

Fachhochschule Aachen
Campus Jülich

Fachbereich 9
Medizintechnik und Technomathematik

Untersuchung der Signalqualität von ballistokardiographischen Signalen mittels Methoden des maschinellen Lernens

Bachelorarbeit
im Studiengang Scientific Programming

von

Cay Jakob Rahn
Matr.-Nr.: 3145495

1. Oktober 2020

1. Prüfer: Prof. Dr. rer. nat. Alexander Voß
2. Prüfer: Dr.-Ing. Christoph Hoog Antink

Erklärung

Diese Arbeit ist von mir selbständig angefertigt und verfasst. Es sind keine anderen als die angegebenen Quellen und Hilfsmittel benutzt worden.

Ort, Datum

Unterschrift

Abstract

Die Beurteilung der Signalqualität bei ballistokardiographischen Signalen ist für eine Anwendung in der Praxis essentiell, wirft aber besonders bei in Betten aufgenommenen Signalen noch große Fragen auf. Ziel dieser Arbeit ist es, Möglichkeiten der Artefakterkennung bei solchen Signalen mittels maschinellen Lernens zu untersuchen. Grundlage sind bereits entwickelte Verfahren und Domainenexpertise.

Schwerpunkte der Arbeit:

- Recherche zu verschiedenen Verfahren des Maschinellen Lernens
- Beurteilung der Qualität bereits existierender Verfahren für Langzeitaufnahmen von bettlägerigen Patient*innen
- Merkmalskonstruktion auf Basis von Domainenexpertise
- Vergleich verschiedener Verfahren und Eingabeparamater
- Untersuchung des Einflusses der Vorverarbeitung des Signals
- Evaluierung und Validierung der Ergebnisse

Inhaltsverzeichnis

Abkürzungsverzeichnis	ix
Abbildungsverzeichnis	xi
1 Einleitung	1
1.1 Motivation	1
1.2 Ziel der Arbeit	2
1.3 Gliederung	2
2 Grundlagen	3
2.1 Medizinische Grundlagen	3
2.1.1 Kardiorespiratorisches System	3
2.1.2 Übersicht Messtechniken	4
2.2 Ballistokardiographie	4
2.2.1 Medizinischer und technischer Hintergrund	4
2.2.2 Einsatzgebiet	5
2.2.3 Signaleigenschaften	6
2.3 Maschinelles Lernen	11
2.3.1 Grundprinzipien	11
2.3.2 Mathematischer Hintergrund	12
2.3.3 Evaluation und Validierung	14
2.3.4 Probleme maschinellen Lernens	15
2.3.5 Weitere Lernmodelle des Überwachten Lernens	16
3 Signalverarbeitung bei ballistokardiographischen Signalen	19
3.1 Grundsätzliches	19
3.2 Detektion von Herzschlägen	20
3.3 Artefakterkennung	22
3.3.1 Allgemein bei kardiorespiratorischen Signalen	22
3.3.2 Bei ballistokardiographischen Signalen	24
3.3.3 Schwellwertbasierte Artefakterkennung	25
3.3.4 Maschinelles Lernen mit statistischen Merkmalen	26
3.3.5 Ähnlichkeit der Intervallschätzer des CLIE-Algorithmus	28

3.4	Messdaten	28
3.4.1	Erfassung	28
3.4.2	Vorliegende Form	29
3.4.3	Verarbeitung und Datenstruktur	29
3.4.4	Annotation der Daten	29
4	Analyse	31
4.1	Anwendung existierender Verfahren	31
4.1.1	Ähnlichkeit der Intervallschätzer nach Brüser	31
4.1.2	Schwellwerte für Standardabweichung, Minimum und Maximum	31
4.1.3	Maschinelles Lernen mittels statistischer Merkmale	31
4.2	Datenanalyse und Merkmalskonstruktion	31
5	Synthese	33
5.1	Eingabeform der Daten	33
5.2	Evaluationsmetriken	33
5.3	Modelle	33
6	Evaluierung der Ergebnisse	35
7	Zusammenfassung und Ausblick	37
7.1	Zusammenfassung	37
7.2	Ausblick	37
	Literatur	39

Abkürzungsverzeichnis

BKG Ballistokardiographie

EKG Elektrokardiographie

PPG Photoplethysmographie

SKG Seismokardiographie

HR Herzrate

HRV Herzratenvariabilität

SQI Signal Quality Index

DTW Dynamic Time Warping

SVM Support Vector Machine

Abbildungsverzeichnis

2.1	Übersicht über die Funktionsweise eines allgemeinen im Bett eingebetteten Ballistokardiographie (BKG)-Systems	6
2.2	Beispiel eines typischen BKG-Signals mit Nomenklatur	7
2.3	Hochpass-gefilterte BKG-Aufnahmen von zwei Herzschlägen zwei verschiedener Personen, jeweils in Rücken- und Seitenlage gemessen, die vertikalen Linien markieren die R-Peaks der EKG-Referenz. ¹	8
2.4	Visualisierung der Variabilität des BKG-Signals	10
2.5	Darstellung von Supervised Learning	11
3.1	Intervallschätzer nach Brüser et al.	22
3.2	Flussdiagramm eines Algorithmus zur Beurteilung der Signalqualität . . .	24
3.3	Artefakte mit hoher Energie	25
3.4	Artefakte mit niedriger Energie	25
3.5	Bland-Altman Graph zwischen von Elektrokardiographie (EKG) und BKG berechneter Herzrate (HR)	26

1 Einleitung

1.1 Motivation

Der derzeitige demographische Wandel stellt das Gesundheitssystem vor eine große Herausforderung: Immer mehr Patient*innen müssen im Alter überwacht und versorgt werden. Eine kontinuierliche autonome Überwachung von Vitalparametern im Krankenhaus oder auch Zuhause erlaubt es, Erkrankungen frühzeitig zu erkennen oder zu beobachten, ohne dass große Personalkapazitäten von Nöten sind.

Für diesen Anwendungszweck eignen sich vor allem Messmethoden, die die Patient*innen im Alltag nicht einschränken und wenig invasiv sind. Im Englischen wird dies mit dem Begriff *unobtrusive* bezeichnet. Da es keine zufriedenstellende deutsche Entsprechung gibt, wird dieser im Folgenden nicht übersetzt verwendet werden. Solche *unobtrusive* Messmethoden beinhalten meist keine Notwendigkeit für direkten Körper- oder Hautkontakt, liefern aber Information über Atmung und Herzschlag. Die Herausforderung bei so ermitteltem Signal besteht in der Signalverarbeitung, da Messungenauigkeiten und Alltagsbewegungen zu Störungen im Signal führen. Nicht informatives, also nicht für die Verarbeitung geeignetes Signal muss aber zwingend identifiziert werden, da die Ergebnisse stark verfälscht werden.

Eine solche *unobtrusive* Messmethode ist die Ballistokardiographie (BKG). Sensoren lassen sich beispielsweise in Betten und Stühlen implementieren. Aufgezeichnet werden Aktivitäten des Herzens und der Atmung. Die Signalmorphologie variiert jedoch sowohl zwischen den Patient*innen als auch innerhalb einer Person sehr stark, wodurch die automatische Beurteilung der Signalqualität erschwert wird. Um eine aussagekräftige Signalverarbeitung zu ermöglichen, ist dies jedoch essentiell. Besonders bei in Betten aufgenommenem Signal ist die Variation des Signals in Kombination mit Artefakten durch Körperbewegungen oder ähnliches problematisch.

