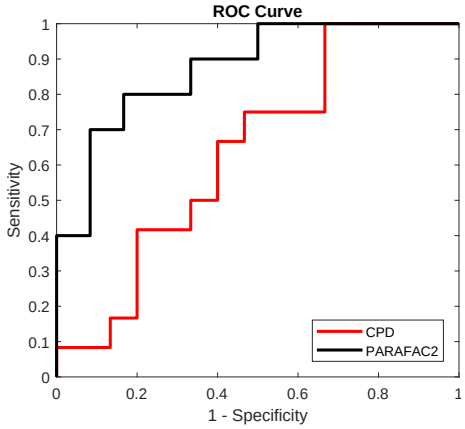


Figure 5.9: ROC curve for classification using CPD (AUC = 0.64) and PARAFAC2 (AUC = 0.88).

number was higher for the TWA group for both methods, but the result was again only significant for PARAFAC2 ($p = 0.03$). On average TWA was detected in double the windows of a signal in the TWA group than in the control group. Since TWA is linked to increased heart rates we also performed an analysis of the heart rate. This did however not reveal significant differences between both groups (mean heart rate of 96.8±9 bpm in the TWA group vs 96.1±13 bpm in the control group). Other signal characteristics such as signal length and number of channels did not differ significantly between both groups either.

To provide a suggestion for a clinically useful TWA threshold, we have used the values obtained by both CPD and PARAFAC2 to classify the signals as either containing TWA or not. The ROC curve of these classifications is shown in Figure 5.9. The AUC values of CPD and PARAFAC2 are respectively 0.64 and 0.88. Further analysis of the results for PARAFAC2 showed that for a threshold of 13 μV the method obtains a sensitivity of 80% and a specificity of 83%, indicating that this might be a good cut-off for further use.

5.5 Discussion and Conclusion

In this Chapter, we presented a method to detect T wave alternans using tensor decomposition methods. When comparing the two proposed tensor decompositions, CPD and PARAFAC2, we can conclude that PARAFAC2 has clear advantages over CPD. While CPD obtained good results in ideal

circumstances (e.g. no noise and perfectly aligned T waves), these conditions are rarely met in real-life situations. This was also demonstrated by the superior results of PARAFAC2 on the non-artificial signals in the Physionet database and the clinical dataset from the University Hospitals Leuven. CPD failed to deliver good results in these circumstances, as was shown by the non-significant results on the patient data from the University Hospitals Leuven. CPD seems however slightly more robust to noise, based on Figures 5.6c and 5.6d. Its inability to deal with changing heart rates and moving T waves makes it nevertheless unacceptable for real-life applications. While it is possible to control changing heart rates up to a certain degree, the results in Figure 5.6b show that CPD results deteriorate quickly and become unreliable even with a minor change. Since small heart rate differences are unlikely to be avoided in real-life circumstances (even with control for heart rate), CPD would not be a good choice for use in clinical practice.

PARAFAC2 on the other hand yielded good results both on artificial and clinical ECG signals. The difference was especially large in the presence of a T wave shift such as in Figure 5.5h. Since this occurs frequently in clinical signals, where a changing heart rate is expected, it is important that methods can deal with this type of variation. Figure 5.6d shows that the results of PARAFAC2 are less robust under high noise conditions. This indicates that it is presumably not an optimal method for ambulatory measurements. However, since many TWA tests occur in clinical environments where noise can be avoided to a certain extent, this is not likely to be a large obstacle in real life. The results shown on Figure 6.4 and 5.9 confirm this. Special attention for the choice of segmentation window is however needed when dealing with patients that have conditions that can alter the T wave interval and timing, such as Long QT Syndrome (LQTS). The results on the Physionet Database seem to indicate that in extreme cases the proposed segmentation window is not sufficient to detect TWA. It would perhaps be a possibility to increase the length of the segmentation window in patients with known LQTS.

When comparing the results of the Physionet Database with other methods, the tensor-based approach performed better than the other methods found in literature and obtained comparable results with the best-scoring method of the CinC challenge [198] (0.87 vs 0.911). The Kendall coefficient only compares the ranking of the different signals, and gives no information about whether the estimation of the TWA magnitude is correct. For this a fully annotated standardized TWA database with correct estimations of TWA magnitude is necessary which does not exist at this moment.

When analyzing the results of the clinical dataset from the University Hospitals Leuven, the difference between TWA positive signals and controls was statistically significant for PARAFAC2 ($p = 0.008$). Note however that also

in the control group a certain amount of TWA was still detected. This could be explained by a suboptimal choice of detection threshold (e.g. the minimal number of consequent sign changes for TWA presence). Full optimization of this parameter (using a database where the exact amount of TWA is known as explained earlier) could improve results here. Another possible explanation however is that the clinical spectral based TWA test does not detect very low levels of T wave alternans ($< 10\mu V$). The spectral based method commercialized by Cambridge heart was chosen as reference for the clinical dataset because it is one of the 2 official FDA-cleared methodologies and has extensively been studied [230]. To verify such hypothesis the artificial signals containing fixed amounts of TWA could be analyzed by the clinical methods to compare the results. This could be an interesting approach for a future study. The suggested clinical threshold of 13 μV could also be validated in this way.

A limitation of this study is the small size of the clinical database of only nine patients. In order to fully validate the clinical threshold and to verify whether it has prognostic utility, a larger prospective dataset is needed. This would also enable us to examine if the detected estimations of TWA magnitude correlate with clinical findings such as occurrences of arrhythmic events. A dedicated study would furthermore allow us to study the link between TWA levels and heart rate, as multiple studies have shown that TWA increases with increasing heart rate [118, 114, 35]. The limited sample size did not allow us to make strong observations in this sense here.

The obtained results show that the tensor-based approaches are a promising tool for TWA detection. They are fully automated and require only a minimal amount of preprocessing. The presented method is fairly general and can (with slight adaptations) also be utilized in different applications. Other types of alternans such as ST alternans or QRS alternans can be evaluated easily by constructing a tensor that consists of other ECG segments.

Based on the obtained results we can conclude that the proposed methods are a robust way of detecting T Wave Alternans in different conditions. PARAFAC2 is a better choice as tensor decomposition method than CPD since the more general structure supports time shifts in subsequent T waves.

Chapter 6

Analysis of changes in heartbeat morphology prior to in-hospital cardiac arrest

Following a similar approach to the previous Chapters, here we present another algorithm that uses CPD to analyse the ECG signal. The main interest in this Chapter is the analysis of changes in heartbeat morphology prior to in-hospital cardiac arrest. For this study, a dataset with long-term patients in the intensive care unit was collected in the UCSF/UCLA hospital. The data were however characterized by large amounts of noise, diminishing the signal quality. In order to make a robust tensor decomposition feasible, we therefore incorporated information about the signal quality in the CPD optimization problem, which allows more accurate analysis. Preliminary results of this study have been published as "Goovaerts G., Van Huffel S., Hu X., "Tensor-based Analysis of ECG changes prior to in-hospital cardiac arrest", in Proc. of the 44rd Annual Computing in Cardiology, CinC, Rennes, France, Sep. 2017" [86]. The paper describing the full method and results is currently under review in the IEEE Journal of Transactions in Biomedical Engineering.

6.1 Introduction

In-hospital cardiac arrests account for approximately 40% of all cardiac arrests [21]. Although patients who experience a cardiac arrest in a hospital have a

higher chance of experiencing arrhythmia while being monitored and thus have a higher chance of being resuscitated in time, only 20% of patients survives to discharge [84]. Early patient identification would thus be very useful in this environment. One strategy is to continuously monitor vital signs in patients and identify patterns or changes in these signals that are predictive for a later cardiac arrest in that patient. These methods can then be applied to out-of-hospital cardiac arrest populations in the future.

When a patient experiences a cardiorespiratory arrest in a hospital, a '*code blue*' is called, indicating a medical emergency or a person in need of immediate medical attention. When a patient has an increased chance of deterioration and eventually of cardiac arrest, rapid response teams are called in order to start early interventions to avoid the occurrence of an actual code blue. In this Chapter, we will therefore examine how the ECG signal morphology changes prior to a code blue in patients with different causes of cardiac arrest. We make a distinction between different types of cardiac arrest, since the underlying mechanisms preceding different type of arrests are fundamentally different, and therefore different ECG changes can be expected for the different patient groups. The identification of ECG features that change significantly before a cardiac arrest would allow a more precise monitoring and could therefore potentially decrease the number of actual code blues. The ECG signal is chosen over other vital signs since it can be measured non-invasively and monitoring equipment is available in most clinical settings.

There have been many studies investigating the immediate mechanisms visible on ECG signals that lead to a cardiorespiratory arrest [115, 134, 236], but changes over a longer timescale (multiple hours before the cardiac arrest) are less well-studied. Previous studies focused on bradyasystolic cardiac arrests [67, 103] or heavily relied on manual measurements [68]. Our previous work examined changes in ECG parameters between a stable period and the period right before the cardiac arrest [86]. The study described here extends the analysis to other causes of cardiac arrest, using automated methods that need a minimum of user interaction. This makes them more useful for future use in clinical practice.

The method proposed here introduces the use of weighted tensor decompositions, which allow us to incorporate prior knowledge about the signal quality in the tensor decomposition in the form of a weight tensor [167]. When choosing the weights properly, they can automatically deal with the noise that is inherent to biomedical signals, leading to a more accurate analysis and making them better suited to work with real-life signals. The method used in this study furthermore uses a computationally efficient weighting scheme [25], which is essential for real-time processing. As such, this Chapter has the following main contributions:

- The application of computationally efficient weighted tensor decompositions to extract ECG features, which permits to correct for lower signal qualities in quasi real-time.
- Development of a noise model to estimate the SNR value of each individual heartbeat (fully described in Appendix B).
- The use of different analyses of changes in parameter values prior to in-hospital cardiac arrest
- The inclusion of three different patient groups with different preceding rhythms.

The rest of this Chapter is structured as follows: The next Section gives a short description of the dataset used here. The different steps of the developed method are described in Section 6.3. Sections 6.4 and 6.5 contain respectively the results and discussion, after which the conclusion is summarized in Section 6.6.

6.2 Data

The dataset contained multilead ECG signals of patients that experienced a code blue due to a cardiopulmonary arrest in the intensive care units of the Medical Centers of the University of California in San Francisco (UCSF) and Los Angeles (UCLA). All code blues occurred between 2013 and 2015. For all patients, information was provided about the date and time of the code blue, the type of cardiac arrest and whether they survived the code blue/and or survived to discharge. Patients were divided in three different groups depending on the preceding rhythm (directly prior to the arrest):

1. Ventricular tachycardia or ventricular fibrillation (VT/VF)
2. Asystole
3. Pulseless Electrical Activity (PEA)

For each patient, ECG signals from the last 24 hours before the code blue occurred were extracted. However, not all patients were continuously monitored over these 24 hours. Only patients with a total ECG length of more than three hours were included in the study. Additionally, patients of which the cause of cardiac arrest was not known or dubious were also excluded for analysis.

Table 6.1: Number of patients with preceding rhythms in the dataset.

Cardiac arrest	# patients	%
VT/VF	33	20.6%
Asystole	28	17.5%
PEA	99	61.9%

The final dataset contained 160 recordings with a sampling frequency of 240 Hz. The signals from UCSF contain 7 channels: I, II, III, V, aVR, aVL and aVF, those from UCLA were measured with four leads. The number of patients in each cardiac arrest group is summarized in Table 6.1.

6.3 Methods

Our method consists of five different steps: 1) The signals were first preprocessed to remove major noise sources and detect R peaks. 2) The two-dimensional signals were then transformed into a third-order tensor, which was subsequently 3) decomposed to extract a template heartbeat. 4) This heartbeat was used to derive seven parameters 5) which were finally analysed in two different ways.

6.3.1 Preprocessing

Some ECG recordings contained corrupted portions in the beginning of the recording, which were visually detected and deleted. All signals were then preprocessed channel-by-channel. First a Butterworth high-pass filter with order 2 and cutoff frequency 0.5 Hz was applied to remove baseline wander. The filter was applied twice, first in forward and afterwards in backward direction to remove phase distortion so the order is effectively doubled to 4. The filter parameters were chosen in order to minimize deformations of the ECG signals, as shown in [127]. Then the signals were normalized by calculating the modified z-score z', which is more robust to outliers than the standard z-score [106]

$$z'_i = \frac{0.6745(x_i - \tilde{x})}{MAD} \tag{6.1}$$

where $\tilde{x}$ denotes the median and MAD the median absolute deviation of the signal. R peaks were detected using a multilead extension to the Pan-Tompkins algorithm that fuses R-peaks detected in separate channels [125].

6.3.2 Tensorization

We tensorized the preprocessed ECG data matrices into third-order tensors with dimensions *channel* x *time* x *heartbeats* where each mode-two fiber represents one heartbeat in one channel. First, the signals were segmented in individual heartbeats of length I, obtaining J heartbeats for all M channels. Next, we stacked all heartbeats in the third mode, resulting in a third-order tensor $\mathcal{T} \in \mathbb{R}^{M \times I \times J}$. Segmentation was done here by taking a fixed-size segmentation window symmetrical around the R peak. The window length is $1.6RR$ (with RR the mean RR interval length).

All signals were processed with a sliding window of 100 heartbeats without overlap between subsequent windows. We constructed, decomposed and analysed a tensor for each window, which allowed to capture the dynamic ECG changes over time. The result of the tensor construction was thus one tensor $\mathcal{T}$ for each 100 beat window. Each tensor had three modes: the first mode was the spatial mode which corresponded to the different ECG leads in the signal. The second temporal mode showed the time course of the individual heartbeat and had a varying length equal to the length of the segmentation window for each tensor. Finally, the third mode was the heartbeat dimension with a fixed length of 100.

6.3.3 Tensor decomposition

Weighted Canonical Polyadic Decomposition

As explained before, CPD writes a tensor $\mathcal{T}$ as a sum of rank-one terms. Here we used an alternative way to compute the CPD, namely Weighted CPD (WCPD) which uses *weighted* least-squares instead of regular least-squares. WCPD permits the incorporation of prior knowledge about the signal quality in the tensor decomposition, giving lower weight to entries with higher noise levels. This is done by introducing a weight tensor $\mathcal{W} \in \mathbb{R}^{M \times I \times J}$ with the same dimensions as $\mathcal{T} \in \mathbb{R}^{M \times I \times J}$ into the standard CPD optimization problem 2.3. Each entry of $\mathcal{W}$ contains the weight for the corresponding entry of $\mathcal{T}$. A detailed explanation of the weight tensor construction follows in the next paragraph. The new optimization problem is then:

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \frac{1}{2} ||\mathcal{W} * (\mathcal{T} - [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!])||_F^2 \tag{6.2}$$

The optimization problem was solved using a novel Weighted Least Squares approach (WLS) where the weight tensor is modelled by a polyadic decomposition (as in Equation 2.3), enabling efficient weighting. The computational details of the WLS algorithm can be found in [25].

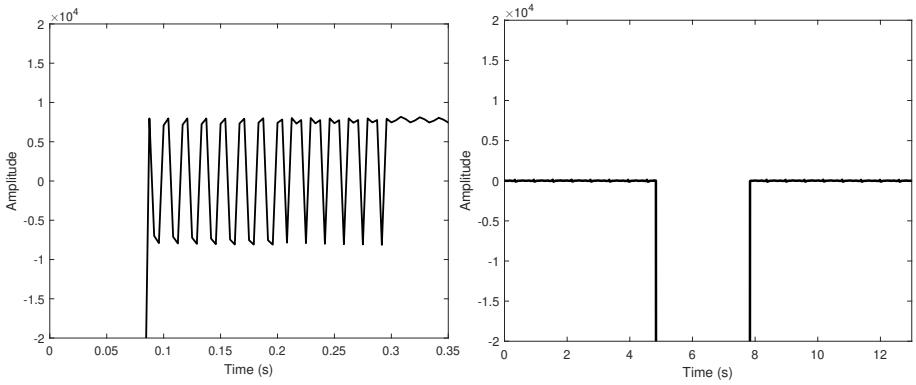
In the previous Chapters, a CPD rank of one was shown to lead to good results [87, 88]. The same approach was used here. This means that the tensor $\mathcal{T}$ was compressed into a single rank-one term, consisting of factor vectors $\mathbf{a}_1 \in \mathbb{R}^M$, $\mathbf{b}_1 \in \mathbb{R}^I$ and $\mathbf{c}_1 \in \mathbb{R}^J$. While the use of a different rank value was explored, increasing the rank did not lead to improved results. This further confirmed our initial choice.

Construction of weight tensor

The weight tensor $\mathcal{W} \in \mathbb{R}^{M \times I \times J}$ contains information about the signal quality, e.g. entries with higher quality receive higher weights. The quality of ECG signals is reduced by artifacts, which are technical or physiological. Technical artifacts can be caused by equipment malfunctioning or electrode loosening. During a technical artifact no ECG signal is measured; the corresponding entries in $\mathcal{T}$ therefore receive a weight of 0, effectively eliminating them from further analysis. Physiological artifacts are caused by for example muscle contractions and are superimposed onto the ECG signal, reducing the Signal-to-Noise Ratio (SNR). For signals which do not contain technical artifacts an estimate of the SNR (calculated using a novel SNR model described in the next paragraph and Appendix A) was therefore used as weight. We calculated one weight for each complete mode-2 fiber, e.g. each full heartbeat in each channel. It was thus assumed that the signal quality was the same during the time course of one heartbeat but could differ from channel to channel or between different heartbeats. Hence, the resulting weight tensor $\mathcal{W}$ has rank M, by construction, with M equal to the number of channels in the ECG signal.

Technical artifacts Some typical examples of technical artifacts are shown in Figure 6.1. They are characterized by extreme signal amplitudes (Fig 6.1b) and/or very structured patterns (Fig 6.1a) that are substantially different from physiological signals. Figure 6.1b includes three seconds of normal ECG signal in the beginning and end as an amplitude reference. Since the artifacts were extremely dissimilar from normal ECG signals, they could be detected with the following straightforward heuristic. A heartbeat was determined to have a technical artifact if at least five samples had either:

1. Modified z-score > 5
2. Modified z-score $= 0$
3. Constant derivative



(a) Technical artifact with structured pattern

(b) Technical artifact with extreme amplitude

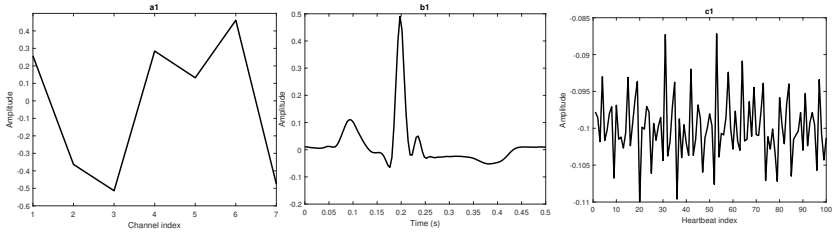
Figure 6.1: Two examples of typical technical artifacts found in the dataset. They can be easily detected and removed by setting their corresponding weights in the weight tensor to zero.

The weights of technical artifacts were fixed to zero, as described in the previous paragraph.

Physiological artifacts As explained earlier, we used an estimate of the SNR of a heartbeat as weight in the case of physiological artifacts. We developed a SNR model based on the MIT-BIH Noise Stress Test Database [156] which contains ECG signals with standardized SNR values. The model contained eight different features found in the literature that have been proven successful in calculating the SNR of an ECG signal [45]:

1. Skewness
2. Kurtosis
3. Power in 6 different sub bands: 0-10 Hz, 10-20 Hz, 20-48 Hz, 48-52 Hz, 52-100 Hz, 100-120 Hz

A detailed explanation of the construction and validation of the SNR model can be found in Appendix B.



(a) Mode-1 spatial vector $\mathbf{a}_1$

(b) Mode-2 temporal vector $\mathbf{b}_1$

(c) Mode-3 heartbeats vector $\mathbf{c}_1$

Figure 6.2: Illustration of the factor vectors of the weighted CPD for a typical ECG signal. Each vector corresponds to one mode of the original tensor $\mathcal{T}$.

6.3.4 Parameter calculation

The output of the weighted tensor decomposition are 3 factor vectors $\mathbf{a}_1$, $\mathbf{b}_1$ and $\mathbf{c}_1$, corresponding to each mode of the tensor. Examples of each factor vector for a typical ECG signal are depicted in Figure 6.2. They are identical to the factor vectors used for irregular heartbeat detection in Chapter 4. The mode-1 vector $\mathbf{a}_1$ (Fig 6.2a) shows the distribution over the different ECG channels. The mode-2 vector of the temporal dimension $\mathbf{b}_1$ can be seen as a template heartbeat for all heartbeats in that window (Fig 6.2b). The mode-3 factor vector $\mathbf{c}_1$ shows the differences among all heartbeats in a window (Fig 6.2c). Since the analysis focused on the changes in ECG morphology, the mode-2 factor vector was used for further analysis. All factor vectors are by definition unique up to scaling. Since there is only one term here, there is no permutation ambiguity. In order to remove differences in scaling between different tensor decompositions, all vectors were first normalized to unit length, b_1 was then rescaled by multiplying it with λ [119]

$$\lambda = ||\mathbf{a}_1||.||\mathbf{b}_1||.||\mathbf{c}_1|| \tag{6.3}$$

with $||\mathbf{a}_1||$, $||\mathbf{b}_1||$ and $||\mathbf{c}_1||$ the original norms of the vectors.

As explained earlier, the mode-2 factor vector corresponds to the model heartbeat in the ECG signal. Standard techniques could thus be used to automatically detect the individual ECG waves. Here, we used the wavelet-based ECG delineator developed by Martinez et al. [143] to extract the locations of the P wave peak, Q wave, R peak, S peak and T peak. Only the peak of the P and T wave were detected since the begin and end points of these waves are often hard to distinguish. All annotations were inspected visually and adjusted where necessary. 7 different metrics were calculated for each analysed window:

1. Mean heart rate in BPM (HR), calculated from the length of $\mathbf{b}_1$ (l) which is equal to the segmentation window for each tensor (see Section 6.3.2):

$$\left.\begin{array}{l} l = 1.6RR \\ HR = \frac{60.fs}{RR} \end{array}\right\} HR = \frac{60.fs}{l/1.6}$$

   with RR the mean RR interval length in samples and fs the sampling frequency of the recording.

2. PR interval in ms (PR)

3. QRS interval in ms (QRS)

4. QT interval in ms (QTc), corrected for heart rate using Fridericia's correction formula [159]

$$QTc = \frac{QT}{\sqrt[3]{RR}} \tag{6.4}$$

5. Amplitude of the P wave ($pAmp$)

6. Amplitude of the R wave ($rAmp$)

7. Amplitude of the T wave ($tAmp$)

All parameters could be computed from the detected waves by either calculating the differences between the corresponding time stamps (metrics 2–4) or by measuring the amplitude of $\mathbf{b}_1$ at that point (metrics 5–7). Missing values, for example in windows which fully consist of technical artifacts, were imputed by linearly interpolating the parameter values between the last known value and the next known value. Note that since all signals were normalized in the preprocessing stage, the final three parameters do not present the ECG signal in μV, but rather the normalized amplitudes. This facilitates comparison among subjects.

6.3.5 Analysis of changes in parameter values

Two different analyses were performed. The first one examined the changes of the different parameters compared to a baseline value in the beginning of the recording. The second analysis looked at the dominant trend in the parameter values.

Changes from baseline

The first analysis detected significant changes from baseline. For each parameter in each patient, an individual baseline value was defined as the median of the first five available values. Each parameter in each analysis window was compared to the baseline value, and changes were considered significant if they exceeded a threshold value for at least 20 consecutive windows or if they were sustained until the time of cardiac arrest, as was defined in [68]. Increases and decreases in parameter values were examined separately. For *PR*, *QRS* and *QTc* a significant change was a >20ms change from baseline value. For *pAmp*, *rAmp* and *tAmp* a change was significant if the amplitude rises or falls larger than 20% from baseline value. For heart rate measurements a difference of 25BPM was detected. All baseline values and threshold values for PR, QRS and QTc interval lengths were chose identical to [68] so results can be compared easily. Threshold values for the other parameters (not analysed in [68]), were defined based on the work by [37] which studies normal variations in the amplitudes of ECG waves during stress testing.

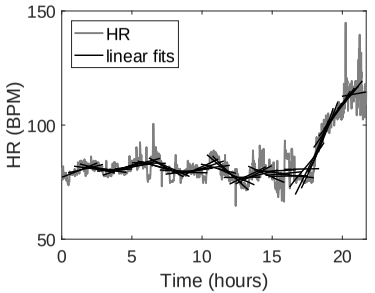
Comparisons between groups (e.g. patients with different causes of cardiac arrest) was performed with a Chi-Square test, which compares proportions between different groups (in this case patients with different causes of cardiac arrest). Since changes of 7 different parameters were tested, correction for multiple comparisons was done with the Benjamini-Hochberg procedure to decrease the false discovery rate [22]. All p-values were ranked in ascending order, and each p-value was compared with its Benjamini-Hochberg critical value c_i to determine significance. c_i was defined as:

$$c_i = \frac{i}{m}Q \tag{6.5}$$

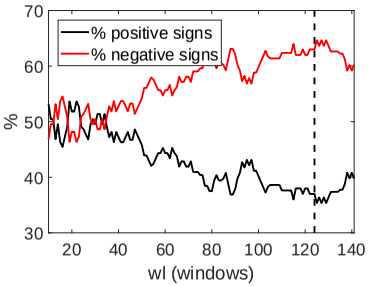
with i the rank of the p-value, m the total number of parameters and Q the false discovery rate, set to a standard value of 10% in this case as suggested in [145].

Dominant trend analysis

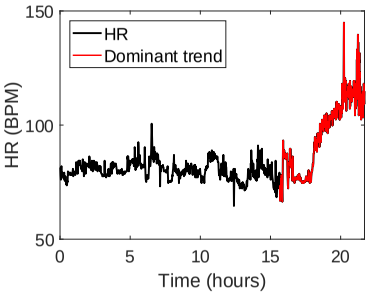
Determination of the dominant trend in a signal is a multi-scale approach to the analysis of trends in a signal. The dominant trend is defined as the longest monotonically increasing or decreasing duration for each metric [67]. After detection of the dominant trend, explained in the next paragraph, we analysed its duration, start and terminal value and slope, and compared values for all groups. To detect the dominant trend in the values of all different parameters, the procedure as outlined in [67] was followed. It consists of four different



(a) Linear fit for one sub-window length wl in black, superimposed on the heart rate signal (grey).



(b) Percentage of positive vs negative signs for all values of wl together with the optimal wl (dotted line).



(c) Final result of dominant trend detection, with the dominant trend indicated in red.

Figure 6.3: Illustration of the different steps in the detection of the dominant trend applied to a heart rate signal.

steps, which are applied to each series of parameter values. Since a tensor was constructed for each window of 100 heartbeats, the time scale of the parameter values is expressed in windows.

1. Let the sub-window length wl range from 10 to 150 windows. Create a set of overlapping sub-windows for each wl that span the entire recording with an increment of one window.

2. Perform robust linear fitting on each sub-window and record the signs of the slope of the linear fit. The sign that occurred most is the dominant slope sign. Robust linear fitting determines the best linear fit between the begin and endpoint of the sub-window while disregarding outliers in the signal.

3. Define the optimal sub-window length by calculating the sub-window length that contained the largest number of dominant slope signs.

4. The dominant trend is the interval that has the most number of consecutive dominant slope signs for the optimal window length.

When the dominant trend was detected, we easily determined its duration, slope, start and terminal value. The differences between values for different groups were tested with one-way Analysis of Variance (ANOVA) testing, followed by a post-hoc Tukey's honest significant difference test. ANOVA verifies whether means of parameter values in different categories are statistically different or not. If ANOVA yields statistically significant results, this indicates that at least one mean is significantly different. Tukey's honest significant difference test is a one-step multiple comparison test to compare all pairs of means that controls the family-wise error rate. For all tests a p-value <0.05 was considered significant.

Figure 6.3 illustrates the different steps in detection of the dominant trend. The signal depicted is a heart rate signal from the dataset with an obvious positive trend in the last quarter of the signal. On Figure 6.3a, the different robust fits for one sub-window length wl are plotted in black together with the heart rate signal. Figure 6.3b shows the percentage of positive and negative signs for each wl. Based on this figure the optimal window length is determined, indicated with a dotted line. The dominant trend is then detected in the signal, and shown on Figure 6.3c in red. Note that while the dominant trend is increasing, there can be periods within the trend where the heart rate instantly decreases (for instance around hour 16).

Table 6.2: Results of the analysis of changes of feature values from baseline. The table represents the percentage (%) of patients in each group with significant increases or decreases in each parameter, together with the p-values from the corresponding Chi-Square test. Two parameters show significant differences between the three patient groups: PR-interval and QRS-interval prolongation. The parameters related to the changes in ECG wave amplitude change significant in a majority of patients over all groups.

Feature	All	VT/VF	Asystole	PEA	p-value
HR+	61,25	57,58	60,71	62,63	0,8737
HR-	25,00	24,24	25,00	25,25	0,9933
PR+	**45,00**	**63,64**	**50,00**	**37,37**	**0,0268***
PR-	38,13	39,39	50,00	34,34	0,3172
QRS+	**27,50**	**27,27**	**53,57**	**20,20**	**0,0023***
QRS-	15,63	15,15	25,00	13,13	0,3105
QTc+	65,00	66,67	67,86	63,64	0,8951
QTc-	58,13	57,58	67,86	55,56	0,5061
pAmp+	80,00	75,76	89,29	78,790	0,3732
pAmp-	74,38	72,73	85,71	71,72	0,3162
rAmp+	81,25	78,79	85,71	80,81	0,7748
rAmp-	66,88	57,58	60,71	71,72	0,2447
tAmp+	85,63	81,82	89,29	85,86	0,7055
tAmp-	72,50	69,70	67,86	74,75	0,7105

* denotes statistically significant results after correction for multiple comparisons with the Benjamini-Hochberg procedure.

6.4 Results

The results of both analyses are summarized in this Section. All calculations were performed in Matlab 2017b using Tensorlab for all tensor operations [231].

6.4.1 Changes from baseline

Table 6.2 shows the result for the first analysis, e.g. the changes compared to baseline values. All parameters changed significantly in at least 15% of all patients. QRS interval shortening was least often detected, T wave amplitude increase most often (85% of all patients). Remarkably, T wave amplitude increase was also detected in 72.5%. Chi-square testing revealed significant differences between the 3 considered groups in two parameters: PR interval prolongation and QRS interval prolongation. Patients with VT/VF cardiac arrests had more increases in PR interval duration. Patients with asystolic arrests showed more increases in QRS interval length compared to other groups. PEA patients had a lower chance of PR and QRS prolongation compared to the other groups. For all groups, both increases and decreases in P, QRS and T wave amplitude were highly prevalent, with significant changes in more than $\frac{2}{3}$ of all patients overall. There were however no statistically significant differences between the different groups.

6.4.2 Dominant trend analysis

After detection of the dominant trend in all seven parameters, four characteristics of each dominant trend were compared between groups: the start value, terminal value, slope and duration. ANOVA testing showed significant differences for 6 parameters, presented in Figure 6.4. The terminal value and duration of the PR interval trend, the terminal value of the QRS interval trend and the begin- and terminal value and duration of the QTc interval trend. Tukey's honest significant difference test was used to examine which groups differed significantly from each other, results are indicated in Figure 6.4 with * (for $p<0.05$) and ** (for $p<0.01$). Results showed that all parameters were significantly different between either the VT/VF and PEA group (PR interval trend duration and all parameters derived from the QTc trends) or the asystole and PEA group (terminal values of PR and QRS interval trends).

6.5 Discussion

This Chapter presented a method to automatically analyse changes in ECG morphology over time. The method was applied to a dataset of ECG signals of patients in the ICU to determine whether ECG morphology changes differently in patients with different causes of cardiac arrest.

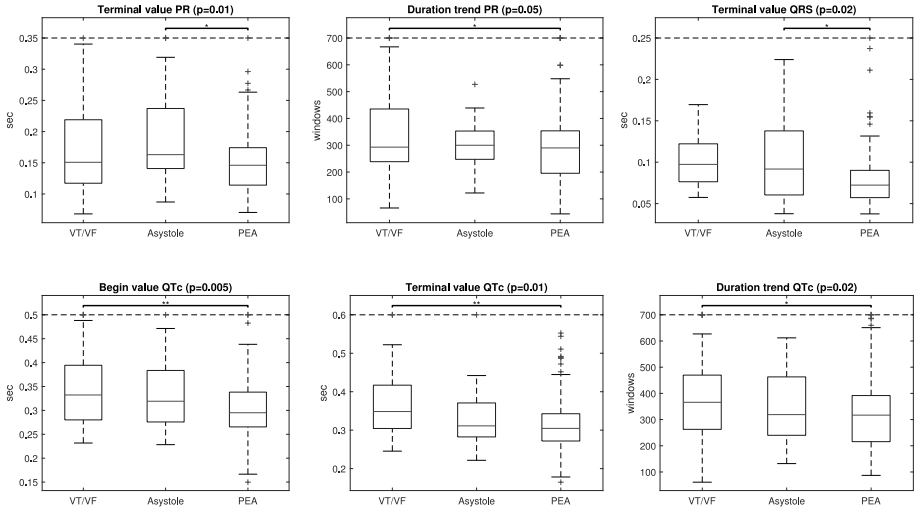


Figure 6.4: Parameters derived from dominant trend analysis which showed statistically significant differences between all groups, together with the p-value results from one-way ANOVA. Results from Tukey's hsd test are indicated with * ($p<0.05$) and ** ($p<0.01$). Boxplots show median values and interquartile ranges. Outliers are clipped to improve visualization.

First the changes compared to a baseline value in the beginning of the recording were analysed. The parameters that showed significant changes, PR and QRS interval prolongation, correspond with the findings of [68] where a similar analysis was done with manual measurements on a different dataset. Both baseline values and significant differences were defined identically to [68] so a valid comparison could be made. While QTc interval prolongation is no significant parameter in this dataset, the QTc interval increases in the majority of patients. This also matches the results in [68] and other similar studies [134]. When further analysing the changes in wave amplitude, it is remarkable that many signals showed both significant increases and decreases in the amplitudes of all ECG waves. This indicates that the ECG amplitude experiences many changes in the period before cardiac arrest that are larger than physiologically normal changes. While they were not discriminative between the groups considered in this experiment, they changed in a large majority of patients with cardiac arrests and might therefore be of use for general patient monitoring.

When analysing the dominant trends in all parameters, six derived features showed significant differences between groups. The parameters where the dominant trend differed between the three groups largely correspond with the

results of the first analysis: PR and QRS interval length. Differences in QTc interval trends were also significant, while this parameter was not picked up in the first analysis.

The start- and endpoint value as well as the duration of the trend in QTc interval length were all significantly larger in VT/VF patients than PEA patients. Since QTc prolongation is a known risk factor for Torsade de Pointes [17], a dangerous form of ventricular tachycardia, this result is not surprising. It is also supported by the findings in [134, 68].

When combining the results of both analyses, it can be concluded that changes in interval length are more significant between groups of patients with different causes of cardiac arrests than changes in amplitude. Changes in amplitude, especially for P wave and T wave amplitude, are however more prevalent overall and might thus be useful to monitor overall patient deterioration. To confirm this, it would be beneficial to collect data from a group of control patients that do not experience a code blue to verify the normal physiological variations in different parameters. This could be the subject of a follow-up study in order to transform the developed methods in useful clinical tools.

The Weighted Least-Squares approach used to compute the WCPD optimization problem is computationally much more efficient than previous weighted tensor decomposition methods. This is extremely beneficial for use in continuous patient monitoring, where the delay between the physiological change and the algorithmic output should be minimal. The current method gives an output for each sliding window of 100 heartbeats (1–1.5 minutes), resulting in a quasi-real time monitoring of ECG morphology. Recent advances in tensor methodology however focus on the development of efficient tensor *updating* methods [220, 221]. This class of methods calculate a new tensor decomposition every time new data arrives (for example after every heartbeat) in a time- and memory-efficient way. Their use would lead to an improvement in the time resolution of the results which could help in gaining a deeper understanding in the timing of the changes in ECG morphology before cardiac arrests.

A first experiment was performed where CPD updating with low-rank weights [221] was tested on a subset of the dataset considered here. The CPD was updated for each new heartbeat (e.g. each new frontal slice of the tensor $\mathcal{T}$. Analysis of the time required to calculate the updated CPD showed that the average computation time was less than 0.5 seconds, which is lower than the normal RR-interval length. This indicates that the computationally efficient approach to calculate both the WCPD and the updating can be performed during the 'waiting time' for a new heartbeat, making this class of methods convenient to use in real-time applications.

6.6 Conclusion

We have developed a tensor-based method to analyse changes in ECG morphology over time and used it to analyse changes prior to in-hospital cardiac arrests in three groups of patients, each with a different cardiac arrest mechanism. The results showed significant differences in multiple parameters with both analyses that largely confirm findings from previous studies. The developed methods work fully automated and are not dependent on the number of channels or measurement devices. Although a follow-up analysis using signals of control patients without cardiac arrests is necessary for validation, the current results show that the proposed tensor-based method leads to clinically valid results that are useful for identification of patients at risk for in-hospital cardiac arrest.

Chapter 7

Automatic detection of atrial fibrillation in single and multilead ECG signals

While the previous Chapters presented methods that used Canonical Polyadic Decomposition as a basis, this Chapter introduces an application where the Multilinear Singular Value Decomposition is used to analyse ECG signals. Also, the previous Chapter exclusively applied tensor methods on multilead ECG signals, since tensorization can be done in an evident way. Many datasets however also contain single lead signals, and therefore the added value of tensor methods for single ECG processing is explored here. This Chapter proposes a method to detect atrial fibrillation, which is the most frequent arrhythmia. Unlike many ventricular arrhythmia it is not immediately fatal, but it may cause pulmonary embolisms or stroke, making timely detection and treatment essential. This Chapter is largely based on the master thesis of Simon Geirnaert, "Detectie van voorkamerfibrillatie: een tensorgebaseerde methode" [83]. The main concept of this method was developed by myself in the context of the 2017 Computing in Cardiology challenge, Simon Geirnaert worked out the full method and finetuned the final classification.