1.2 Ziel der Arbeit

Das Ziel dieser Arbeit ist es, Möglichkeiten der Beurteilung der Signalqualität von BKG-Signalen mittels maschinellen Lernens zu untersuchen. Im besonderen Fokus liegen dabei Langzeitaufnahmen von bettlägerigen Patient*innen, da diese sich in der Vergangenheit als besonders anfällig für geringe Signalqualität gezeigt haben.

Dafür werden zunächst existierende Verfahren der Artefakterkennung für die vorliegenden Daten getestet und bewertet. Anschließend wird auf Basis von Domainenexpertise Merkmalskonstruktion betrieben und verschiedene Verfahren und Eingabeparameter verglichen. Auch wird der Einfluss der Vorverarbeitung des Signals untersucht.

Langfristig soll ermöglicht werden, Ballistokardiographie (BKG) im medizinischen Alltag anzuwenden.

1.3 Gliederung

2 Grundlagen

2.1 Medizinische Grundlagen

Die vorliegende Arbeit beschäftigt sich mit der Beurteilung der Signalqualität in ballistokardiographischen Signalen. Zum Verständnis der gemessenen Vorgänge und der Problematik in Bezug auf die Signalqualität und dessen Beurteilung ist grundlegendes medizinisches Wissen über die gemessenen Vorgänge und messtechnisches Verständnis nötig. Aufgrund dessen wird hier eine kurze Übersicht über die medizinischen Grundlagen gegeben.

2.1.1 Kardiorespiratorisches System

Das kardiorespiratorische System (zusammengesetzt aus *kardia*, deutsch 'Herz' und *respiratio*, deutsch 'Atmung') setzt sich aus zwei Teilsystemen zusammen, dem kardiovaskulären und dem respiratorischen System, die zusammen die Versorgung der Organe mit sicherstellen.

Das kardiovaskuläre System umfasst das Herz, die Arterien und die Venen. In einem Zyklus wird das sauerstoffreiche Blut von der linken Herzkammer durch die Arterien zu den Organen gepumpt, wo sich der Sauerstoff zur Versorgung dieser vom Blut löst. Die Venen transportieren das nun sauerstoffarme Blut in die rechte Herzkammer. Von dort wird es zur Lunge geführt, mit Sauerstoff angereichert und in die linke Herzkammer geleitet. Damit schließt sich der Zyklus. Die Herzfrequenz ist hierbei und relevanter messbarer Vitalparameter.

Ein Herzschlag selbst besteht aus zwei Phasen: einer füllenden und einer auswerfenden Phase. Während der Diastole, der Erschlaffungs- und Bluteinströmungsphase, füllen sich die Herzkammern mit Blut. Diese Phase endet mit dem Schließen der Herzklappen und die Systole beginnt. Die Systole ist die Anspannungs- und Blutausströmungsphase: Die Herzklappen öffnen sich durch Kontraktion des Herzmuskels und das Blut kann ausströmen.

Das respiratorische System umfasst die Lungen und den Lungenkreislauf. In einem Atemzyklus wird durch gezielte Muskelbewegungen Luft aus der Umgebung eingeatmet. Mit dem eingeatmeten Sauerstoff wird sauerstoffarmes Blut angereichert und anschließend die nun sauerstoffarme Luft ausgeatmet. Hier ist der Vitalparameter der Atemfrequenz messbar.

2.1.2 Übersicht Messtechniken

Die untersuchte BKG wird zur Untersuchung oft mit anderen Messmethoden als Referenz aufgenommen. Im Folgenden werden diese kurz vorgestellt. BKG selbst wird im nächsten Abschnitt separat betrachtet.

Die EKG zeichnet die elektrischen Aktivitäten des Herzmuskels auf, indem mit mehreren Elektroden die Spannungsänderung gemessen wird. Hier ist die Herzfrequenz sehr gut ablesbar.

Die Photoplethysmographie (PPG) ist ein optisches Messverfahren, bei dem die Menge des von der Haut reflektierten bzw. transmittierten Lichtes gemessen wird. Dadurch kann die Änderung des Blutvolumens gemessen werden; die Lichtmenge nimmt bei Durchlaufen einer Pulswelle durch die Arterie deutlich ab. Dieses Signal bietet Rückschluss auf Atmung und Herzschlag.

Oft gemeinsam mit dem BKG betrachtet wird die Seismokardiographie (SKG), bei der die Vibration der Wand des Brustkorbs durch den Herzschlag aufgezeichnet wird.

2.2 Ballistokardiographie

Im Folgenden wird die Ballistokardiographie eingeführt. Das beinhaltet den medizinischen und technischen Hintergrund, das Einsatzgebiet und die Signaleigenschaften. Des Weiteren wird näher beleuchtet, welche Probleme sich bei der Beurteilung der Signalqualität durch die Eigenschaften des Signals ergeben.

2.2.1 Medizinischer und technischer Hintergrund

Ballistokardiographie (zusammengesetzt aus altgriechisch *ballein*, deutsch ‘werfen’, *kardía*, deutsch ‘Herz’ und *graphein*, deutsch ‘schreiben’) ist die graphische Darstellung der wiederholten, durch den Herzschlag verursachten Bewegungen des menschlichen

Körpers. Erstmals schon im 19. Jahrhundert beobachtet¹, ermöglicht der technische Fortschritt in der Sensortechnik heute aussagekräftige Messungen. Das BKG liefert durch die Aufzeichnung von zirkulierendem Blut und mechanischer Herzaktivität Informationen über die Gesamtleistung des kardiovaskulären Systems.² Konkret gemessen wird eine Massenbewegung, die durch die schnelle Beschleunigung des Blutes entsteht, wenn es während des Herzschlages durch die großen Arterien bewegt wird: Bei der Verteilung des Bluts in die peripheren Blutgefäße verschiebt sich das Zentrum der Körpermasse in Richtung der Füße und während der atrialen Systole Richtung Körpermitte. Die BKG-Wellenform entsteht durch diese Schwerpunktverschiebung.

Die Messung dieser Bewegung ist mit verschiedenen Sensortypen, die z.B. hydraulisch oder elektromechanisch auf Druck reagieren, möglich. Sensoren können unter anderem in Waagen, Stühlen und Betten eingebaut werden. Besonders bei im Bett gemessenen Signalen kann oft nicht klar zwischen SKG und BKG unterschieden werden, da sich myokardiale Vibrationen und Massverschiebungen durch den Blutfluss überlagern. Diese gemischten Signale werden in der Literatur teils auch als *cardiac vibration signals* bezeichnet.³ Da im Bereich der Signalverarbeitung oft nicht zwischen reinem BKG und gemischten Signalen unterschieden wird, wird dies in der vorliegenden Arbeit ebenfalls nicht.

Verschiedene Studien kommen zu unterschiedlichen Ergebnissen bezüglich der Frage, welchen kardiovaskulären Ursprung die einzelnen Signalteile haben. Aufgrund dessen gestaltet sich die detaillierte Interpretation des BKG-Signals als schwierig. Da es neben Informationen zur HR und Herzratenvariabilität (HRV) ein genauerer Indikator für das Alter des Herzens als Lebensalter ist, hat es trotzdem klinische Relevanz. Außerdem lassen sich durch abnormale Ballistokardiogramme Herzerkrankungen voraussagen, bevor Symptome auftreten. Besonders bei älteren Personen sind diese also eine wichtige Warnung.⁴

2.2.2 Einsatzgebiet

Durch diese Beschreibung wird schon deutlich, dass BKG anders als das sehr bekannte EKG ist. Der entscheidende Vorteil des BKGs liegt darin, dass kein einschränkender Körperkontakt wie z.B. aufgeklebte Elektroden nötig ist: Es lässt sich in Alltagsgegenständen wie Stühlen aber vor allem auch Betten implementieren, ohne dass während der

¹Vgl. Gordon 1877.

²Vgl. Pinheiro et al. 2010.

³Vgl. Brüser et al. 2013.

⁴Vgl. zu diesem Absatz Pinheiro et al. 2010.

Messung zu Einschränkungen im alltäglichen Leben kommt oder medizinisches Fachpersonal anwesend sein muss. Damit gehört es zu den *unobtrusive* Messmethoden und eignet sich gut zur Langzeit- und Trendbeobachtung des Gesundheitszustandes - sowohl im klinischen Kontext als auch Zuhause. Besonders für Patient*innen mit chronischen Krankheiten und zur Früherkennung krankhafter Veränderungen bietet eine gesundheitliche Überwachung von Zuhause großes Potential.⁵ Zusätzlich zu Informationen der Herzaktivitäten bietet in Betten eingebautes BKG auch Informationen über das allgemeine Aktivitätslevel und somit auch über die Schlafqualität.⁶ In dieser Arbeit wird es um die Aufzeichnung von BKG-Signalen in Betten gehen.

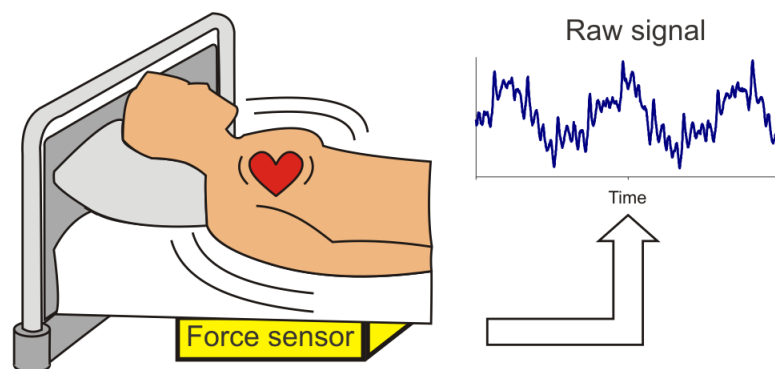


Abbildung 2.1: Übersicht über die Funktionsweise eines allgemeinen im Bett eingebetteten BKG-Systems.⁷

Allerdings ergeben sich neben diesen umfassenden Möglichkeiten auch Nachteile gegenüber konventionellen Messmethoden. Die größte Herausforderung ist eine stark variierende Signalqualität, die sich durch das unkontrollierte Umfeld und die Art der Messung ergibt.

2.2.3 Signaleigenschaften

Das gemessene BKG-Signal setzt sich aus Herzaktivitäten, Atmungsaktivitäten und Körperbewegungen zusammen. Gegebenenfalls wird es noch durch Störungen der Messung beeinflusst. Bei einer gesunden Person ohne Störeinflüsse wird die in 2.2 abgebildete Wellenform erwartet. Diese Idealform lässt sich in 3 Gruppen unterteilen: Die präsysstolische, wobei diese häufig nicht beachtet wird, die systolische und die diastolische Gruppe unterteilen. Die mit H bis K markierten Extremwerte gehören bei dieser Unterteilung zur

⁵Vgl. Inan et al. 2015.

⁶Vgl. Brüser et al. 2011.

⁷Entnommen aus Brüser et al. 2011

systolischen Gruppe, die Wellen L bis N zur diastolischen Gruppe. Die präsysstolische Gruppe, die aus den Wellen F und G besteht, ist in hier nicht abgebildet. I und J werden auch als *ejection waves* bezeichnet. In Bezug auf andere Messmethoden ist zu bemerken, dass die H-Welle nahezu synchron mit dem ersten Herzgeräusch ist. Der Abstand des R-Peaks, des Hochpunkts eines EKGs zur H-Welle variiert im Bereich von 0,2 bis 0,3 Sekunden.⁸ Die Amplitude der Wellen ohne Störeinflüsse ist hauptsächlich abhängig von dem Herzzeitvolumen, der Herzkraft und der Geschwindigkeit des Auswurfs.⁹

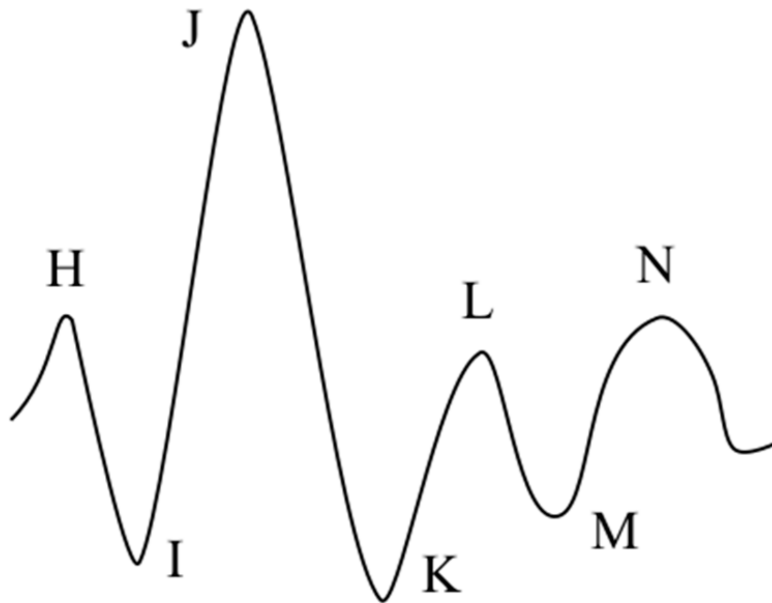


Abbildung 2.2: Beispiel eines typischen BKG-Signals mit Nomenklatur¹⁰

Im Idealfall wird zwar die oben beschriebene Wellenform erwartet, bei der die Wellen H bis L eine deutliche W-Form bilden, allerdings ist es trotz dieser typischen Form selten, dass alle nicht-systolischen Komponenten sichtbar sind.¹¹ Es gibt eine starke Variation der Signalmorphologie sowohl zwischen als auch innerhalb von Individuen. Der größte Einfluss ergibt sich durch die verwendeten Sensoren und die Position der Person, also zum Beispiel ob im Stehen, Sitzen oder Liegen gemessen wird.¹² Es gibt Studien die zeigen, dass die intraindividuelle Varianz über serielle Messungen hinweg niedrig ist.¹³ Allerdings gilt das nicht, wenn sich die Position der Person verändert. Hierbei reicht es schon, wenn die Person in Rückenlage statt Seitenlage liegt.¹⁴ Aufgrund dieser Variationen in der Signalmorphologie wurden schon in den 1950er Jahren 3 Achsen für die

⁸Vgl. de Lalla et al. 1950.