7.1 Introduction

Atrial fibrillation (AF) is one of the most common cardiac arrhythmia, affecting approximately one percent of the general population [82]. During AF, electrical chaos originates in the atria which is added to the electrical signal in the sinus node. The small electrical atrial impulses result in uncoordinated and unsynchronized muscle contractions, which cause the atria to flutter or fibrillate. The atrioventricular node filters these additional impulses so the ventricles still contract normally, although at a highly irregular rate. Since the cardiac pump function is mainly driven by ventricular contraction, it remains unaltered, but the accumulation of blood in the atria can cause the formation of blood clots which can lead to pulmonary embolisms or stroke [239].

Early detection is essential to start treatment which is often done with medication [82]. Since AF has a profound effect on the electrical activity of the heart, it can be diagnosed on the ECG signal, where it will affect both the ECG morphology and the heart rate. During AF, the contraction of the atria is replaced with fibrillations, eliminating the P wave (which is generated during atrial contraction). In some cases, the P wave is replaced by high frequency electrical activity, f-waves. The irregular heart rhythm is furthermore visible in the tachogram, which will show more variations compared to normal sinus rhythm.

This Chapter presents a method to automatically detect atrial fibrillation in both single and multilead ECG signals. The method combines morphology-based features and rhythm-based features in a SVM classifier. The morphology features are calculated using tensors, with an approach similar to the method described in Boussé et al. [26] to detect irregular heartbeats. Here, the latter was extended and optimized for specific detection of atrial beats. As such it can be seen as complementary to the method described in Chapter 4.

The next Sections first describe the datasets used in this Chapter. Afterwards the proposed methods are described, first for single lead and afterwards for multilead signals. Finally, the results are described and discussed in Section 7.4.

7.2 Data

The methods presented in this Chapter were applied to two separate datasets. The dataset for AF detection in single lead ECG signals was taken from the 2017 Physionet/CinC Challenge. A multilead dataset was constructed by combining

signals from the MIT-BIH AFIB dataset and the AFTDB dataset, both available on Physionet. All signals were split in a model set, training set and test set.

7.2.1 Physionet/Computing in Cardiology Challenge 2017

The 2017 Physionet/Computing in Cardiology challenge aimed at tackling the problem of automatic detection of atrial fibrillation in single lead ECG signals [46]. A set of 12186 single lead ECG signals was made available by AliveCor®. The signals were measured using their KardiaMobile device, which is a smartphone-based monitoring system that measures single lead ECG that correspond with lead I-signals (i.e. measurements between left and right arm).

3658 out of the 12186 signals were kept in a hidden test set that was used as an external validation set during the competition. Since this test set has not been made publicly available, the total dataset used here consists of 8528 signals in total. All signals have been annotated on their underlying rhythm by AliveCor® and adjusted afterwards by experts where necessary. The dataset consists of four classes: normal signals, AF, rest and noisy. The normal class and AF class consist of signals where the underlying rhythm is respectively normal sinus rhythm (NSR) and atrial fibrillation. The rest class contains signals with other abnormal rhythms such as bradycardia or tachycardia. Signals where annotators disagreed about the label were also labelled as rest. Finally, the noisy class contains signals that were too noisy to rate. Since noise detection is an inherently different problem which lies beyond the scope of this work, this class was not considered here.

The single lead ECG signals have a length between nine seconds and one minute. The sampling frequency is 300 Hz. Table 7.1 shows how the signals are divided over the different classes. Since the detection is done in a supervised way, the available data were further divided in a model set, a training set and a test set. 60% of the data were used in the modelling stage, 20% were used to train the SVM classifier and 20% as independent test set. The ratios between the different classes were balanced between all three sets.

7.2.2 MIT-BIH AFIB & AFTDB dataset

The multilead dataset was constructed by combining two publicly available databases: the MIT-BIH Atrial Fibrillation dataset [155] and the AF Termination Challenge dataset [154, 175], both available on Physionet [85].

Set	Class			
	Normal	AF	Rest	Total
Model set	3062	413	1471	5117
Training set	994	163	493	1707
Test set	994	162	492	1704
Total	5050	738	2456	**8528**

Table 7.1: The data profile of the model, training and test set of the single lead Physionet/CinC Challenge 2017 dataset.

MIT-BIH AFIB dataset The first dataset consists of 23 Holter signals measured over ten hours with two channels. The sampling frequency is 250 Hz. The dataset mainly contains signals of patients with paroxysmal or intermittent atrial fibrillation. The different episodes of atrial fibrillation are annotated together with normal signals. From these signals, 80 segments of one minute were extracted at random, all containing portions of the signal with *normal sinus rhythm*. Due to the small size of the original dataset, the short segments do not all come from separate test persons, meaning that there is no full independence between training and test set. In general however the variation between normal signals from different patients is very small, meaning that this set can be considered a representative dataset for normal signals. Note that while segments of AF signals can technically also be extracted from this dataset, analysis of the results and annotations indicated that the labelling of AF segments was not trustworthy in all cases. Therefore a second dataset is used for this purpose.

AFTDB dataset The AF Termination Challenge dataset contains 80 Holter ECG signals. All signals have two channels, identical to the MIT-BIH AFIB dataset, sampled at 128 Hz. All signals are atrial fibrillation signals with a fixed length of one minute. The dataset was collected from a larger dataset of Holter signals for a competition in 2004. This assures that all annotations are checked manually, ensuring the quality and correctness of the labels. The dataset is split into three subsets, used here as model, training and test set. All sets contain signals from different patients. Within the model and training set there may be signals from the same patient, but the signals in the test set are all from different patients. This ensures that there is independence *between* the different sets and *within* the test set.

Set	Class		
	Normal	AF	Total
Model set	30	30	60
Training set	20	20	40
Test set	30	30	60
Total	80	80	**160**

Table 7.2: The data profile of the model, training and test set of the MIT-BIH AFIB & AFTDB dataset.

Final dataset The final dataset is a combination of all signals from the AFTDB datasets and 80 randomly extracted normal segments of the MIT-BIH AFIB dataset, with equal numbers of normal and AF signals. The signals from the MIT-BIH dataset were resampled to obtain the same sampling frequency. Table 7.2 gives a summary of all signals in the final dataset.

7.3 Methods

This section presents three different approaches for detection of atrial fibrillation in single and multilead ECG signals. The three methods are all variations of one main concept, which uses the (ML)SVD to model the morphology of heartbeats and uses this as input for a supervised classification problem.

7.3.1 SVD-based detection in single lead ECG signals

This section describes the first method, which detects AF in single lead ECG signals starting from the SVD. The next subsections describe the different parts of the method: 1) preprocessing, 2) calculation of a representative heartbeat, 3) modelling and 4) classification. A flowchart depicting the workflow of the individual steps is shown in Figure 7.1.

Preprocessing

Signals measured with wearable devices are known to suffer a lot from noise. This is not different for ECG signals. Preprocessing is therefore required to improve

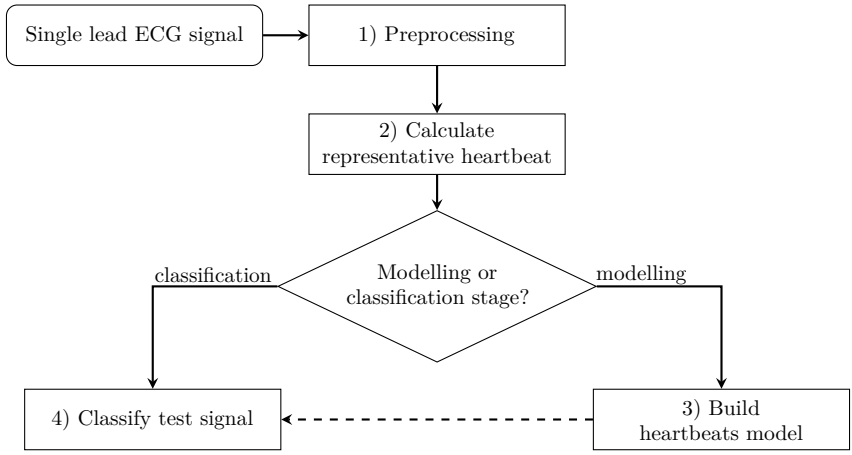


Figure 7.1: Flowchart of the method to classify ECG signals and detect AF in single lead ECG signals starting from the SVD.

the signal quality. More specifically baseline wander, powerline interference and high frequency noise have to be removed. First, powerline interference is eliminated with a notch filter with cut-off frequency of 60 Hz, where an extra pole is added at radius $r = 0.4$ to improve stability. Baseline wander and high frequency noise are removed with respectively Quadratic Variation Reduction [78] and wavelet-based filtering [111], identically to previous Chapters. Multilead signals are processed channel-by-channel. Finally, all signals are normalized by extracting the mean of the signal and dividing by the standard deviation.

Calculation of representative beat

The next step consists of compressing the single lead ECG signal into one representative heartbeat. Pan-Tompkins [173] is used to detect the R peaks in the ECG signal. It is then segmented into different heartbeats similarly to the previous Chapters. A fixed-length segmentation window of 580 ms, starting 250 ms before each R peak, is defined to ensure that the P wave and T wave are both included. Similar values for segmentation windows can be found in literature [26]. In a second step, alignment of different windows is optimized by maximizing the cross-correlation between the different windows. Figure 7.2 visualises the segmentation of the heartbeats on the left-hand side, the alignment is shown on the right.

In a next step, the aligned heartbeats $\mathbf{X} \in \mathbb{R}^{K \times N}$ (with K the number of

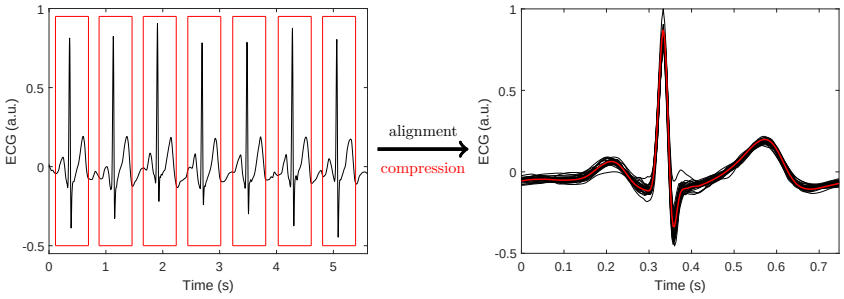


Figure 7.2: The left part of the Figure shows the segmentation of the ECG signal in individual heartbeats by cutting out a window around each R peak. Then, all heartbeats are aligned, shown on the right. The representative beat obtained by compressing all beats with the SVD is superimposed on the aligned beats in red.

beats and N the number of samples in a heartbeat) are compressed into one representative heartbeat $\mathbf{x}_\mathrm{r} \in \mathbb{R}^N$. For this the Singular Value Decomposition of $\mathbf{X}$ is used. The Eckart-Young theorem [74] states that the best rank-r approximation of a matrix $\mathbf{X}$ can be calculated by taking the first r terms of its SVD (see also Section 2.4.3). For r equal to one, this becomes:

$$\mathbf{X} \approx \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\mathrm{T}. \tag{7.1}$$

Vector $\mathbf{v}_1 \in \mathbb{R}^N$ serves as the basis vector of the heartbeats space. This vector is rescaled to norm one and represents the representative heartbeat for the signal. It is also ensured that the sign of the peak is positive. Approximating the signal to one representative beat is meaningful since the variation between beats within a (short term) signal is normally quite small, resulting in one dominant singular value. Figure 7.3 shows the singular values for a random signal in the dataset. There is one very dominant singular value, making a good rank-1 approximation feasible. The result of the compression of the heartbeats is shown on the right-hand side of Figure 7.2.

Model construction

In the next step, a distinction is made between the modelling and classification stage. During the modelling stage, a model set of representative heartbeats is used to construct a basis for the heartbeats space. The labels of the signals in the model set (normal, AF or other) are considered known. The truncated SVD forms the base for the model.

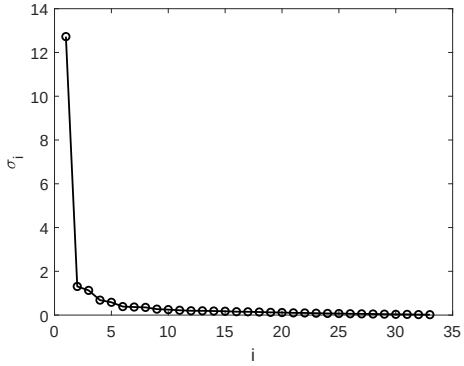


Figure 7.3: The singular values of a random heartbeat matrix $\mathbf{X}$. There is one very dominant singular value, making a good rank-one approximation feasible [83].

$\mathbf{D}_{\text{model}} \in \mathbb{R}^{N \times M}$ is a matrix that contains all M representative heartbeats $\mathbf{x}_r \in \mathbb{R}^N$ from a model set of M ECG signals. The truncated SVD models this matrix with a low-rank approximation:

$$\mathbf{D}_{\text{model}} \approx \hat{\mathbf{U}}_{\text{time}} \hat{\mathbf{S}} \hat{\mathbf{U}}_{\text{heartbeat}}^{\text{T}}, \tag{7.2}$$

with $\hat{\mathbf{U}}_{\text{time}} \in \mathbb{R}^{N \times r}$ an orthonormal basis for the time subspace, $\hat{\mathbf{U}}_{\text{heartbeat}} \in \mathbb{R}^{M \times r}$ an orthonormal basis for the heartbeats subspace and $\hat{\mathbf{S}} \in \mathbb{R}^{r \times r}$ the core matrix. In Equation 7.2, r stands for the rank of the truncated SVD. The choice of r is crucial to obtain a model with good generalisation properties. When r is too small, the model does not describe enough information, leading to *underfitting*. A value that is too high will result in *overfitting*, where too many non-relevant characteristics are modelled, leading to suboptimal detection. Here, 5-fold cross-validation on the model set is used to optimize the choice of r.

Equation 7.2 expresses each representative heartbeat $\mathbf{x}_{\text{r}}^{(i)} \in \mathbb{R}^N, i \in \{1, \ldots, M\}$ (e.g. each column from $\mathbf{D}_{\text{model}}$) as:

$$\mathbf{x}_{\text{r}}^{(i)} \approx \hat{\mathbf{U}}_{\text{time}} \hat{\mathbf{S}} \mathbf{c}_h^{(i)} = \mathbf{B} \mathbf{c}_h^{(i)}, \tag{7.3}$$

with $\mathbf{c}_{\text{h}}^{(i)^{\text{T}}} \in \mathbb{R}^r$ the i^{th} row from $\hat{\mathbf{U}}_{\text{heartbeat}}$. This can be interpreted as follows: the column vectors of $\mathbf{B}$ form an orthogonal basis for the heartbeats space. Column vector $\mathbf{c}_h^{(i)}$ is the coefficient vector that expresses the corresponding heartbeat in this basis. These column vectors are unique up to sign and scaling, and they will therefore be normalised. All normalised coefficient vectors $\tilde{\mathbf{c}}_{\text{h}}^{(i)} = \text{sign}(\mathbf{c}_{\text{h},1}^{(i)}) \frac{\mathbf{c}_{\text{h}}^{(i)}}{\left|\left|\mathbf{c}_{\text{h}}^{(i)}\right|\right|_2}$ then form a database, analogue to the work in [26].

Each vector of this database has a known label, and is afterwards used for classification of new signals.

Classification of new heartbeats

Algorithm 1 explains how the vector with morphological features is calculated for a *new* representative heartbeat $\mathbf{x}_{\mathrm{r}}^{(\mathrm{test})}$, given an annotated model matrix $\mathbf{D}_{\mathrm{model}}$. After calculating the basis and constructing the database of normalised coefficient vectors, a new coefficient vector $\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}$ can be calculated by projecting the new heartbeat onto the heartbeats space (spanned by $\mathbf{B}$). This comes down to solving:

$$\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})} = \hat{\mathbf{S}}^{-1}\hat{\mathbf{U}}_{\mathrm{time}}^{\dagger}\mathbf{x}_{r}^{(\mathrm{test})} = \hat{\mathbf{S}}^{-1}\hat{\mathbf{U}}_{\mathrm{time}}^{\mathrm{T}}\mathbf{x}_{r}^{(\mathrm{test})}, \tag{7.4}$$

Comparison between the normalised (new) coefficient vector $\tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})}$ and the coefficient vectors in the database is done by calculating the inner product:

$$s_i = \left\langle \tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})}, \tilde{\mathbf{c}}_{\mathrm{h}}^{(i)} \right\rangle = \tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})^{\mathrm{T}}} \tilde{\mathbf{c}}_{\mathrm{h}}^{(i)} = \cos(\theta_i), \forall i \in \{1, \ldots, M\}.$$

This score $-1 \leq s_i \leq 1$ measures the angle θ_i between two vectors as a measure for the similarity between the vectors. The result of Algorithm 1, a vector of morphological features $\mathbf{f}^{(\mathrm{test})} \in \mathbb{R}^C$ with C the number of classes, represents the resemblance between the new heartbeat and each of the classes. It is determined by calculating a weighted average of the best W scores s_i of each class c. Both the value of W (equal to 15) and the weight vector $\mathbf{w}$ were empirically determined.

Finally, different classification algorithms can be used to obtain a classification from the feature vector $\mathbf{f}^{(\mathrm{test})}$. The most straightforward method is to assign the test signal the label of the class which has the highest score. This method is for example used during cross-validation to determine the optimal rank. In the remainder of the Chapter, SVMs will be used for classification to simplify the integration of the morphological feature vector with other HRV features.

7.3.2 MLSVD based detection for single lead ECGs

This Section describes a second method to detect AF in single lead ECG signals. It is a higher-order extension of the previous method, relying on the MLSVD instead of the SVD. The framework is similar to the flowchart shown in Figure 7.1, with an extra step to create the higher-order tensor: After preprocessing, the ECG signal is compressed into one representative heartbeat. This is done

Algorithm 1 Calculation of morphological features using SVD

Input: New representative heartbeat $\mathbf{x}_{\mathrm{r}}^{(\mathrm{test})} \in \mathbb{R}^N$, annotated model matrix $\mathbf{D}_{\mathrm{model}} \in \mathbb{R}^{N \times M}$, rank r, weights $\mathbf{w} \in \mathbb{R}^W$
Output: Vector with morphological features $\mathbf{f}^{(\mathrm{test})} \in \mathbb{R}^C$, with C the number of classes

1: Calculate the truncated SVD with rank r:

$$\mathbf{D}_{\mathrm{model}} \approx \hat{\mathbf{U}}_{\mathrm{time}} \hat{\mathbf{S}} \hat{\mathbf{U}}_{\mathrm{heartbeat}}^{\mathrm{T}}$$

2: Construct database $\{\tilde{\mathbf{c}}_{\mathrm{h}}^{(i)}\}$, with $\tilde{\mathbf{c}}_{\mathrm{h}}^{(i)} = \mathrm{sign}(\mathbf{c}_{\mathrm{h},1}^{(i)}) \frac{\mathbf{c}_{\mathrm{h}}^{(i)}}{\left|\left|\mathbf{c}_{\mathrm{h}}^{(i)}\right|\right|_2}$, of normalised rows $\mathbf{c}_{\mathrm{h}}^{(i)^{\mathrm{T}}}$ from $\hat{\mathbf{U}}_{\mathrm{heartbeat}}, \forall i \in \{1, \ldots, M\}$
3: Construct basis matrix $\mathbf{B} = \hat{\mathbf{U}}_{\mathrm{time}} \hat{\mathbf{S}}$
4: Solve

$$\mathbf{x}_{\mathrm{r}}^{(\mathrm{test})} = \mathbf{B}\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}$$

5: Normalise: $\tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})} \leftarrow \mathrm{sign}(\mathbf{c}_{\mathrm{h},1}^{(\mathrm{test})}) \frac{\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}}{\left|\left|\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}\right|\right|_2}$
6: **for** $i = 1 \ldots M$ **do**
7: Calculate similarity s_i:

$$s_i = \tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})^{\mathrm{T}}} \tilde{\mathbf{c}}_{\mathrm{h}}^{(i)}$$

8: **end for**
9: **for** $c = 1 \ldots C$ **do**
10: Calculate $\mathbf{f}_c^{(\mathrm{test})}$ as weighted average $\mathbf{w}$ of the best W scores s_i, for $i \in C_c$, the set of labels c
11: **end for**
12: **return** $\mathbf{f}^{(\mathrm{test})}$

in an identical way as the previous Section. Afterwards, the discrete wavelet transform (DWT) is used to create an extra mode, leading to a model *tensor*, which is used to construct a model to classify new ECG signals.

Tensorisation

The result of the first two steps, preprocessing and compression, is one representative heartbeat $\mathbf{x}_r \in \mathbb{R}^N$ for each single lead ECG signal. Instead of collecting all heartbeats from the model set in a matrix, tensorisation is used to add an extra dimension. As explained in Chapter 2.4.1, many tensorisation approaches exist. Here the Redundant Wavelet Transform (RWT) is chosen

to decompose the heartbeat in different frequency bands. The RWT converts the first-order representative heartbeat vectors to second-order matrices, which are then stacked in a third-order model tensor. The RWT is a variant of the well-known Discrete Wavelet Transform (DWT), which has been used before in combination with the MLSVD to characterize myocardial infarctions [169]. The main difference between DWT and RWT is that there is no down sampling after each filter bank. The resulting scales thus automatically have the same number of coefficients, leading to a straightforward tensorisation. The algorithm is applied with a biorthogonal mother wavelet and five decomposition scales. The appropriate number of scales J was determined with a heuristic equation [8]:

$$J = \lfloor \log_2 f_s - 2.96 \rfloor. \tag{7.5}$$

which leads to a value of five for signals sampled at 300 Hz.

Model construction

The next step is similar to Section 7.3.1 and consists of constructing a model of the heartbeats space. A model tensor $\mathcal{D}_{\text{model}} \in \mathbb{R}^{J \times N \times M}$ contains all wavelet-transformed representative heartbeats $\mathbf{X}_{\text{r}}^{(i)} \in \mathbb{R}^{J \times N}$ from a model set with known labels. Figure 7.4 visualises this tensor, with modes *scale* $\times$ *time* $\times$ *signal index*. The truncated MLSVD is used here as modelling tool. By truncating in each mode individually, an orthonormal base is estimated for each subspace (scale, time, heartbeats). The truncation leads to the following rank-$(r_{\text{scale}}, r_{\text{time}}, r_{\text{heartbeat}})$ approximation of $\mathcal{D}_{\text{model}}$, with $r_{\text{scale}} < J, r_{\text{time}} < N, r_{\text{heartbeat}} < M$:

$$\mathcal{D}_{\text{model}} \approx \hat{\mathcal{S}} \cdot_1 \hat{\mathbf{U}}_{\text{scale}} \cdot_2 \hat{\mathbf{U}}_{\text{time}} \cdot_3 \hat{\mathbf{U}}_{\text{heartbeat}}, \tag{7.6}$$

with $\hat{\mathbf{U}}_{\text{scale}} \in \mathbb{R}^{J \times r_{\text{scale}}}, \hat{\mathbf{U}}_{\text{time}} \in \mathbb{R}^{N \times r_{\text{time}}}$ en $\hat{\mathbf{U}}_{\text{heartbeat}} \in \mathbb{R}^{M \times r_{\text{heartbeat}}}$ orthonormal matrices that form approximations of the corresponding subspaces. The core tensor $\hat{\mathcal{S}} \in \mathbb{R}^{r_{\text{scale}} \times r_{\text{time}} \times r_{\text{heartbeat}}}$ explains the interaction between the three modes.

The optimal rank can, similarly to Section 7.3.1, be determined using cross-validation. The main difference is that here, a multilinear rank $(r_{\text{scale}}, r_{\text{time}}, r_{\text{heartbeat}})$ has to be determined, compared to one rank-value in the previous case. Again, an interval of rank values can be determined based on analysis of the multilinear singular values. To restrict the complexity of the rank optimisation, the three values will be optimised sequentially. This iterative optimisation can be visualised as follows:

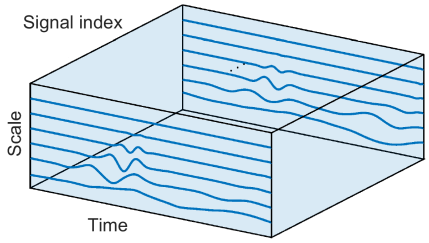


Figure 7.4: A visualisation of the model tensor $\mathcal{D}_{\text{model}}$. The first mode is the scale mode, and the mode-2 vectors are the wavelet coefficients in time. The third mode collects all representative heartbeats from the model sets [83].

$$\left.\begin{matrix} r_{\text{scale}}^{(i)} \\ r_{\text{time}}^{(i)} \\ r_{\text{heartbeat}}^{(i)} \end{matrix}\right\} \xrightarrow[r_{\text{scale}}^{(i)}]{\text{Vary}} \left.\begin{matrix} r_{\text{scale}}^{(i+1)} \\ r_{\text{time}}^{(i)} \\ r_{\text{heartbeat}}^{(i)} \end{matrix}\right\} \xrightarrow[r_{\text{time}}^{(i)}]{\text{Vary}} \left.\begin{matrix} r_{\text{scale}}^{(i+1)} \\ r_{\text{time}}^{(i+1)} \\ r_{\text{heartbeat}}^{(i)} \end{matrix}\right\} \xrightarrow[r_{\text{heartbeat}}^{(i)}]{\text{Vary}} \left.\begin{matrix} r_{\text{scale}}^{(i+1)} \\ r_{\text{time}}^{(i+1)} \\ r_{\text{heartbeat}}^{(i+1)} \end{matrix}\right\} \longrightarrow$$

This method changes one rank at a time while fixing the other two. In practice, the method converges fast with good initial values. The disadvantage is that the method does not guarantee an optimal solution since not all possible combinations are evaluated. It however leads to satisfactory results in practice.

Given the low multilinear rank approximation of a model tensor $\mathcal{D}_{\text{model}}$, then for each tensorized representative heartbeat $\mathbf{X}_{\text{r}}^{(i)} \in \mathbb{R}^{J \times N}$, it holds that:

$$\mathbf{X}_{\text{r}}^{(i)} \approx \underbrace{\hat{\mathcal{S}} \cdot_1 \hat{\mathbf{U}}_{\text{scale}} \cdot_2 \hat{\mathbf{U}}_{\text{time}}}_{\mathcal{B}} \cdot_3 \mathbf{c}_{\text{h}}^{(i)^{\text{T}}}, \tag{7.7}$$

with $\mathbf{c}_{\text{h}}^{(i)^{\text{T}}} \in \mathbb{R}^{r_{\text{heartbeat}}}$ the i^{th} row from $\hat{\mathbf{U}}_{\text{heartbeat}}$ and $\mathcal{B} \in \mathbb{R}^{J \times N \times r_{\text{heartbeat}}}$. Applying Equation (2.1) on Equation (7.7) results in:

$$\begin{aligned} \text{vec}\left(\mathbf{X}_{\text{r}}^{(i)}\right)^{\text{T}} &\approx \mathbf{c}_{\text{h}}^{(i)^{\text{T}}} \mathbf{B}_{(3)} \\ \Leftrightarrow \text{vec}\left(\mathbf{X}_{\text{r}}^{(i)}\right) &\approx \mathbf{B}_{(3)}^{\text{T}} \mathbf{c}_{\text{h}}^{(i)}, \end{aligned} \tag{7.8}$$

with $\mathbf{B}_{(3)}^{\mathrm{T}} \in \mathbb{R}^{JN \times r_{\text{heartbeat}}}$ equal to:

$$\begin{aligned} \mathcal{B} &= \hat{\mathcal{S}} \cdot_1 \hat{\mathbf{U}}_{\text{scale}} \cdot_2 \hat{\mathbf{U}}_{\text{time}} \\ \Rightarrow \mathbf{B}_{(3)} &= \hat{\mathbf{S}}_{(3)} (\hat{\mathbf{U}}_{\text{time}} \otimes \hat{\mathbf{U}}_{\text{scale}})^{\mathrm{T}} \\ \Leftrightarrow \mathbf{B}_{(3)}^{\mathrm{T}} &= (\hat{\mathbf{U}}_{\text{time}} \otimes \hat{\mathbf{U}}_{\text{scale}}) \hat{\mathbf{S}}_{(3)}^{\mathrm{T}}, \end{aligned} \tag{7.9}$$

$\mathbf{B}_{(3)}^{\mathrm{T}}$ serves as a basis for the heartbeats space over all scales. Similar as in Section 7.3.1, the normalised coefficient vectors $\tilde{\mathbf{c}}_{\mathrm{h}}^{(i)} = \text{sign}(\mathbf{c}_{\mathrm{h},1}^{(i)}) \frac{\mathbf{c}_{\mathrm{h}}^{(i)}}{\left|\left|\mathbf{c}_{\mathrm{h}}^{(i)}\right|\right|_2}$ form a database with known labels.

Classification of new heartbeats

After construction of the model, a new (wavelet-transformed) representative heartbeat $\mathbf{X}_{\mathrm{r}}^{(\text{test})}$ can be classified similarly to the method explained in Section 7.3.1. Algorithm 2 shows how to calculate the morphological feature vector $\mathbf{f}^{(\text{test})} \in \mathbb{R}^C$, with C the number of heartbeat classes. The representative heartbeat (decomposed in different scales) is projected onto the new heartbeat basis. The steps after this projection are identical to Algorithm 1. An SVM can then use the morphological feature values as input for further classification.

7.3.3 MLSVD-based detection for multilead ECGs

In a clinical setting, ECG is usually measured with multiple leads. This section therefore describes a method to detect AF in multilead ECG signals using the same principles as the previous methods. Another difference is that here, several approaches to perform the final classification are considered. Since multilead ECG signals can be tensorized in a natural way, no additional tensorization step is required, making the multilead approach similar to the method described in Section 7.3.1. A last variant describes an approach to classify single lead ECGs using a database of multilead ECG. This *coupling* can be useful when an annotated multilead database is available (for example after a clinical study), which can then be used with signals measured by wearable devices.

Algorithm 2 Calculation of morphological features using MLSVD

Input: New representative heartbeat $\mathbf{x}_{\mathrm{r}}^{(\mathrm{test})} \in \mathbb{R}^N$, annotated model tensor $\mathcal{D}_{\mathrm{model}} \in \mathbb{R}^{J \times N \times M}$, multilinear rank $(r_{\mathrm{scale}}, r_{\mathrm{time}}, r_{\mathrm{heartbeat}})$, weights $\mathbf{w} \in \mathbb{R}^W$
Output: Vector with morphological features $\mathbf{f}^{(\mathrm{test})} \in \mathbb{R}^C$, with C the number of classes

1: $\mathbf{X}_{\mathrm{r}}^{(\mathrm{test})} \leftarrow \mathrm{RWT}(\mathbf{x}_{\mathrm{r}}^{(\mathrm{test})}) \in \mathbb{R}^{J \times N}$, with J levels and the `bior4.4`-wavelet
2: Calculate the truncated MLSVD with multilinear rank $(r_{\mathrm{scale}}, r_{\mathrm{time}}, r_{\mathrm{heartbeat}})$:

$$\mathcal{D}_{\mathrm{model}} \approx \hat{\mathcal{S}} \cdot_1 \hat{\mathbf{U}}_{\mathrm{scale}} \cdot_2 \hat{\mathbf{U}}_{\mathrm{time}} \cdot_3 \hat{\mathbf{U}}_{\mathrm{heartbeat}}$$

3: Construct database $\{\tilde{\mathbf{c}}_{\mathrm{h}}^{(i)}\}$, with $\tilde{\mathbf{c}}_{\mathrm{h}}^{(i)} = \mathrm{sign}(\mathbf{c}_{\mathrm{h},1}^{(i)}) \frac{\mathbf{c}_{\mathrm{h}}^{(i)}}{\left|\left|\mathbf{c}_{\mathrm{h}}^{(i)}\right|\right|_2}$, of normalised rows $\mathbf{c}_{\mathrm{h}}^{(i)^{\mathrm{T}}}$ rom $\hat{\mathbf{U}}_{\mathrm{heartbeat}}, \forall i \in \{1, \ldots, M\}$
4: Construct basis matrix $\mathbf{B} = (\hat{\mathbf{U}}_{\mathrm{time}} \otimes \hat{\mathbf{U}}_{\mathrm{scale}})\hat{\mathbf{S}}_3^{\mathrm{T}}$
5: Solve

$$\mathrm{vec}\left(\mathbf{X}_{\mathrm{r}}^{(\mathrm{test})}\right) = \mathbf{B}\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}$$

in a least-squares way
6: Normalise: $\tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})} \leftarrow \mathrm{sign}(\mathbf{c}_{\mathrm{h},1}^{(\mathrm{test})}) \frac{\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}}{\left|\left|\mathbf{c}_{\mathrm{h}}^{(\mathrm{test})}\right|\right|_2}$
7: **for** $i = 1 \ldots M$ **do**
8: Calculate similarity s_i:

$$s_i = \tilde{\mathbf{c}}_{\mathrm{h}}^{(\mathrm{test})^{\mathrm{T}}} \tilde{\mathbf{c}}_{\mathrm{h}}^{(i)}$$

9: **end for**
10: **for** $c = 1 \ldots C$ **do**
11: Calculate $\mathbf{f}_c^{(\mathrm{test})}$ as weighted average $\mathbf{w}$ of the best W scores s_i, for $i \in C_c$, the set of labels c
12: **end for**
13: **return** $\mathbf{f}^{(\mathrm{test})}$

Model construction

A model set of multichannel ECG signals is used as basis for the model. As in the previous methods, the ECG signals are first preprocessed channel-by-channel to improve the signal quality. They are then compressed into one representative heartbeat per channel. Since a multilead ECG signal naturally has a second-order structure, stacking the representative beats as frontal slices of a model tensor $\mathcal{D}_{\text{model}}$ automatically leads to a tensorisation which preserves the structural information between all channels. $\mathcal{D} \in \mathbb{R}^{n_k \times N \times M}$ has modes channels $\times$ time $\times$ heartbeats, with n_k the number of leads per signal, N the number of samples per representative heartbeat and M the number of signals in the model set. The truncated MLSVD can again be used to calculate a low multilinear rank approximation to obtain the following model:

$$\mathcal{D}_{\text{model}} \approx \hat{\mathcal{S}} \cdot_1 \hat{\mathbf{U}}_{\text{channel}} \cdot_2 \hat{\mathbf{U}}_{\text{time}} \cdot_3 \hat{\mathbf{U}}_{\text{heartbeat}}, \tag{7.10}$$

where $\hat{\mathbf{U}}_{\text{channel}} \in \mathbb{R}^{n_k \times r_{\text{channel}}} (r_{\text{channel}} < n_k)$, $\hat{\mathbf{U}}_{\text{time}} \in \mathbb{R}^{N \times r_{\text{time}}} (r_{\text{time}} < N)$ and $\hat{\mathbf{U}}_{\text{heartbeat}} \in \mathbb{R}^{M \times r_{\text{heartbeat}}} (r_{\text{heartbeat}} < M)$ form an orthonormal basis for the corresponding subspaces. The optimal multilinear rank is determined using cross-validation. Each multilead ECG signal with k channels and thus k representative heartbeats $\mathbf{X}_{\text{r}}^{(i)} \in \mathbb{R}^{n_k \times N}$ can then be written as

$$\text{vec}\left(\mathbf{X}_{\text{r}}^{(i)}\right) \approx \mathbf{B}\mathbf{c}_{\text{h}}^{(i)},$$

with $\mathbf{B} = \mathbf{B}_{(3)}^{\text{T}} \in \mathbb{R}^{n_k N \times r_{\text{heartbeat}}}$ the mode-3 unfolding of the tensor $\mathcal{B} = \hat{\mathcal{S}} \cdot_1 \hat{\mathbf{U}}_{\text{channel}} \cdot_2 \hat{\mathbf{U}}_{\text{time}}$:

$$\mathbf{B} = (\hat{\mathbf{U}}_{\text{time}} \otimes \hat{\mathbf{U}}_{\text{channel}})\hat{\mathbf{S}}_{(3)}^{\text{T}}.$$

The columns of $\mathbf{B}$ then form a basis for all channels. The coefficient vector expresses a multilead ECG signal in all channels simultaneously.

The normalised coefficient vectors $\{\tilde{\mathbf{c}}_{\text{h}}^{(i)}\}$, with $\tilde{\mathbf{c}}_{\text{h}}^{(i)} = \text{sign}(\mathbf{c}_{\text{h},1}^{(i)}) \frac{\mathbf{c}_{\text{h}}^{(i)}}{\left|\left|\mathbf{c}_{\text{h}}^{(i)}\right|\right|_2}$, make up a labelled database which can be used for classification.