⁹Vgl. Pinheiro et al. 2010.

¹⁰Entnommen aus Albukhari et al. 2019 nach Starr et al. 1939.

¹¹Vgl. Pinheiro et al. 2010.

¹²Vgl. Sadek, Biswas und Abdulrazak 2019.

¹³Vgl. Inan et al. 2015.

¹⁴Vgl. Brüser et al. 2011.

Aufzeichnung des BKGs definiert: Die longitudinale (Kopf-Fuß), die transversale (Seite-Seite) und die dorsoventrale (Rücken-Brust).¹⁵ Zu Beginn maßen die meisten Systeme entlang der longitudinalen Achse, die z.B. der Messung auf einer Waage entspricht. *Unobtrusive* Messsysteme, wie die hier betrachtete Messung in Betten, messen entlang einer Kombination der transversalen und der dorsoventralen Achse - abhängig von der Position der Person. Besonders diese Kombination sorgt für eine große intra- und individuelle Variation des Signals. Abbildung 2.3 verdeutlicht dies durch den direkten Vergleich von BKG-Aufzeichnungen zweier Herzschläge von 2 Personen. Bei jedem Proband wurde in 2 verschiedenen Positionen gemessen.¹⁶ Auch der Ursprung des Signals ist abhängig von der Messachse. Bei longitudinale gemessenem BKG ist der Einfluss des Herzzeitvolumens schon seit 1929 beobachtet.¹⁷ Im Gegensatz dazu ist der Ursprung des in Betten gemessenen BKG-Signals nicht genau bekannt. Das liegt unter anderem daran, dass mechanische Komponenten wie z.B. die Matratze einen schwer zu modellierenden Einfluss haben.

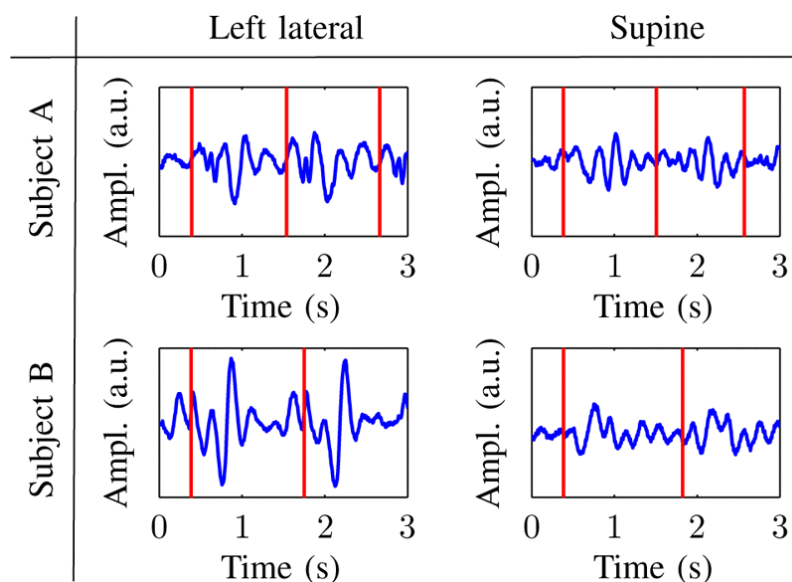


Abbildung 2.3: Hochpass-gefilterte BKG-Aufnahmen von zwei Herzschlägen zweier verschiedener Personen, jeweils in Rücken- und Seitenlage gemessen. die vertikalen Linien markieren die R-Peaks der EKG-Referenz.¹⁸

Neben Einflüssen der verwendeten Messachse und der Körperposition beeinflusst auch die Atmung die Signalform. Normale Atmung beeinflusst die Amplitude der *ejection waves* I und J. Bei Atemstillstand dagegen werden die H und J Wellen verzerrt. Auch bei einer gesunden, sich nicht bewegenden Person, die ihre Atmung kontrolliert, wird kein

¹⁵Brüser et al. 2011; Inan et al. 2015, Vgl.

¹⁶Brüser et al. 2011.

¹⁷Vgl. Starr et al. 1939.

¹⁸Entnommen aus Brüser et al. 2011.

exakt Schlag für Schlag reproduzierbares Signal erzeugt werden.¹⁹ Von Zink et al. werden die Einflüsse der Atmung in der vertikalen Achse eines dorsoventralen BKGs als große Schwingungen einer Wellenlänge von fünf bis zehn Sekunden beschrieben. Innerhalb dieser sind kleinere Schwingungen mit höherer Frequenz sichtbar, die jedoch keiner bestimmten Sequenz folgen.²⁰ Zusätzlich zu dieser schon beschriebenen Variabilität kommt es sehr leicht zum Entstehen von Artefakten. Ursprung ist entweder das Messsystem selbst oder Körperbewegungen. Insgesamt führt Bewegung der Patient*innen, auch die der Atmung, zu einem *baseline drift*. Stärkere Bewegungen führen zu einer Massenschiebung, die um ein Vielfaches größer als die gemessenen Vorgänge ist. Aufgrund dessen führt sie immer dazu, dass das Signal stark verzerrt oder sogar vollständig überlagert wird.

Besonders im Vergleich zu anderen kardiorespiratorischen Signalen wie dem EKG und PPG wird deutlich, dass BKG-Signale auch in konsekutiven Messungen deutlich variabler sind. Abbildung 2.4 zeigt dies am Beispiel von BKG-Aufnahmen eines im Bett integrierten Messsystems im Vergleich zum parallel aufgenommenen EKG. Es zeigt sich, dass selbst nach Entfernung von Überlagerungen von Atmung und Bewegung das BKG-Signal eine höhere Variabilität in Bezug auf Amplitudenhöhe, Reihenfolge der Extremwerte und der gesamten Form aufweist.²¹ Es wird allerdings angenommen, dass aufeinander folgende Herzschläge sich ähneln. Diese Eigenschaft wird Selbstähnlichkeit genannt. Brüser et al. nennt als eine mögliche Ausnahme den Fall, dass ein unregelmäßiger Herzschlag mit sehr niedrigem Schlagvolumen einem regulären Herzschlag folgt. In dem Fall ist es möglich, dass die Amplitude im Vergleich so klein ist, dass sie verdeckt wird. Dies ist z.B. bei Vorhofflimmern möglich. Eine Untersuchung von Rosales et al. zeigt dieses Verhalten der Selbstähnlichkeit nicht bei den kleineren Extremwerten die J umgeben. Dass die Ähnlichkeit um J am größten ist zeigt auch Abbildung 2.4.

¹⁹Vgl. Pinheiro et al. 2010.

²⁰Vgl. Zink et al. 2017.

²¹Vgl. Zink et al. 2017.

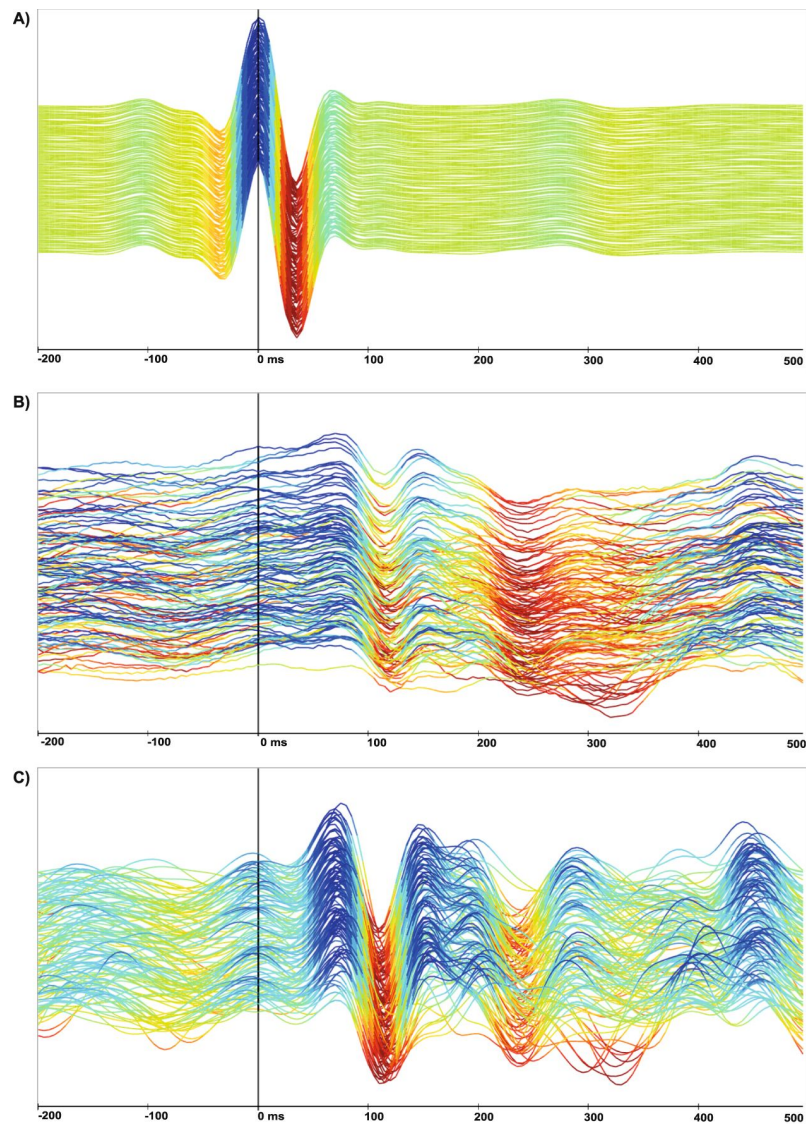


Abbildung 2.4: Diagramm aus 128 konsekutiven Herzschlägen im EKG (A) und BKG (B,C), segmentiert durch das EKG. Die Farben dienen der besseren Visualisierung der Amplituden. (A) EKG-Signal; (B) BKG-Signal mit Überlagerungen durch Atmung und Bewegung; (C) BKG-Signal ohne Bewegungsartefakte und Atmung.²²

Zusammengefasst lässt sich sagen, dass es sich bei ballistokardiographischen Signalen um nichtlineare, nichtstationäre Signale handelt, dessen Ursprung nicht genau bekannt ist. Die Signalform wird von der Messachse, der Position und Körperhaltung der Proband*innen und dem Messsystem selbst beeinflusst. Besonders bei dem hier im Fokus liegenden Anwendungsfall Bett kommt es sowohl durch die unkontrollierbare Umgebung als auch die Signaleigenschaften selbst zu einer starken Variation der Morphologie und vielen Artefakten im Signal. Trotz dieser Einschränkungen ist die Ballistokardiographie

²²Entnommen aus Zink et al. 2017.

eine Messtechnik, die sich einfach *unobtrusive* in den Alltag einbauen lässt und Aussagen über die Herzrate und die Herzratenvariabilität ermöglicht.

2.3 Maschinelles Lernen

2.3.1 Grundprinzipien

- „künstliche“ Generierung von Wissen auf Basis von Erfahrung: Lernen aus Beispielen -> verallgemeinern (Erkennung von Mustern, statistisches Modell basierend auf Daten mit denen trainiert wurde)
- 3 Arten: supervised, unsupervised, reinforcement
- Trainingsdaten X : Merkmale in Vektoren $x \in X \subseteq \mathbb{R}^n$
- supervised, zB für Klassifizierung von Daten
- Ziel Y bekannt
- Daten sind Eingabe-Ausgabe-Paare $(x_1, y_1), \dots, (x_n, y_n), x_i \in X$
- nur Funktion $f : X \rightarrow Y$ unbekannt
- gesucht: g dass f approximiert

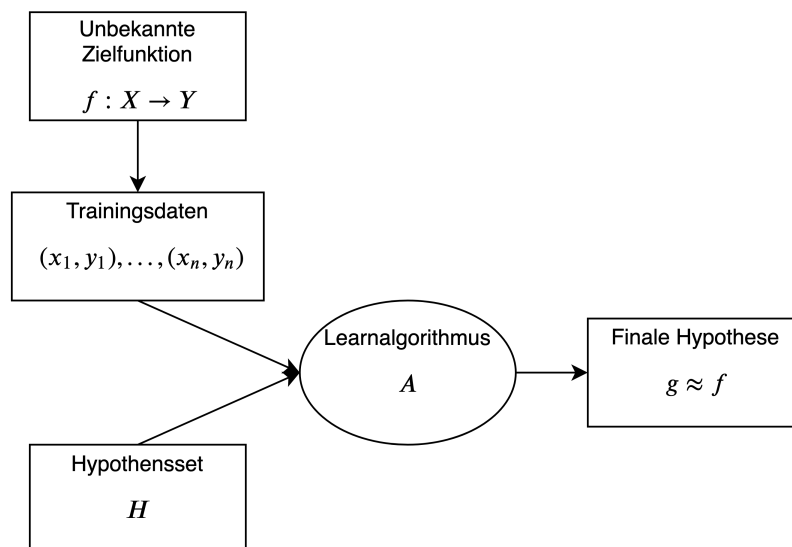


Abbildung 2.5: Darstellung von Supervised Learning

- Unterscheidung zwischen Regression, Klassifikation
- Klassifikation: unterteilt Daten in Klassen

- Regression: beschreibt Verteilung der Daten, kann z.B. genutzt um Verkaufszahlen von Büchern vorauszusagen
- logistische Regression: Wahrscheinlichkeit Angehörigkeit Klasse
- unsupervised, Beschreibung von Daten
- keine Labels -> Form der Eingabedaten (x_1, \dots, x_n)
- $f : X \rightarrow Y$
- gesucht: f , das Daten gut beschreibt
- -> erstellt selbstständig Klassifikationen
- Beispiel: Clustering zur Einteilung der Daten in Kategorien
- reinforcement: Strategie gesucht, die eine Belohnung maximiert
- in Arbeit supervised Learning betrachtet -> Daten müssen annotiert sein