Classification of new heartbeats

The classification of an unseen multilead test signal is done in a similar way as described in Algorithm 2, but $\mathbf{X}_{\text{r}}^{(\text{test})} \in \mathbb{R}^{n_k \times N}$ now follows directly from the calculation of the representative heartbeats. The coefficients vector $\mathbf{c}_{\text{h}}^{(\text{test})}$ can

be calculated by solving the following linear system:

$$\text{vec}\left(\mathbf{X}_{\text{r}}^{(\text{test})}\right) = \mathbf{B}\mathbf{c}_{\text{h}}^{(\text{test})}. \tag{7.11}$$

Solution per channel In (7.11), the coefficient vector $\mathbf{c}_{\text{h}}^{(\text{test})}$ is calculated coupled over all channels. Another possibility is to calculate an individual coefficient vector for each channel. For representative heartbeat $\mathbf{x}_{\text{r}}^{(i,k)}$ in channel $k \in \{1, \ldots, n_{\text{k}}\}$ (the k^{th} row from $\mathbf{X}_{\text{r}}^{(i)}$), we can write (using Equation 7.10):

$$\begin{aligned} \mathbf{x}_{\text{r}}^{(i,k)} &\approx \hat{\mathcal{S}} \cdot_1 \mathbf{c}_{\text{k}}^{(k)^{\text{T}}} \cdot_2 \hat{\mathbf{U}}_{\text{time}} \cdot_3 \mathbf{c}_{\text{h}}^{(i)^{\text{T}}} \\ \Leftrightarrow \mathbf{x}_{\text{r}}^{(i,k)^{\text{T}}} &\approx \underbrace{(\hat{\mathbf{U}}_{\text{time}} \otimes \mathbf{c}_{\text{k}}^{(k)^{\text{T}}})\hat{\mathbf{S}}_{(3)}^{\text{T}}}_{\mathbf{B}_k} \mathbf{c}_{\text{h}}^{(i)}, \end{aligned} \tag{7.12}$$

with $\mathbf{c}_{\text{k}}^{(k)^{\text{T}}} \in \mathbb{R}^{r_{\text{channel}}}$ the k^{th} row from $\hat{\mathbf{U}}_{\text{channel}}$.

The solution of (7.12) results after comparison of each vector $\tilde{\mathbf{c}}_{\text{h}}^{(\text{test},k)}$ with the database, in a new vector of morphological features $\mathbf{f}^{(\text{test})} \in \mathbb{R}^{n_{\text{k}}C}$. This vector represents the morphological similarity of each lead in the test signal with each class. It is however expected that these features are less informative than the features obtained by Equation 7.11 since the coefficient vectors in the database are calculated for all channels simultaneously. It can even be seen as a step backwards towards the matrix-case, where the combination of information from all channels is only done in a later phase (*late integration*), failing to exploit the structure present in the data. We therefore do not consider this an appropriate solution here. It will however be picked back up in the next Section to couple single- and multilead recordings.

SVD-based variant The modelling of the data using the MLSVD results in an *independent* modelling of the time- and channel-mode. It is however also possible to use the mode-3 unfolding $\mathbf{D}_{\text{model},(3)} \in \mathbb{R}^{M \times n_{\text{k}}N}$ to perform the modelling with the truncated SVD, calculating a *simultaneous* model of the time and channel space:

$$\mathbf{D}_{\text{model},(3)}^{\text{T}} \approx \hat{\mathbf{U}}_{\text{tk}}\hat{\mathbf{S}}\hat{\mathbf{U}}_{\text{heartbeat}}^{\text{T}},$$

with $\hat{\mathbf{U}}_{\text{tk}} \in \mathbb{R}^{n_{\text{k}}N \times r}, \hat{\mathbf{U}}_{\text{heartbeat}} \in \mathbb{R}^{M \times r}$ and $\hat{\mathbf{S}} \in \mathbb{R}^{r \times r}$. Algorithm 1 can then classify a new ECG signal after concatenation of each channel (vec $\left(\mathbf{X}_{\text{r}}^{(\text{test})}\right)$).

Comparison of the results using the MLSVD and SVD-approach gives information how much performance is added by modelling the channel space individually, and quantifies the added value of tensor methods in this application.

Combination of single- and multilead ECG signals

Single-lead ECG signals are rapidly gaining importance in the context of digital health since they can be more easily recorded with wearable devices. Contrary to multilead measurements they however do not contain any spatial information, and only reflect the cardiac electrical activity from one single angle. This section briefly describes a framework to detect AF in single lead channels using a model set obtained from a multilead dataset. The combination of datasets allows the use of a larger amount of data, leading to more reliable and better classification methods. One can also imagine scenarios where an annotated dataset has been collected and is supplemented with unannotated single lead recordings. The proposed method allows to combine these datasets without requiring extra labelling effort.

In general, the channel of the single lead ECG signal is known beforehand. In the KardiaMobile signals of the Physionet database for example, measurements are done between the left and right hand, corresponding to lead I. In that case, the procedure of Section 7.3.3 (*Solution per channel*) can be used to classify the representative heartbeat of a single lead ECG signal. The model set of multilead signals then forms the corresponding basis matrix $\mathbf{B}_k$ from (7.12), with the correct channel index k.

The methods described in the previous Sections use model and test signals measured with similar devices in a similar context, which makes a comparison between coefficient vectors feasible. When combining datasets with signals from different devices, this comparison is less meaningful since there is an increased chance that measurement circumstances change: differently-placed electrodes, technical errors or inverted leads are all possible. When there is uncertainty about the lead, comparing the signal with a labelled database is less meaningful, and we therefore propose to use the coefficient vector directly as a feature vector: $\mathbf{f} = \tilde{\mathbf{c}}_{\mathrm{h}}^{(k)} \in \mathbb{R}^{r_{\mathrm{heartbeat}}}$.

7.3.4 Combination of morphological and HRV characteristics

For the final method, the morphological features derived with one of the methods described in the previous Sections are combined with features based on heart rate variability. Figure 7.5 shows how the integration between the (ML)SVD-

based detection and the HRV characteristics leads to a global method to detect AF in single- or multilead ECG signals. The HRV characteristics used are:

- **AVRR**: Average RR interval
- **RMSSD**: Root Mean Square of the differences between subsequent RR intervals (ΔRR)
- **pRR50**: Fraction of subsequent RR intervals that differ more than 50ms
- **Cov(ΔRR)**: Coefficient of variance of the ΔRR intervals
- **min(RR)**: Minimal RR interval
- **HTI**: HRV Triangular Index, which characterizes the concentration of the RR interval histogram
- **SD1/SD2 ratio**: Characterizes the Poincaré plot of RR intervals
- **AFEvidence**: Characterizes the Poincaré plot of ΔRR intervals
- **ApEn**: Approximate entropy, which is equivalent to the complexity of the signal
- **Toeplitz distance**: Quantifies the uniformity of RR intervals

A more extensive explanation of the HRV features and their calculation can be found in Appendix A.

Figure 7.5 shows a flowchart of the global method which integrates the ML(SVD)-based methods and the HRV features with a SVM. The full method determines the class of a new test signal using the combined feature vector. Depending on the database, this can be a binary or multiclass classification problem. Many databases define a 'rest' class, which contains all signals that do not show sinus rhythm or AF. They can be signals with other arrhythmia or signals where no consensus on the underlying rhythm could be reached. The multiclass classification problem is solved using a one-versus-one approach. Optimization of hyperparameters is done with ten-fold cross validation.

The SVM is trained using the training sets defined in both databases. The training sets contain signals that are not included in the model set (used to determine the morphological features) or the test set (used to obtain the final classification results). To optimize classification performance, different SVM kernels were tested and compared.

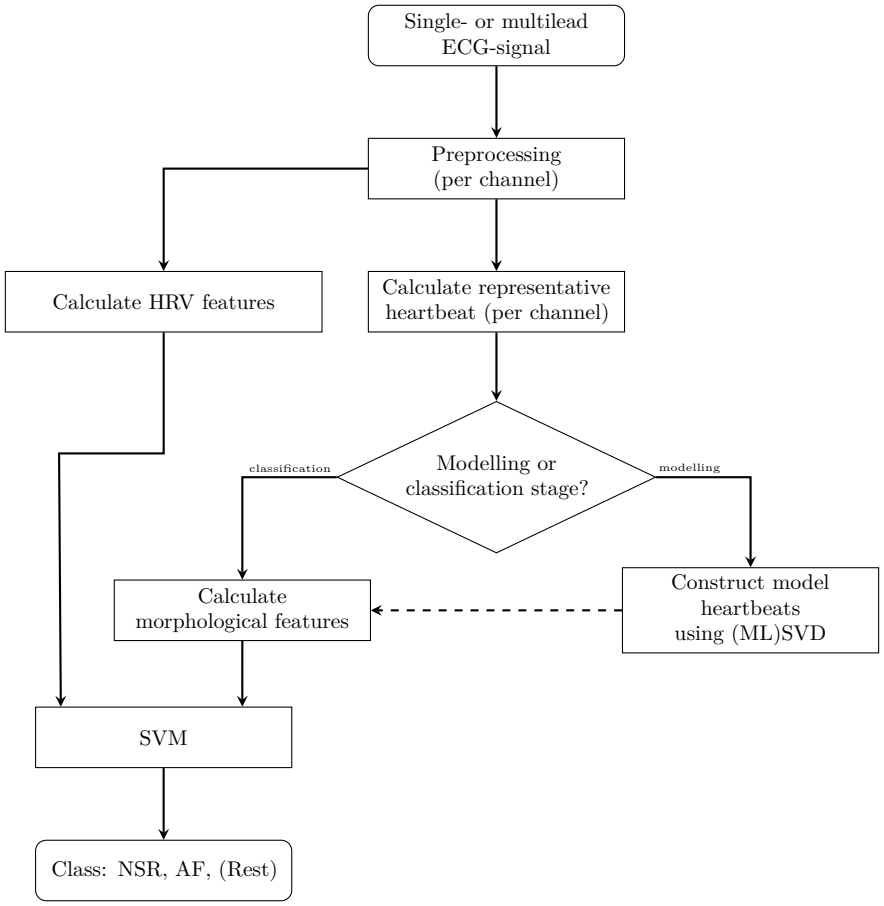


Figure 7.5: Flowchart of the global AF detection method. The method integrates the HRV features described in Appendix A with the morphological features obtained using (ML)SVD in a SVM.

7.3.5 Evaluation of results

After the modelling and training stage, the performance of the algorithm is determined by applying the method on an independent test set. Dependent on the dataset used, the methods either perform binary or multiclass classification: the Physionet/CinC dataset has 3 classes (normal, AF and other), the multilead database has only two classes (normal and AF). To evaluate the results, we will mainly use the F_1-score and AUC value as defined in Section 3.4. The F_1 score is first calculated for each class individually, resulting in F_{1n}, F_{1a}, F_{1n}, for respectively the normal class, AF class and rest class. The global F_1-score is then defined as the average of the individual scores of the different classes. The AUC value is only used to evaluate the results of binary classification problems, and can thus not be used in the single-lead dataset.

7.4 Results and Discussion

First, the results for AF detection in single lead ECG signals are summarized and discussed. Both the SVD- and MLSVD-based approaches are evaluated and compared. Furthermore, the influence of parameters such as the size of the model set is examined. The results for multilead signals and the coupling between single- and multilead are described next.

7.4.1 AF detection in single lead ECG signals

Table 7.3 gives a summary of the results on the CinC/Physionet dataset, obtained with weight vector $\mathbf{w} \in \mathbb{R}^{15}$ (from Algorithm 1) equal to:

$$\begin{bmatrix} 0.3478 & 0.1739 & 0.1304 & 0.0870 & 0.0435 & 0.0217 & 0.0217 & 0.0217 & 0.0217 & 0.0217 & 0.0217 & 0.0217 & 0.0217 & 0.0217 & 0.0217 \end{bmatrix}^{\mathrm{T}}.$$

SVD-based detection

Figure 7.6a shows the singular values of the model matrix $\mathbf{D}_{\text{model}} \in \mathbb{R}^{176\times 4946}$, which were used to get a first estimate for an appropriate truncation rank. The singular values decrease quickly, as predicted in Section 7.3.1. Based on this Figure, the search interval for the rank was defined as $[1, 30]$, since these singular values contain most of the energy. After 10-fold cross-validation, the optimal rank was found to be 22. Figure 7.6b shows the evolution of the cross-validation performance in function of the rank r. When the rank is too small (≤ 10), the

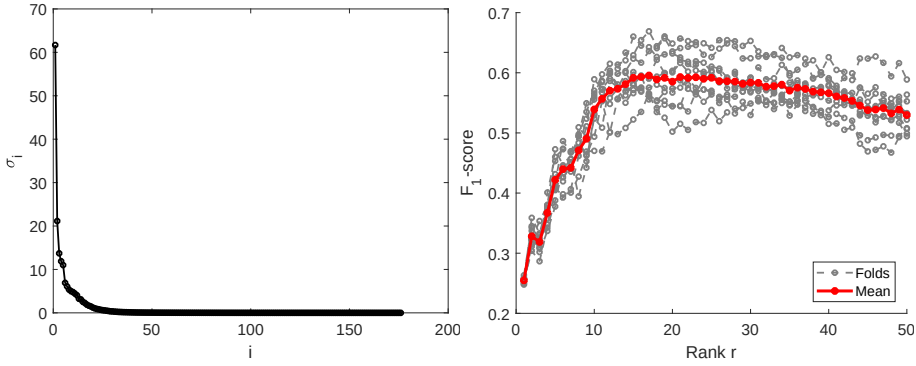
Method	Kernel	$P(\%)$	F_{1n}, F_{1a}, F_{1r}	F_1
SVD, without HRV	Linear	70.02	0.813,0.570,0.396	0.593
SVD, without HRV	RBF	69.60	0.812,0.561,0.381	0.585
SVD, without HRV	Polynomial	69.78	0.812,0.563,0.389	0.588
MLSVD, without HRV	Linear	70.45	0.815,0.585,0.404	0.601
MLSVD, without HRV	RBF	70.75	0.815,0.590,0.402	0.602
MLSVD, without HRV	Polynomial	70.75	0.817,0.589,0.407	0.604
SVD + HRV	RBF	80.22	0.866,0.796,0.647	**0.770**
MLSVD + HRV	RBF	79.85	0.865,0.796,0.640	0.767
Only HRV	RBF	77.67	0.853,0.760,0.585	0.732
Only HRV	Polynomial	74.70	0.840,0.673,0.529	0.681

Table 7.3: Results for the different methods on the Physionet/CinC Challenge 2017 dataset. The MLSVD has an optimal multilinear rank of $(16, 3, 14)$ while the SVD has rank 22.

model does not contain all relevant information and is thus underfitted. The performance is maximal for $15 \leq r \leq 30$. When the rank increases further, the performance starts to drop again due to overfitting.

The influence of the size of the model set is also evaluated. Figure 7.7 shows how the results for the different classes change when changing the number of signals used to construct the SVD model. For each size of model set, a random model set was selected and results were tested on the remainder of the dataset. Each experiment was repeated 20 times with different random sets, the average over all runs is shown on the Figure. The rank is fixed to 22. The largest increase in performance is seen for the class of AF signals. This is not coincidentally the class with the least number of signals: more AF signals lead to a better model and better performance. For other classes the results increase less, since they already have a considerable number of signals in the model set to begin with. Figure 7.7 confirms that the initial choice for size of model set (which was chosen based on the standard 60/40 distribution for training/test set) forms a good trade-off between size of model set and performance.

After construction of the model, the morphological features are calculated for both the training and test set Figure 7.8 plots the two-dimensional projections of $\mathbf{f} \in \mathbb{R}^3$ for each combination of classes. From the Figure it is clear that there is a clear separation between the normal and AF classes on one hand and AF and rest classes on the other hand. The classes of normal and rest signals are



(a) The singular values of the model matrix.

(b) The F_1-scores in the different folds of the cross-validation, with the mean value in red. There is clear underfitting when the rank is too low. Vice versa, there is overfitting when the rank increases too much.

Figure 7.6: The singular values of the model matrix and results of cross-validation on the model set for different ranks r [83].

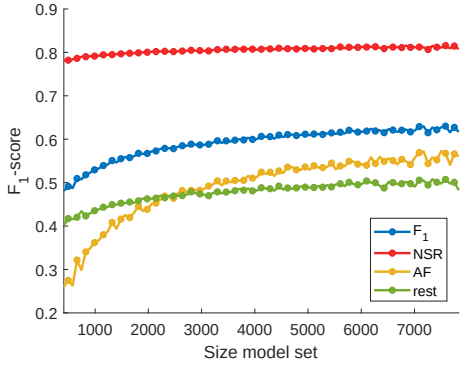


Figure 7.7: The F_1-scores for different sizes of model set for the SVD-based method. Both the overall F_1-score and the scores for the individual classes are shown. The performance of the AF class depends more on the model set size than the performance of the other classes [83].

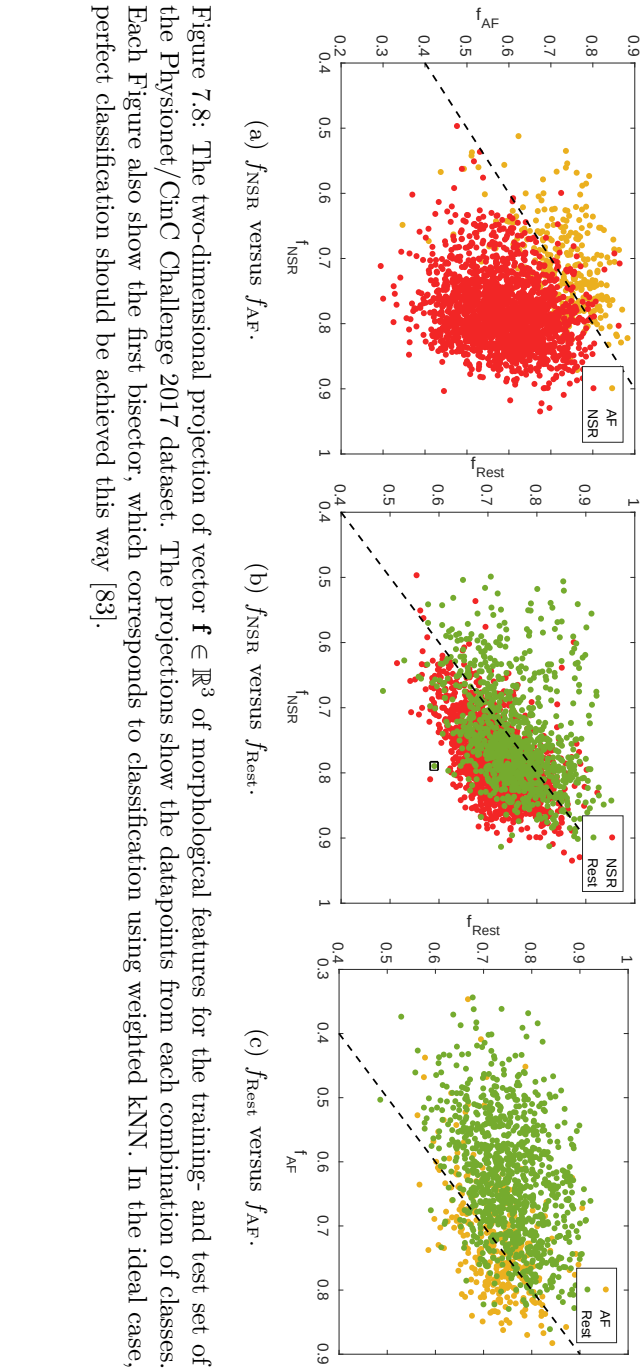
		True label		
		NSR	AF	Rest
Predicted label	NSR	1755	18	215
	AF	94	145	86
	Rest	521	49	415

Table 7.4: Confusion matrix of the best-scoring SVD-based method, combining morphological and HRV features ($F_1 = 0.770$).

however strongly overlapping: many signals in the rest category show a large (morphological) resemblance to normal signals. To illustrate this, Figure 7.9 shows an example of such signal from the rest class that resembles a normal signal. It corresponds to the circled data point on Figure 7.8b.

Table 7.3 indicates that, without using HRV features, the SVM with linear kernel obtains the best results with an F_1 score of 0.593. Changing the kernel does not seem to have a significant influence on the results. Adding HRV features however improves the results drastically, since the features were selected to discriminate AF signals (see Appendix A for a full analysis): Using *only* HRV features, a global F_1 score of 0.732 can be obtained. The optimal score overall ($F_1 = 0.77$) is achieved by combination of the morphological and HRV features. This confirms that the proposed method does have a significant added value for identification of abnormal rhythms. Note that the highest increase in performance is accomplished in the rest class. The normal class has the best classification performance for all classifiers.

Table 7.4 shows the confusion matrix, with the predicted labels versus the true labels for all classes for the SVD-based classification method. The number of AF signals that are classified as normal is minimal. This is clinically relevant, since the argument can be made that these false negatives (e.g. AF signals that are missed) are more dangerous than normal signals that are classified as abnormal. As expected, most missclassifications are seen between the normal and rest class. The HRV features used here were specifically selected to distinguish AF signals and Figure 7.8 already showed that also the morphological features were more discriminative for AF and normal signals. It could be helpful to divide this rest class further in more specific subclasses to obtain a better performance. Furthermore it should be mentioned that the rest class does not only contain signals with a different cardiac rhythm, but also signals where no consensus could be reached between different raters. This means that this class most likely also contains normal and AF signals, adding further to the difficulty of classifying them correctly.



(a) f_{NSR} versus f_{AF}. (b) f_{NSR} versus f_{Rest}. (c) f_{Rest} versus f_{AF}.

Figure 7.8: The two-dimensional projection of vector $\mathbf{f} \in \mathbb{R}^3$ of morphological features for the training- and test set of the Physionet/CinC Challenge 2017 dataset. The projections show the datapoints from each combination of classes. Each Figure also show the first bisector, which corresponds to classification using weighted kNN. In the ideal case, perfect classification should be achieved this way [83].

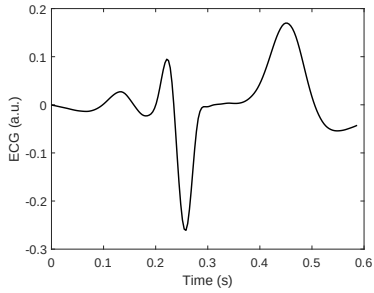
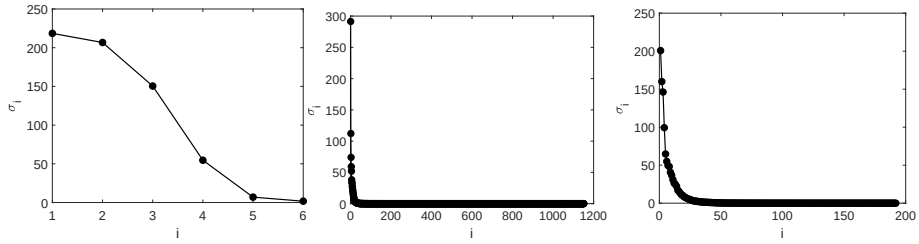


Figure 7.9: An example of a representative heartbeat that strongly resembles a normal beat, corresponding to the circled datapoint in Figure 7.8. The large resemblance with a normal beat is probably due to the presence of a clear P- and T-wave, resulting in $f_{\text{NSR}} > f_{\text{Rest}}$ [83].



(a) Mode-1 singular values (scale).

(b) Mode-2 singular values (time).

(c) Mode-3 singular values (heartbeat).

Figure 7.10: The multilinear singular values of the model tensor $\mathcal{D}_{\text{model}}$, after tensorization with the RWT [83].

MLSVD-based detection

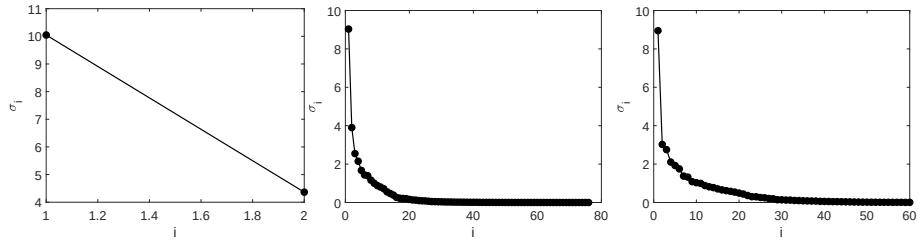
Similarly to Figure 7.6a, Figure 7.10 shows the multilinear singular values of the model tensor $\mathcal{D}_{\text{model}} \in \mathbb{R}^{6\times 176\times 4946}$. As predicted in Section 7.3.2, the rank of the scale mode is lower than six. After sequential 10-fold cross-validation, the optimal multilinear rank is fixed to $(16, 3, 14)$.

Table 7.3 indicates that the MLSVD-based approach does not have any added value compared to the SVD approach for single lead ECG signals. Without integration of the HRV features, the performance seems slightly better but the increase is negligible. The results did not improve when changing the multilinear rank. Adding the HRV features causes the MLSVD performance to drop under the SVD performance; especially for the rest class the SVD approach seems to work better. The wavelet-based tensorization thus does not achieve the desired

effect: for this application, time-frequency analysis does not add any useful information to the representative heartbeat. Therefore we can conclude that, to detect AF in single lead signals, the method that combines morphological features obtained by the SVD approach with the HRV features is preferred over a multilinear method. While the results are very comparable, calculating the RWT and MLSVD is computationally more expensive, making it less suitable for long-term signals or real-time applications.

As explained in Section 7.2.1, the single lead dataset was collected in the context of the 2017 Computing in Cardiology/Physionet Challenge, an annual ECG processing competition. The best algorithms developed during the competition obtained a global F_1-score of 0.83 [138, 240, 209, 102, 54] while the worst performance had an F_1-score of 0.25 [46]. The proposed method obtains a result of 0.76, which would have lead to a shared 38^{th} place out of 67 official participants. It is important to note that a large part of the top-scoring methods used deep learning methods for classification [209, 102, 54]. While obviously powerful, they have one big disadvantage: since the features are learned from the data, they are rarely interpretable, which is important when communicating the results with both cardiologists and patients. While the SVM used here is also considered a *black-box* method, both the morphology and HRV features have physiological meaning.

Another important remark is that the comparison is done based on results from a different test set. The test set used during the competition was hidden and has not been made publicly available. Using a separate test set would mean that the full public dataset of 8528 could be used to model the dataset and train the SVM. Per Figure 7.7, this could mean an extra increase in performance. Finally, the best result achieved on this dataset was a global F_1-score of 0.83. This performance can be seen as the best possible performance, since combining the results of different top-scoring methods did not increase performance further, according to the challenge organisers. Since annotation is done manually, errors in the labelling cannot be avoided especially taking into account the size of the dataset used here. This is especially harmful for supervised methods, where errors in the training set can seep through to the test set. It effectively puts an upper limit to the maximal detection performance meaning that methods that obtain significantly higher scores are probably due to overfitting. It is an important factor to keep into account when evaluating the performance on manually annotated datasets.



(a) Mode-1 singular values (channel).

(b) Mode-2 singular values (time).

(c) Mode-3 singular values (heartbeat).

Figure 7.11: The multilinear singular values of the model tensor $\mathcal{D}_{\text{model}} \in \mathbb{R}^{2\times 76\times 60}$ of the MIT-BIH AFIB & AFTDB dataset [83].

7.4.2 AF detection in multilead ECG signals

In this Section, the results of the methods described in Section 7.3.3 are evaluated on a dataset of multilead ECG signals.

Table 7.5 shows the results on the combined MIT-BIH AFIB & AFTDB dataset, obtained with the following weight vector:

$$\begin{bmatrix} 0.3902 & 0.1951 & 0.1463 & 0.0976 & 0.0488 & 0.0244 & 0.0244 & 0.0244 & 0.0244 & 0.0244 \end{bmatrix}^{\mathrm{T}}$$

Section 7.4.2 described three different ways to classify a new signal: the first method searches for a coupled coefficient vector, the second method finds a solution per channel and the final variant uses an SVD-variant to the algorithm

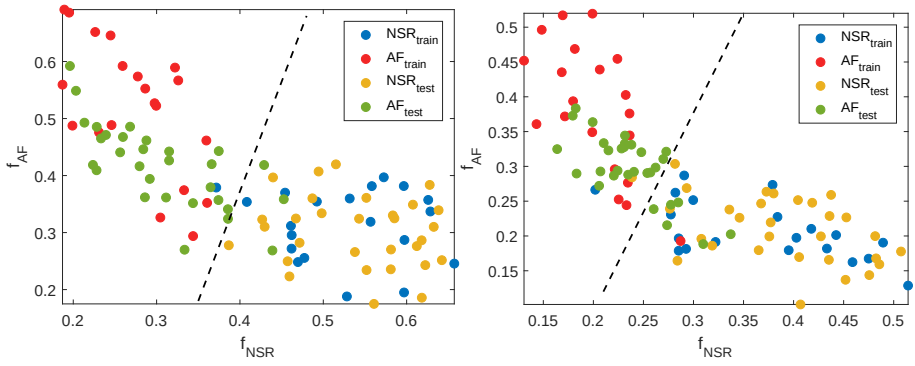
MLSVD, coupled over all channels

Prior to MLSVD, the multilinear rank of the decomposition has to be defined. This is again done with 20-fold cross validation. The multilinear values of the model tensor are shown in Figure 7.11 for all three modes. The optimal rank is estimated to be $(1, 23, 23)$.

The results in Table 7.5 indicate that only using the morphological features $\mathbf{f} \in \mathbb{R}^2$ a good discrimination is possible between normal and AF signals. Figure 7.12a shows the features for all signals in both the training and test set together with the SVM decision boundary. Two clusters can be clearly distinguished: the samples in the upper left part of the plot correspond with signals which show a large morphological similarity with AF signals ($f_{\text{AF}} > f_{\text{normal}}$) while the points in the lower right part show a larger correspondence

Method	Kernel	$P(\%)$	AUC	F_{1n}, F_{1a}	F_1
MLSVD, coupled, without HRV	Linear	93.33	0.990	0.935,0.931	0.933
SVD, without HRV	Linear	85.00	0.969	0.866,0.830	0.848
MLSVD, per channel (all), without HRV	Linear	86.67	0.970	0.882,0.846	0.864
MLSVD, per channel (channel 1), without HRV	Linear	86.67	0.911	0.879,0.852	0.865
MLSVD, per channel (channel 2), without HRV	Linear	76.67	0.911	0.806,0.708	0.757
MLSVD, coupled, with integration HRV	RBF	100	1.000	1.000,1.000	1.000
Only HRV	RBF	98.33	1.000	0.984,0.983	0.983

Table 7.5: Results for the different multilead methods on the MIT-BIH AFIB & AFTDB dataset. The MLSVD has optimal rank $(23, 1, 23)$, while the SVD-based variant uses rank 43 [83].



(a) MLSVD-base. There is only one error in the training set (1 FP detection) and three errors in the test set (3 FN detections).

(b) SVD-based variant. The classifier wrongly labels two datapoints (1 FP, 1 FN) in the training set and seven (1 FP, 6FN) in the test set.

Figure 7.12: The morphological features $\mathbf{f} \in \mathbb{R}^2$ for the training- and test set of the MIT-BIH AFIB & AFTDB dataset, for the SVD- and MLSVD-based approaches. The linear decision boundary obtained by the SVM is shown as well [83].

with normal signals ($f_{\text{normal}} > f_{\text{AF}}$). The linear decision boundary is capable of separating these points: there is only one FP detection in the training set and three FNs in the test set.

Adding the HRV features results in a perfect classification: the two classes are completely separable. Note that only using HRV features the AUC is also 1, meaning that this feature set is very informative to separate normal and AF signals. While the morphological features thus not add extra performance per se, they do contain a lot of information. This is confirmed by training a linear SVM to the complete feature set (morphological + HRV) and analysing the weights of the different features:

$$\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} f_{\text{NSR}} & f_{\text{AF}} & \text{AVRR} & \text{RMSSD} & \text{pRR50} & \text{HTI} & \min(\text{RR}) & \text{SD1/SD2} & \text{ApEn} & \text{Toeplitz} & \text{CoV}(\Delta_{\text{RR}}) & \text{AFEvidence} \\ -0.5466 & 0.4227 & -0.2158 & 0.2619 & 0.4972 & 0.4466 & -0.1354 & -0.1181 & 0.4549 & 0.2688 & 0.2871 & 0.5059 \end{bmatrix},$$

The coefficients corresponding to the morphological features have the highest value, and are thus most important to verify whether AF is present or not. Note that the other coefficients are consistent with Figure A.2: large values for AVRR, min(RR) or SD1/SD2-ratio indicate normal signals while the other features imply the presence of atrial fibrillation.

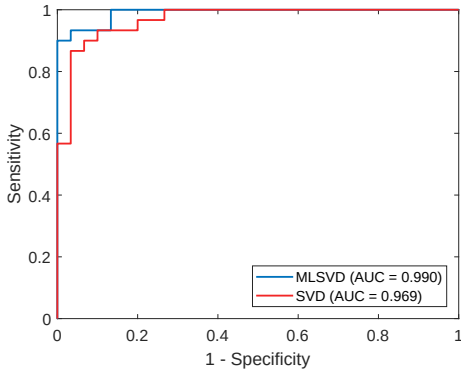


Figure 7.13: The ROC curves for the results of the linear SVM on MIT-BIH AFIB & AFTDB test set, for the SVD- and MLSVD-based approach. The MLSVD obtains a higher specificity for the same sensitivity and has a larger area under the curve [83].

MLSVD-based, solution per channel

The coefficient vector can also be determined for each channel individually (see Section 7.3.3). Table 7.5 confirms the hypothesis from Section 7.3.3 that the performance is lower compared to the coupled coefficient vector. The first channel seems to be more informative than the second channel when visually comparing the ROC curves, even though they have the same AUC value.

SVD-based solution

Figure 7.12b shows the morphological features $\mathbf{f} \in \mathbb{R}^2$ and linear decision boundary after calculating the coefficient vector with the SVD-based approach (see Section 7.3.3). Visually we can already determine that the two clusters are more difficult to separate: there is more overlap between the classes in the area around the decision boundary. Table 7.5 and Figure 7.13 validate this confirmation. The AUC is smaller and while the sensitivity is the same, the specificity is smaller for the MLSVD-based method. This confirms the utility of using multilinear methods for AF classification for multilead ECG signals.

Set	Class		
	NSR	AF	Total
Model set (multilead)	80	80	160
Training set (single lead)	3030	443	3473
Test set (single lead)	2020	295	**2315**

Table 7.6: Data profile of the model, training and test set of combination of the Physionet/CinC Challenge 2017 and the MIT-BIH AFIB & AFTDB dataset.

7.4.3 Detection of AF in single lead ECG signals in combination with multilead ECG signals

Data

To evaluate the performance of the combination of single- and multilead signals, the model set was constructed using 160 two-channel ECG signals from the MIT-BIH AFIB& AFTDB dataset (see Table 7.2), while the training and test set consist of the signals from the Physionet/CinC Challenge dataset, without the rest class (since this class is not present in the model set, and can thus not be classified). Table 7.6 shows the number of signals in each set.

Results and discussion

The feature vector is now equal to the normalised coefficient vector: rhythms are not defined by the similarity with signals from the model set, but directly by the coefficients. The multilinear rang is still equal to $(1, 23, 23)$. This implies that the feature vector is now a 23-dimensional vector: $\mathbf{f} = \tilde{\mathbf{c}}_{\mathrm{h}}^{(k)} \in \mathbb{R}^{23}$. Since the training set is sufficiently large, this is however not a problem. Adding the HRV features results in a total feature vector of length 33.
Table 7.7 shows the results on the independent test set, after training the SVM. Using only the morphological coefficient vector, a strong discrimination between normal and AF signals is possible.

It is remarkable that combining the HRV and morphological features leads to better results than solely using the HRV features. Deeper analysis of the results show multiple signals which are wrongly classified using only HRV features. Figure 7.15 shows an example of the tachogram and representative heartbeat of such signal: using only the HRV features, the signal is classified as normal while

Method	Kernel	$P(\%)$	AUC	F_{1n}, F_{1v}	F_1
MLSVD	RBF	93.56	0.955	0.964,0.707	0.836
MLSVD, HRV	RBF	97.75	0.991	0.987,0.908	**0.948**
HRV	RBF	97.02	0.974	0.983,0.875	0.929

Table 7.7: The results for different methods on the combination of the Physionet/CinC 2017 Challenge dataset and MIT-BIH AFIB & AFTDB dataset.

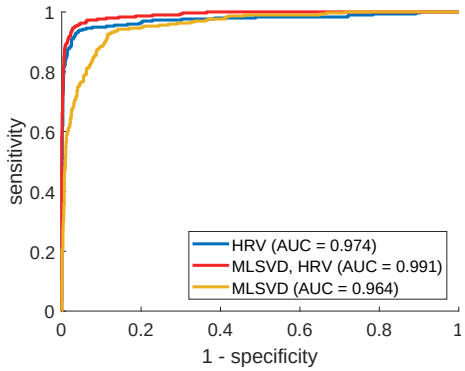
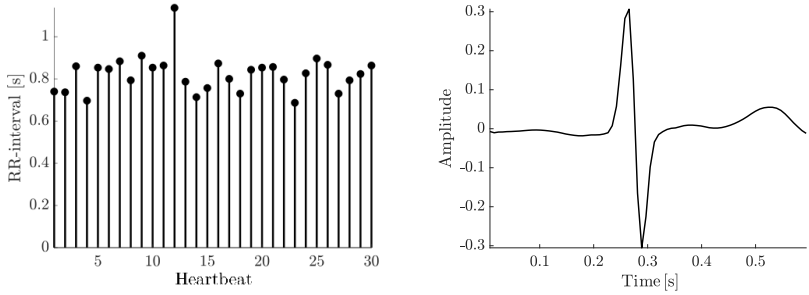


Figure 7.14: ROC curves for the SVMs on the combination of the Physionet/CinC Challenge 2017 test set and the MIT-BIH AFIB & AFTDB dataset.

the true label is AF. Adding the morphological features corrects the classification. The tachogram in Figure 7.15a shows some irregularities but not enough to be classified as AF. Figure 7.15b shows the representative heartbeat, where the P wave is clearly absent. The morphological features are thus decisive in this case. Other classification errors using HRV features are the result of incorrect R peak detection (such as for example cases where an R peak is missed). The morphological features are able to correct these errors. The combination of the morphological and HRV features results in a global F_1-score of 0.948 with AUC of 0.991. This is better than the HRV features and morphological features individually. The ROC-curves of the three classifiers are shown in Figure 7.14.



(a) The tachogram. (b) The representative heartbeat.

Figure 7.15: The tachogram and representative heartbeat of a signal that was wrongly classified based on the HRV, but correctly classified when adding morphology information [83].