2.3.2 Mathematischer Hintergrund

- auch hier nur supervised betrachtet
- Verteilung der Daten
- linear separierbar oder nicht: 3 Fälle (linear, linear mit Rauschen, nicht linear)
- Abbildung dazu?
- Ziel ist, eine Funktion zu finden, die die verschiedenen Klassen voneinander möglichst genau separiert, wodurch y ermittelt werden kann bzw. bei einer Regression den Wert möglichst genau vorhersagt, Entscheidungsfunktion genannt
- Lineare Modelle kombinieren Merkmale linear miteinander
- einfachstes Modell für Klassifikation von linear separierbare Daten: Perzeptron
- Komponenten Merkmalsvektor x unterschiedlich gewichtet
- mit Gewichtungsvektor $w \in \mathbb{R}^n$
- Ausgabe binärer Klassifikator $y \in Y = \{-1, 1\}$

- Ermittlung Schwellwert *Bias* b mit

$$y = +1, \text{ falls } \sum_{i=1}^n w_i \cdot x_i > b$$

$$y = -1, \text{ falls } \sum_{i=1}^n w_i \cdot x_i < b$$

- wenn $w_0 = b$ und $x_0 = 1$ durch Umformen Perzeptron Entscheidungsfunktion $h(x) = \text{sign}(w^T x)$
- w so gewählt, dass Daten korrekt klassifiziert
- wenn Perzeptron künstliches Neuron in Neuronalem Netz wird diese Funktion auch Aktivierungsfunktion genannt und kann variiert werden
- Signum Funktion hat Nachteil des "hartenSSchwellwerts, oft Funktionen wie logistische Funktion mit "weichemSSchwellwert besser
- Lineare Regression: Gerade $h(x)$ finden, die Datenpunkte beschreibt
- Methode der kleinsten Quadrate, also quadratische Abweichungen zwischen Gerade und Punkten minimieren
- gesucht sind Koeffizienten w ,

$$\underset{w}{\operatorname{argmin}} \sum_{i=1}^n (w^T x_i - y_i)^2$$

- durch Auflösen $w = (X^T X)^{-1} X^T$
- bis jetzt nur lineare einfache Modelle betrachten
- bei komplexeren Minimierungsproblemen Methode kleinster Quadrate nicht ausreichend, andere Methoden zur Fehlerminimierung -> Gradientenabstiegsverfahren
- graphische Vorstellung als Hügellandschaft
- In jedem Schritt Ableitung der Kostenfunktion nach jedem Gewicht und Bias berechnet -> Schritt wählen, der Kostenfunktion am stärksten minimiert
- wenn Daten nicht linear separierbar: Kerneltrick möglich
- Transformation (Ersetzen) des Skalarprodukts transformiert implizit den Variablenraum

- Beispiel Gaußscher RBF-Kernel

$$K_{RBF}(x, x') := \exp(-\gamma \|x - x'\|^2) \text{ mit } \gamma > 0$$

- lineare Trennung in höher-dimensionalem Variablenraum
- Berücksichtigung aller Nicht-Linearitäten in der Transformation -> Trennfläche kann wiederum linear sein

2.3.3 Evaluation und Validierung

- Modell auf unbekannten Daten validieren
- Hold-Out-Validierung: zufällige Verteilung in Trainings- und Testset
- Kreuzvalidierung
- Daten werden auf v gleich große Mengen (Folds) verteilt
- V Modelle auf allen möglichen Folds trainiert
- validierung auf ausgeschlossenen Fold
- extrem: Leave-One-Out Kreuzvalidierung mit v = n (Menge der Datenpunkte)
- üblich: v-fache Kreuzvalidierung
- typische Werte v=5 oder 10
- auch abhängig von verfügbarer Rechenleistung
- Wofür
- Schätzung von Fehler auf unbekannten Daten
- Wahl von Modellen/Bestimmung von Hyperparametern
- Parameter, die Modellarchitektur bestimmen
- Parameter, die Lernalgorithmus betreffen
- Regularisierungsparameter
- Ergebnis: 3 Datensets
- training, validierung, test
- Validierungsset beeinflusst Lernalgorithmus
- Testset nachdem alle Entscheidungen getroffen wurden

- typisch: Validierung durch zB Kreuzvalidierung auf Testset
- Retraining auf ganzem Trainingsset
- Ablaufdiagramm
- dafür evaluationsmetrik nötig -> Übergang
- verschiedenste Metriken
- begriffe TP, FP, TN, FN
- Accuracy $ACC = \frac{TP+TN}{TP+TN+FN+FP}$ einfaches Gütemaß, problematisch bei ungleich großen Klassen, da eine hohe Genauigkeit erreicht wird, wenn immer die größere Klasse vorausgesagt wird
- Precision, "bestraft" Falsch-Positive $PPV = \frac{TP}{TP+FP}$
- Recall, auch Sensitivity genannt, "bestraft" Falsch-Negative $TPR = \frac{TP}{TP+FN}$
- Oft Recall und Precision in einem Maß zusammengefasst: F1 score harmonisches Mittel aus beiden $F_1 = 2 * \frac{precision * recall}{precision + recall}$
- nicht balancierte Datensets
- balanced accuracy
- roc kurz für Receiver Operating Characteristic. basiert auf True Positive Rate TPR und False Positive Rate FPR, charakterisiert, wie gut beide Verteilungen durch Schwellwert trennbar sind
- auc Fläche unter der ROC Kurve

2.3.4 Probleme maschinellen Lernens

- offensichtlich: Underfitting
- Overfitting -> gute Approximation aber schlechte Generalisierung auf unbekannte Daten
- Einfluss von: Komplexität des Lernmodells, Datenmenge, Kontamination der Daten mit Rauschen
- mehr Datenpunkte mehr gut, Rauschen und Komplexität schlecht
- TODO: Bild
- Regularisierung gegen Overfitting: TODO näher nachlesen

- Stichprobenverzerrung

2.3.5 Weitere Lernmodelle des Überwachten Lernens

Entscheidungsbäume

- Bild mit Thresholds und Features
- an jedem Knoten Entscheidung ob bestimmtes Feature über Threshold liegt
- generiert z.B. durch rekursives binäres Teilen
- bei Regression mittlere quadratische Abweichung MSE minimieren, bei Klassifikationsbäumen Messung der "Unreinheit" der Blätter (Fehler = 0 wenn Blatt nur Punkte derselben Klasse, Fehler maximal wenn keine Klasse Mehrheit)
- Optimierung rechnerisch nicht effizient lösbar
- greedy den Schritt, also die Kombination aus Feature und Threshold wählen, der den Fehler am Stärksten reduziert
- Abbruchkriterien: maximal erlaubte Tiefe, Mindestanzahl Datenpunkte in einem Blatt
- keine Featureskalierung nötig

Random Forest

- besteht aus mehreren unkorrelierten Entscheidungsbäumen
- Randomisierung bei Erstellung der Bäume
- Mehrheitsentscheidung aller Bäume
- Nach Breiman für jeden Baum im Wald
- n Bootstrap Samples ziehen, also mit Zurücklegen aus dem Trainings-Datensatz gezogen
- von M Merkmalen $m \ll M$ Merkmale zufällig gewählt, die als Kriterium für Split infrage kommen
- voller Ausbau des Baums