7.5 Conclusion

This Chapter presented a series of methods to detect atrial fibrillation from single and multilead ECG signals. While the use of tensors did not provide added value for AF detection in single lead signals, we also presented a method where single and multilead signals can be combined, leading to excellent results on the single lead signals. The combination of morphological features and heart rate variability features leads to stronger results than the use of either feature set individually.

Part III

QRS Fragmentation

Chapter 8

Detection and Quantification of QRS Fragmentation

The third part of this research focuses on QRS fragmentation, which is a promising SCD risk predictor that can be derived from the 12-lead ECG signal. Currently, detection of QRS fragmentation is done visually, however this is a time-consuming process which may lead to subjective results. We therefore propose a method to detect and quantify QRS fragmentation in ECG signals in an automated way using machine learning approaches. The method described in this Chapter will be published as: Goovaerts G., Padhy S., Vandenberk B., Varon C., Willems R., and Van Huffel S., A Machine Learning Approach for Detection and Quantification of QRS Fragmentation. Accepted for publication in IEEE Journal of Biomedical and Health Informatics. Preliminary results were published earlier as [89].

8.1 Introduction

Another promising marker for prediction of sudden cardiac death that can be detected in the 12-lead ECG is QRS fragmentation (fQRS). It is defined by Das et al. [53] as the presence of an additional R wave (R') or a notch in the nadir of the S wave in at least two contiguous leads on the 12-lead ECG. Das et al. showed that fQRS can be a convenient marker of myocardial scar, which may lead to high-risk cardiac events like heart failure, need for revascularization, or sudden cardiac death [177, 53]. Other studies have linked the presence of QRS

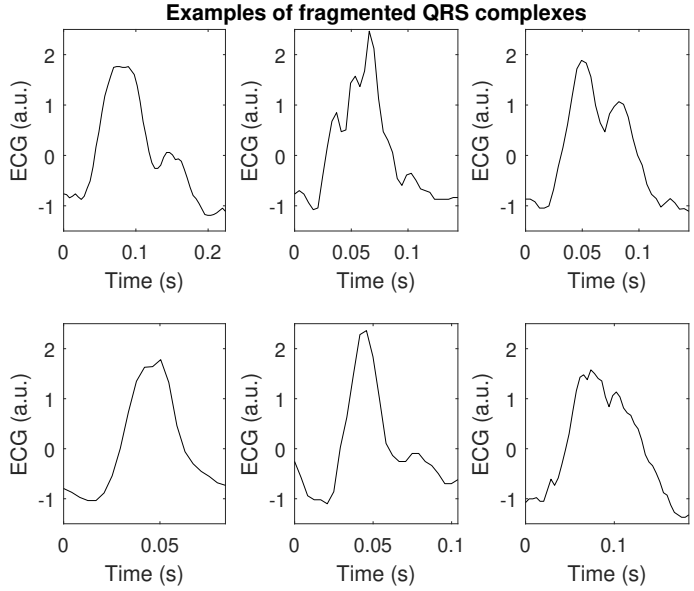


Figure 8.1: Examples of different subtypes of fragmented QRS complexes.

fragmentation to adverse outcome in patients with Brugada syndrome [157] and defibrillator shocks [223].

While the presence of fQRS was first noticed in 1969 by Flowers et al. [80], to the best of our knowledge, there have only been a few studies that focus on *automated* fQRS detection [139, 110]. Until now, visual assessment of each lead individually is considered to be the gold standard in clinical practice. Interpreting ECG signals is however time-consuming, expensive, and most importantly requires adequate training of clinicians in order to get reliable results. Recently, we have shown that visual fQRS assessment is not ideal as the inter-observer variability differs significantly depending on the expertise level of the observers [223]. Therefore, automated methods for fQRS detection should be considered and can serve as a complementary tool for the clinician.

Currently, fQRS *detection* is done where each lead is given a score of 0 or 1 depending on the absence or presence of fragmentation. Fragmentation can however take many forms as the number and location of deflections varies. This is illustrated in Figure 8.1 which shows several examples of fragmented QRS complexes, each with a different number of deflections or notches. Binary scoring might therefore not be optimal since it fails to capture differences between fQRS subtypes. It is furthermore expected that the spatial and

temporal characteristics of the deflections can be important prognostic factors in determining patient outcome [100]. This Chapter therefore proposes an automatic and objective method to determine a fQRS score that quantifies the certainty of the presence of QRS fragmentation in each lead.

For this, a necessary step is an accurate segmentation of the QRS complex. QRS segmentation is a mature technology and nowadays a number of methods have been developed, including time-domain, frequency-domain and transform-domain methods. A broad methodological review on QRS segmentation can be found in [144]. In the past two decades, wavelet transform methods such as [143] have been widely adopted. Recently, empirical mode decomposition (EMD) has also been used to segment the QRS complex [11, 172]. EMD is an empirical algorithm with lack of a theoretical basis that is known to be sensitive to noise. Additionally, it has been shown that EMD introduces distortions in the beginning and end of the QRS complex, which may cause erroneous results in this specific application [192]. The majority of methods including the wavelet- and EMD-based approaches are non-adaptive and hence cannot be used when signal characteristics change extensively. To overcome these limitations, Dragomiretskiy and Zosso proposed variational mode decomposition (VMD) [72]. VMD is based on the framework of variational theory and adaptively determines the relevant frequency band. In the last four years, it has been successfully used in different fields including ECG applications like arrhythmia characterization [147, 140] or detection of shockable ventricular arrhythmia [210]. In this work, we apply VMD to segment the QRS complex and extract features that help quantify the certainty of fQRS.

In this Chapter, machine learning is used to detect fQRS and quantify its degree of certainty. In our previous study [89], phase-rectified signal averaging (PRSA) was used to extract features that characterize QRS fragmentation. Here, additional features are computed from the results of VMD and combined in a classifier. The main goal here is to assess whether the combination of this novel feature set with different machine learning approaches succeeds in both detecting and quantifying fQRS in a continuous way. The comprehensive approach conducted in this study has the following main contributions:

- Accurate segmentation of the QRS complex using VMD
- Extracting VMD- and PRSA-based features that characterize the certainty of fragmentation in a lead
- Comparison of different machine learning methods for fQRS detection and quantification
- The assessment of the certainty of the presence of QRS fragmentation in a continuous way as opposed to a binary classification

The remainder of the Chapter is organized as follows. The database used in this work is described in Section 8.2. The different steps required for calculating the fQRS score are described in Section 8.3. Results of the proposed method (including accuracy of the QRS segmentation) with comparison of evaluation performance among different classifiers are presented in Section 8.4. The experimental results are discussed in 8.5. Finally, conclusions of the proposed work are given in Section 8.6.

8.2 Data

For this study, a dataset of 12-lead ECG signals recorded in 723 patients before the implantation of a cardiac defibrillator (ICD) in the prevention of sudden cardiac death due to cardiac arrhythmia was used. All signals were recorded in the University Hospitals Leuven, Belgium. The ECG signals are digitized with a sampling frequency of 250 Hz and have a duration of ten seconds. ECGs from patients with ventricular pacing or cardiac arrhythmia like atrial fibrillation were excluded since they have altered ECG morphologies. While these patients might exhibit QRS fragmentation, they were insufficiently represented in the dataset in order for the machine learning algorithms to learn their specific ECG characteristics. We therefore decided to remove them. A total of 616 records were included in the final dataset. Clinically, lead aVR is not used for fQRS analysis [223, 225]. This lead was therefore excluded from all records in this analysis. The study was approved by the ethical committee of the University Hospitals Leuven (S56074/ML9965).

The database was fully annotated by five readers on the presence of fQRS in each lead: they individually gave a score of 1 if fQRS was present and 0 otherwise. The persons who scored the signals were all clinicians: four cardiologists in training and one cardiology laboratory technician. Two of the readers were considered experienced observers in fQRS analysis, due to their involvement in previous studies and the three remaining ones were novice in fQRS analysis prior to this project but received training prior to scoring the signals. All five readers are experienced in research. The scores of all raters can be combined by summing them, resulting in a total score. If all readers agree that a lead does not show fQRS, the total score is 0; similarly, the total score is 5 if all readers agree on the presence of fQRS in a lead. Leads where some raters disagree have intermediate values. The total score given by all raters is thus related to the uncertainty of the raters about the presence of fQRS. In other words, a higher score means more raters have scored the signal as fragmented and the fragmentation is assumed to be more prominent in that lead. Signals

Table 8.1: Frequency of occurrence of different scores, obtained by summing the annotations from all five readers, in the database.

Total score	Counts	Percent
0	2775	40.95
1	894	13.19
2	490	7.23
3	390	5.76
4	553	8.16
5	1674	24.7

with certain absence of fragmentation will have lower scores. The frequency of occurrence of the scores (0 to 5) in the dataset is shown in Table 8.1. A full description of the scoring process, including results on inter- and intra-rater variability can be found in [223].

8.3 Methods

The block diagram of the proposed method for fQRS quantification in ECG signals is depicted in Figure 8.2. It consists of four main steps: 1) preprocessing, 2) QRS segmentation using VMD, 3) feature extraction and 4) classification. Each step is described in detail in the following subsections.

8.3.1 Preprocessing

In the preprocessing step, baseline wander and high-frequency noise were removed from ECG signals. Baseline wander was removed by passing each lead through a digital fourth-order Butterworth high-pass filter with cut-off frequency of 0.5 Hz and high-frequency noise components using a fourth-order Butterworth low-pass filter with cut-off frequency of 70 Hz. The filters were applied both in forward and backward direction to get zero-phase distortion. Then, the signal was normalized by calculating the z-score of each lead in order to remove the amplitude differences between different recordings.

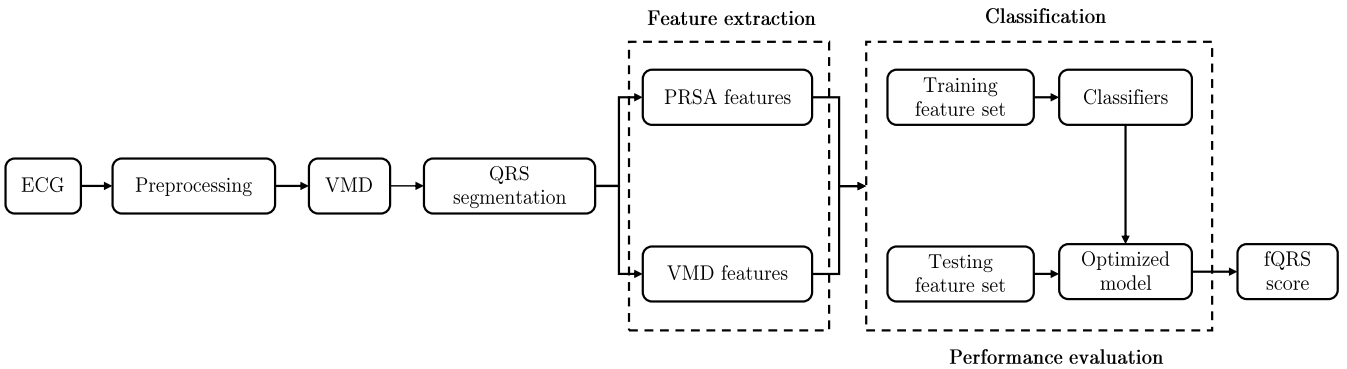


Figure 8.2: Block diagram of the proposed method for fQRS detection and quantification. After preprocessing and QRS segmentation using VMD, features were extracted using PRSA and VMD after which a classifier was trained and evaluated. The final fQRS score represents the certainty of QRS fragmentation in an ECG lead.

8.3.2 VMD-based QRS segmentation

Variational Mode Decomposition

In this study, variational mode decomposition was applied to segment the QRS complexes. VMD adaptively decomposes a real-valued multi-component signal $x(\mathrm{t})$ into K discrete number of modes or components $u_k(t)$ with $k = 1, ..., K$. All components have certain sparsity properties, and the bandwidth in spectral domain is considered as the sparsity prior of each mode. All modes are mostly compact around their center frequencies ω_k.

The bandwidth of a mode is evaluated in three steps: (i) the analytic signal of the real-valued signal is computed using the Hilbert transform such that the frequency spectrum becomes unilateral, (ii) the frequency spectrum of the analytic signal corresponding to each mode is shifted to baseband regions by multiplying it with the factor $e^{-i\omega_k t}$, (iii) then, the bandwidth is estimated through the squared L^2-norm of the gradient of the demodulated signal. Mathematically, this constrained variational problem can be expressed as follows

$$\begin{aligned} \min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \tfrac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_k u_k(t) = x(t) \end{aligned} \tag{8.1}$$

where δ is the Dirac-Delta function, $*$ is the convolution operator, and $\{u_k\}$ and $\{\omega_k\}$ represent the set of modes and center frequencies, respectively.

The constrained problem in Equation (8.1) can be solved by converting it into an unconstrained problem using the balancing parameter α and the Lagrangian multiplication parameter $\lambda(t)$. The modified equation with the augmented Lagrangian multiplier is expressed as

$$\begin{aligned} \mathcal{L}(\{u_k\},\{\omega_k\},\lambda) := \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \tfrac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ + \left\| x(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), x(t) - \sum_k u_k(t) \right\rangle . \end{aligned} \tag{8.2}$$

The above Equation (8.2) is solved using the alternate direction method of multipliers [28]. We have skipped the intermediate steps to solve this equation and some other related algorithms as these are beyond the scope of this work. The readers are encouraged to refer to the original article for the full

implementation of the VMD method [72]. The solutions to Eq. (8.2) in the Fourier domain represent all modes and their central frequencies where each mode and its center frequency are updated iteratively as

$$\hat{U}_k^{n+1}(\omega) = \frac{\hat{X}(\omega) - \sum_{i \neq k} \hat{U}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{8.3}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left|\hat{U}_k(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{U}_k(\omega)\right|^2 d\omega} \tag{8.4}$$

where $\hat{X}(\omega)$, $\hat{U}_i(\omega)$, $\hat{\lambda}(\omega)$, and $\hat{U}_k^{n+1}(\omega)$ are the Fourier transforms of the respective time domain representations. The Wiener filter structure in Equation (8.3) makes the VMD method more robust to noise and sampling [93]. Finally, the real part of the inverse Fourier transform of $\hat{U}_k^{n+1}(\omega)$ gives the time domain representation of the modes. The center frequency in each iteration in Equation (8.4) is the center of gravity of the corresponding positive part of mode's power spectrum.

QRS segmentation

From the VMD algorithm discussed above, it is clear that VMD is a parametrized signal decomposition method. VMD requires two parameters to be fixed: the number of modes K and the bandwidth control parameter α. Initialization of these parameters is a non-trivial problem, since wrong parameter choices may create problems like mode splitting (where one component is shared by several modes) or mode mixing (where multiple components are decomposed into one mode) [93]. To avoid this problem, we adopted the optimization technique developed by Guo et al. [93] to fix the number of decomposition modes K and balancing parameter α prior to applying VMD. After optimization, these values were fixed to $K = 5$ and $\alpha = 100$. The other input VMD parameters, namely the time step of the dual ascent and the tolerance of the convergence criterion were set to standard values 0 and $1e^{-6}$ [72]. All central frequencies ω_k were uniformly distributed at initialization.

VMD was applied to each ECG lead $x(t)$ resulting in a decomposition into 5 modes:

$$x(t) = \sum_{k=1}^{5} u_k(t) \tag{8.5}$$

where $u_k(t)$ is the kth mode. Figure 8.3 shows an example of an ECG segment after preprocessing with the corresponding modes u_1–u_5 in different rows. The

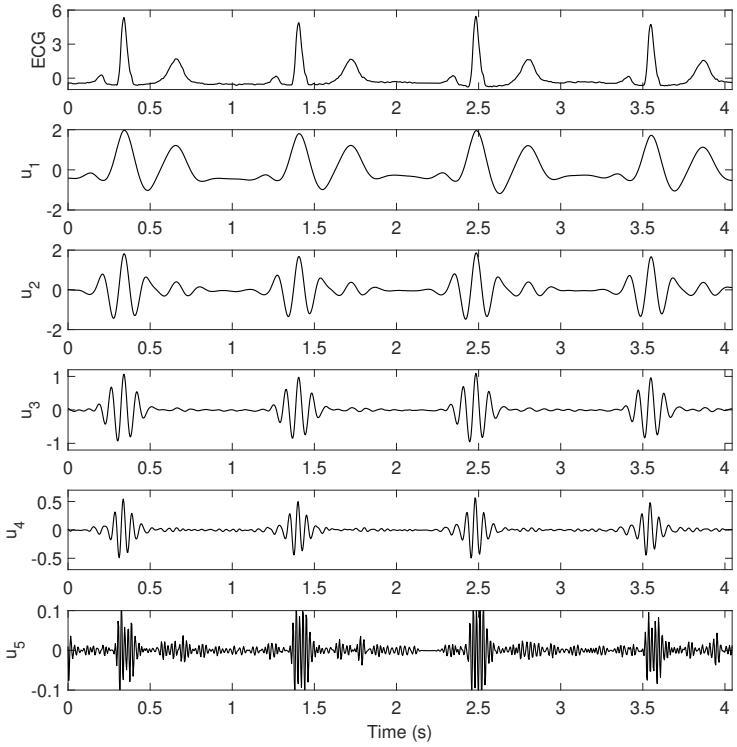


Figure 8.3: Example of ECG signal with the corresponding modes of the output of Variational Mode Decomposition with $k = 5$ and $\alpha = 100$.

modes are sorted by their central frequency in ascending order. It can be observed that different components (or characteristic waves) of the ECG signal are decomposed into different modes. The low-frequency components appear in lower modes and vice versa. The QRS complex is a high frequency wave with sharp amplitude and thus appears in higher modes (u_2-u_4). Based on the fact that the modes are mostly compact around their center frequencies, the QRS complex and hence the Q and S points can best be detected using u_3, which is amplitude-normalized to obtain $u_{3,norm}$. The complete QRS segmentation algorithm consists of three main steps, which are illustrated in the different rows of Figure 8.4:

1. *R peak detection*
   Square $u_{3,norm}$ in order to obtain $u^2_{3,norm}$ to enhance the main peaks in $u_{3,norm}$ (corresponding to QRS complexes) and minimize noise peaks. Detect the R peaks in $u^2_{3,norm}$ by finding local maxima with an amplitude

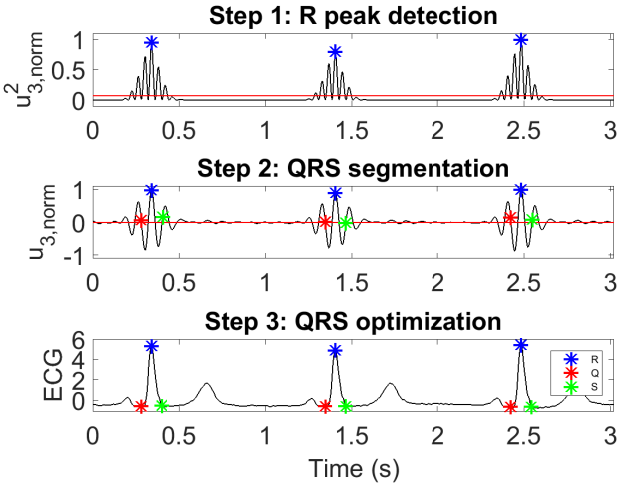


Figure 8.4: Illustration of the three main steps for QRS segmentation, applied to the signal shown in Figure 8.3. First the R peaks are detected in $u^2_{3,norm}$. Then the QRS complex is segmented in $u_{3,norm}$. Finally the QRS locations are optimized in the original ECG signal. The R peaks are depicted in blue, the beginning and end of the QRS complex in respectively red and green.

higher than 0.075. The minimal distance between two consecutive peaks must be larger than 0.5 seconds. This corresponds to a maximal heart rate of 120 beats per minute which is realistic for ECG signals measured in rest. For signals with potentially increased heart rates, this parameter can easily be changed to accommodate them.

2. *QRS segmentation*
   Transfer the R peaks to $u_{3,norm}$. The beginning of the QRS complex corresponds to the first zero-crossing *before* the first local minimum *preceding* the R-peak in $u_{3,norm}$. Similarly, the end of the QRS complex can be found by detecting the first zero-crossing *after* the first local minimum *after* the R peak.

3. *QRS optimization*
   Optimize the locations of the Q and S points by finding a local minimum in the ECG signal in the neighbourhood ($\pm$ 5 samples) of the points detected in $u_{3,norm}$.

QRS segmentation was done in each lead separately. In a final stage, segmentations in all leads were combined in an automated way to remove false

detections. False positive detections can occur when the ECG signal contains noise with characteristics similar to QRS complexes. They were removed by first calculating in how many leads a QRS complex was detected at a certain time instance and subsequently removing QRS complexes that were detected in less than half of the ECG leads. False negative detections were solved similarly by automatically adding QRS complexes in leads where a complex was missed by the algorithm. The start- and end points for these additional complexes were selected as the mean start- and end points of the complexes detected by the algorithm in the other leads of the signal.

Finally all points between the end of the QRS complex and the beginning of the next complex are set to zero since they will not be analyzed further.

8.3.3 Feature extraction

In this subsection, feature extraction using PRSA and VMD is discussed. PRSA feature extraction was used in our earlier study [89], and the same approach has been adopted in this work. All features were extracted from the QRS complexes segmented using the method described in the previous subsection.

PRSA-based features

Phase-rectified signal averaging aims to detect and quantify quasi-periodic oscillations masked by multi-component non-stationary signals [16]. The method consists of three steps:

(i) *Selection of anchor points*
First, two sets of anchor points q_i were selected according to certain properties in the QRS complex. The anchor points can be selected in different ways [16]; here, a first set of anchor points consists of all points located on the increasing part of the QRS complex (q_i^+), a second set contains all points on decreasing parts of the QRS complex(q_i^-).

(ii) *Window selection*
Then, temporal windows of length $2L$ were selected around each anchor point. The anchor points which were too close ($< L$ ms) to the beginning or end of the complex were discarded since for these points no such window can be selected. The choice of L is dependent on the characteristics of the signal: it must be larger than the period of the slowest oscillation that should be detected. Here L is fixed to 20 ms, similarly to [89].

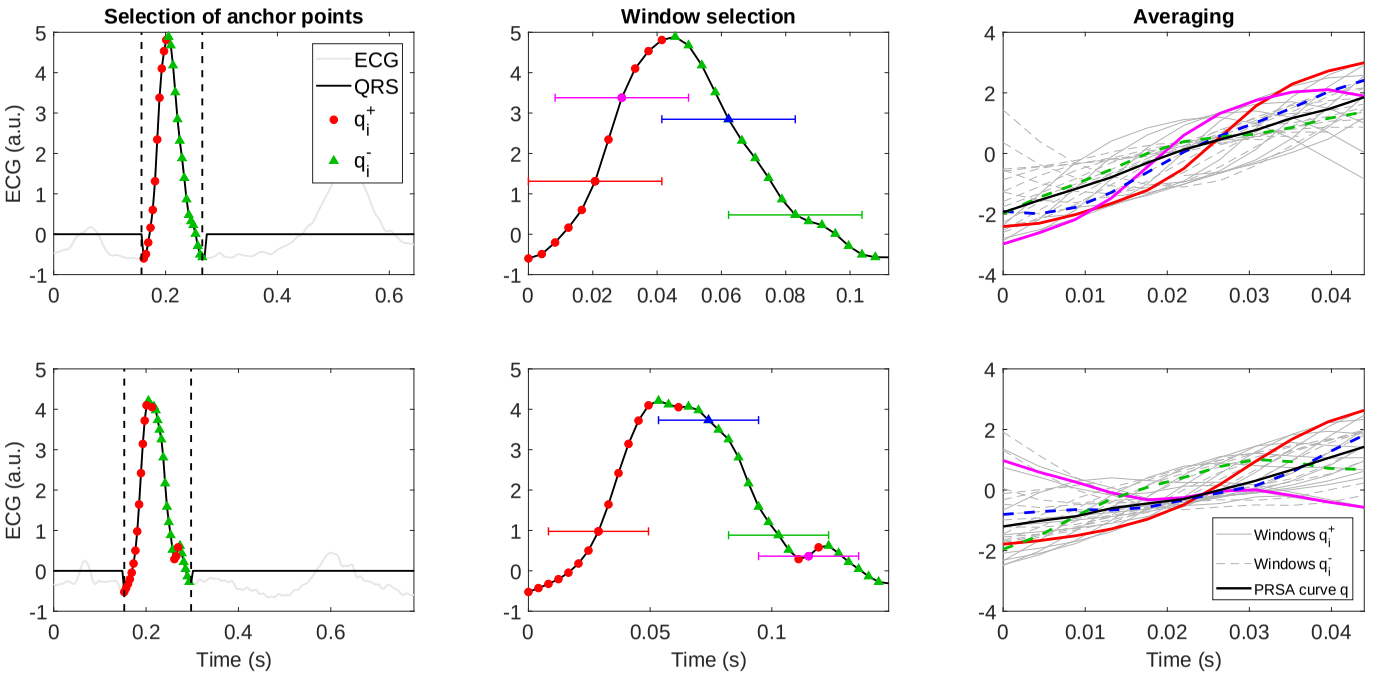


Figure 8.5: Illustration of the three steps to create the PRSA curve for a normal heartbeat (top) and a heartbeat with fragmentation (bottom). In the middle panel, two increasing and two decreasing anchor points are highlighted in respectively red, cyan and green, blue. Their corresponding windows are also indicated in the right panel.

(iii) *Averaging*
In the final step, the PRSA curve $\tilde{q}$ was obtained by aligning all windows and averaging them. The windows corresponding to decreasing anchor points q_i^- were inverted before calculating the average, e.g. $window(q_i^-) = -window(q_i^-)$. Inverting the windows is necessary to ensure that all windows have similar slopes.

The different steps to calculate the PRSA curve for a normal and a fragmented beat are shown in Figure 8.5. Two positive and two negative anchor points are highlighted in respectively red, cyan and green, blue to highlight the changes in the signals. For a normal beat (depicted in the top row), all increasing anchor points are located in the first part of the QRS complex and all decreasing points in the second part due to the very simple morphology of the wave. Most windows are thus similar in shape and the PRSA curve can be approximated by a straight line. When a signal shows fragmentation (shown in the bottom row), the anchor points are not contained in one part of the QRS complex. The surrounding windows thus exhibit more differences which result in a PRSA curve with a smaller inclination, which can be seen by comparing the rightmost panels of Figure 8.5.

To extract features from the PRSA curve, it was approximated with a linear fit. Three parameters were extracted to quantify the curve:

(i) Mean derivative of the PRSA curve (sl): $\frac{1}{2L}\sum_{t=0}^{2L}\frac{d\tilde{q}(t)}{dt}$

(ii) Slope of the linear fit (m)

(iii) Intercept of the linear fit with the y-axis (c).

VMD-based features

Fragmentation introduces additional high-frequency components in the ECG signal, which in turn causes extra oscillations in the VMD modes that contain QRS information. Therefore the central frequencies of these modes also increase. As discussed in Subsection 8.3.2, modes u_1 and u_2 contain the low-frequency ECG components such as P- and T-waves, and other modes mostly carry the high-frequency components such as the QRS complex. Due to the presence of notches or additional waves, it is expected that the number of local optima per QRS complex will increase with respect to the certainty of fQRS. Hence, two different VMD-based features were extracted from modes u_3, u_4 and u_5: the average number of local optima per QRS complex (pks_3, pks_4, pks_5) and the central frequency of the QRS complexes in the considered modes ($\omega_3^{QRS}, \omega_4^{QRS}, \omega_5^{QRS}$), leading to a total of six VMD-based features.

One additional feature was extracted directly from the QRS complex: the number of local peaks in the ECG signal per QRS complex, pks_{ECG}. This feature was combined with the three PRSA features and six VMD features to obtain a total of ten features for each lead.

8.3.4 Classification

The extracted features were used as input to a classifier in order to detect and quantify QRS fragmentation. The performance of SVMs with a linear, polynomial and radial basis function kernel was compared to select the kernel which obtains the best performance. In order to define the optimal SVM parameters, automatic Bayesian optimization of hyperparameters (both the soft margin constant and kernel parameters) using 10-fold cross-validation was performed. Additionally, to illustrate the effectiveness of the proposed method, the performance of three different classifiers (K-nearest neighbors (kNN), Naive Bayes classifier (NB), and TreeBagger (TB)) was also evaluated and compared with the results obtained by SVMs. The output of the classifiers (e.g. the score belonging to the positive class) was transformed to a score between 0 and 1 through the use of Platt scaling, which fits a logistic regression model to the output [178]. The obtained score is taken as fQRS score and is expected to reflect the certainty of fQRS in a channel.

The dataset used in this study contained 6776 signals in total (616 records with 11 leads per record). These signals were divided into one training and two test sets. To train the classifier, only signals where all experts agreed on the presence of QRS fragmentation were used. There were 2775 normal signals and 1674 fragmented signals without disagreement (see Table 8.1). These were randomly split in 80% training set and 20% test set. Both sets had equal ratios of normal and fragmented signals. A second test set contained all signals that were not used in the training stage. The first test set was used to evaluate the performance of the method for fQRS detection: since all raters agreed on the label for all signals in this set, this label was considered to be correct and binary classification could be done. The second test set on the other hand was used to assess whether the developed score reflected the certainty of QRS fragmentation in an ECG signal.

The training set and the first test set are fully independent: signals from the same patient are only present in one of both sets. Since the second test set contains all ECG channels were no perfect agreement was reached, it also contains channels from patients in the training set, and both sets are not fully independent.

8.3.5 Performance evaluation metrics

Different measures were used to evaluate the performance of the developed method. Statistical analyses of parameter values were done with one-way Analysis of Variance (ANOVA) with the F-test [186] followed by a post-hoc test analysis based on Tukey's honest significant difference (hsd) test. These tests were carried out to verify whether means of different categories are significantly different or not. ANOVA assesses the relative size of variance among category means compared to the average variance within categories and is an appropriate test for evaluating the effect of categorical independent variables on a continuous response variable. The F-ratio, the value obtained from the ratio of the variance between categories and the variance within categories, was used as the parameter to decide the statistically significant performance. The F-test represents the determination of the significance of the F-ratio by comparing it to a critical value derived from the probability distribution. The null hypothesis that the means of all categories are equivalent is rejected if the F-ratio is greater than the critical value. If the F-test is statistically significant, then there is, in principle, at least one significant difference in means. Then, a post-hoc test e.g. Tukey's hsd test was executed to perform specific comparisons to discover the origin(s) of the difference. In all tests, a p-value < 0.01 was considered statistically significant.

Performance of the different classifiers was evaluated by constructing Receiver Operator Characteristic curves (ROC) and calculating the corresponding Area Under The Curve (AUC). The results of fQRS detection were further quantified by calculating the number of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) detections and computing the sensitivity (Se), specificity (Sp) and accuracy (Acc).

8.4 Results

The full dataset was analyzed using the proposed method depicted in Figure 8.2. In this section, we present the results of the different analyses. All experiments in this work were performed using MATLAB R2017b.

8.4.1 QRS segmentation

To evaluate the results of the proposed approach to QRS segmentation, the publicly available QT database was used, which contains 105 ECG signals of 15 minutes with manually determined wave boundaries [122]. The database

Table 8.2: Accuracy results of QRS segmentation of the proposed method and three state-of-the art alternative approaches on the QT database. Mean and standard deviation of the difference between the provided annotations and the segmentation obtained by the algorithm are given in ms.

	QRS_{on}		QRS_{off}	
Method	mean	std	mean	std
Proposed method	-3.6	11.16	6.67	17.28
Martínez et al. [143]	4.6	7.7	0.8	8.7
Madeiro et al. [137]	-3.4	11.6	-6.5	12.3
Akhbari et al. [6]	-5	10	1.5	11.5

has been used extensively as benchmark tool for ECG delineation algorithms. The performance is quantified by calculating the mean deviation and standard deviation between the VMD-based segmentation and the manual annotations.

The proposed R peak detection algorithm detected 98.11% of all R peaks correctly. This is equivalent to the results found in literature [143, 137, 6]. The accuracies of the delineation of the onset and offset of the QRS complex are summarized in Table 8.4. The results for segmentation of the QRS onset are comparable to the state-of-the art algorithms. The standard deviation of the QRS offset detection is slightly worse than the results reported by Madeiro et al. [137] and Akhbari et al. [6], but the difference is limited to five ms, which is equal to one sample for signals sampled at 250 Hz.

8.4.2 Analysis of feature values

Figure 8.6 shows box plots for all ten features. Each set of box plots is calculated by computing the particular feature value for all leads of all signals and grouping them per total score (e.g. the sum of the scores by all raters). As discussed in Section 8.2, the total score is a substitute for the (un)certainty of fQRS in an ECG lead. We observe that for all features the feature values gradually change with changing total score. The average slope of the PRSA curve and the slope of the linear fit are both inversely related to fQRS certainty, which is expected based on the examples in Figure 8.5. Similarly, the intercept of the linear fit increases with increasing certainty. The number of peaks per QRS complex in the VMD modes regularly increases with the level of fragmentation. The increase in the median values in modes u_4 and u_5 is more prominent compared

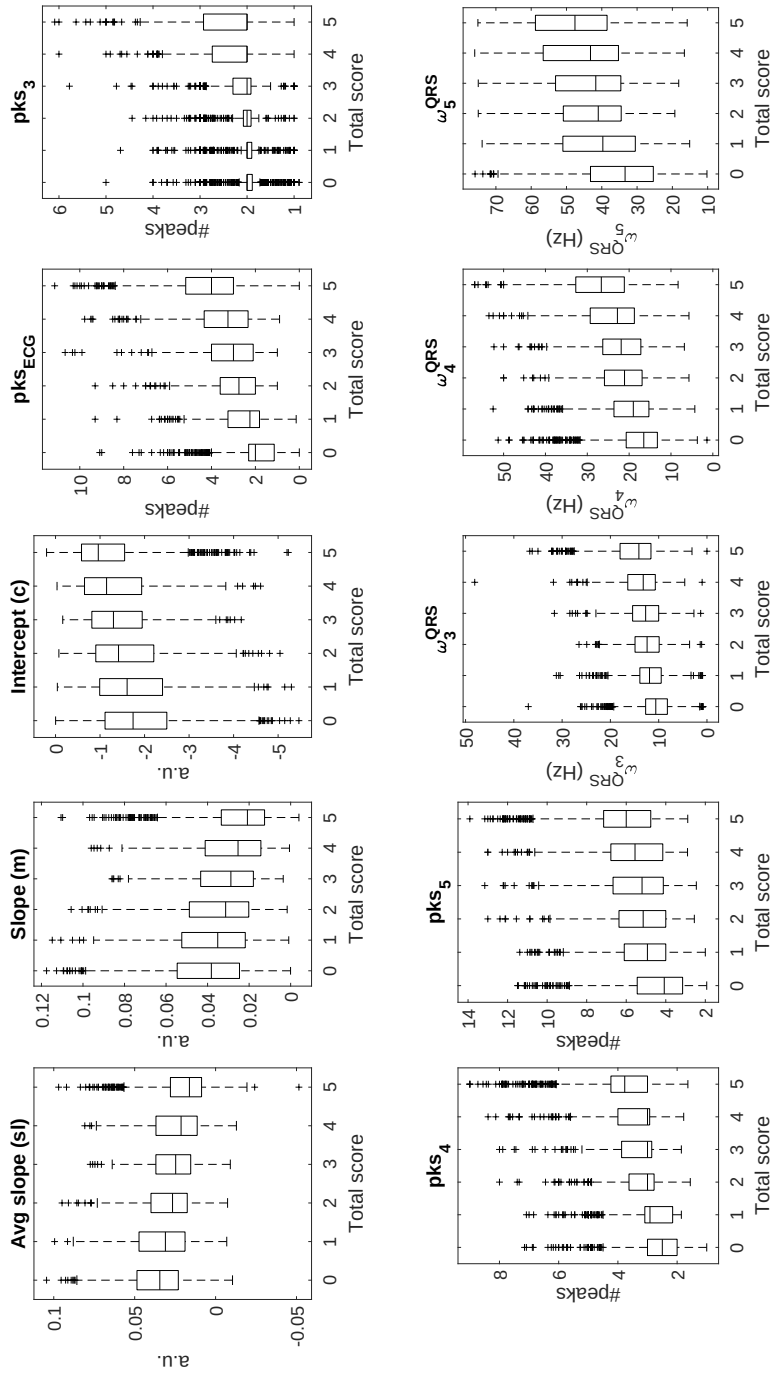


Figure 8.6: Box plots showing all feature values grouped by the total score given by five experts. The box plots represent the median value in a group together with the interquartile range.

Table 8.3: Significance results of the post-hoc analysis for the comparison of the VMD- and PRSA-based features using Tukey's hsd test. Results are shown only between the consecutive categories. 'Y' stands for statistically significant with $p < 0.01$. p-values are mentioned for all non-significant results.

Compared categories:	**0-1**	**1-2**	**2-3**	**3-4**	**4-5**
PRSA-based features					
Avg. slope (sl)	Y	Y	0.465	0.064	Y
Slope (m)	Y	Y	0.453	0.112	Y
Intercept (c)	Y	Y	0.387	0.165	Y
VMD-based features					
ω_3^{QRS}	Y	0.014	0.999	0.013	Y
ω_4^{QRS}	Y	Y	0.788	Y	Y
ω_5^{QRS}	Y	0.212	0.904	0.277	Y
pks_3	Y	0.020	0.774	0.015	Y
pks_4	Y	Y	0.410	Y	Y
pks_5	Y	0.201	0.541	0.337	Y
pks_{ECG}	Y	Y	0.821	Y	Y

to the medians in u_3. It is interesting to note that the central frequency of the QRS complexes in these modes also follows the same trend.

ANOVA is used to statistically compare the means of the values in different score groups. Here, the null hypothesis is that the mean of a feature is identical for all six categories. The F-ratio results reveal that for *all* features, there is at least one mean that is significantly different from one or more categories. As this happens for all features, we only present the post-hoc analysis.

Table 8.3 shows the post-hoc analysis based on the Tukey's hsd test that shows which categories differ from each other by comparing the means between consecutive categories. Significant p-values ($p < 0.01$) are indicated with 'Y'.

8.4.3 Classifier performance

In this subsection, we present the performance of different classifiers for fQRS detection and quantification. As mentioned in Section 8.3.4, the classifiers were trained on 80% of the signals without disagreement between the raters. The other signals were used for evaluation during the testing stage.

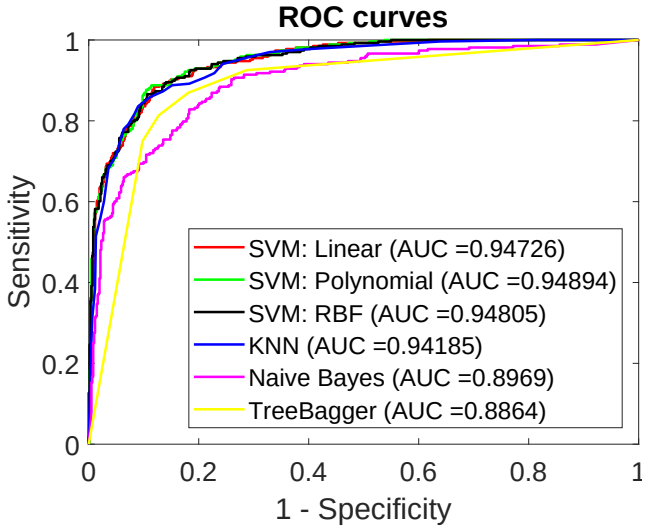


Figure 8.7: ROC curves for fQRS detection with all classifiers together with the corresponding AUC values.

Table 8.4: Comparison of the fQRS classification results from the proposed method with methods from literature. The obtained sensitivity, specificity and accuracy on the first test set are mentioned.

Method	**Se (%)**	**Sp (%)**	**Acc (%)**
Proposed method	**86%**	**89%**	**88%**
Jin et al. [110]	72%	78%	75%
Maheshwari et al. [139]	84%	87%	85%
Bono et al. [24]	75%	78%	77%

Figure 8.7 shows the ROC curves with corresponding AUC values for all classifiers. Here, the test set consists of the 20% remaining signals with perfect agreement. The true class of these signals is considered to be certain so classification can be performed. The differences between the SVMs with different kernels are negligible (AUC = 0.947 (lin), 0.948 (pol) and 0.948 (RBF)). The KNN classifier (AUC = 0.941) performs slightly worse than the SVMs while the performance of the NB (AUC = 0.896) and TB (AUC = 0.886) classifiers are comparatively worse.

Table 8.4 finally summarizes the sensitivity, specificity and accuracy results obtained by applying the proposed method (SVM with RBF kernel) and the three

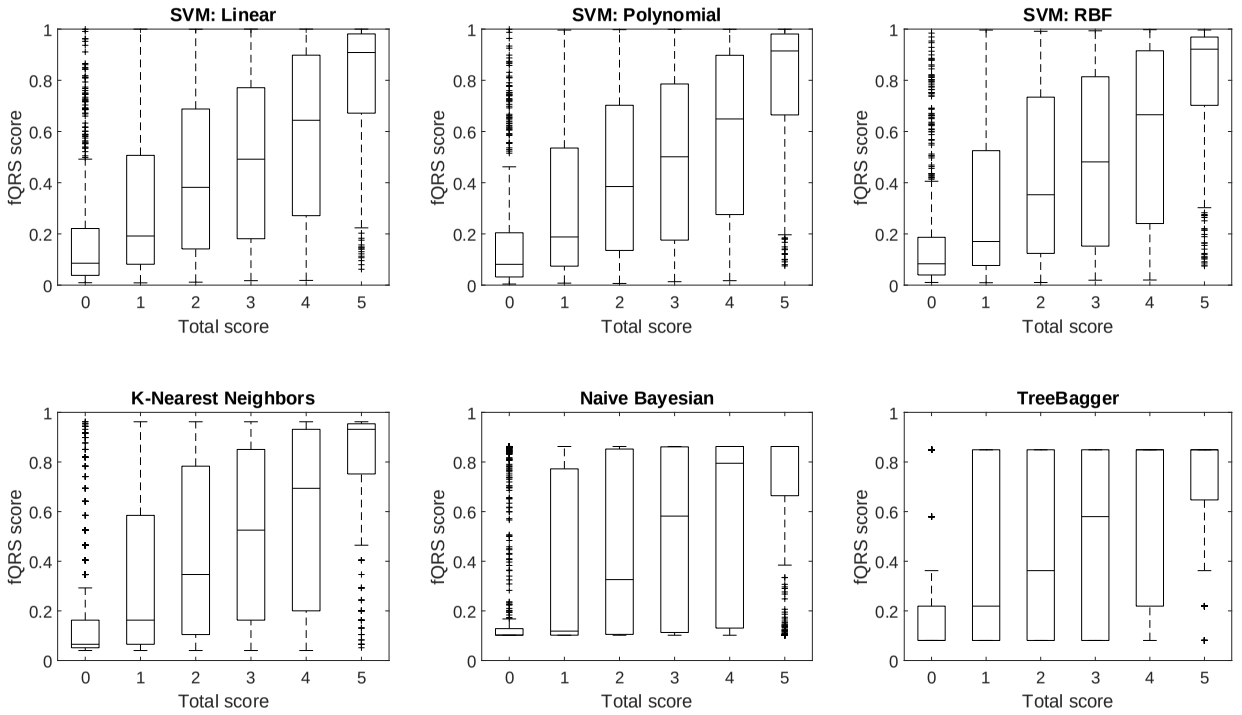


Figure 8.8: Illustrating the fQRS quantification score for the second group (Section 8.3.4) using SVM (Linear, Polynomial and RBF kernels), KNN, NB and TB classifiers.

alternative algorithms [110, 139, 24] on the first test set. The other approaches use the wavelet transform [138, 24] or intrinsic time-scale decomposition [110] to decompose the ECG signal and detect fQRS from features obtained from the decomposed signal.

The box plots in Figure 8.8 show the fQRS quantification results. Here, the test dataset contains all instances that are not considered in the training stage. The box plot for each category is computed by combining the fQRS scores for all signals with the same total score. The output scores from most classifiers follow an increasing trend with respect to the different categories. The results for the NB and TB classifiers are considerably worse than the other classifiers. Results for SVM and KNN classifiers are again comparable, but the interquartile range for the different SVM box plots is smaller than for KNN. Furthermore the range of SVM scores spans the full range (0–1) while the range for the other classifiers is limited to a smaller interval.

8.5 Discussion

This study proposes an innovative method to detect and quantify fQRS. In the first step, the QRS complexes are segmented using a novel VMD-based algorithm. According to Table 8.2, the approach gives results that are comparable to many state-of-the-art algorithms. While the widely-used wavelet-based method from Martínez et al. [143] obtains superior results, the differences are limited to a few samples.

Figure 8.6 and Table 8.3 summarize results for the feature values in different groups. The values for all features change gradually with increasing raters' scores, indicating that the reasoning in Section 8.3.3 is correct: fragmentation has a significant influence on the PRSA curve and introduces extra high frequency components in the results of VMD, which in turn give rise to extra peaks. Table 8.3 shows that all features are significantly different between the most extreme groups (0-1 and 4-5). The median differences between other groups are not always statistically significant, but there is a clear trend for all features which can also be noticed in Figure 8.6. The feature values for signals with a total score of 2 and 3 are similar, with p-values larger than 0.35. This is not unexpected since the signals in these groups are comparable: they represent the signals which are scored as fragmented by approximately half of the experts and normal by the other half.

Once the PRSA- and the VMD-based features have been extracted from the QRS complexes, they were fed into different classifiers in order to evaluate and compare their performance. Experimental results presented in Figure 8.7

show that all classifiers give good results for detecting signals with clear QRS fragmentation. SVMs are slightly superior to other classifier types. The results in Table 8.4 show that our approach outperforms the other methods, although the results obtained by [139] are in the same range. A second test set was used to evaluate the performance for fQRS quantification. Here the goal was to verify whether the developed score is representative for the certainty of the presence of fQRS in an ECG lead. Platt scaling transforms the classifier output to a score between 0 and 1, which corresponds to the posterior probability of the classifier. This results in the box plots shown in Figure 8.8. For the different types of SVMs, the fQRS score varies between 0 and 1 and linearly changes with the total score from all experts. As explained in Section II, the total score is related to the certainty of fQRS in a lead. It is therefore expected that leads with a questionable fQRS presence (represented by a total score of 1–2, where only one or two raters observe fQRS in that lead) have lower fQRS scores than leads with a probable (total score of 3–4) or certain presence (total score = 5). Figure 8.8 confirms this hypothesis. If the boxplots of the different scores are compared, the boxplot heights of values 0 and 5 are significantly smaller than the heights of the intermediate values. This confirms that the classifier output for values 1–4 is indeed more uncertain than for the extreme values. This is also expected, since the true label contains more uncertainty (for the intermediate values, the presence or absence of QRS fragmentation is after all not certainly known). While the number of outliers in the boxplot corresponding to a total score of 0 seems large at first sight, they represent less than 2% of all signals in the second test set. These outliers can have different explanations: For some signals, the QRS segmentation is not 100% accurate, which can have an influence on the feature values. For other signals, an increased noise level can change the ECG characteristics and have an influence on the final classification. The fact that the number of outliers is small and the good AUC scores obtained on the first test set however confirms that this is only the case in a small minority of the data.

KNN also results in a stepwise increasing score, while the results of NB and TB are considerably worse. The scores obtained by KNN, however, do not span the full range between 0 and 1 but are limited from 0.05 to 0.95. It has been shown before that Platt calibration of KNN classifiers can lead to diminished results while the results of SVM classifiers are significantly boosted [39, 164]. Since SVMs also result in superior detection results (shown on Figure 8.7), SVM classifiers are preferred over other classification methods.

In this study the performances of three different SVM kernels (linear, polynomial and RBF kernels) were compared. They all have similar results both in fQRS detection and quantification. Since the numerical results are almost identical, the choice of kernel does not appear to have a significant effect in this application.

In general, however, the RBF kernel is preferred in such cases since it is known to be a universal approximator.

The results presented in this Chapter demonstrate that the proposed fQRS score is an effective way of detecting and quantifying QRS fragmentation. Comparison with existing techniques shows that the proposed method outperforms other methods found in literature. The extensive scoring done by five independent raters is a strong indicator that the results are robust to inter-rater variability and can be generalized to other datasets. This should, however, be validated with other datasets scored in a similar way. The availability of scores by different raters also allows us to evaluate the quantification of fQRS certainty in an objective way, which is a novel approach to examine the biomarker.

Clinically, QRS fragmentation is scored on a per-lead basis, after which the detections in separate leads are combined per cardiac region. Future work includes combining the fQRS scores in a similar way to achieve a score per cardiac region which can be used for clinical validation. The final goal is to verify whether a continuous way of scoring QRS fragmentation also leads to superior results in clinical studies that focus on analyzing patient outcomes. This will be explored in the next Chapter.

8.6 Conclusion

This Chapter presents a novel and precise way to detect and quantify QRS fragmentation in ECG signals using machine learning techniques. The fQRS score represents the certainty of QRS fragmentation in a continuous way based on fQRS annotations from five experts. The features extracted using VMD and PRSA show values that change gradually with increasing fQRS certainty. The fQRS score obtained using SVM classifiers is closely related to the total score given by all raters which is representative for the fragmentation in a lead.

The developed method is an automated and objective way of characterizing fQRS and will thus lead to results that are more reliable than scoring based on visual analysis which is known to be dependent on rater experience. It is therefore expected that the proposed fQRS score will also benefit the practical use of the parameter in clinical practice.

Chapter 9

Risk Assessment of All-Cause Mortality using a QRS fragmentation score

The previous Chapter proposed a method to automatically detect and quantify QRS fragmentation. The goal of this Chapter is to examine the clinical utility of this novel parameter in the assessment of the risk on all-cause mortality. Additionally, we will use a combination of standard statistical models and more advanced SVM approaches to verify whether the use of machine learning methods can provide further added value here. The first part of this Chapter, related to the determination of optimal cut points, was published as Goovaerts G., Padhy S., Varon C., Vandenberk B., Willems R., Van Huffel S., "Risk on all-cause mortality in ICD patients using a novel QRS fragmentation score", in Proc. of the 45th Annual Computing in Cardiology, CinC 2018, Maastricht, The Netherlands, Sep. 2018, where it was nominated as one of the finalists for the Rosanna Degani Young Investigator Award.

9.1 Introduction

QRS fragmentation was defined in the previous Chapter as the presence of extra deflections or notches on the QRS complex [53]. It is known to be an independent predictor for ICD shocks and mortality in certain patient groups [223]. More specifically, in primary prevention the presence of fQRS in the

anterior channels can predict all-cause mortality while inferior fQRS is associated with appropriate ICD shocks [223]. The previous Chapter presented a method to automatically determine an objective fQRS score that reflects the certainty of QRS fragmentation in an ECG lead. We showed that the proposed algorithm obtains good detection results and that the quantification is in line with the combined score from five experts. The objective of this Chapter is now to examine whether this novel parameter also has potential to be used as risk factor for all-cause mortality in ICD patients. Since previous research showed that the location of fQRS contains important diagnostic information, we will furthermore examine whether these findings can be replicated using our novel approach.

In a first study [222], the goal was to verify whether the PRSA-features presented in Chapter 8 can lead to better predictions of appropriate ICD shocks and all-cause mortality than traditional visual assessment. There, we showed that if PRSA-features are used as input to an SVM-classifier, the predictive value improved significantly for the prognosis of appropriate shocks. For all-cause mortality, the dataset considered in [222] was too small and contained not enough events to make strong conclusions. The obtained results however reassured us of the power of the developed approach and inspired the work described in this Chapter.

Since a larger dataset was collected for the purpose of developing the fQRS detection method, this dataset will be further used here to assess the risk on all-cause mortality, which is a prime example of survival analysis. Support vector machines can be used for the analysis of survival data by considering survival up to a certain cut-off time t as a binary variable, the approach taken in [222]. It however leads to a significant loss of information, as outlined in [77]: First, in order to include as many events as possible, the cut-off t should be as long as possible. On the other hand, in many clinical studies patients are only followed for a certain period, and no further information is available about the patient status after that period, they are lost to follow-up. For binary classification of survival data, all patients lost to follow-up before the cut-off will be discarded, favouring a low t. Second, when dichotomizing the survival time, all events that occur before the cut-off time t are handled in the same way, regardless of the moment when they occurred. There may however be valuable information in the exact time of the event, which is lost this way. In the past years however, some extensions of SVMs that are able to handle the specific characteristics of survival data have been developed and will be evaluated here.

In this Chapter, we will use a combination of standard statistical methods and more advanced machine learning methods to examine to what extent the developed score is useful as a prognostic risk factor in a population of ICD patients. The goal of this Chapter is therefore twofold:

1. Verify whether the novel QRS fragmentation score has added value compared to standard clinical characteristics to predict mortality in a group of ICD patients
2. Verify whether advanced machine learning techniques have added value compared to standard statistical survival models

Since survival analysis methods have some specific differences compared to standard regression and machine learning techniques, the next Section first gives a summary of the most important concepts and methods, largely based on Bradburn et al. [29] and Clark et al. [43]. Section 9.3 provides an overview of the dataset. The methods to analyse this data are described in Section 9.4, and the results and discussion follow in Sections 9.5 and 9.6 respectively.

9.2 Survival analysis

Chapter 3 gave an overview of some important concepts in machine learning, which can be used to classify samples in different categories. A specific subset of methods deals with survival analysis, which involves the description and modelling of time-to-event data [29, 43]. Another goal is to identify risk factors that influence the risk on experiencing the event under analysis such that a negative event can be avoided. As such it is closely related to regression, but the main difficulty in survival analysis lies in the presence of *censored* data. Censored data are data which are not completely observed. In the context of survival analysis, where the event under study is the time of death of patients, not all patients are followed until their time of death. For those patients, the only information that is available is the fact that they were alive at the end of the study period, and that their survival time is at least equal to some time t. Their data are *right-censored*, and the censoring time represents a lower bound on the event time. There are generally speaking three reasons why censoring occurs:

1. A subject does not experience the event before the end of the study
2. A subject is lost to follow-up during the study period
3. A subject withdraws from the study

Apart from right-censoring, data can also be left-censored, where only an upper bound of the event time is known. It is however less common in the datasets used here. Regardless of the type of censoring, we have to assume that the

censoring is non-informative about the event. When patients for example drop out of a medical study because their condition is worsening, the censoring might give information about the endpoint and is considered informative.

Censored data are generally described as 3-tuples of a binary event variable δ which is set to 1 if the event occurs and is thus equal to 0 for censored subjects, the time-to-event t and a feature vector of covariates $\mathbf{x} \in \mathbb{R}^D$ with D the number of features considered. The time-to-event is also called the failure or survival time. Note that the events studies in survival analysis do not have to be related to survival. The same techniques can be used to study the reliability of products or phenomena such as employment and inflation.

9.2.1 Kaplan-Meier analysis

The survival function $S(t)$ is a function that gives the probability that a subject will survive past time t, or that the event did not occur before time t. Kaplan-Meier (KM) estimators are non-parametric estimators that can be used to model $S(t)$, meaning that they do not assume that the survival times follow a known probability distribution [112]. It is calculated by ranking the survival times of all patients in decreasing order:

For a dataset with N patients, let the observed survival times for each patient be:

$$t_1 \leq t_2 \leq t_3 ... \leq t_N$$

Then, let n_i be the total number of non-censored subjects until time t_i, and d_i the total number of events happening until time t_i. The Kaplan-Meier estimator is then expressed as:

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{(n_i - d_i)}{n_i}, \text{for } 0 \leq t \leq t_N \tag{9.1}$$

The estimation of the survival function $\hat{S}(t)$ is thus a stepwise decreasing function that changes value only at the time of an event. The results of the Kaplan-Meier estimator are typically expressed in a curve which shows the probability of survival over time. These *Kaplan-Meier plots* form a straightforward way to compare survival functions between different patient groups.

Evaluation

The Kaplan-Meier curves of two patient groups can be statistically compared with the log-rank test, in order to verify whether there are statistically significant

differences between the survival functions of both groups [174]. The groups may be for example patients undergoing different treatments or different prognostic groups. The log-rank test is a variant of the chi-square test and measures the observed number of events O_i versus the expected number of events E_i for each time instant. The test statistic is calculated by:

$$X^2 = \sum_{i=1}^{g} \frac{(O_i - E_i)^2}{E_i} \tag{9.2}$$

with g the number of groups. This value is then compared to a χ^2 distribution with $(g-1)$ degrees of freedom, which results in a p-value that expresses the statistical significance.

Another way to compare survival functions between two groups is by calculating the hazard ratio (HR), a measure of the relative survival in the two groups. It is estimated by:

$$\text{HR} = \frac{O_1/E_1}{O_2/E_2} \tag{9.3}$$

It can be interpreted as the chance of an event occurring in group 1 divided by the chance of an event occurring in group 2. An HR of 1 indicates that there is no difference in survival between the groups.

9.2.2 Cox proportional-hazards model

The hazard function $h(t)$ is similar to the survival function, but describes the probability that an individual has an event at time t. It can be estimated with the Cox proportional hazards model (Cox PH model), one of the most well-known and widely used methods to model survival data [50]. While it does not assume a statistical distribution of survival time, its main assumption is the *proportional hazards assumption*: it assumes that the effect of variables on survival time is time-independent, and that the hazards of two subjects are thus proportional over time.

The hazard function of the Cox PH model is as follows:

$$h(t) = h_0(t) * exp(\sum_{i=1,...,D} b_i x_i) \tag{9.4}$$

Here, $h_0(t)$ is the baseline hazard, which is multiplied by a factor for each covariate x_i, for $i = 1, ..., D$. The baseline hazard corresponds to the value of the hazard if all variables x_i are equal to zero. The coefficients $exp(b_i)$ represent the hazard ratios for all parameters. A hazard ratio $exp(b_i) > 1$ implies that as the value of x_i increases, the hazard will increase in proportion and the

probability of survival will decrease. Note that the baseline hazard may change in time while the hazard ratios of the variables are constant.

The coefficients b_i are estimated using partial likelihood estimators, which is similar to the maximal likelihood estimators used in linear and logistic regression.

Evaluation

Cox PH models can be evaluated and compared using the c-index or concordance index, which is defined as a measure of the amount of agreement between the predicted and observed survival [98]. It can be considered as an alternative to the AUC-value for survival models. It is calculated as the ratio of the number of concordant pairs and the number of comparable pairs. Two observations are *comparable* if the ranking of their event times is known. In the context of right-censored data, a pair of events is comparable if both events are non-censored or for a pair with one censored observation and one event, if the time of the event comes before the total observation time for the censored observation. A comparable pair is *concordant* if the survival order as predicted by the model is the same as the observed survival order.

9.2.3 Survival SVM

The general framework of SVM-based classification was described in Section 3.3. It is however also possible to apply SVMs to regression problems, where the desired result is no longer a discrete class label but rather a continous output value. This is referred to as Support Vector Regression (SVR) [73]. Recently, several methods to extend SVM formulations for survival analysis were proposed [77, 196, 217]. We will now explain three different approaches that use either regression constraints, ranking constraints or a hybrid combination of both.

Regression approach

The regression approach directly starts from the SVR formulation and aims at finding a function to estimate the survival time $\tilde{t}$ using covariates $\mathbf{x} \in \mathbb{R}^D$, taking into account censoring information [196].

The standard SVR problem is similar to the classification SVM problem described in Section 3.3.1 [73]. It predicts a continuous value $\hat{y}$ as the linear combination of the kernel-transformed variables $\phi(\mathbf{x})$ and a constant b: $\hat{y} = \mathbf{w}^T\phi(\mathbf{x}) + b$. Similarly to Equation 3.1, this is stated as an optimization

problem where a loss function needs to be minimized, subject to a number of constraints. The constraints are chosen to make sure that the predicted values $\hat{y}$ have at most ϵ deviation from the actual values y_i obtained from the training data, while keeping ϵ as small as possible. This can be written as the following convex optimization problem:

$$\begin{aligned} \min_{\mathbf{w},b,\epsilon,\epsilon^*} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma\sum_{i=1}^{n}(\epsilon_i + \epsilon_i^*) & \\ s.t. \quad & \mathbf{w}^T\phi(\mathbf{x}_i) + b \geq y_i - \epsilon_i, & \forall i = 1, ..., n \\ & -\mathbf{w}^T\phi(\mathbf{x}_i) - b \geq -y_i - \epsilon_i^*, & \forall i = 1, ..., n \\ & \epsilon_i \geq 0 & \forall i = 1, ..., n \\ & \epsilon_i^* \geq 0 & \forall i = 1, ..., n \end{aligned} \tag{9.5}$$

For censored observations, the exact survival time is not known, but is known to be larger than the censoring time. This implicates that predicted survival times that are greater than the censoring time do not need to be penalized. Similarly, predicted times lower than the survival time require penalization as usual. For uncensored observations, the same constraints as in the standard SVR model still hold. The support vector regression model for censored data can therefore be formulated as:

$$\begin{aligned} \min_{\mathbf{w},b,\epsilon,\epsilon^*} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma\sum_{i=1}^{n}(\epsilon_i + \epsilon_i^*) & \\ s.t. \quad & \mathbf{w}^T\phi(\mathbf{x}_i) + b \geq y_i - \epsilon_i, & \forall i = 1, ..., n \\ & -\delta_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq -\delta_i y_i - \epsilon_i^*, & \forall i = 1, ..., n \\ & \epsilon_i \geq 0 & \forall i = 1, ..., n \\ & \epsilon_i^* \geq 0 & \forall i = 1, ..., n \end{aligned} \tag{9.6}$$

Here, survival predictions for censored and non-censored data are penalized similarly. Alternatives which used different penalizations for censored and non-censored data exist [116], but require extra parameters to be tuned which strongly increases computation time [217].

Ranking approach

The goal of the ranking approach is to predict risk ranks between individuals instead of calculating an estimate of the survival time. As such, the aim is comparable to maximizing the c-index described earlier. The methods proposed by Van Belle et al. [214] and Evers and Messow [77] penalize a comparable pair of observations for which the order as predicted by the SVM model is different from the observed order. The model is therefore formulated by Van Belle et al. [214] as:

$$\begin{aligned} \min_{\mathbf{w},\epsilon} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma \sum_{i=1}^{n} \sum_{j:y_i>y_j} \epsilon_{ij} \\ s.t. \quad & \mathbf{w}^T(\phi(\mathbf{x}_{(j)}) - \phi(\mathbf{x}_{(i)})) \geq 1 - \epsilon_{ij}, \forall i = 1, ..., n; \forall j > i \\ & \epsilon_{ij} \geq 0, \qquad \forall i = 1, ..., n; \forall j > i \end{aligned} \tag{9.7}$$

Only comparisons between comparable pairs ($\mathbf{x}_i$, y_i, δ_i) and ($\mathbf{x}_j$, y_j, δ_j) are included. This model will be referred to as *vanbelle1*.

The major disadvantage is that a quadratic programming problem of $\mathcal{O}\left(n^2\right)$ needs to be solved. Therefore, an alternative formulation of this problem, from hereon called *vanbelle2*, was proposed that reduced the number of constraints to $\mathcal{O}(n)$ [215]. The reduction is done by comparing each data point i with only its nearest comparable neighbour $\bar{j}(i)$ instead of comparing it with all available comparable points j:

$$\begin{aligned} \min_{\mathbf{w},\epsilon} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma \sum_{i=1}^{n} \epsilon_i \\ s.t. \quad & \mathbf{w}^T(\phi(\mathbf{x_i}) - \phi(\mathbf{x}_{\bar{j}(i)})) \geq y_i - y_{\bar{j}(i)} - \epsilon_i, \forall i = 1, ..., n \\ & \epsilon_i \geq 0, \qquad \forall i = 1, ..., n \end{aligned} \tag{9.8}$$

Hybrid approach

The hybrid approach combines the regression and ranking models and includes the constraints from both Eq. 9.6 and Eq. 9.8 [217]:

$$
\begin{aligned}
\min_{\mathbf{w},\epsilon} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma \sum_{i=1}^{n} \epsilon_i + \mu \sum_{i=1}^{n} (\xi_i + \xi_i^*) \\
s.t. \quad & \mathbf{w}^T(\phi(\mathbf{x_i}) - \phi(\mathbf{x}_{\bar{j}(i)})) \geq y_i - y_{\bar{j}(i)} - \epsilon_i, \forall i = 1, ..., n \\
& \mathbf{w}^T\phi(\mathbf{x}_i) + b \geq y_i - \xi_i, \qquad \forall i = 1, .., n \\
& -\delta_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq -\delta_i y_i - \xi_i^*, \qquad \forall i = 1, ..., n \\
& \epsilon_i, \xi_i, \xi_i^* \geq 0, \qquad \forall i = 1, ..., n
\end{aligned}
\tag{9.9}
$$

All survival SVM models explained above are implemented in the R-package *survivalsvm*, which can be used to fit and evaluate several survival models [81]. It includes four different kernels that can be used with all models: linear, additive, RBF and polynomial. As with Cox PH models, different models can be compared by calculating and comparing their respective c-indices.

9.3 Data

The dataset used in this Chapter consists of data from the same patients used in the previous Chapter. It contains 616 patients who were followed in the University Hospitals Leuven between October 1996 until December 2013, excluding patients with pacemakers or abnormal heart rhythms such as atrial fibrillation. Demographic, clinical and electrocardiographic data before the first implantation of the ICD were collected from their electronic medical record. The dataset has no variables with missing values. The ECG signals from all patients were used to calculate an fQRS score for each lead with the method described in Chapter 8.

The full dataset is composed of more than 20 clinical variables including QTc duration, age at the moment of ICD implantation, clinical history,... From this list, nine features were selected to construct a baseline clinical model. The model includes age, gender, etiology of cardiac disease, LVEF, the presence of left bundle branch block, renal function and use of beta-blockers, ACE-inhibitors and anti-arrhythmic drugs. The included features were selected because they

were pointed out as strong prognostic factors in either previous examination of this patient group or large population studies.

The primary endpoints included in the database are first appropriate shock and all-cause mortality. Endpoints were collected until 31-12-2014. The endpoint considered here is all-cause mortality at 7.5 years (the mean follow-up time of all patients). Out of the 616 patients, 126 died during the follow-up period.

9.4 Methods

In this Chapter, we analysed the available data in two separate ways. The first method is based on the Kaplan-Meier analysis and aims at verifying if the proposed fQRS score can be used to dichotomize the patients in a high- and low-risk group. In the second part, the goal is to verify whether the it can improve the accuracy of risk prediction when combined with survival SVM methods.

9.4.1 Optimal cut point determination

The first method aims to examine whether the proposed fQRS score can be used to divide the patients in a low- and high-risk group. A second goal is to analyse which channels are more useful to predict all-cause mortality. We dichotomized the fQRS score of each lead by determining an optimal cut point θ_{ch} for each channel to distinguish a high-risk and low-risk group. The patients were first divided in a training and test set: $\frac{2}{3}$ of the patients were used to determine the cut points, the remainder was used to test the results. Training and test groups contained equal ratios of censored and non-censored patients.

Since the focus here is to determine the usability of the fQRS score for risk assessment, only *univariate* analyses were considered here. Kaplan-Meier analysis was used to calculate survival curves, risk tables and hazard ratios [112]. Statistical differences between the curves of two groups were analysed with log-rank tests.

In order to get a robust estimate of the optimal cut point that is less dependent on the choice of training set, we used *bootstrapping* to generate 2500 bootstrap samples. Each bootstrap sample was drawn from the training set. Optimal cut points for each channel ch and each bootstrap sample i were determined with the minimum p-value approach: for each possible threshold a log-rank test was performed, and the threshold which generated the lowest p-value was selected as cut point $\hat{\theta}_{i,ch}$.

The optimal cut point θ_{ch} for each channel was then defined as the median of the cut points $\hat{\theta}_{i,ch}$ of all bootstrap samples. 95% confidence intervals (CI) for the median were calculated as described in [146]:

$$\text{CI} = \theta_{ch} \pm 1.7 \frac{1.25R}{1.35\sqrt{N}} \tag{9.10}$$

with R the interquartile range and N the number of bootstrap samples.

The optimal cutpoints θ_{ch} were finally used to dichotomize the test set and construct the corresponding survival curves. Differences were again evaluated with the log-rank test, with $p<0.05$ considered statistically significant.

9.4.2 Comparison of survival models

A second part of this chapter consists of the comparison and evaluation of different methods to estimate a survival model, in order to evaluate whether machine learning approaches provide added value compared to standard statistical methods. Furthermore, the standard clinical baseline model described in Section 9.3 was compared to the baseline model combined with fQRS scores to examine the utility of using the proposed score in survival analysis.

The Cox PH model was considered as reference model, since it is the most commonly used multivariate approach in medical research [29]. It was compared with the four survival SVM models described in Section 9.2.3 (regression, vanbelle1, vanbelle2 and hybrid) using linear, clinical and RBF kernels. The clinical kernel was developed specifically for use with clinical data and makes a distinction between continuous, ordinal and nominal variables [51].

For all models, parameters were tuned with 5 × 10-fold nested cross-validation: the data were randomly divided in five equally-sized subsamples, and in each iteration, one of the five samples was used as independent test set. The models were then trained on the remaining samples using 10-fold cross validation. The c-index was used to evaluate all models. To assess whether the model including both the clinical variables and fQRS scores is significantly better than the baseline model, the methods from Hanley and McNeil [97] were used. They define a critical ratio z, calculated as:

$$z = \frac{c_1 - c_2}{\sqrt{SE_1^2 + SE_2^2 - 2rSE_1SE_2}} \tag{9.11}$$

where c_1 and SE_1 refer to the average observed c-statistic and estimated standard error of the first model, c_2 and SE_2 describe the second model and r stands for the correlation between c_1 and c_2. The standard error is calculated as

	Training set		Test set	
Channel	$\boldsymbol{\theta_{opt}}$	**95%C.I.**	**p-value**	**HR**
I	0.47	0.4574-0.4856	0.484	1.25
II	0.24	0.2239-0.2560	0.112	1.76
III	0.44	0.4368-0.4431	0.887	0.96
aVL	0.43	0.4189-0.4410	0.589	1.18
aVF	0.65	0.6355-0.6644	0.464	0.80
V1	0.52	0.5137-0.5263	0.757	1.15
V2	0.66	0.6531-0.6669	0.218	1.58
V3	**0.58**	**0.5633-0.5967**	**0.014**	**2.23**
V4	**0.25**	**0.2446-0.2553**	**0.028**	**1.98**
V5	0.77	0.7671-0.7728	0.102	1.73
V6	**0.68**	**0.6690-0.6910**	**0.047**	**1.94**

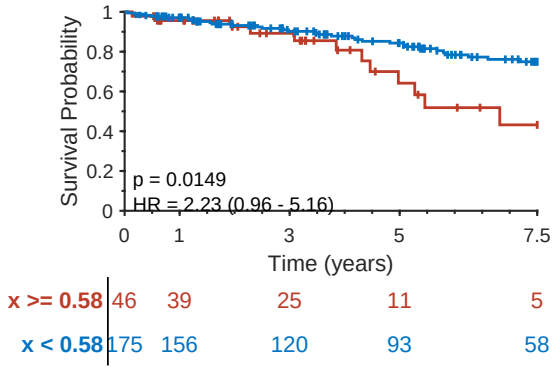
Table 9.1: Optimal cut points, 95% confidence intervals and results on the test set for each channel.

the standard deviation divided by the square root of the number of observations. The z-statistic can then be used to derive a p-value, and $p < 0.05$ was considered significant.

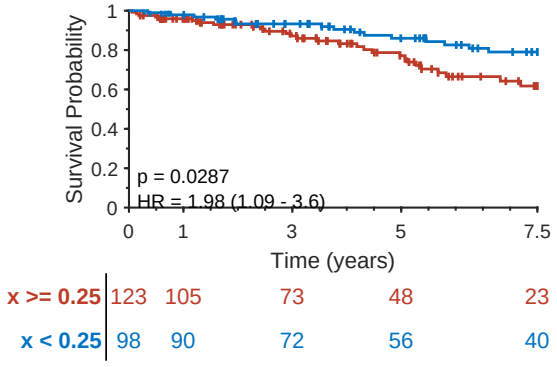
9.5 Results

9.5.1 Optimal cut point determination

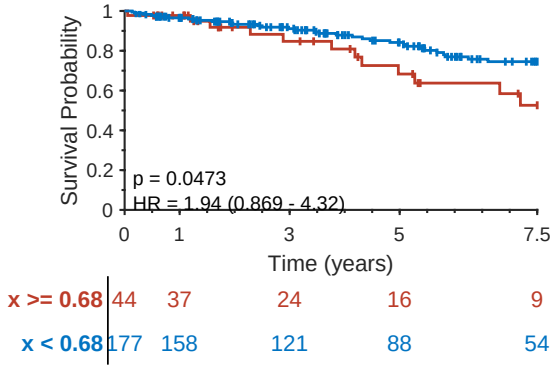
Table 9.1 shows the value of the optimal cut point and confidence interval (determined on the training set) for each channel and the results on the test set. The optimal threshold varied greatly between channels: the lowest threshold is reached for channel V4 (= 0.25), the highest for channel V5 (= 0.77). For three channels, applying the optimal cut point on the independent test set lead to statistically significant differences in survival times between both groups: V3, V4 and V6. The corresponding Kaplan-Meier curves for these channels can be seen in Figure 9.1. They include risk tables and hazard ratios. Two additional channels, II and V5 show notable trends (p $\approx$ 0.1).



(a) Channel V3



(b) Channel V5



(c) Channel V6

Figure 9.1: Kaplan-Meier plots of channels with statistically significant differences ($p < 0.05$) between survival curves. The risk tables represent the number of patients alive at different time instances for fQRS scores higher and lower than θ_{ch}.

		Baseline model		Baseline + fQRS	
Method	**Kernel**	**C-index**	**Rank**	**C-index**	**Rank**
vanbelle1	linear	$\mathbf{0.59 \pm 0.04}$	**10**	$\mathbf{0.66 \pm 0.08}$	**3***
	additive	0.58 ± 0.04	11	0.58 ± 0.05	11
	RBF	$\mathbf{0.60 \pm 0.05}$	**8**	$\mathbf{0.65 \pm 0.03}$	**4***
vanbelle2	linear	0.62 ± 0.06	7	0.62 ± 0.10	7
	additive	0.57 ± 0.05	12	0.56 ± 0.08	13
	RBF	0.57 ± 0.04	12	0.6 ± 0.10	10
regression	linear	0.66 ± 0.04	3	0.69 ± 0.07	2
	additive	0.65 ± 0.06	5	0.58 ± 0.14	11
	RBF	0.64 ± 0.05	3	0.65 ± 0.09	4
hybrid	linear	$\mathbf{0.65 \pm 0.05}$	**5**	$\mathbf{0.70 \pm 0.07}$	**1***
	additive	0.67 ± 0.07	1	0.62 ± 0.11	7
	RBF	NA	-	NA	-
Cox PH model		0.67 ± 0.07	1	0.65 ± 0.04	4

Table 9.2: Results of the different survival models using only clinical parameters (baseline model) or the combination of clinical parameters and fQRS scores (baseline + fQRS). C-indices from the five subsamples are reported as mean and standard deviation, together with the overall rank for each model. Models where the combination of baseline and fQRS resulted in significantly increased c-index ($p<0.05$) are indicated with *.

9.5.2 Comparison of survival models

The results of the different survival models are summarized in Table 9.2. The c-indices for all models are reported as the mean value and standard deviation obtained from the results of the subsamples. Each model also received a rank based on the average c-index, with 1 assigned to the best-scoring model and 13 to the worst-scoring method. The result for the hybrid model with RBF kernel is not mentioned since calculation time for this classifier exceeded acceptable lengths, and the computation was interrupted after five days.