Nächste Nachbarn Modelle

- sehr einfache Modelle, öfters zur Schätzung einer Baseline im Einsatz
- Klassifikation anhand der Datenpunkte, die zu klassifizierenden Datenpunkt am nächsten liegen
- dafür Ähnlichkeit quantifizieren
- kein Training, Modell direkt durch Trainingsdaten definiert
- Voronoi-Regionen werden erzeugt: Menge aller Punkte die näher an einem Zentrum liegen als an allen anderen Zentren
- daraus Voronoi-Diagramm (gesammelte Grenzen)
- k-Nächste Nachbarn: das Label der Mehrheit der k nächsten Punkte wird zurückgegeben
- Regression: Mittelwert-Bildung über die Label der k nächsten Nachbarn

Support Vector Machines

- Rauschen in Daten
- statt einfache Gerade wie bei Perzeptron wird Hyperebene mit maximalem Rand gesucht, die Daten korrekt klassifiziert
- TODO: Bild von Gerade mit Rand
- Ausgangspunkt Perzeptron, Umformulierung in Optimierung mit Nebenbedingung, dass Hyperebene größten Rand besitzt
- bei SVMs oft oben schon erwähnter Kernel-Trick, also Transformation des Skalarprodukts, damit nicht-lineare Separierung möglich

Mehrlagiges Perzeptron

- Klasse von künstlichen neuronalen Netzwerken
- 3 Schichten: Eingabe-Schicht, Versteckte Schicht, Ausgabe-Schicht
- außer den Eingabeneuronen wird nicht-lineare Aktivierungsfunktion verwendet

3 Signalverarbeitung bei ballistokardiographischen Signalen

3.1 Grundsätzliches

- Signale sind durch quasiperiodische Natur des Herzens selbst quasiperiodisch
- gleichzeitige Messung von zwei verschiedenen Vorgängen gleichzeitig
- Morphologie variiert im Verlauf eines Atemzyklus
- Filterung nach Frequenzen durch Bandpass-Filterung möglich
- Normbereich Atemfrequenz: 12 bis 25 Atemzüge pro Minute, ober- und unterhalb abnormal
- Herzfrequenz: 30 bis 210 Schläge pro Minute, dabei 30 Schläge pro Minute Pulsabsenkung nachts, Ruhepuls ist höher
- durch Eigenschaften von BKG-Signalen schwieriger in Signalverarbeitung als andere kardiorespiratorische Signale
- Paalasmaa et al. Zitat:

The properties of the BCG signal vary so much in practice that no simple filtering rule can be devised for an accurate and reliable beat-to-beat interval detection¹

- verschiedene Arten Signalverarbeitung: Arbeit im Zeitbereich, Arbeit im Frequenzbereich
- Zeitbereich: oft basierend auf existierendem Wissen über Morphologie des physiologischen Signals -> bei BKG schwierig, da Morphologie sehr variabel
- frequenzbasiert: Analyse von spektralen Eigenschaften
- Analysen im Frequenzbereich erstmal nur durchschnittliche Frequenzen
- für einige medizinische Anwendungen ausreichend, für andere, zB HRV nicht

¹Paalasmaa et al. 2015.

3.2 Detektion von Herzschlägen

In dieser Arbeit wird der von Brüser et al. entwickelte Algorithmus, der *Continuous Local Interval Estimator*, kurz *CLIE* verwendet. Aus diesem Grund wird er hier vorgestellt. Der Algorithmus beruht auf der in 2.2 erwähnten Annahme, dass sich aufeinander folgenden Herzschläge ähneln und schätzt die Herzrate anhand der Selbstähnlichkeit des Signals.

Der Algorithmus iteriert mit einem *Moving window* über das mit einem Bandpass gefilterte Signal. Es werden zwei Schwellwerte für die Intervalllänge T genutzt, T_{min} und T_{max} , basierend auf bekanntem Bereich der Herzrate von 30 bis 200. Die Länge des Analysefensters w_i entspricht $2 * T_{max}$, sodass mindestens zwei vollständige Herzschläge enthalten sind.

$$w_i[v] = x[n_i + v], v \in \{-T_{max} * f_s, \dots, T_{max} * f_s\}$$

In jedem Fenster wird die lokale Intervalllänge T_i geschätzt und anschließend das Zentrum des Fensters n_i weiterbewegt.

$$n_{i+1} = n_i + \Delta t * f_s$$

Die Schätzung der Intervalllänge beruht auf drei Selbstähnlichkeitsmaßen, die wie folgt definiert sind:

$$\begin{aligned} E_{\text{Corr}}[N] &= \frac{1}{N} \sum_{v=0}^N w[v]w[v-N], \\ E_{\text{AMDF}}[N] &= \left(\frac{1}{N} \sum_{v=0}^N |w[v] - w[v-N]| \right)^{-1}, \\ E_{\text{MAP}}[N] &= \max_{v \in \{0, \dots, n\}} (w[v] + w[v-N]). \end{aligned}$$

Dabei berechnet E_{corr} eine modifizierte Autokorrelationsfunktion, mit E_{AMDF} wird die Differenz des Signals zueinander miteinbezogen und mit E_{MAP} werden die maximale Amplituden von beliebigen 2 Samples über das ganze Fenster berechnet. AMDF steht für *modified average magnitude difference function* und MAP für *maximum amplitude of pairs*. Diese Schätzer entsprechen jeweils einer Wahrscheinlichkeitsfunktion, die beschreibt wie wahrscheinlich es ist, dass n dem tatsächlichen Schlag-zu-Schlag-Intervall

entspricht. Durch Skalierung können sie in Wahrscheinlichkeitsdichtefunktionen verwandelt werden. Durch Kombination dieser drei Funktionen wird nun der wahrscheinlichste Wert für N ermittelt:

$$N_i = \underset{N}{\operatorname{argmax}} p(N|E_{\text{Corr}}, E_{\text{AMDF}}, E_{\text{MAP}})$$

Nach dem Satz von Bayes kann die Wahrscheinlichkeit, dass N der tatsächlichen Intervalllänge entspricht auch wie folgt ausgedrückt werden:

$$p(N|E_{\text{Corr}}, E_{\text{AMDF}}, E_{\text{MAP}}) = \frac{p(E_{\text{Corr}}, E_{\text{AMDF}}, E_{\text{MAP}}|N)p(N)}{p(E_{\text{Corr}}, E_{\text{AMDF}}, E_{\text{MAP}})}$$

Durch Anwenden der Theoreme von Bayes unter der Annahme, dass die Ergebnisse der 3 Schätzer nicht voneinander, sondern nur von n abhängen und dass n gleichverteilt ist, erhält man folgende Wahrscheinlichkeitsfunktion, mit der man die wahrscheinlichste Intervalllänge erhalten kann. Auch diese Funktion entspricht mit Skalierung einer Wahrscheinlichkeitsdichtefunktion.

$$\begin{aligned} E_f[n] &= E_{\text{Corr}}[n] \cdot E_{\text{AMDF}}[n] \cdot E_{\text{MAP}}[n], \\ n_{\text{opt}} &= \underset{n}{\operatorname{argmax}} E_f[n] \end{aligned}$$

Wie die 3 einzelnen Schätzer und die fusionierte Funktion aussehen kann, ist in 3.1 gezeigt.