The results for the clinical baseline model were compared with the models combining clinical parameters and fQRS scores. Models where adding fQRS scores resulted in significantly higher c-indices ($p<0.05$) are marked in bold.

This was the case for three models: *vanbelle1* with linear and RBF kernel and the hybrid model with linear kernel. The other models with linear and RBF kernel also showed increases in c-index after adding fQRS scores, but the increases were not significant. The SVM models with additive kernels often resulted in a lower c-index when fQRS scores were included. A similar result can be seen for the Cox PH model, where the c-index decreased from 0.67 to 0.65.

For the clinical baseline model, the Cox PH model and the hybrid model with additive kernel obtained the highest rank, corresponding to a c-index of 0.67. For the combined model, different SVMs scored better than Cox PH. Here the hybrid model with linear kernel resulted in an average c-index of 0.70.

9.6 Discussion

In this Chapter, the potential to use the fQRS score as prognostic risk factor for all-cause mortality was explored. We analysed this in two different ways, using a combination of standard statistical models and SVM-based machine learning algorithms.

In the first part of this study, the fQRS scores in different channels were used to divide a dataset of ICD patients into two groups in order to assess their risk on all-cause mortality. Dichotomization of continuous variables in survival analysis is a debatable subject since the choice of optimal cut point should be done in a way that results can be generalized. The use of bootstrapping on the training set and evaluation of the performance on an independent test set however ensures that the optimal cut points determined here are minimally dependent on the choice of training set.

Results on the test set indicate that the fQRS score in three different channels (V3, V4 and V6) can be used as an indication of the risk on all-cause mortality in ICD patients. Hazard ratios derived from the Kaplan Meier plots shown in Figure 9.1 are approximately 2 (1.94-2.23). This means that the probability of all-cause mortality for patients with fQRS in these channels is roughly double the probability for other patients. Similar conclusions can be drawn from the risk tables. The table for channel V3 shows that only 5/46 patients (10%) with a fQRS score larger than 0.58 in V3 is alive at the end of the analysis period compared to 58/175 patients (33%) with fQRS score lower than the threshold.

V3 and V4, the channels with the lowest p-values are located in the anterior regions of the heart. This corresponds with findings in [223], where the presence of fQRS in anterior channels was an independent risk factor for mortality in a

subset of the same patient population. This confirms that the proposed method to detect fQRS is not only capable of reproducing the results of visual fQRS detection, but also reaches results that are comparable to clinical studies.

Kaplan-Meier analysis is an example of univariate analysis, where the effect of one variable at a time is examined. While this is valuable to confirm the potential of the fQRS score, using the parameter in a multivariate method such as the Cox proportional-hazards model is more meaningful, since risk assessment is mostly done by combining multiple parameters. This was done in the second part of this Chapter, where the fQRS scores of the different channels were added to a baseline model with nine clinical parameters.

The results in Table 9.2 show that the fQRS scores do not improve the result of the Cox PH model, and even lead to a slight decrease in the average c-index. This can be explained by the large number of variables: the baseline model has 9 variables, the fQRS score adds 11 extra features resulting in a model with 20 variables in total. In Harrell et al. [98], a rule of thumb is defined that states that "in order to have predictive discrimination that validates on a new sample, no more than $m/10$ predictors should be examined,..., where m is the number of uncensored event times in the training set" [98]. When too many variables are included, the risk on overfitting increases. Here, the total number of events in the dataset was 126 (see Section 9.3). This is sufficient to fit a survival model with 9 variables, but clearly not enough for a model that includes 20 covariates.

SVMs are widely used to deal with high-dimensional data in classification or regression problems. It has been shown that survival SVMs show the same characteristic and also have good applicability and performance in problems involving high-dimensional data [212]. Table 9.2 confirms that this finding also holds for this dataset: for many of models based on survival SVM, adding fQRS scores to the baseline model does increase the c-index of the model.

There are however large differences between the different SVM models and even more striking contrasts between the different kernels. The additive kernel consistently leads to poor results for the baseline + fQRS model, and even leads to decreased performance in most cases. One possible explanation is that the additive kernel was designed specifically to be used with clinical data, which often consist of binary, ordinal or categorical variables [51]. The baseline model uniquely contains clinical variables, and here the additive kernel obtains good results: the hybrid SVM model with additive kernel leads to the highest c-index. The fQRS scores are however all continuous variables, which are not targeted specifically by the additive kernel. The combined baseline + fQRS feature set therefore contains mostly continuous features, and can therefore perhaps be better modelled by linear and RBF kernels. Similar findings are reported in Van Belle et al. [213], where the additive kernel led to good models in many

datasets, but performed worse than linear kernels in datasets where the number of continuous variables greatly exceeded the number of categorical variables.

The best overall performance was obtained by the hybrid SVM model with linear kernel, using the combination of the baseline model and fQRS scores. This model obtained an average c-index of 0.70, which is considered a good model. This indicates that both the proposed fQRS score and the use of machine learning models have added value in assessing the risk on cardiac mortality. The addition of extra features and the use of SVM models however come at the cost of increased computation time. Practically, computation time was mostly influenced by the number of parameters that needed to be optimized. The hybrid model needs optimization of an extra hyperparameter and the RBF kernel also requires tuning of an extra kernel parameter. The hybrid model with RBF kernel therefore needed an enormous calculation time, and computation was ceased after five days. This is an important consideration to keep in mind when determining the optimal approach.

Clinically, the presence of fQRS in a cardiac region is determined based on a combination of channels rather than single channels. Combining scores per cardiac region is therefore a logical extension of this study. This can be done by simply summing scores of individual channels or by more advanced machine learning techniques. For instance, Interval Coded Survival methods [218], which are based on SVMs and capable of modelling both linear and non-linear trends in data in an interpretable way could be used for this purpose. ICS models have been extended to deal with censored data [216], which would be very useful in this application.

Finally, the only endpoint considered here was all-cause mortality at 7.5 years. Results presented in [223] and [222] however indicated that fQRS is even more strongly associated with the risk on appropriate ICD shocks. Applying the methods proposed on this Chapter to alternative endpoints could therefore potentially lead to even stronger findings.

9.7 Conclusion

This Chapter aimed at assessing whether the fQRS score proposed in Chapter 8 could be used to predict the risk of all-cause mortality in a dataset of ICD patients. A second aim was focused on exploring the added value of using machine learning techniques based on SVM in survival analysis. We presented two different ways to tackle these goals. Both methods showed the potential of the fQRS score as a prognostic risk factor, and the second method furthermore

showed that several SVM models add significant performance compared to the Cox proportional hazards model generally used in clinical research.

In the previous Chapter, we showed that the proposed fQRS score performs well in both detection and quantification of QRS fragmentation. The additional results presented here further indicate the power of our proposed approach, and indicate that the final fQRS score can have significant added value in prediction of all-cause mortality. This novel way of detecting and quantifying QRS fragmentation is therefore a promising way to promote the clinical usefulness of the parameter.

Chapter 10

Conclusions and Future directions

In the introduction, the three major research objectives for this thesis were defined as 1) the development of algorithms for extraction of SCD risk factors, 2) exploration of the use of tensors in ECG analysis and 3) risk assessment. This final Chapter summarizes the conclusions related to each of these principal goals in Section 10.1. Finally, Section 10.2 provides suggestions for potential future research directions.

10.1 Conclusions

Sudden cardiac death is one of the main causes of death worldwide with a severe socio-economic impact. Patients at risk for SCD can often be treated to prevent arrhythmic events, but current risk assessment strategies are far from ideal, as described in Chapter 1. The development of novel ECG analysis methods that can extract risk factors in an automated and objective way has the potential to improve the current risk stratification. Therefore, we have focused the research in this PhD on three main goals:

1. Develop algorithms for extraction of SCD risk factors
2. Explore the use of tensors in ECG analysis
3. Assess the mortality risk in ICD patients using novel machine learning methods

The methods presented in the previous Chapters all contributed to one or more of these research objectives. The next Sections summarize the major findings related to each objective and list the most important conclusions.

10.1.1 Algorithms for SCD risk factor extraction

The first objective of this PhD was the development of algorithms to extract features from the ECG that can potentially assess the risk on sudden cardiac death. Section 1.3 listed a number of risk factors currently used in SCD risk prediction. Chapters 4–8 each presented a method to extract and analyse a different risk factor from the multilead ECG.

The method presented in Chapter 4 aimed at separating normal and abnormal beats in an unsupervised manner. Results on two separate public datasets indicate that while the method does not succeed in detecting atrial beats, the performance for ventricular beat detection is comparable or better than state-of-the-art unsupervised approaches. Since ventricular arrhythmia are mostly initiated by this type of irregular beat, and the frequency of PVCs is thought to be useful as prognostic risk factor, this is nevertheless an important result. While the method is not suitable as general irregular heartbeat detection, it would be effective as PVC detection algorithm. The combination of the morphological tensor-based features could furthermore be combined with heart rate variability features in a supervised classifier to further improve performance.

Chapter 5 aimed at detecting and quantifying T wave alternans. The proposed method is capable of dealing with changing heart rates through the use of PARAFAC2. Since in clinical practice TWA is often detected with a bicycle test, this is an important practical consideration. The extensive validation on multiple databases indicates that the method is capable of producing precise results in many different circumstances.

The study described in Chapter 6 did not focus on one particular risk factor but rather aspired to characterize changes in overall ECG morphology prior to in-hospital cardiac arrest. Different features that have the potential to distinguish different types of cardiac arrest were identified and could be used in the future for patient monitoring. Further validation of the results with a control database containing patients in the intensive care *without* cardiac arrest is however required to confirm that the obtained results can be generalized to other risk groups.

Atrial fibrillation is a frequent arrhythmia that is not immediately fatal but that is nevertheless linked to an increased risk of SCD. In Chapter 7 we studied different ways to detect AF in single and multilead ECG signals. The proposed

algorithm successfully managed to identify AF patterns in both signal types using a combination of HRV and morphological features. In the final step, the coupling of single and multilead databases was explored which permits the combination of clinical and wearable ECG signals. As digital health becomes increasingly important, incorporating signals of wearable devices in clinical practice will become more and more important, making the method a promising way to analyse signals from mobile devices in the future.

The final risk factor considered in this work is QRS fragmentation. A method to detect and quantify fQRS was proposed in Chapter 8. The quantification of fQRS is a novel approach to examining this feature that can however be perceived as more realistic as fQRS is not necessarily a binary phenomenon. We showed that the developed fQRS score correlates closely to visual scores given by five experts and that the method detects signals with certain fQRS with high accuracies. Furthermore the extensive dataset that was collected for this purpose can be used to advance future research in this area.

10.1.2 Tensors in ECG analysis

The second goal of the thesis was to explore the use of tensors in cardiac applications. This objective was tackled in the different Chapters contained in Part II, and some of the conclusions were already summarized in the previous Section. Here we will discuss the main findings related to the tensor part.

Tensors are a rather novel concept in ECG signal processing, as illustrated by the limited number of tensor-related cardiac applications described in Section 2.5. The development of the tensor-based methods described in Part II can therefore be considered as one of the main achievements of this research. We showed that tensors can be applied in a straightforward way on different types of real-life ECG signals to extract clinically meaningful components. We furthermore highlighted the flexible nature of different tensor decompositions that allowed to adapt the decompositions to the characteristics of the signals at hand. The proposed framework can easily be extended to other types of applications, opening the door for the development of additional tensor-based approaches.

Chapters 4–6 used Canonical Polyadic Decomposition as main building block to decompose and analyse the ECG signal. Although CPD can be considered as one of the more simple tensor methods, we have showed throughout the dissertation that it is nevertheless capable of reaching strong results. CPD decomposes a tensor in a sum of rank-1 components. Since the structure of the components is fixed, the decomposition rank is the only parameter that requires optimization. This makes the method straightforward to use. Additionally, CPD mostly produces results that are easy to interpret and can be linked to physiology. In

biomedical signal processing, where the result of tensor decomposition is used afterwards for clinical analyses, this is a noteworthy advantage.

CPD is a widely-used method, and many adaptations exist that can be used to adjust algorithms to specific aspects of the problem under analysis. In Chapter 5 for example, we showed that allowing variations in the factor vectors of one mode lead to improved results in real-life signals where the heart rate might change over the course of a signal. Additionally, the use of weighted tensor decompositions was explored in Chapter 6, where incorporating prior knowledge about the signal quality in the cost function led to more robust analyses in the presence of heavy noise.

While the simplicity of CPD has many benefits, it also comes with the disadvantage of reduced flexibility. The algorithm is very powerful for assessment of global features of the ECG, but is less optimal for a lead-per-lead analysis. In those cases, other approach like the multilinear singular value decomposition are better suited. This was shown in Chapter 7, where we used the MLSVD in various ways for atrial fibrillation detection. Compared to CPD, a multilinear decomposition rank however had to be determined which requires the use of more advanced optimization techniques.

10.1.3 Risk assessment in ICD patients

Finally, the third objective was to assess the cardiac risk in ICD patients using machine learning methods. This objective was the focus of Chapter 9, where the goal was to use the fQRS score to assess the risk of all-cause mortality.

We showed that machine learning methods can provide added value to risk assessment in two different ways. First, the fQRS score was calculated by using a list of feature vectors as input to an SVM classifier. Adding these objectively determined features to a list of clinical features improved the accuracy of most survival models. Second, we also used survival SVMs to build the survival model itself. Comparison with the standard Cox proportional hazard model indicated that SVMs also provide added value here: the best performance was reached with a hybrid survival SVM model, combining regression and ranking constraints. Especially in high-dimensional feature spaces, where standard statistical models are easily overfitted, the advantage was clear.

The major drawback of using survival SVMs to build a survival model is the increased computation time. Depending on the constraints and kernels used, one to three SVM and kernel parameters have to be optimized in order to to train the classifier. Especially for the hybrid SVM with RBF kernel, which has three additional parameters, the training process lasted excessively long and

had to be interrupted after five days. This is an important obstacle to keep in mind, in particular when working with extensive datasets that contain a large number of variables.

10.2 Future directions

10.2.1 Algorithms for SCD risk factor extraction

The algorithms presented in this work focused on the detection of specific risk factors. However, many more potential features exist that can be determined using the ECG signal. A straightforward future direction therefore consists of the development of additional algorithms to extract more potential risk factors

The majority of the methods presented in this work can be used to detect other irregularities after minor modifications. Depending on which part of the signal is used to construct the tensor, different ECG segments can be analysed. Changes in the ST segment, which is known to change after ischemia, could for example be examined by constructing a tensor which contains different ST segments, allowing analysis of ST segments in time.

One of the main bottlenecks for algorithm development is however the availability of high-quality labelled databases. While many publicly available databases exist, the quality of both the signals and annotations is often not sufficient to perform accurate validations. Furthermore, many of them have been recorded many decades ago with outdated equipment. Making more datasets publicly available would benefit the development of novel algorithms, since comparison with a golden standard is an essential step in the development process.

The algorithms described in this thesis are all available as Matlab-based algorithms. In order to be used in a practical manner in clinical practice, it is necessary to present both the methods and results in a user-friendly way in graphical user interfaces. Recently, a graphical user interface for QT interval analysis was proposed, which contains the irregular heartbeat detection method described in Chapter 4 as one of the building blocks [203]. The other methods proposed in this thesis could be included in a similar way, leading to a comprehensive toolbox that can be used for different aspects of ECG analysis.

Finally, the development and continuous rise of novel ECG technologies will lead to an enormous expansion of data that will become available in the coming years. Many of these signals will be measured by wearable devices, which will bring on very specific challenges in ECG analysis. The development of ECG patches that can be worn for many days in a row will allow patients to be followed

closely for many days or even weeks in a row. On the other hand, including ECG measurement technologies in wearable devices such as smart watches gives patients the opportunity of recording short-term ECG recordings whenever they have unfamiliar sensations. While these short-term and long-term signals are very different from each other, they can both be used to assess the risk of sudden cardiac death on an even larger scale. The development of dedicated ECG analysis methods that can handle these immense amounts of signals and process them in a efficient and fully automated way will surely facilitate the extraction of features from these novel data sources.

10.2.2 Tensors in ECG analysis

The tensor-based methods proposed in this research primarily used CPD as building block for further analysis. CPD leads to interpretable components in the form of factor vectors which were used to detect useful features. In most algorithms however, only one feature vector corresponding to one tensor mode was used for further analysis. It would however be interesting to analyse the feature vectors corresponding to the other modes since they might also contain valid diagnostic information that is not considered now. An example is the application handled in Chapter 6, where changes in ECG morphology prior to cardiac arrest were characterized. The analysis uniquely focused on the morphology that could be examined from the second factor vector. The first and third vectors, equivalent to the spatial mode and heartbeats mode, could however be used to analyse how the signal changes in space and whether there is increased variance between different heartbeats in a window.

CPD is a powerful method which can be easily adapted to deal with specific signal characteristics. It is however also limited in the sense that the rank of the components is restricted to 1. Exploring the use of more advanced tensor decompositions in cardiac applications is therefore an additional suggestion for future research. A block term decomposition approximates a tensor by a sum of low multilinear rank terms [58]. The ranks are however not restricted to 1 here, allowing a greater flexibility. A special case is the $(L, L, 1)$-decomposition, which has been successfully used in a number of EEG applications [105, 243]. Furthermore, exploring the use of tensor decompositions for features that are detected in a single channel is another topic that was not handled in this dissertation but that is a logical extension of the presented work. The detection of QRS fragmentation is for example done for each lead individually, and tensor methods could be applied here as well.

The use of tensor decompositions for single lead ECG signal is another future direction that has a lot of potential. As more and more wearable devices are

equipped with ECG recording devices, the amount of single lead ECG signals is expected to increase drastically in the next years. Singe lead ECG signals require an extra tensorization step prior to tensor decomposition methods. In Chapter 7, the discrete wavelet transform was used to transform the signals to a tensor, which did unfortunately not create any added value. Exploring different tensorization techniques could however improve results. Löwnerization is an example of a tensorization that enable separation of signals into rational functions. It has been successfully applied for the separation of fetal and maternal ECG and could therefore be further applied in additional cardiac applications [65].

The use of tensor updating was briefly explored in Chapter 6. It is a promising approach for most real-time monitoring methods, where an immediate output is required. Here, tensor updating was applied in a very straightforward way, by updating the result of the tensor decomposition for each new heartbeat. The updating could however be taken even further, by for example not only updating the results of the tensor decomposition but also updating the rank of the decomposition as the signals are changing over the course of time.

Tensors can also be applied for data fusion, where signals from different sources are analysed simultaneously. While joint analysis of measurements from different signal types has not been explored in this thesis, it has the potential to reveal additional features that describe the *interaction* between different biosignals. It is for example known that the cardiorespiratory coupling, which describes the influences of respiration on heart rate and blood pressure, is different for healthy and sick people. Using data fusion approaches on the combination of ECG and respiration or blood pressure would open the door to the development of algorithms to analyse a whole new class of biomarkers, potentially leading to new insights in the mechanisms of sudden cardiac death.

10.2.3 Risk assessment in ICD patients

In Chapter 9, the fQRS score was used to assess the risk of all-cause mortality. Including additional risk factors calculated using the algorithms from Chapters 4–7 is a logical future step in order to examine whether the other methods developed in the context of this PhD can be used in a similar way. Additionally, using extra endpoints would be interesting to verify whether the other findings from Vandenberk et al. [223] related to appropriate ICD shocks could be replicated.

Chapter 9 also presented a method to determine optimal cut-off values to dichotomize the patient group in a low- and high risk group. There is however no reason to restrict this analysis to one cut-off threshold: If multiple cut-offs

can be determined that divide the patients in multiple risk groups, this might lead to a more accurate partition of the patient group. The ultimate goal is to determine a continuous risk score that presents the SCD risk as a value between 0 and 1. This could be done through the use of interval-coded scoring models [218], which are models that calculate a risk score using SVMs and that present the output in an interpretable way. Since adaptations for survival models have been developed [216], they can be straightforwardly applied in future research.

Finally, validation of the results on larger independent datasets is required to ensure that the developed models are also applicable on different datasets and that the findings can be replicated in patients from other hospitals. Together, the collection of algorithms developed in this research present different ways to analyse the ECG signals in an innovative way. Including the proposed features in larger clinical risk prediction models can potentially open the door to new insights in clinical diagnostics and sudden cardiac death prediction.

Appendix A

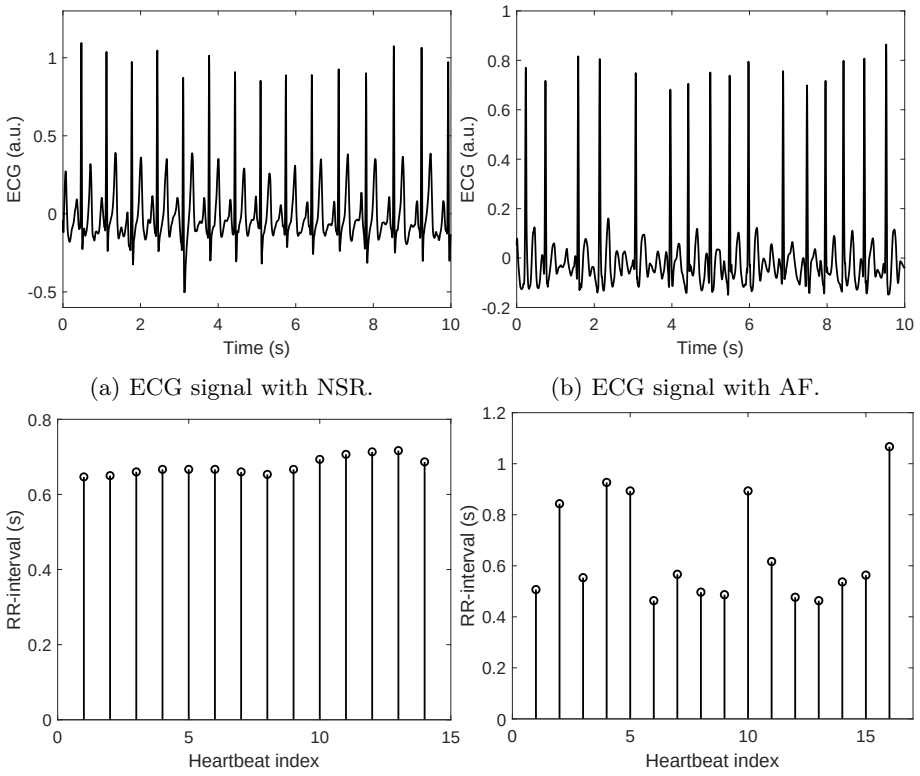
List of Heart Rate Variability features

The method for detection of atrial fibrillation proposed in Chapter 7 combines morphological features (calculated using tensor-based approaches) with HRV features. The included HRV features were selected after a review of the available literature. This Appendix first describes the features used here in Section A.1, and analyses the correlation between the different features in Section A.1.2. All results and Figures in the Appendix are taken from [83].

A.1 Feature Description

One of the main characteristics of atrial fibrillation is the irregular contraction of the atria. Features that characterize the heart rhythm are therefore a popular tool to describe and detect AF. A (visual) tool to analyse the RR intervals in time is the tachogram. Figure A.1 shows a normal ECG signal with corresponding tachogram on the left hand side. The right hand side shows an example of an AF signal and tachogram. It is clear that the variability in the latter is bigger.

This section explains some of the most widely-used HRV features that can make a distinction between AF and NSR. They are divided in statistical, geometrical and non-linear characteristics. The difference between normal and AF signals is illustrated with the Physionet/CinC Challenge 2017 dataset, described in Section 7.2.1. The different features are calculated for all normal and AF signals to emphasize the differences between both rhythms. Afterwards, a



(a) ECG signal with NSR.

(b) ECG signal with AF.

(c) Tachogram of NSR signal.

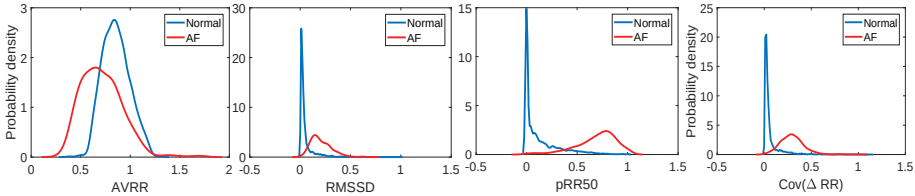
(d) Tachogram of AF signal.

Figure A.1: An ECG signal with NSR (left) and AF (right) with their corresponding tachograms. There is clearly a larger variability in the RR intervals of the AF signal [83].

non-parametrized probability distribution for each class is defined by a kernel density estimator. This produces a probability estimate from the sample data, similarly to a histogram. The result is however not divided into discrete bins, but rather smoothed to obtain a continuous probability curve [76].

A.1.1 Statistical features

The first class of HRV features consists of statistical measures, derived from the statistical moments of RR intervals and ΔRR-intervals (defined as differences between RR intervals) [194, 117].

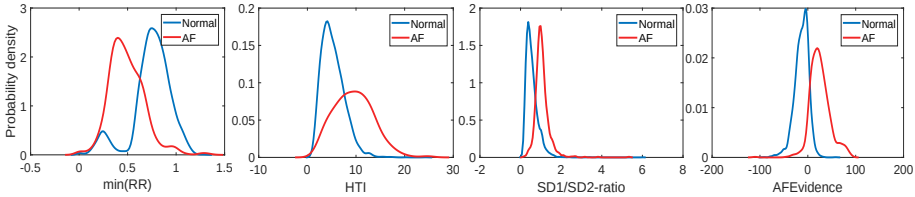


(a) The average RR interval. A slight overlap can be noticed.

(b) The *Root Mean Square of Successive Differences*. The short-term variability is larger for AF.

(c) The fraction of consecutive RR-intervals that differ more than 50 ms. A strong discrimination can be achieved.

(d) The coefficient of variance of the ΔRR-values. This parameter correlates strongly with the RMSSD.

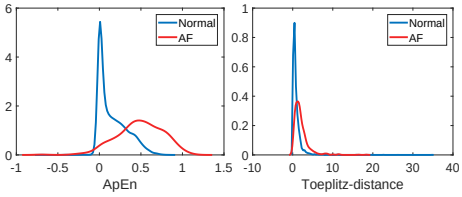


(e) The minimal RR-interval. Signals with AF often contain closely-spaced heartbeats.

(f) The *HRV Triangular Index*. The RR-intervals are more strongly concentrated in the NSR histogram.

(g) The SD1/SD2-ratio. It is clear that the RR-intervals are less predictable in the case of AF.

(h) AFEvidence. There is a large difference between both classes.



(i) *Approximate entropy*. The complexity is significantly higher for AF.

(j) The Toeplitz-distance. This feature is also higher for AF, although the difference is less pronounced.

Figure A.2: The estimated non-parametric probability distributions for all ten HRV-features calculated using the Physionet/CinC Challenge 2017 dataset [83].

AVRR

One of the most straightforward features is the average RR interval length (AVRR) since AF is associated with increased heart rates [82]. Figure A.2a shows the probability distributions of AVRR for normal (blue) and AF signals (red). While both distributions show some overlap, discrimination is possible.

RMSSD

The *Root Mean Square of Successive Differences* (RMSSD) calculates the RMS value of the ΔRR-values and gives an indication of short-term variations in the heart rate [194]:

$$\mathrm{RMSSD} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}\Delta\mathrm{RR}_i^2} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(\mathrm{RR}_{i+1}-\mathrm{RR}_i)^2}, \qquad \text{(A.1)}$$

with N the number of RR intervals in the signals. Figure A.2b shows that the RMSSD is higher for AF signals.

pRR50

pRR50 measures the fraction of subsequent RR intervals that differ more than 50 ms:

$$\mathrm{pRR50} = \frac{\#(\Delta\mathrm{RR}_i > 50\,\mathrm{ms})}{N-1}. \qquad \text{(A.2)}$$

This feature strongly correlates with RMSSD [194]. Figure A.2c shows a large difference between the feature values of AF and normal signals.

In principle, only normal heartbeats are used to calculate this feature. This however requires a separate heartbeat classification stage which was beyond the scope of this study. Figure A.2c indicates that this feature can also discriminate between these two classes without extra classification.

CoV(ΔRR)

Another popular feature is the coefficient of variance (CoV) of the ΔRR-intervals (CoV(ΔRR)) [208]:

$$\mathrm{CoV}(\Delta\mathrm{RR}) = \frac{\sigma_{\Delta\mathrm{RR}}}{\mathrm{AVRR}}, \qquad \text{(A.3)}$$

with $\sigma_{\Delta RR}$ the standard deviation of the ΔRR-intervals. Figure A.2d indicates another strong correlation with RMSSD with similar large differences between both classes.

min(RR)

A non-linear statistical feature which is not strongly correlated with the previous features is the minimal RR interval length (min(RR)). Figure A.2e indicates that in signals with AF the shortest difference between RR intervals is smaller than in normal signals.

Other statistical features

Many other statistical features exist, such as the standard deviation of the RR intervals, the number of RR intervals *larger* than 50ms, the difference between the minimal and maximal heart rate, ... They were not included here since they show very strong correlations with some of the previous features or can only be measured confidently on long-term signals [194].

A.1.2 Geometric and other non-linear characteristics

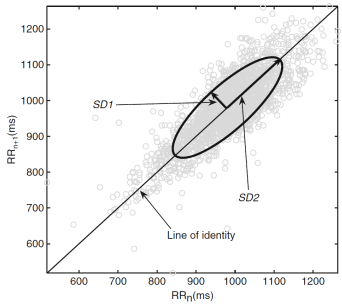
Geometric features represent and describe the heart rate in a geometrical way [117]. Popular techniques are the RR-interval histogram and Poincaré plots.

HTI

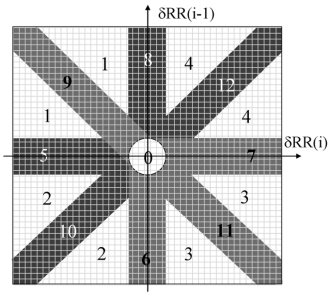
The *HRV Triangular Index* (HTI) is defined based on the RR-interval histogram as the inverse of the relative frequency of the most frequent bin [194, 117, 234]:

$$\mathrm{HTI} = \frac{N}{f_{\max}}. \tag{A.4}$$

It describes how concentrated the RR-interval histogram is, i.e. how many of the RR-intervals are located in the same bin. Based on the probability distributions shown in Figure A.2f, we can conclude that for normal signals the histogram of the heart rate is more concentrated.



(a) The Poincaré plot of RR intervals. For a normal signal, the RR intervals are clustered around the first bisector. (Figure from [117])



(b) Poincaré plot of the ΔRR-intervals, with projected 2D-histogram. (Figure from [190])

Figure A.3: Poincaré plots of the RR- and ΔRR-interval required for calculation of the SD1/SD2-ratio and AFEvidence.

SD1/SD2-ratio

A popular tool in non-linear HRV analysis is the Poincaré plot, which plots the RR intervals compared to the next value (e.g. RR_n versus RR_{n+1}). Figure A.3a shows an example of a Poincaré plot of an ECG signal. For normal signals, the cloud is concentrated around the first bisector. To analyse the plot, an ellipse can be fitted to the plot. SD1 and SD2 are then defined as the standard deviations of the projections on the major ellipsoidal axes [117]:

$$\mathrm{SD1} = \frac{1}{\sqrt{2}}\mathrm{RMSSD} \text{ and } \mathrm{SD2} = \sqrt{2\mathrm{SDRR}^2 - \frac{1}{2}\mathrm{RMSSD}^2},$$

with SDRR the standard deviation of the RR-intervals.

An interesting derived feature is then the ratio between SD1 and SD2, which characterizes the unpredictability of the RR-intervals [194]. Figure A.2g confirms that the instability is higher in the case of AF.

AFEvidence

AFEvidence uses a 2D-histogram of the Poincaré plot of the *ΔRR-intervals* [190]. Figure A.3b shows the projection of a Poincaré plot on a 2D histogram. The histogram consists of a 2D raster of bins (shown as squares on the Figure), where each bin contains the number of points with corresponding (ΔRR(i),

ΔRR(i-1))-values. The histogram is divided into twelve segments, as shown on the Figure. Each segment contains similar (ΔRR(i), ΔRR(i-1))-pairs. Pairs in section 11, with ΔRR(i) > 0 and ΔRR(i-1) < 0 for example contain segments with a sequence of 'long-short-long' RR intervals [190]. For a normal signal, segment zero contains most points, since the RR-interval length is normally rather stable (and ΔRR is thus not expected to change much). AFEvidence is therefore defined as:

$$\text{AFEvidence} = \text{IrregularityEvidence} - \text{OriginCount} - 2\text{PACEvidence}, \quad \text{(A.5)}$$

with

$$\begin{cases} \text{IrregularityEvidence} = \sum_{n=1}^{12} \text{BinCount}_n, \\ \text{PACEvidence} = \sum_{n=1}^{4} (\text{PointCount}_n - \text{BinCount}_n) + \\ \qquad \sum_{n=5,6,10} (\text{PointCount}_n - \text{BinCount}_n) - \\ \qquad \sum_{n=7,8,12} (\text{PointCount}_n - \text{BinCount}_n), \end{cases}$$

In these Equations, PointCount_n is the number of points in segment n, BinCount_n the number of bins in segment n that contain at least one point and OriginCount the number of points in segment zero.

ApEn

Approximate entropy (ApEn) is a measure for short-term signals based on entropy that measures the complexity of a signal: the higher ApEn, the more insecurity between different segments of the signal [124, 194, 4]. It is thus expected that AF signals have a larger ApEn value in the RR intervals. This is confirmed by Figure A.2i, which shows that there is a significant difference between the entropy of normal and AF signals.

Toeplitz-distance

A final feature, specifically developed for this study, quantifies the uniformity in the RR-intervals of the complete signal (compared to local measures such as RMSSD). Algorithm 3 summarizes the algorithm to calculate the feature [64].

Algorithm 3 Calculation of Toeplitz-distance

Input: Vector $\mathbf{q} \in \mathbb{N}^N$ containing R peak positions
Output: Toeplitz-distance t_d

1: Construct the symmetric distance matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$, with $\mathbf{D}_{ij} = \mathbf{q}_j - \mathbf{q}_i, 1 \leq i, j \leq N$
2: Calculate $\mathbf{g} \in \mathbb{R}^{2N-1}$, with $\mathbf{g}_i = \text{average}(\begin{bmatrix} \mathbf{D}_{N-i+1,1} & \mathbf{D}_{N-i+2,2} & \cdots & \mathbf{D}_{N,i} \end{bmatrix})$ and $\mathbf{g}_{2N-i} = \mathbf{g}_i, 1 \leq i \leq N$
3: Construct the Toeplitz matrix $\mathbf{D}_{\text{toep}} \in \mathbb{R}^{N \times N}$ with generating vector $\mathbf{g}$
4: Calculate the Toeplitz-distance as: $t_d = \frac{1}{N} \|\mathbf{D} - \mathbf{D}_{\text{toep}}\|_{\text{F}}$
5: **return** t_d

If the RR-intervals are perfectly stable, the distance matrix $\mathbf{D}$ would be a perfect Toeplitz-matrix with constant diagonals. The Toeplitz-distance quantifies the distance between $\mathbf{D}$ and the closest Toeplitz-matrix, which is expected to be larger for AF signals with a larger RR-interval variability. Figure A.2j shows indeed that the RR-intervals for normal signals show a more uniform character compared to AF signals.

Other characteristics

Heart rate variability analysis can be useful for many applications, including stress measurements, sleep studies, monitoring of athletes,... Therefore many additional features can be found in literature. Some of most widely used characteristics include frequency-based features, which measure the energy in different spectral intervals, and additional non-linear approaches such as fractal measures. These features were not used in this study, since they often require long-term ECG signals to make a reliable estimate. Since the signals in the available AF datasets were limited in length, they could therefore not be used here.

A.2 Linear analysis of features

The covariance matrix of the (normalised) feature set, which contains the Pearson correlation coefficients between the different features, shows to what extent the different features are linearly related. It is shown in Table A.1.

The strongest linear relations, indicated in bold, are between: AVRR and min(RR), RMSSD and pRR50, RMSSD and Cov(ΔRR) and pRR50 and CoV(ΔRR). This confirms the observations from the previous Section.

	AVRR	RMSSD	pRR50	HTI	min(RR)	SD1/SD2	ApEn	Toeplitz	Cov	AFEvidence
AVRR	1.00	-0.01	-0.06	-0.05	**0.72**	0.01	-0.10	-0.23	-0.25	-0.02
RMSSD	-0.01	1.00	**0.73**	0.40	-0.51	0.52	0.42	0.50	**0.91**	0.48
pRR50	-0.06	0.73	1.00	0.60	-0.41	0.55	0.65	0.39	**0.75**	**0.72**
HTI	-0.05	0.40	0.60	1.00	-0.28	0.11	0.64	0.44	0.43	0.63
min(RR)	0.72	-0.51	-0.41	-0.28	1.00	-0.22	-0.36	-0.50	-0.65	-0.28
SD1/SD2	0.01	0.52	0.55	0.11	-0.22	1.00	0.27	0.06	0.51	0.34
ApEn	-0.10	0.42	0.65	0.64	-0.36	0.27	1.00	0.42	0.47	**0.71**
Toeplitz	-0.23	0.50	0.39	0.44	-0.50	0.06	0.42	1.00	0.58	0.32
CoV(ΔRR)	-0.25	0.91	0.75	0.43	-0.65	0.51	0.47	0.58	1.00	0.52
AFEvidence	-0.02	0.48	0.72	0.63	-0.28	0.34	0.71	0.32	0.52	1.00

Table A.1: Covariance table of the (normalised) HRV feature set, showing the Pearson correlation coefficients between the different features.

Appendix B

Construction and Validation of Noise Model

In order to construct the weight tensor required for calculating the weighted tensor decomposition used in Chapter 6, a method to reliably estimate the signal quality of each heartbeat is necessary since it is used as additional input in the optimization problem. For this we constructed a model that estimates the Signal-to-Noise ratio (SNR) of each heartbeat in each channel based on different parameters.

In literature, many methods can be found that estimate the ECG signal quality, however a large number of these methods is not applicable in this case. For some quality estimators, such as the method described in Behar et al. [19], the goal is not to estimate the SNR, but mainly to classify ECG signals as good or bad quality (e.g. clinically usable or not), which is not applicable in this case. Other methods require a longer period of ECG signal, using for example measures based on heart rate or heart rate variability. Furthermore, methods that compare heartbeats to each other or to a template heartbeat are not desired either, since in the context of prediction of cardiac arrest we expect the heartbeat morphology to change over time due to changes in patient condition.

We therefore opted to develop a novel SNR method, taking into account these considerations. The next Sections describe the construction and validation of the proposed model.

B.1 Construction of SNR model

To construct the model, we used the MIT-BIH Noise Stress Test Database [156] (available on Physionet [85]), which contains 12 half-hour ECG recordings from two subjects with known SNR levels varying from -6dB to 24dB. The recordings were created by starting from two clean recordings and adding calibrated amounts of realistic ECG noise. The noise consists of a mixture of baseline wander and high-frequency noise, similar to muscle artifacts. Half of the available data (e.g. all heartbeats from all SNR levels from one subject) was used as a training set, the other half, consisting of all heartbeats from the second subject, as a test set. Based on the literature, eight features were selected that have been shown to work effectively in estimating ECG signal quality, and that can be easily calculated based on one single-lead heartbeat [45]:

1. Sample skewness s of a heartbeat $\mathbf{x}$:
$$s = \frac{E(\mathbf{x} - \mu)^3}{\sigma^3} \tag{B.1}$$
with μ and σ respectively the mean and standard deviation of the heartbeat $\mathbf{x}$.
2. Sample kurtosis k of a heartbeat $\mathbf{x}$:
$$k = \frac{E(\mathbf{x} - \mu)^4}{\sigma^4} \tag{B.2}$$
3. Power in 6 different subbands: 0–10 Hz, 10–20 Hz, 20–48 Hz, 48–52 Hz, 52–100 Hz, 100–120 Hz

All features were derived from the preprocessed ECG signal, e.g. after baseline wander removal and normalization (using the methods described in Section 6.3.1).

A linear regression model was fitted using all heartbeats of the available training data, using the SNR labels provided in the dataset as predictor variables.

B.2 Validation of SNR model

The SNR model was first tested on the heartbeats in the test set of the MIT-BIH database. Afterwards it was further validated using a subset of the clinical dataset used in Chapter 6 to verify whether the SNR model could be generalized for use on other real-life datasets.

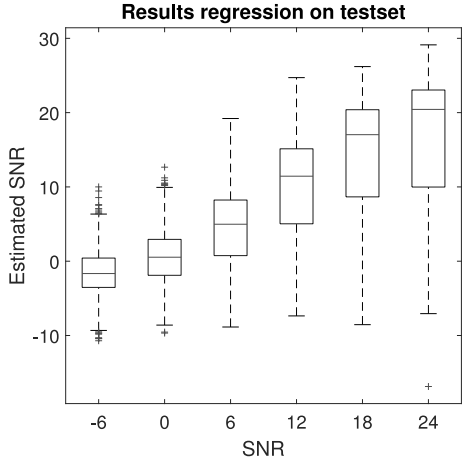


Figure B.1: Boxplots depicting the results of applying the SNR model on the test set. There is a clear correlation between the SNR values estimated by the proposed noise model and the true noise levels.

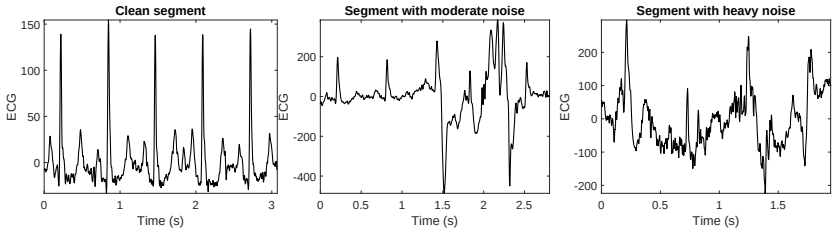


Figure B.2: Examples of segments of the same patient without noise (left), moderate noise (middle) and heavy noise (right). The clean segment contains no visual noise, the segment with moderate noise contains a considerate amount of baseline wander and the noisy segments is corrupted by both low- and high frequency artifacts.

B.2.1 MIT-BIH Noise Stress Database

Figure B.1 shows the result of applying the SNR model on the test set grouped by SNR level. The Root Mean Square Error is 4.6dB. There is a clear correlation between the estimated SNR levels and the true SNR levels.

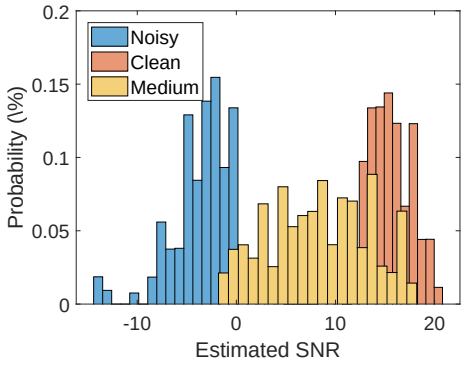


Figure B.3: Histograms of SNR values for each noise class: noisy (blue), medium (yellow) and clean (orange) heartbeats: The SNR values of noisy segments are clearly lower than clean segments, with intermediate values for medium quality segments.

B.2.2 Clinical database

In order to verify whether the developed model can be applied on the current dataset (for which no reference SNR values are available), we selected 15 segments from five signals: five with good quality and no detectable artifacts, five with medium quality (some artifacts present, but ECG waves still clearly recognizable in the majority of the signal) and five with bad quality signal. Each segment contains approximately 100 heartbeats, resulting in 1500 heartbeats overall. Excerpts of all segments are shown in Figure B.2. The SNR model was then used to calculate a quality estimate for each heartbeat in each segment. The results are presented as histograms for each class (clean, medium and noisy) and shown in Figure B.3. It is clear that the different segments resulted in different SNR values. The heartbeats in the clean segment obtained a SNR value between 10 and 22, with an average of 15 dB. The heartbeats in the segment with medium quality on the other hand had SNR values around 7 dB and ranged between -2 and 18. Finally the noisy segment had much lower SNR values, between -15 and 5.

To summarize, we developed a linear regression model to estimate the SNR value of individual heartbeats. The model uses eight statistical and spectral features that have lead to good results in previous studies. The model was trained and tested on the MIT-BIH Noise Stress Database and validated on the clinical database. Results on both databases show that the model achieves good agreement with either reference SNR scores or visual analysis.

Bibliography

[1] AAMI. *ANSI/AAMI EC57/Ed.3, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms.*

[2] Abadi, B. M., Jarchi, D., and Sanei, S. "Simultaneous localization and separation of biomedical signals by tensor factorization". *15th IEEE Workshop on Statistical Signal Processing*. Cardiff, UK, 2009, pp. 497–500.

[3] Acar, E., Aykut-Bingol, C., Bingol, H., Bro, R., and Yener, B. "Multiway analysis of epilepsy tensors". *Bioinformatics* 23.13 (2007), pp. 110–118.

[4] Acharya, U. R., Suri, J. S., Spaan, J. A. E., and Krishnan, S. M. *Advances in Cardiac Signal Processing*. 2007.

[5] Akbari, H., Shamsollahi, M., and Phlypo, R. "Fetal ECG extraction using Tucker decomposition". *International Conference on Systems, Signals and Image Processing (IWSSIP)*. 2015, pp. 174–178.

[6] Akhbari, M., Shamsollahi, M. B., Jutten, C., Armoundas, A. A., and Sayadi, O. "ECG denoising and fiducial point extraction using an extended Kalman filtering framework with linear and nonlinear phase observations". *Physiological Measurement* 37.2 (2016), pp. 203–226.

[7] Akhbari, M., Niknazar, M., Jutten, C., Shamsollahi, M., and Rivet, B. "Fetal Electrocardiogram R-peak Detection using Robust Tensor Decomposition and Extended Kalman Filtering". *Computing in Cardiology*. Zaragoza, Spain, 2013, pp. 189–192.

[8] Al-Fahoum, A. "Quality Assessment of ECG Compression Techniques Using a Wavelet-Based Diagnostic Measure". *IEEE Transactions on Information Technology in Biomedicine* 10.1 (2006), pp. 182–191.

[9] Andersson, C. A. and Bro, R. "The N-way Toolbox for MATLAB". *Chemometrics and Intelligent Laboratory Systems* 52.1 (2000), pp. 1–4.

[10] Ankerst, M., Breunig, M. M., Kriegel, H.-P., and Sander, J. "OPTICS: Ordering Points To Identify the Clustering Structure". *ACM Sigmod record* 28.2 (1999), pp. 49–60.

[11] Arafat, M. A. and Hasan, M. K. "Automatic detection of ECG wave boundaries using empirical mode decomposition". *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. Taipei, Taiwan, 2009, pp. 461–464.

[12] Bardy, G. H., Lee, K. L., Mark, D. B., Poole, J. E., Packer, D. L., Boineau, R., Domanski, M., Troutman, C., Anderson, J., Johnson, G., McNulty, S. E., Clapp-Channing, N., Davidson-Ray, L. D., Fraulo, E. S., Fishbein, D. P., Luceri, R. M., and Ip, J. H. "Amiodarone or an Implantable Cardioverter–Defibrillator for Congestive Heart Failure". *New England Journal of Medicine* 352.3 (2005), pp. 225–237.

[13] Bastiaenen, R., Batchvarov, V., and Gallagher, M. M. "Ventricular automaticity as a predictor of sudden death in ischaemic heart disease". *Europace* 14.6 (2012), pp. 795–803.

[14] Bauer, A., Kantelhardt, J. W., Barthel, P., Schneider, R., Mäkikallio, T., Ulm, K., Hnatkova, K., Schömig, A., Huikuri, H., Bunde, A., Malik, M., and Schmidt, G. "Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study". *The Lancet* 367.9523 (2006), pp. 1674–1681.

[15] Bauer, A., Malik, M., Schmidt, G., Barthel, P., Bonnemeier, H., Cygankiewicz, I., Guzik, P., Lombardi, F., Müller, A., Oto, A., Schneider, R., Watanabe, M., Wichterle, D., and Zareba, W. "Heart Rate Turbulence: Standards of Measurement, Physiological Interpretation, and Clinical Use". *Journal of the American College of Cardiology* 52.17 (2008), pp. 1353–1365.

[16] Bauer, A., Kantelhardt, J. W., Bunde, A., Barthel, P., Schneider, R., Malik, M., and Schmidt, G. "Phase-rectified signal averaging detects quasi-periodicities in non-stationary data". *Physica A: Statistical Mechanics and its Applications* 364 (2006), pp. 423–434.

[17] Beach, S. R., Celano, C. M., Noseworthy, P. A., Januzzi, J. L., and Huffman, J. C. "QTc Prolongation, Torsades de Pointes, and Psychotropic Medications". *Psychosomatics* 54.1 (2013), pp. 1–13.

[18] Becker, H., Albera, L., Comon, P., Haardt, M., Birot, G., Wendling, F., Gavaret, M., Bénar, C., and Merlet, I. "EEG extended source localization: Tensor-based vs. conventional methods". *NeuroImage* 96 (2014), pp. 143–157.

[19] Behar, J., Oster, J., Qiao Li, and Clifford, G. D. "ECG Signal Quality During Arrhythmia and Its Application to False Alarm Reduction". *IEEE Transactions on Biomedical Engineering* 60.6 (2013), pp. 1660–1666.

[20] Behar, J., Oster, J., and Clifford, G. D. "Non-Invasive FECG Extraction from a Set of Abdominal Sensors". *Computing in Cardiology*. Zaragoza, Spain, 2013, pp. 297–300.

[21] Benjamin, E. J. et al. "Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association." *Circulation* 135.10 (2017), pp. 146–603.

[22] Benjamini, Y. and Hochberg, Y. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". *Journal of the Royal Statistical Society* 57.1 (1995), pp. 289–300.

[23] Bharath, H., Sima, D., Sauwen, N., Himmelreich, U., De Lathauwer, L., and Van Huffel, S. "Nonnegative Canonical Polyadic Decomposition for Tissue-Type Differentiation in Gliomas". *IEEE Journal of Biomedical and Health Informatics* 21.4 (2017), pp. 1124–1132.

[24] Bono, V., Mazomenos, E. B., Taihai Chen, Rosengarten, J. A., Acharyya, A., Maharatna, K., Morgan, J. M., and Curzen, N. "Development of an Automated Updated Selvester QRS Scoring System Using SWT-Based QRS Fractionation Detection and Classification". *IEEE Journal of Biomedical and Health Informatics* 18.1 (2014), pp. 193–204.

[25] Boussé, M. and De Lathauwer, L. "Nonlinear Least Squares Algorithm for Canonical Polyadic Decomposition Using Low-Rank Weights". *IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*. Curaçao, Dutch Antilles, 2017, pp. 39–43.

[26] Boussé, M., Goovaerts, G., Vervliet, N., Debals, O., Van Huffel, S., and De Lathauwer, L. "Irregular Heartbeat Classification Using Kronecker Product Equations". *39th IEEE Engineering in Medicine and Biology Conference*. Jeju Island, Korea, 2017, pp. 438–441.

[27] Bousseljot, R., Kreiseler, D., and Schnabel, A. "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet". *Biomedizinische Technik/Biomedical Engineering* 40.s1 (2009), pp. 317–318.

[28] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers". *Foundations and Trends in Machine Learning* 3.1 (2010), pp. 1–122.

[29] Bradburn, M. J., Clark, T. G., Love, S. B., and Altman, D. G. "Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods". *British Journal of Cancer* 89.3 (2003), pp. 431–436.

[30] Bro, R., Andersson, C., and Kiers, H. "Parafac2—part ii. modeling chromatographic data with retention time shifts". *Journal of Chemometrics* 309.July 1998 (1999), pp. 295–309.

[31] Bro, R. and Kiers, H. A. L. "A new efficient method for determining the number of components in PARAFAC models". *Journal of Chemometrics* 17 (2003), pp. 274–286.

[32] Bruyne, M. C. de, Hoes, A. W., Kors, J. A., Hofman, A, Bemmel, J. H. van, and Grobbee, D. E. "Prolonged QT interval predicts cardiac and all-cause mortality in the elderly. The Rotterdam Study." *European heart journal* 20.4 (1999), pp. 278–84.

[33] Burattini, L and Giuliani, C. "T-Wave Frequency Content Evaluation in Healthy Subjects and Patients Affected By Myocardial Infarction". *Signal Processing: New Research* (2013), pp. 79–93.

[34] Burattini, L., Zareba, W., and Moss, A. J. "Correlation Method for Detection of Transient T-Wave Alternans in Digital Holter ECG Recordings". *Annals of Noninvasive Electrocardiology* 4.4 (1999), pp. 416–424.

[35] Burattini, L., Man, S., Ottaviano, G., Fioretti, S., Di Nardo, F., and Swenne, C. "T-wave alternans rate of change with exercise for cardiac risk assessment". *Computing in Cardiology*. Boston, USA, 2014, pp. 177–180.

[36] Burke, A. P., Farb, A., Virmani, R., Goodin, J., and Smialek, J. E. "Sports-related and non-sports-related sudden cardiac death in young adults". *American Heart Journal* 121.2 (1991), pp. 568–575.

[37] Cammarota, C. and Curione, M. "Trend Extraction in Functional Data of Amplitudes of R and T Waves in Exercise Electrocardiogram". *Fluctuation and Noise Letters* 16.02 (2017), p. 1750014.

[38] Cardoso, J.-F. "Multidimensional independent component analysis". *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Vol. 4. Seattle, USA: EE, 1998, pp. 1941–1944.

[39] Caruana, R. and Niculescu-Mizil, A. "An empirical comparison of supervised learning algorithms". *Proceedings of the 23rd international conference on Machine learning - ICML '06*. New York, New York, USA, 2006, pp. 161–168.

[40] Chow, T., Kereiakes, D. J., Bartone, C., Booth, T., Schloss, E. J., Waller, T., Chung, E. S., Menon, S., Nallamothu, B. K., and Chan, P. S. "Prognostic utility of microvolt T-wave alternans in risk stratification of patients with ischemic cardiomyopathy." *Journal of the American College of Cardiology* 47.9 (2006), pp. 1820–7.

[41] Chugh, S. S., Jui, J., Gunson, K., Stecker, E. C., John, B. T., Thompson, B., Ilias, N., Vickers, C., Dogra, V., Daya, M., Kron, J., Zheng, Z.-J., Mensah, G., and McAnulty, J. "Current burden of sudden cardiac death: Multiple source surveillance versus retrospective death certificate-based review in a large U.S. community". *Journal of the American College of Cardiology* 44.6 (2004), pp. 1268–1275.

[42] Cichocki, A., Mandic, D, Caiafa, C, Phan, A.-h., Zhou, G, Zhao, Q, and De Lathauwer, L. "Tensor Decompositions for Signal Processing Applications From Two-way to Multiway Component Analysis". *IEEE Signal Processing Magazine* 32.2 (2015), pp. 145–163.

[43] Clark, T. G., Bradburn, M. J., Love, S. B., and Altman, D. G. "Survival Analysis Part I: Basic concepts and first analyses". *British Journal of Cancer* 89.2 (2003), pp. 232–238.

[44] Clifford, G, Azuaje, F, and McSharry, P. *Advanced Methods and Tools for ECG Data Analysis*. 2006.

[45] Clifford, G. D., Behar, J, Li, Q, and Rezek, I. "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms". *Physiological Measurement* 33.9 (2012), pp. 1419–1433.

[46] Clifford, G., Liu, C., Moody, B., Lehman, L.-w., Silva, I., Li, Q., Johnson, A., and Mark, R. "AF Classification from a Short Single Lead ECG Recording: the Physionet Computing in Cardiology Challenge 2017". *Computing in Cardiology*. Rennes, France, 2017, pp. 1–4.

[47] Clifford, G., Nemati, S., and Sameni, R. "An artificial multichannel model for generating abnormal electrocardiographic rhythms". *Computing in Cardiology*. Bologna, Italy, 2008, pp. 773–776.

[48] Cong, F., Nandi, A. K., He, Z., Cichocki, A., and Ristaniemi, T. "Fast and effective model order selection to determine the number of sources in a linear transformation model". *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. Bucharest, Romania, 2012, pp. 1870–1874.

[49] Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., and Ristaniemi, T. "Tensor decomposition of EEG signals: A brief review". *Journal of Neuroscience Methods* 248 (2015), pp. 59–69.

[50] Cox, D. R. "Regression Models and Life-Tables". Springer, New York, NY, 1992, pp. 527–541.

[51] Daemen, A. and De Moor, B. "Development of a kernel function for clinical data". *IEEE Engineering in Medicine and Biology Society*. Minneapolis, USA, 2009, pp. 5913–5917.

[52] Dandapat, S. and Padhy, S. "Exploiting multi-lead electrocardiogram correlations using robust third-order tensor decomposition". *Healthcare Technology Letters* 2.5 (2015), pp. 112–117.

[53] Das, M. K., Saha, C., El Masry, H., Peng, J., Dandamudi, G., Mahenthiran, J., McHenry, P., and Zipes, D. P. "Fragmented QRS on a 12-lead ECG: A predictor of mortality and cardiac events in patients with coronary artery disease". *Heart Rhythm* 4.11 (2007), pp. 1385–1392.

[54] Datta, S., Puri, C., Mukherjee, A., Banerjee, R., Choudhury, D., Singh, R., Ukil, A., Bandyopadhyay, S., Pal, A., and Khandelwal, S. "Identifying Normal, AF and other Abnormal ECG Rhythms using a Cascaded Binary Classifier". *Computing in Cardiology*. Vol. 44. Rennes, France, 2017, pp. 1–4.

[55] Daubert, J. P., Zareba, W., Cannom, D. S., Mcnitt, S., Rosero, S. Z., Wang, P., Schuger, C., Steinberg, J. S., Higgins, S. L., Wilber, D. J., Klein, H., Andrews, M. L., Hall, W. J., and Moss, A. J. "Inappropriate Implantable Cardioverter-Defibrillator Shocks in MADIT II Frequency, Mechanisms, Predictors, and Survival Impact". *Heart Rhythm Disorders* 51.14 (2008), pp. 1357–1365.

[56] De Cooman, T., Goovaerts, G., Varon, C., Widjaja, D., Willemen, T., and Van Huffel, S. "Heart beat detection in multimodal data using automatic relevant signal detection". *Physiological Measurement* 36.8 (2015), pp. 1691–.

[57] De Cooman, T., Goovaerts, G., Varon, C., Widjaja, D., and Van Huffel, S. "Heart beat detection in multimodal data using signal recognition and beat location estimation". *Computing in Cardiology*. Vol. 41. January. Boston, USA, 2014, pp. 257–260.

[58] De Lathauwer, L. "Decompositions of a Higher-Order Tensor in Block Terms—Part I: Lemmas for Partitioned Matrices". *SIAM Journal on Matrix Analysis and Applications* 30.3 (2008), pp. 1022–1032.

[59] De Lathauwer, L. "Signal Processing Based on Multilinear Algebra". PhD thesis. Department of Electrical Engineering, KU Leuven, 1997.

[60] De Lathauwer, L. and De Moor, B. "From Matrix to Tensor: Multilinear Algebra and Signal Processing 1". *Mathematics in Signal Processing IV* (1997), pp. 17–19.

[61] De Lathauwer, L., De Moor, B., and Vandewalle, J. "A Multilinear Singular Value Decomposition". *Society for Industrial and Applied Mathematics* 21.4 (2000), pp. 1253–1278.

[62] De Lathauwer, L., De Moor, B., and Vandewalle, J. "Fetal Electrocardiogram Extraction by Blind Source Subspace Separation". *IEEE Transactions in Biomedical Engineering* 47 (2000), pp. 567–572.

[63] De Vos, M., De Lathauwer, L., Vanrumste, B., Van Huffel, S., and Van Paesschen, W. "Canonical decomposition of ictal scalp EEG and accurate source localisation: principles and simulation study." *Computational intelligence and neuroscience* 2007 (2007), p. 58253.

[64] Debals, O. and De Lathauwer, L. "Stochastic and Deterministic Tensorization for Blind Signal Separation". *Latent Variable Analysis and Signal Separation*. 2015, pp. 3–13.

[65] Debals, O., Van Barel, M., and De Lathauwer, L. "Löwner-Based Blind Signal Separation of Rational Functions With Applications". *IEEE Transactions on Signal Processing* 64.8 (2016), pp. 1909–1918.

[66] DeChazal, P., O'Dwyer, M., and Reilly, R. "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features". *IEEE Transactions on Biomedical Engineering* 51.7 (2004), pp. 1196–1206.

[67] Ding, Q., Bai, Y., Tinoco, A., Mortara, D., Do, D., Boyle, N. G., Pelter, M. M., and Hu, X. "Developing new predictive alarms based on ECG metrics for bradyasystolic cardiac arrest". *Physiological Measurement* 36.12 (2015), pp. 2405–2422.

[68] Do, D. H., Hayase, J., Tiecher, R. D., Bai, Y., Hu, X., and Boyle, N. G. "ECG changes on continuous telemetry preceding in-hospital cardiac arrests". *Journal of Electrocardiology* 48.6 (2015), pp. 1062–1068.

[69] Domanov, I. and De Lathauwer, L. "On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors—Part I: Basic Results and Uniqueness of One Factor Matrix". *SIAM Journal on Matrix Analysis and Applications* 34.3 (2013), pp. 855–875.

[70] Domanov, I. and De Lathauwer, L. "On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors—Part II: Uniqueness of the Overall Decomposition". *SIAM Journal on Matrix Analysis and Applications* 34.3 (2013), pp. 876–903.

[71] Dower, G. E., Yakush, A, Nazzal, S. B., Jutzy, R. V., and Ruiz, C. E. "Deriving the 12-lead electrocardiogram from four (EASI) electrodes." *Journal of electrocardiology* 21 Suppl (1988), S182–7.

[72] Dragomiretskiy, K. and Zosso, D. "Variational Mode Decomposition". *IEEE Transactions on Signal Processing* 62.3 (2014), pp. 531–544.

[73] Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., and Vapoik, V. "Support Vector Regression Machines". *Advances in Neural Information Processing System* (1997), pp. 155–161.

[74] Eckart, C. and Young, G. "The approximation of one matrix by another of lower rank". *Psychometrika* 1.3 (1936), pp. 211–218.

[75] Elavarasi, S, Akilandeswari, J, and Sathiyabhama, B. "A survey on partitioning clustering techniques". *International Journal of Enterprise Computing and Business Systems* 1.1 (2011).

[76] Epanechnikov, V. A. "Non-Parametric Estimation of a Multivariate Probability Density". *Theory of Probability & Its Applications* 14.1 (1969), pp. 153–158.

[77] Evers, L. and Messow, C.-M. "Sparse kernel methods for high-dimensional survival data". *Bioinformatics* 24.14 (2008), pp. 1632–1638.

[78] Fasano, A. and Villani, V. "Baseline wander removal for bioelectrical signals by quadratic variation reduction". *Signal Processing* 99 (2014), pp. 48–57.

[79] Fishman, G. I., Chugh, S. S., Dimarco, J. P., Albert, C. M., Anderson, M. E., Bonow, R. O., Buxton, A. E., Chen, P.-S., Estes, M., Jouven, X., Kwong, R., Lathrop, D. A., Mascette, A. M., Nerbonne, J. M., O'Rourke, B., Page, R. L., Roden, D. M., Rosenbaum, D. S., Sotoodehnia, N., Trayanova, N. A., and Zheng, Z.-J. "Sudden cardiac death prediction and prevention: report from a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop." *Circulation* 122.22 (2010), pp. 2335–48.

[80] Flowers, N. C., Horan, L. G., Thomas, J. R., and Tolleson, W. J. "The anatomic basis for high-frequency components in the electrocardiogram". *Circulation* 39.4 (1969), pp. 531–539.

[81] Fouodo, C. J. K., König, I. R., Weihs, C., Ziegler, A., and Wright, M. N. "Support Vector Machines for Survival Analysis with R". *The R Journal* 10.1 (2018), pp. 412–423.

[82] Fuster, V, Rydén, L., Cannom, D., Crijns, H., Curtis, A., Ellenbogen, K., Halperin, J., Kay, G., Le Huezey, J., Lowe, J., Olsson, S., Prystowsky, E., Tamargo, J., and Wann, L. "2011 ACCF/AHA/HRS Focused Updates Incorporated Into the ACC/AHA/ESC 2006 Guidelines for the Management of Patients With Atrial Fibrillation". *Journal of the American College of Cardiology* 57.11 (2011), pp. 1330–1337.

[83] Geirnaert, S. "Detectie van voorkamerfibrillatie: een tensorgebaseerde methode". Master thesis. KU Leuven, 2018.

[84] Girotra, S., Nallamothu, B. K., Spertus, J. A., Li, Y., Krumholz, H. M., and Chan, P. S. "Trends in Survival after In-Hospital Cardiac Arrest". *New England Journal of Medicine* 367.20 (2012), pp. 1912–1920.

[85] Goldberger, A. L., Amaral, L. A., Glass, L, Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., and Stanley, H. E. "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals." *Circulation* 101.23 (2000), E215–20.

[86] Goovaerts, G., Van Huffel, S., and Hu, X. "Tensor-based analysis of ECG changes prior to in-hospital cardiac arrest". *Computing in Cardiology*. Vol. 44. Rennes, France, 2017, pp. 1–4.

[87] Goovaerts, G., Vandenberk, B., Willems, R., and Van Huffel, S. "Automatic detection of T wave alternans using tensor decompositions in multilead ECG signals". *Physiological Measurement* 38.8 (2017), pp. 1513–1524.

[88] Goovaerts, G., De Wel, O., Vandenberk, B., Willems, R., and Van Huffel, S. "Detection of irregular heartbeats using tensors". *Computing in Cardiology*. Vol. 42. Nice, France, 2015, pp. 573–576.

[89] Goovaerts, G., Vandenberk, B., Varon, C., Willems, R., and Van Huffel, S. "Phase-Rectified Signal Averaging for automatic detection of QRS fragmentation". *Computing in Cardiology*. Vol. 43. Vancouver, Canada, 2016, pp. 637–640.

[90] Goovaerts, G., Varon, C., Vandenberk, B., Willems, R., and Van Huffel, S. "Tensor-based detection of T wave alternans in multilead ECG signals". *Computing in Cardiology*. Vol. 41. Boston, USA, 2014, pp. 185–188.

[91] Goovaerts, G., Vandenberk, B., Willems, R., and Van Huffel, S. "Tensor-based detection of T wave alternans using ECG". *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan, Italy, 2015, pp. 6991–6994.

[92] Greenwald, S. D. "The development and analysis of a ventricular fibrillation detector". PhD thesis. Massachusetts Institute of Technology, Cambridge, MA, USA, 1986.

[93] Guo, Y., Zhang, Z., Cao, J., Gong, T., and Yang, W. "An optimized variational mode decomposition for extracting weak feature of viscoelastic sandwich cylindrical structures". *Measurement Science and Technology* 29.3 (2018), p. 035006.

[94] Guzik, P. and Malik, M. "ECG by mobile technologies". *Journal of Electrocardiology* 49 (2016), pp. 894–901.

[95] Haigney, M. C., Zareba, W., Gentlesk, P. J., Goldstein, R. E., Illovsky, M., McNitt, S., Andrews, M. L., Moss, A. J., and Multicenter Automatic Defibrillator Implantation Trial II investigators. "QT interval variability and spontaneous ventricular tachycardia or fibrillation in the Multicenter Automatic Defibrillator Implantation Trial (MADIT) II patients". *Journal of the American College of Cardiology* 44.7 (2004), pp. 1481–1487.

[96] Haim, M., Hoshen, M., Reges, O., Rabi, Y., Balicer, R., and Leibowitz, M. "Prospective National Study of the Prevalence, Incidence, Management and Outcome of a Large Contemporary Cohort of Patients With Incident Non-Valvular Atrial Fibrillation". *Journal of the American Heart Association* 4.1 (2015), e001486–e001486.

[97] Hanley, J. A. and McNeil, B. J. "A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases'". *Radiology* 148.3 (1983), pp. 839–843.

[98] Harrell, F. E., Lee, K. L., Califf, R. M., Pryor, D. B., and Rosati, R. A. "Regression modelling strategies for improved prognostic prediction". *Statistics in Medicine* 3.2 (1984), pp. 143–152.

[99] Harshman, R. A. "Foundations of the PARAFAC Procedure: Models and Conditions for an "Explanatory" Multimodal Factor Analysis". *UCLA Working Papers in Phonetics* 16 (1970), pp. 1–84.

[100] Haukilahti, M. A. E., Eranti, A., Kenttä, T., and Huikuri, H. V. "QRS Fragmentation Patterns Representing Myocardial Scar Need to Be Separated from Benign Normal Variants: Hypotheses and Proposal for Morphology based Classification." *Frontiers in physiology* 7 (2016), p. 653.

[101] Hayn, D, Kollmann, A, and Schreier, G. "Automated QT interval measurement from multilead ECG signals". *Computing in Cardiology*. Valencia, Spain, 2006, pp. 381–384.

[102] Hong, S., Wu, M., Zhou, Y., Wang, Q., Shang, J., Li, H., and Xie, J. "ENCASE: an ENsemble ClASsifiEr for ECG Classification Using Expert Features and Deep Neural Networks". *Computing in Cardiology*. Rennes, France, 2017, pp. 1–4.

[103] Hu, X., Do, D., Bai, Y., and Boyle, N. G. "A case–control study of non-monitored ECG metrics preceding in-hospital bradyasystolic cardiac arrest: Implication for predictive monitor alarms". *Journal of Electrocardiology* 46.6 (2013), pp. 608–615.

[104] Huang, K. and Zhang, L. "Cardiology knowledge free ECG feature extraction using generalized tensor rank one discriminant analysis". *Journal on Advances in Signal Processing* 2 (2014), pp. 1–15.

[105] Hunyadi, B., Signoretto, M., Van Paesschen, W., Suykens, J. A., Van Huffel, S., and De Vos, M. "Incorporating structural information from the multichannel EEG improves patient-specific seizure detection". *Clinical Neurophysiology* 123.12 (2012), pp. 2352–2361.

[106] Iglewicz, B. and Hoaglin, D. "Volume 16: How to Detect and Handle Outliers". *The ASQC Basic References in Quality Control: Statistical Techniques*. Vol. 16. 1993.

[107] Ikeda, T., Saito, H., Tanno, K., Shimizu, H., Watanabe, J., Ohnishi, Y., Kasamaki, Y., and Ozawa, Y. "T-Wave Alternans as a Predictor for Sudden Cardiac after Myocardial Infarction". *The American Journal of Cardiology* 89.01 (2002), pp. 79–82.

[108] Ishteva, M., Absil, P.-A., Van Huffel, S., and De Lathauwer, L. "Best Low Multilinear Rank Approximation of Higher-Order Tensors, Based on the Riemannian Trust-Region Scheme". *SIAM Journal on Matrix Analysis and Applications* 32.1 (2011), pp. 115–135.

[109] Jain, A. K. "Data clustering: 50 years beyond K-means". *Pattern Recognition Letters* 31 (2010), pp. 651–666.

[110] Jin, F., Sugavaneswaran, L., Krishnan, S., and Chauhan, V. S. "Quantification of fragmented QRS complex using intrinsic time-scale decomposition". *Biomedical Signal Processing and Control* 31 (2017), pp. 513–523.

[111] Kania, M, Fereniec, M, and Maniewski, R. *Wavelet Denoising for Multi-lead High Resolution ECG Signals*. Tech. rep. 2. 2007.

[112] Kaplan, E. L. and Meier, P. "Nonparametric Estimation from Incomplete Observations". *Journal of the American Statistical Association* 53.282 (1958), pp. 457–481.

[113] Kaufman, L. and Rousseeuw, P. J. *Finding groups in data : an introduction to cluster analysis*. Wiley-Interscience, 1990, p. 342.

[114] KAVESH, N. G., SHOROFSKY, S. R., SARANG, S. E., and GOLD, M. R. "Effect of Heart Rate on T Wave Alternans". *Journal of Cardiovascular Electrophysiology* 9.7 (1998), pp. 703–708.

[115] Kempf, F. C. and Josephson, M. E. "Cardiac arrest recorded on ambulatory electrocardiograms". *The American Journal of Cardiology* 53.11 (1984), pp. 1577–1582.

[116] Khan, F. M. and Zubek, V. B. "Support Vector Regression for Censored Data (SVRc): A Novel Tool for Survival Analysis". *IEEE International Conference on Data Mining*. Pisa, Italy, 2008, pp. 863–868.

[117] Khandoker, A. H., Karmakar, C., Brennan, M., Palaniswami, M., and Voss, A. *Poincaré Plot Methods for Heart Rate Variability Analysis*. Boston, MA, 2013.

[118] Kitamura, H., Ohnishi, Y., Okajima, K., Ishida, A., Galeano, E., Adachi, K., and Yokoyama, M. "Onset heart rate of microvolt-level T-wave alternans provides clinical and prognostic value in nonischemic dilated cardiomyopathy". *Journal of the American College of Cardiology* 39.2 (2002), pp. 295–300.

[119] Kolda, T. G. and Bader, B. W. "Tensor Decompositions and Applications". *SIAM Review* 51.3 (2008), pp. 455–500.

[120] Kowey, P. R. "Ambulatory Electrocardiographic Recording". *Circulation* 108.5 (2003), 31e–33.

[121] Kuzilek, J., Kremen, V., and Lhotska, L. "Comparison of JADE and Canonical Correlation Analysis for ECG de-noising". *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Chicago. USA, 2014, pp. 3857–3860.

[122] Laguna, P., Mark, R., Goldberg, A., and Moody, G. "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG". *Computers in Cardiology*. Lund, Sweden, 1997, pp. 673–676.

[123] Laguna, P., Jane, R., and Caminal, P. "Automatic Detection of Wave Boundaries in Multilead ECG Signals: Validation with the CSE Database". *Computers and Biomedical Research* 27 (1994), pp. 45–60.

[124] Lake, D. E. and Moorman, J. R. "Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices". *American Journal of Physiology-Heart and Circulatory Physiology* 300.1 (2011), H319–H325.

[125] Ledezma, C. A., Severeyn, E., Perpin, G., Altuve, M., and Wong, S. "A new on-line electrocardiographic records database and computer routines for data analysis". *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference* 2014 (2014), pp. 2738–2741.

[126] Lee, H., Kin, Y.-D., Cichocki, A., and Choi, S. "Non-negative tensor factorization for continuous EEG classification". *International Journal of Neural Systems* 17.04 (2007), pp. 305–317.

[127] Lenis, G., Pilia, N., Loewe, A., Schulze, W. H. W., and Dössel, O. "Comparison of Baseline Wander Removal Techniques considering the Preservation of ST Changes in the Ischemic ECG: A Simulation Study". *Computational and Mathematical Methods in Medicine* 2017 (2017), pp. 1–13.

[128] Li, D., Huang, K., Zhang, H., and Zhang, L. "UMPCA Based Feature Extraction for ECG". *ISNN2013: Advances in Neural Networks*. Springer, Berlin, Heidelberg, 2013, pp. 383–390.

[129] Lin, C.-C. and Yang, C.-M. "Heartbeat Classification Using Normalized RR Intervals and Morphological Features". *Mathematical Problems in Engineering* 2014 (2014), pp. 1–11.

[130] Liu, Y., Li, Z., Xiong, H., Gao, X., and Wu, J. "Understanding of Internal Clustering Validation Measures". *2010 IEEE International Conference on Data Mining*. Sydney, Australia, 2010, pp. 911–916.

[131] Llamedo, M. and Martinez, J. P. "An Automatic Patient-Adapted ECG Heartbeat Classifier Allowing Expert Assistance". *IEEE Transactions on Biomedical Engineering* 59.8 (2012), pp. 2312–2320.

[132] Llamedo, M. and Martínez, J. P. "Heartbeat Classification Using Feature Selection Driven by Database Generalization Criteria". *IEEE Transactions on Biomedical Engineering* 58.3 (2011), pp. 616–625.

[133] Llamedo, M., Martín-Yebra, A., Laguna, P., and Martínez, J. P. "Noninvasive Fetal ECG Estimation Based on Linear Transformations". *Computing in Cardiology*. Zaragoza, Spain, 2013, pp. 285–288.

[134] Luna, A. B. de, Coumel, P., and Leclercq, J. F. "Ambulatory sudden cardiac death: Mechanisms of production of fatal arrhythmia on the basis of data from 157 cases". *American Heart Journal* 117.1 (1989), pp. 151–159.

[135] Luz, E. J.d. S., Schwartz, W. R., Cámara-Chávez, G., and Menotti, D. "ECG-based heartbeat classification for arrhythmia detection: A survey". *Computer Methods and Programs in Biomedicine* 127 (2016), pp. 144–164.

[136] Lyon, A, Mincholé, A, Martínez, J. P., Laguna, P, and Rodriguez, B. "Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances". *Journal of the Royal Society, Interface* 15.138 (2018).

[137] Madeiro, J. P., Cortez, P. C., Oliveira, F. I., and Siqueira, R. S. "A new approach to QRS segmentation based on wavelet bases and adaptive threshold technique". *Medical Engineering & Physics* 29.1 (2007), pp. 26–37.

[138] Mahajan, R., Kamaleswaran, R., Howe, J. A., and Akbilgic, O. "Cardiac Rhythm Classification from a Short Single Lead ECG Recording via Random Forest". *Computing in Cardiology*. Rennes, France, 2017, pp. 1–4.

[139] Maheshwari, S., Acharyya, A., Puddu, P. E., Mazomenos, E. B., Leekha, G., Maharatna, K., and Schiariti, M. "An automated algorithm for online detection of fragmented QRS and identification of its various morphologies." *Journal of the Royal Society, Interface* 10.89 (2013), p. 20130761.

[140] Maji, U., Mitra, M., and Pal, S. "Characterization of cardiac arrhythmias by variational mode decomposition technique". *Biocybernetics and Biomedical Engineering* 37.3 (2017), pp. 578–589.

[141] Mar, T., Zaunseder, S., Martinez, J. P., Llamedo, M., and Poll, R. "Optimization of ECG Classification by Means of Feature Selection". *IEEE Transactions on Biomedical Engineering* 58.8 (2011), pp. 2168–2177.

[142] Martínez, J. P. and Olmos, S. "Methodological Principles of T Wave Alternans Analysis : A Unified Framework". *IEEE Transactions on Biomedical Engineering* 52.4 (2005), pp. 599–613.

[143] Martínez, J. P., Almeida, R., Olmos, S., Rocha, A. P., and Laguna, P. "A wavelet-based ECG delineator: evaluation on standard databases." *IEEE transactions on bio-medical engineering* 51.4 (2004), pp. 570–81.

[144] Martis, R. J., Acharya, U. R., and Adeli, H. "Current methods in electrocardiogram characterization". *Computers in Biology and Medicine* 48 (2014), pp. 133–149.

[145] Mcdonald, J. H. *Handbook of biological statistics*. 2009.

[146] McGill, R., Tukey, J. W., and Larsen, W. A. "Variations of Box Plots". *The American Statistician* 32.1 (1978), p. 12.

[147] Mert, A. "ECG feature extraction based on the bandwidth properties of variational mode decomposition". *Physiological Measurement* 37.4 (2016), pp. 530–543.

[148] Meyer, D., Leisch, F., and Hornik, K. "The support vector machine under test". *Neurocomputing* 55.1-2 (2003), pp. 169–186.

[149] Milligan, G. W. and Cooper, M. C. "An examination of procedures for determining the number of clusters in a data set". *Psychometrika* 50.2 (1985), pp. 159–179.

[150] Moeyersons, J., Goovaerts, G., Huijghebaert, S., Vandenberk, B., Willems, R., and Van Huffel, S. "Automated T Wave End Detection Methods: Comparison of Four Different Methods for T Wave End Detection". *BIOSIGNALS*. Porto, Portugal, 2017, pp. 92–98.

[151] Monasterio, V., Laguna, P., and Martinez, J. "Multilead Analysis of T-Wave Alternans in the ECG Using Principal Component Analysis". *IEEE Transactions on Biomedical Engineering* 56.7 (2009), pp. 1880–1890.

[152] Moody, G. B. "The physionet / computers in cardiology challenge 2008: T-wave Alternans". *Computers in Cardiology*. Vol. 35. Bologna, Italy, 2008, pp. 505–508.

[153] Moody, G. B. and Mark, R. G. "The Impact of the MIT-BIH Arrhythmia Database". *IEEE Engineering in Medicine and Biology* 20 (2001), pp. 45–50.

[154] Moody, G. "Spontaneous termination of atrial fibrillation: a challenge from physionet and computers in cardiology 2004". *Computers in Cardiology*. Chicago. USA, 2004, pp. 101–104.

[155] Moody, G. and Mark, R. "A new method for detecting atrial fibrillation using R-R intervals". *Computers in Cardiology*. Aachen, Germany, 1983, pp. 227–230.

[156] Moody GB, Muldrow WE, and Mark RG. "The MIT-BIH Noise Stress Test Database". *Computers in Cardiology*. Park City, USA, 1984, pp. 381–384.

[157] Morita, H., Kusano, K. F., Miura, D., Nagase, S., Nakamura, K., Morita, S. T., Ohe, T., Zipes, D. P., and Wu, J. "Fragmented QRS as a Marker of Conduction Abnormality and a Predictor of Prognosis of Brugada Syndrome". *Circulation* 118.17 (2008), pp. 1697–1704.

[158] Mørup, M., Hansen, L., Arnfred, S., Lim, L., and Madsen, K. "Shift-invariant multilinear decomposition of neuroimaging data". *NeuroImage* 42.4 (2008), pp. 1439–1450.

[159] Moss, A. J. "Measurement of the QT interval and the risk associated with QTc interval prolongation: A review". *The American Journal of Cardiology* 72.6 (1993), B23–B25.

[160] Moss, A. J., Zareba, W., Hall, W. J., Klein, H., Wilber, D. J., Cannom, D. S., Daubert, J. P., Higgins, S. L., Brown, M. W., Andrews, M. L., and Multicenter Automatic Defibrillator Implantation Trial II Investigators. "Prophylactic Implantation of a Defibrillator in Patients with Myocardial Infarction and Reduced Ejection Fraction". *New England Journal of Medicine* 346.12 (2002), pp. 877–883.

[161] Myerburg, R. J., Mitrani, R., Interian, A., and Castellanos, A. "Interpretation of Outcomes of Antiarrhythmic Clinical Trials Design Features and Population Impact". *Circulation* 97.15 (1998), pp. 1514–1521.

[162] Nearing, B. D. and Verrier, R. L. "Modified moving average analysis of T-wave alternans to predict ventricular fibrillation with high accuracy." *Journal of applied physiology (Bethesda, Md. : 1985)* 92.2 (2002), pp. 541–549.

[163] Nichol, G., Thomas, E., Callaway, C. W., Hedges, J., Powell, J. L., Aufderheide, T. P., Rea, T., Lowe, R., Brown, T., Dreyer, J., Davis, D., Idris, A., and Stiell, I. "Regional Variation in Out-of-Hospital Cardiac Arrest Incidence and Outcome". *JAMA* 300.12 (2008), p. 1423.

[164] Niculescu-Mizil, A. and Caruana, R. "Predicting good probabilities with supervised learning". *Proceedings of the 22nd international conference on Machine learning - ICML '05*. New York, New York, USA: ACM Press, 2005, pp. 625–632.

[165] Niknazar, M., Rivet, B., and Jutten, C. "Fetal QRS Complex Detection Based on Three-Way Tensor Decomposition". *Computing in Cardiology*. Zaragoza, Spain, 2013, pp. 185–188.

[166] Oliveira, P. M. R. de and Zarzoso, V. "Source Analysis and Selection Using Block Term Decomposition in Atrial Fibrillation". *Latent Variable Analysis and Signal Separation*. Springer, Cham, 2018, pp. 46–56.

[167] Paatero, P. "A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis". *Chemometrics and Intelligent Laboratory Systems* 38.2 (1997), pp. 223–242.

[168] Padhy, S. and Dandapat, S. "A new multilead ECG data compression method using Higher-Order Singular Value Decomposition". *Twentieth National Conference on Communications (NCC)*. Kanpur, India, 2014, pp. 1–4.

[169] Padhy, S. and Dandapat, S. "Third-order tensor based analysis of multilead ECG for classification of myocardial infarction". *Biomedical Signal Processing and Control* 31 (2017), pp. 71–78.

[170] Padhy, S. and Dandapat, S. "Validation of μ-volt T-wave alternans analysis using multiscale analysis-by-synthesis and higher-order SVD". *Biomedical Signal Processing and Control* 40 (2018), pp. 171–179.

[171] Pal, S. "ECG Monitoring: Present Status and Future Trend". *Reference Module in Biomedical Sciences*. Elsevier, 2017.

[172] Pal, S. and Mitra, M. "Empirical mode decomposition based ECG enhancement and QRS detection". *Computers in Biology and Medicine* 42.1 (2012), pp. 83–92.

[173] Pan, J, Eng, W. T.I.T. B., and 1985, U. "A real-time QRS detection algorithm". *IEEE Transactions in Biomedical Engineering* 32.3 (1985), pp. 230–236.

[174] Peto, R, Pike, M. C., Armitage, P, Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N, McPherson, K, Peto, J, and Smith, P. G. "Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples". *British Journal of Cancer* 35.1 (1977), pp. 1–39.

[175] Petrutiu, S., Sahakian, A. V., and Swiryn, S. "Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans". *EP Europace* 9.7 (2007), pp. 466–470.

[176] Piepoli, M. F. et al. "2016 European Guidelines on cardiovascular disease prevention in clinical practice". *European Heart Journal* 37.29 (2016), pp. 2315–2381.

[177] Pietrasik, G. and Zaręba, W. "QRS fragmentation: Diagnostic and prognostic significance". *Cardiology Journal* 19.2 (2012), pp. 114–121.

[178] Platt, J. C. "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods". *Advances in Large Margin Classifiers1* 10.3 (1999), pp. 61–74.

[179] Priori, S. G., Blomström M-Lundqvist, C., Mazzanti, A., Blom, N., Borggrefe, M., Camm, J., Elliott, P., Fitzsimons, D., Hatala, R., Hindricks, G., Kirchhof, P., Kjeldsen, K., Kuck, K.-H., Hernandez-Madrid, A., Nikoloau, N., Norekval, T., Spaulding, C., and Van Veldhuisen, D. "ESC GUIDELINES 2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death The Task Force for the Management of Patients with Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death". *European Heart Journal2* 36 (2015), pp. 2793–2867.

[180] Qibin Zhao and Liqing Zhang. "ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines". *International Conference on Neural Networks and Brain*. Vol. 2. Beijing, China: IEEE, 2005, pp. 1089–1092.

[181] Rai, H. M., Trivedi, A., and Shukla, S. "ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier". *Measurement* 46.9 (2013), pp. 3238–3246.

[182] Ramasamy, S. and Balan, A. "Wearable sensors for ECG measurement: a review". *Sensor Review* 38.4 (2018), pp. 412–419.

[183] Razavipour, F., Haghpanahi, M., and Sameni, R. "Fetal QRS Complex Detection using Semi-Blind Source Separation Framework". *Computing in Cardiology*. Zaragoza, Spain, 2013, pp. 181–184.

[184] *Recording a 12 lead ECG – First aid for free*.

[185] Ribeiro, L. N., Hidalgo-Munoz, A. R., Favier, G., Mota, J. C. M., Almeida, A. L. F. de, and Zarzoso, V. "A tensor decomposition approach to noninvasive atrial activity extraction in atrial fibrillation ECG". *23rd European Signal Processing Conference (EUSIPCO)*. Nice, France, 2015, pp. 2576–2580.

[186] Rosner, B. *Fundamentals of Biostatistics*. Belmont Press, 2006.

[187] Salvador, S. and Chan, P. "Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms". *IEEE International Conference on Tools with Artificial Intelligence*. Boca Raton, USA, 2004, pp. 576–584.

[188] Sameni, R. *OSET: The open-source electrophysiological toolbox*.

[189] Samuel, A. L. "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development* 3.3 (1959), pp. 210–229.

[190] Sarkar, S., Ritscher, D., and Mehra, R. "A Detector for a Chronic Implantable Atrial Tachyarrhythmia Monitor". *IEEE Transactions on Biomedical Engineering* 55.3 (2008), pp. 1219–1224.

[191] Sassi, R., Cerutti, S., Lombardi, F., Malik, M., Huikuri, H. V., Peng, C.-K., Schmidt, G., Yamamoto, Y., Gorenek, B., Lip, G. Y., Grassi, G., Kudaiberdieva, G., Fisher, J. P., Zabel, M., and Macfadyen, R. "Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society". *Europace* 17.9 (2015), pp. 1341–1353.

[192] Satija, U., Ramkumar, B., and Manikandan, M. S. "A unified sparse signal decomposition and reconstruction framework for elimination of muscle artifacts from ECG signal". *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Shanghai, China, 2016, pp. 779–783.

[193] Schwartz, P. J., Priori, S. G., Spazzolini, C, Moss, A. J., Vincent, G. M., Napolitano, C, Denjoy, I, Guicheney, P, Breithardt, G, Keating, M. T., Towbin, J. A., Beggs, A. H., Brink, P, Wilde, A. A., Toivonen, L, Zareba, W, Robinson, J. L., Timothy, K. W., Corfield, V, Wattanasirichaigoon, D, Corbett, C, Haverkamp, W, Schulze-Bahr, E, Lehmann, M. H., Schwartz, K, Coumel, P, and Bloise, R. "Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias." *Circulation* 103.1 (2001), pp. 89–95.

[194] Shaffer, F. and Ginsberg, J. P. "An Overview of Heart Rate Variability Metrics and Norms". *Frontiers in Public Health* 5 (2017), p. 258.

[195] Shimizu, W. and Antzelevitch, C. "Cellular and ionic basis for T-wave alternans under long-QT conditions." *Circulation* 99.11 (1999), pp. 1499–1507.

[196] Shivaswamy, P. K., Chu, W., and Jansche, M. "A Support Vector Approach to Censored Targets". *Seventh IEEE International Conference on Data Mining (ICDM 2007)*. Omaha, USA, 2007, pp. 655–660.

[197] Sidiropoulos, N. D., De Lathauwer, L., Fu, X., Huang, K., Papalexakis, E. E., and Faloutsos, C. "Tensor Decomposition for Signal Processing and Machine Learning". *IEEE Transactions on Signal Processing* 65.13 (2017), pp. 3551–3582.

[198] Sieed, J. and Hasan, M. "Automatic Detection and Quantification of T-Wave Alternans". *Computing in Cardiology*. Bologna, Italy, 2008, pp. 1–4.

[199] Silverthorn, D., Ober, W., and Garrison, C. *Human physiology: an integrated approach*. 2004.

[200] Smith, J. M., Clancy, E. a., Valeri, C. R., Ruskin, J. N., and Cohen, R. J. "Electrical alternans and cardiac electrical instability." *Circulation* 77.1 (1988), pp. 110–121.

[201] Stecker, E. C., Reinier, K., Marijon, E., Narayanan, K., Teodorescu, C., Uy-Evanado, A., Gunson, K., Jui, J., and Chugh, S. S. "Public Health Burden of Sudden Cardiac Death in the United States". *Circulation: Arrhythmia and Electrophysiology* 7.2 (2014), pp. 212–217.

[202] Stegeman, A. "Degeneracy in Candecomp/Parafac and Indscal Explained For Several Three-Sliced Arrays With A Two-Valued Typical Rank." *Psychometrika* 72.4 (2007), pp. 601–619.

[203] Suarez, A., Varon, C., Willems, R., Huffel, S. V., and Seisdedos, C. R. V. "PyECG: A software tool for the analysis of the QT interval in the electrocardiogram". *Revista Ingeniería Electrónica, Automática y Comunicaciones* 39.2 (2018), pp. 54–69.

[204] Suárez-León, A., Varon, C., Goovaerts, G., Vázquez-Seisdedos, C., and Van Huffel, S. "Irregular heartbeat detection using sequentially truncated multilinear singular value decomposition". *Computing in Cardiology*. Vol. 44. Rennes, France, 2017, pp. 1–4.

[205] Suárez León, A., Molina, D., Vázquez Seisdedos, C., Goovaerts, G., Vandeput, S., and Van Huffel, S. "Neural network approach for T-wave end detection: A comparison of architectures". *Computing in Cardiology*. Vol. 42. 2015.

[206] Sullivan, T. J., Deiss, S. R., and Cauwenberghs, G. "A Low-Noise, Non-Contact EEG/ECG Sensor". *2007 IEEE Biomedical Circuits and Systems Conference*. IEEE, 2007, pp. 154–157.

[207] Suykens, J. A. K., Van Gestel, T., De Brabanter, J., De Moor, B., and Vandewalle, J. *Least Squares Support Vector Machines*. 2002.

[208] Tateno, K. and Glass, L. "Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and ΔRR intervals". *Medical & Biological Engineering & Computing* 39.6 (2001), pp. 664–671.

[209] Teijeiro, T., García, C. A., Castro, D., and Félix, P. "Arrhythmia Classification from the Abductive Interpretation of Short Single-Lead ECG Records". *Computing in Cardiology*. Rennes, France, 2017, pp. 1–4.

[210] Tripathy, R. K., Sharma, L. N., and Dandapat, S. "Detection of Shockable Ventricular Arrhythmia using Variational Mode Decomposition". *Journal of Medical Systems* 40.4 (2016), p. 79.

[211] Tsipouras, M., Fotiadis, D., and Sideris, D. "An arrhythmia classification system based on the RR-interval signal". *Artificial Intelligence in Medicine* 33.3 (2005), pp. 237–250.

[212] Van Belle, V., Pelckmans, K., Van Huffel, S., and Suykens, J. A. K. "Improved performance on high-dimensional survival data by application of Survival-SVM". *Bioinformatics* 27.1 (2011), pp. 87–94.

[213] Van Belle, V, Pelckmans, K, Suykens, J. A. K., and Van Huffel, S. "On the use of a clinical kernel in survival analysis". *European Symposium on Artificial Neural Networks*. Bruges, Belgium, 2010, pp. 451–456.

[214] Van Belle, V., Pelckmans, K., Suykens, J., and Van Huffel, S. "Support Vector Machines for Survival Analysis". *Proceedings of the Third International Conference on Computational Intelligence in Medicine and Healthcare*. Plymouth, United Kingdom, 2007, pp. 1–8.

[215] Van Belle, V, Pelckmans, K, Suykens, J. A. K., and Van Huffel, S. "Survival SVM: a Practical Scalable Algorithm". *European Symposium on Artifical Neural Networks*. Bruges, 2008, pp. 89–94.

[216] Van Belle, V., Neven, P., Harvey, V., Van Huffel, S., Suykens, J. A., and Boyd, S. "Risk group detection and survival function estimation for interval coded survival methods". *Neurocomputing* 112 (2013), pp. 200–210.

[217] Van Belle, V., Pelckmans, K., Van Huffel, S., and Suykens, J. A. "Support vector methods for survival analysis: a comparison between ranking and regression approaches". *Artificial Intelligence in Medicine* 53.2 (2011), pp. 107–118.

[218] Van Belle, V. M.C. A., Van Calster, B., Timmerman, D., Bourne, T., Bottomley, C., Valentin, L., Neven, P., Van Huffel, S., Suykens, J. A. K., and Boyd, S. "A Mathematical Model for Interpretable Clinical Decision Support with Applications in Gynecology". *PLoS ONE* 7.3 (2012). Ed. by Brusic, V., e34312.

[219] Vandael, E. "The risk of QTc-prolongation and Torsade de Pointes in clinical practice". PhD thesis. KU Leuven, 2017.

[220] Vandecappelle, M., Vervliet, N., and De Lathauwer, L. "Nonlinear least squares updating of the canonical polyadic decomposition". *25th European Signal Processing Conference (EUSIPCO)*. Kos, Greece: IEEE, 2017, pp. 663–667.

[221] Vandecappelle, M., Boussé, M., Vervliet, N., and De Lathauwer, L. "CPD Updating using Low-Rank Weights". *25th European Signal Processing Conference (EUSIPCO)*. Kos, Greece, 2017, pp. 663–557.

[222] Vandenberk, B., Goovaerts, G., Garweg, C., Ector, J., and Willems, R. "Automated quantitative assessment of QRS fragmentation can improve non-invasive risk stratification". *Journal of Electrocardiology* 50.6 (2017), p. 862.

[223] Vandenberk, B., Robyns, T., Goovaerts, G., Van Soest, S., Floré, V., Garweg, C., Van Huffel, S., Ector, J., and Willems, R. "Inferior and anterior QRS fragmentation have different prognostic value in patients who received an implantable defibrillator in primary prevention of sudden cardiac death". *International Journal of Cardiology* 243 (2017), pp. 223–228.

[224] Vandenberk, B. "Non-Invasive Risk Stratification of Ventricular Arrhythmia and Sudden Cardiac Death". PhD thesis. KU Leuven, 2017.

[225] Vandenberk, B., Robyns, T., Goovaerts, G., Claeys, M., Helsen, F., Van Soest, S., Garweg, C., Ector, J., Van Huffel, S., and Willems, R. "Inter- and intra-observer variability of visual fragmented QRS scoring in ischemic and non-ischemic cardiomyopathy". *Journal of Electrocardiology* 51.3 (2018), pp. 549–554.

[226] Vapnik, V. "Pattern recognition using generalized portrait method". *Automation and Remote Control* 24 (1963), pp. 774–780.

[227] Vapnik, V. *The nature of statistical learning theory*. 2nd ed. 1998.

[228] Varanini, M, Tartarisco, G, Billeci, L, Macerata, A, Pioggia, G, and Balocchi, R. "A Multi-step Approach for Non-invasive Fetal ECG Analysis". *Computing in Cardiology*. Zaragoza, Spain, 2013, pp. 281–284.

[229] Varon, C. "Mining the ECG: Algorithms and Applications". PhD thesis. KU Leuven, 2015.

[230] Verrier, R. L., Klingenheben, T., Malik, M., El-Sherif, N., Exner, D. V., Hohnloser, S. H., Ikeda, T., Martínez, J. P., Narayan, S. M., Nieminen, T., and Rosenbaum, D. S. "Microvolt T-Wave Alternans". *Journal of the American College of Cardiology* 58.13 (2011), pp. 1309–1324.

[231] Vervliet, N., Debals, O., and De Lathauwer, L. "Tensorlab 3.0 — Numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization". *50th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2016, pp. 1733–1738.

[232] Vervliet, N., Debals, O., Sorber, L., and De Lathauwer, L. "Breaking the Curse of Dimensionality Using Decompositions of Incomplete Tensors: Tensor-based scientific computing in big data analysis". *IEEE Signal Processing Magazine* 31.5 (2014), pp. 71–79.

[233] Vítek, M., Hrubeš, J., and Kozumplík, J. "A Wavelet-Based ECG Delineation in Multilead ECG Signals: Evaluation on the CSE Database". *World Congress on Medical Physics and Biomedical Engineering*. Springer, Berlin, Heidelberg, 2009, pp. 177–180.

[234] Vollmer, M. "A robust, simple and reliable measure of heart rate variability using relative RR intervals". *Computing in Cardiology*. Nice, France: IEEE, 2015, pp. 609–612.

[235] Von Luxburg, U. "A Tutorial on Spectral Clustering". *Statistics and Computing* 17.4 (2007), pp. 395–416.

[236] Watanabe, E., Tanabe, T., Osaka, M., Chishaki, A., Takase, B., Niwano, S., Watanabe, I., Sugi, K., Katoh, T., Takayanagi, K., Mawatari, K., Horie, M., Okumura, K., Inoue, H., Atarashi, H., Yamaguchi, I., Nagasawa, S., Moroe, K., Kodama, I., Sugimoto, T., and Aizawa, Y. "Sudden cardiac arrest recorded during Holter monitoring: Prevalence, antecedent electrical events, and outcomes". *Heart Rhythm* 11.8 (2014), pp. 1418–1425.

[237] Whang, W., Mittleman, M. A., Rich, D. Q., Wang, P. J., Ruskin, J. N., Tofler, G. H., Muller, J. E., Albert, C. M., and TOVA Investigators. "Heart Failure and the Risk of Shocks in Patients With Implantable Cardioverter Defibrillators: Results From the Triggers Of Ventricular Arrhythmias (TOVA) Study". *Circulation* 109.11 (2004), pp. 1386–1391.

[238] Wilson, P. W. "Established Risk Factors and Coronary Artery Disease: The Framingham Study". *American Journal of Hypertension* 7.2 (1994), pp. 7–12.

[239] Wolf, P. A., Abbott, R. D., and Kannel, W. B. "Atrial Fibrillation as an Independent Risk Factor for Stroke: The Framingham Study". *Stroke* 22.8 (1991), pp. 983–988.

[240] Zabihi, M., Rad, A. B., Katsaggelos, A. K., Kiranyaz, S., Narkilahti, S., and Gabbouj, M. "Detection of Atrial Fibrillation in ECG Hand-held Devices Using a Random Forest Classifier". *Computing in Cardiology*. Rennes, France, 2017, pp. 1–4.

[241] Zarzoso, V. and Nandi, A. "Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation". *IEEE Transactions on Biomedical Engineering* 48.1 (2001), pp. 12–18.

[242] Zarzoso, V. "Parameter estimation in block term decomposition for noninvasive atrial fibrillation analysis". *7th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. Curaçao, Dutch Antilles, 2017, pp. 1–5.

[243] Zink, R., Hunyadi, B., Van Huffel, S., and De Vos, M. "Tensor-based classification of an auditory mobile BCI without a subject-specific calibration phase". *Journal of Neural Engineering* 13.2 (2016), p. 026005.

[244] Zipes, D. P. and Wellens, H. J. J. "Sudden cardiac death". *Circulation* 98.21 (1998), pp. 621–645.

Curriculum vitae

Griet Goovaerts was born in Turnhout on July 31, 1990. In 2013, she obtained the Master of Science in Biomedical Engineering, magna cum laude, from KU Leuven, Belgium. In September 2013, she started her PhD at the STADIUS Center of Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering (ESAT), KU Leuven under the supervision of prof. dr. ir. Sabine Van Huffel and prof. dr. Rik Willems. She received a grant from the Institute for Innovation through Science and Technology (IWT) in 2014. In 2017, she was a visiting researcher in the University of California, San Francisco (UCSF). She received a travel grant from the Flanders Research Foundation (FWO) to support this research. She works on the development of tensor-based signal processing methods for ECG signals. In 2018, she was nominated as finalist for the Rosanna Degani Young Investigator award at the annual Computing in Cardiology conference.

List of publications

Articles in internationally reviewed journals

1. **Goovaerts G.***, Padhy S.*, Varon C., Vandenberk B., Willems R., Van Huffel S., "A machine learning approach to detection of QRS fragmentation in multilead ECG signals", accepted for publication in IEEE Journal of Biomedical and Health Informatics.

2. **Goovaerts G.**, Boussé M, Do D., De Lathauwer L., Van Huffel S., Hu X., "Analysis of changes in ECG morphology prior to in-hospital cardiac arrest using weighted tensor decompositions", submitted to IEEE Journal of Transactions in Biomedical Engineering.

3. Vandenberk B., Robyns T., **Goovaerts G.**, Claeys M., Helsen F., Van Soest S., Garweg C., Ectort J., Van Huffel S., Willems R., "Inter- and intra-observer variability of visual fragemented QRS scoring in ischemic and non-ischemic cardiomyopathy", *Journal of Electrocardiology*, vol. 51, no. 3, Dec. 2017, pp 549-554.

4. Vandenberk B., Robyns T., **Goovaerts G.**, Van Soest S., Flore V., Garweg C., Van Huffel S., Ector J., Willems R., "Inferior and anterior QRS fragementation have different prognostic value in patients who received an implantable defibrillator in primary prevention of sudden cardiac death", *International Journal of Cardiology*, vol. 243, Sep. 2017, pp. 223-228.

5. **Goovaerts G.**, Vandenberk B., Willems R., Van Huffel S., "Automatic detection of T wave alternans using tensor decompositions in multilead ECG signals", *Physiological Measurement*, vol. 38, no. 8, Aug. 2017, pp. 1513-1528.

*Equal contribution

6. Sciot B., Vandenberk B., Huijghebaert S., **Goovaerts G.**, Van Huffel S., Ector J., Willems R., "Influence of food intake on the QT and QT/RR relation", *Journal of Electrocardiology*, vol. 49, no. 5, 2016, pp. 720-727.

7. Vandenberk B., Hnatkova K., **Goovaerts G.**, Garweg C., Ector J., Van Huffel S., Malik M., Willems R., "Inappropriate ICD shocks do not induce pro-arrhythmic electrocardiographic changes in men.", *Scandinavian Cardiovascular Journal*, vol. 51, no. 1, 2016, pp. 47-52.

8. De Cooman T., **Goovaerts G.**, Varon C., Widjaja D., Willemen T., Van Huffel S., "Heart Beat Detection in Multimodal Data Using Automatic Relevant Signal Detection", *Physiological Measurement*, Focus issue on robust detection of heart beats in multimodal data, vol. 36, no. 8, Aug. 2015, pp. 1691-1704.

Papers in proceedings of international conferences

1. Geirnaert S., **Goovaerts G.**, Padhy S., Boussé M., De Lathauwer L., Van Huffel S., "Tensor-based ECG signal processing applied to atrial fibrillation detection", accepted for publication in *Proc. of the 52nd Asilomar Conference on Signals, systems and computers*, Pacific Grove, United States, Oct. 2018.

2. **Goovaerts G.**, Padhy S., Varon C., Vandenberk B., Willems R., Van Huffel S., "Risk on all-cause mortality in ICD patients using a novel QRS fragmentation score", in *Proc. of the 45th Annual Computing in Cardiology* (CinC 2018), Maastricht, The Netherlands, Sep. 2018, 4 p.

3. **Goovaerts G.**, Van Huffel S., Hu X., "Tensor-based Analysis of ECG changes prior to in-hospital cardiac arrest", in *Proc. of the 44rd Annual Computing in Cardiology* (CinC 2017), Rennes, France, Sep. 2017, 4 p.

4. Moeyersons J., **Goovaerts G.**, Huijghebaert S., Vandenberk B., Willems R., Van Huffel S., "Automated T Wave End Detection Methods: Comparison of Four Different Methods for T Wave End Detection", in *Proc. of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies* (BIOSIGNALS 2017), Porto, Portugal, Jan. 2017, pp. 92-98.

5. Suarez L.A.A., **Goovaerts G.**, Varon C., Vazquez Seisdedos C.R., Van Huffel S., "Irregular Heartbeats Detection Using Sequentially Truncated Multi-Linear Singular Value Decomposition", in *Proc. of the 44th Annual Computing in Cardiology* (CinC 2017), Rennes, France, Sep. 2017, 4 p.

6. Boussé M., **Goovaerts G.**, Vervliet N., Debals O., Van Huffel S., De Lathauwer L., "Irregular Heartbeat Classification Using Kronecker Product Equations", in *Proc. of the 39th Annual International Conference of the IEEE Engineering in Medicine & Biology Society* (EMBC 17), Jeju Island, South Korea, Jul. 2017, pp. 438-441.

7. **Goovaerts G.**, Varon C., Vandenberk B., Willems R., Van Huffel S., "Phase-Rectified Signal Averaging for Automatic Detection of QRS Fragmentation", in *Proc. of the 43rd Annual Computing in Cardiology* (CinC 2016), Vancouver, Canada, Sep. 2016, pp. 637-640.

8. Suarez L.A.A., **Goovaerts G.**, Vazquez Seisdedos C.R., Van Huffel S., "Irregular heartbeats detection using tensors and support vector machines", in *Proc. of the 43rd Annual Computing in Cardiology* (CinC 2016), Vancouver, Canada, Sep. 2016, pp. 1037-1040.

9. **Goovaerts G.**, De Wel O., Vandenberk B., Willems R., Van Huffel S., "Detection of Irregular Heartbeats using tensors", in *Proc. of the 42nd Annual Computing in Cardiology* (CinC 2015), Nice, France, Sep. 2015, pp. 573-576.

10. **Goovaerts G.**, Vandenberk B., Willems R., Van Huffel S., "Tensor-based detection of T wave alternans using ECG", in *Proc. of the 37th Annual International Conference of the IEEE Engineering in Medicine & Biology Society* (EMBC 2015), Milan, Italy, Aug. 2015, pp. 6991-6994.

11. Suarez L.A.A., Molina D.M., Vazquez Seisdedos C.R., **Goovaerts G.**, Vandeput S., Van Huffel S., "Neural network approach for T-wave End detection : A comparison of Architectures", in *Proc. of the 42nd Annual Computing in Cardiology* (CinC 2015), Nice, France, Sep. 2015, pp. 589-592.

12. De Cooman T., **Goovaerts G.**, Varon C., Widjaja D., Van Huffel S., "Heart Beat Detection in Multimodal Data Using Signal Recognition and Beat Location Estimation", in *Proc. of the 41st Annual Computing in Cardiology* (CinC 2014), Cambridge, USA, Sep. 2014, pp. 257-260.

13. **Goovaerts G.**, Varon C., Vandenberk B., Willems R., Van Huffel S., "Tensor-based Detection of T Wave Alternans in Multilead ECG Signals", in *Proc. of the 41st Annual Computing in Cardiology* (CinC 2014), Cambridge, USA, Sep. 2014, pp. 185-188.

14. **Goovaerts G.**, Denissen A., Milosevic M., van Boxtel G., Van Huffel S., "Advanced EEG processing for the detection of drowsiness in drivers", in *Proc. of the International Joint Conference on Biomedical Engineering*

Systems and Technologies (BIOSIGNALS 2014), Angers, France, Mar. 2014, pp. 205-212.

Book chapters

1. Padhy S.⋆, **Goovaerts G.**⋆, Boussé M., De Lathauwer L., Van Huffel S., "The Power of Tensor-Based Approaches in Cardiac Applications", accepted for publication in *Biomedical Signal Processing - Advances in Theory, Algorithms and applications.* (G. Naik), Springer

Abstracts in proceedings of (inter)national conferences

1. Vandenberk B., **Goovaerts G.**, Garweg C., Ector J., Van Huffel S., Willems R. "Automated quantitative assessment of QRS fragmentation can improve non-invasive risk stratification", in *International Society of Computerized Electrocardiology*, St. Simons Island, USA, Apr. 2017.

2. **Goovaerts G.**, Vandenberk B., Varon C., Willems R. and Van Huffel S., "Phase-Rectified Signal Averaging for Automatic Detection of QRS Fragmentation", in *Proc. of the 6th Dutch Biomedical Engineering Conference 2017*, Egmond aan Zee, The Netherlands, Jan. 2017.

3. Moeyersons J., **Goovaerts G.**, Huijghebaert S., Vandenberk B., Willems R., Van Huffel S., Aerts J., "Effect of Food and Meals on ECG Dynamics", in *Proc. of the 15th Belgian Day on Biomedical Engineering*, Brussels, Belgium, Nov. 2016.

4. Suárez Léon A. **Goovaerts G.**, Vázquez Seisdedos C., Van Huffel S., "Algorithm for Detecting Abnormal Heartbeats Using Tensors and Support Vector Machines", in *Proc. of of the 15th Belgian Day on Biomedical Engineering*, Brussels, Belgium, Nov. 2016.

5. **Goovaerts G.**, De Wel O., Vandenberk B., Willems R., Van Huffel S., "Irregular Heartbeat Classification Using Tensors", in *Proc. of the 13th Belgian Day on Biomedical Engineering*, Brussels, Belgium, Dec. 2015.

6. De Cooman T., **Goovaerts G.**, Varon C., Widjaja D., Willemen T., Van Huffel S., "Heart Beat Detection in Multimodal Data Using

⋆Equal contribution

Automatic Signal Type Recognition", in *Proc. of the 5th Dutch Biomedical Engineering Conference 2015*, Egmond aan Zee, The Netherlands, Jan. 2015.

7. De Cooman T., **Goovaerts G.**, Varon C., Widjaja D., Willemen T., Van Huffel S., "Heart Beat Detection in Multimodal Data Using Automatic Signal Type Recognition", in *Proc. of the 13th Belgian Day on Biomedical Engineering*, Brussels, Belgium, Nov. 2014.

8. **Goovaerts G.**, Varon C., Vandenberk B., Willems R., Van Huffel S., "Tensor-based Detection of T-wave Alternans in Multilead ECG Signals", in *Proc. of the 13th Belgian Day on Biomedical Engineering*, Brussels, Belgium, Nov. 2014.

9. Vandenberk B., **Goovaerts G.**, Garweg C., Ector J., Van Huffel S., Willems R., "Inappropriate defibrillator shocks are not associated with higher pro-arrhythmic risk", in *Acta Cardiologica*, vol. 69, no.5., Belgian Heart Rhythm Meeting, Brussels, Belgium, Oct. 2014, pp. 581.

10. **Goovaerts G.**, Denissen A., Milosevic M., van Boxtel G., Van Huffel S., "Advanced EEG processing for the Detection of Drowsiness in Drivers", in *Proc. of the 12th Belgian Day on Biomedical Engineering*, Brussels, Belgium, Dec. 2013.



FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
ESAT-STADIUS
Kasteelpark Arenberg 10 - box 2446
B-3001 Leuven
griet.goovaerts@esat.kuleuven.be
http://stadius.esat.kuleuven.be