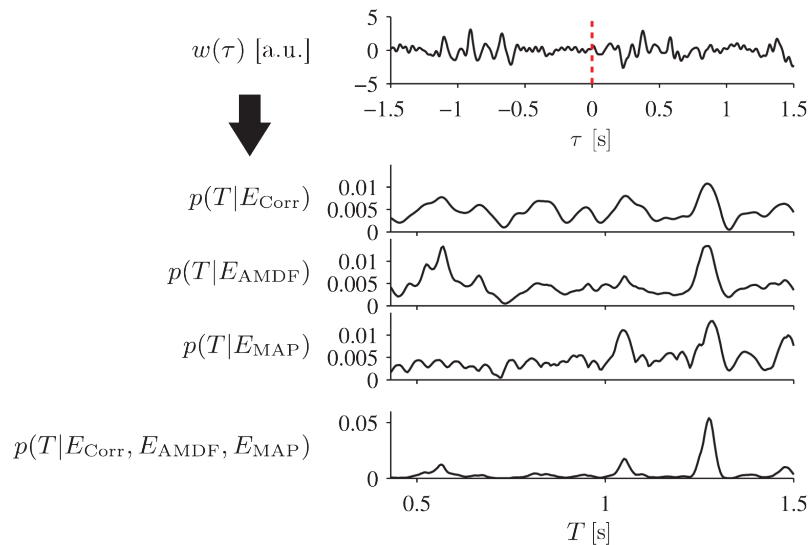


Abbildung 3.1: Die drei Intervallschätzer und ihre Fusionierung

Nun gibt es für jeden Punkt im Signal eine Schätzung der Intervalllänge. Mit Hilfe der Fenstergröße und dieser Länge können die zu einem Herzschlag gehörenden Hochpunkte ermittelt werden, nämlich die, die die größte kombinierten Amplitude mit dem durch die Intervalllänge gegebenen Abstand besitzen. Für jeden dieser Punkte P_k existiert nun eine Menge an Schätzungen T_k , mit der eine robuste Intervallschätzung $\bar{T}_k = \text{median}(T_k)$ ermittelt werden kann.²

3.3 Artefakterkennung

3.3.1 Allgemein bei kardiorespiratorischen Signalen

- Artefakte = irrelevante Signalteile mit variierender Amplitude, Frequenz und Dauer, die physiologisches Signal stören³
- Ziel Beurteilung Signalqualität: nur die Teile des Signals, die Vitalparameter enthalten verarbeiten
- Bewegungsartefakte, Sensorstörungen etc nicht verarbeiten
- Quelle Störung irrelevant, aber medizinische Abnormalitäten dürfen nicht als gestörtes Signal klassifiziert werden
- Signalqualität oft mit so genannten Signal Quality Indices gemessen

²Vgl. zu diesem Kapitel Brüser et al. 2013.

³Nizami et al. 2013.

- je nach SQI und Anwendungsfall verschiedene Aussagen
- Sadek, Biswas, Yongwei et al.: Unterscheidung in Bezug auf Signalqualität zwischen informativ und nicht informativ
- informativ: *noise* und Signal von guter Qualität, Features können ohne weitere Verarbeitung extrahiert werden
- nicht informativ: Informationen mit Artefakten und noise vermischt, weitere Verarbeitung vor Extraktion Vitalparameter nötig oder Extraktion von physiologischen Eigenschaften unmöglich
- im klinischen Kontext genutzte Artefakterkennung oft durch relativ einfaches Pre-processing⁴
- oft bestimmte Informationen direkt oder indirekt *hard coded*, kann zum einen etwas wie Typ oder Frequenz der Daten sein, aber auch demographische Informationen über die Patient*innen wie Alter, Gewicht oder medizinischer Zustand⁵
- Beispiel für Beurteilung der Signalqualität von anderen kardiorespiratorischen Signalen, in diesem Fall EKG und PPG in Abbildung, betrachtet werden 10-Sekunden-Fenster

⁴Vgl. Nizami et al. 2013.

⁵Vgl. Nizami et al. 2013.

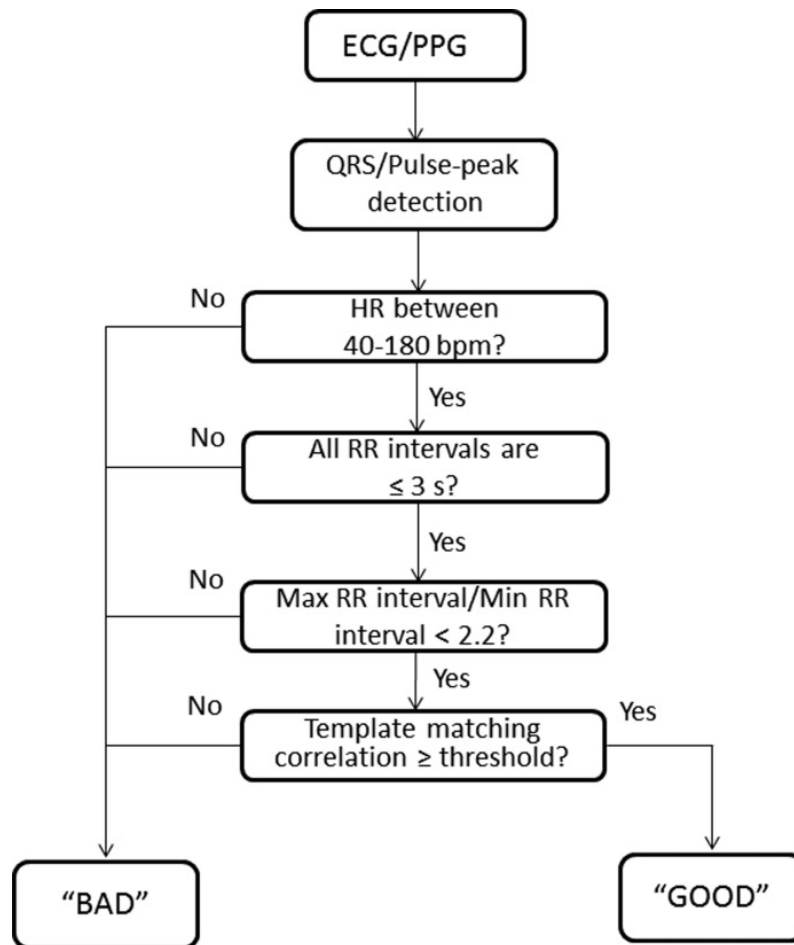


Abbildung 3.2: Flussdiagramm eines Algorithmus zur Beurteilung der Signalqualität⁶

- erst Segmentierung der Herzschläge und dann 4 Kriterien überprüft -> reichen jeweils um Signal als schlecht zu klassifizieren
- Herzrate zwischen 40 und 180 Schlägen pro Minute? Abstand der Hochpunkte aufeinander folgender Herzschläge unter 3 Sekunden (kein Schlag fehlt?), Verhältnis maximales und minimales Schlag-zu-Schlag-Intervall kleiner als 2.2 (Begrenzung Veränderung Herzrate in untersuchtem 10-Sekunden-Fenster), Korrelation zu erstelltem Template höher als Schwellwert?

3.3.2 Bei ballistokardiographischen Signalen

- bei BKG auch Artefakterkennung schwieriger als bei anderen kardiorespiratorischen Signalen: zusätzlich zur Variabilität bei Atmung Veränderungen bei Positionsänderungen -> vorher erstellte Templates werden obsolet
- Beispiele für Artefakte, einmal mit hoher, einmal mit niedriger Energie

- hohe Energie größere Bewegungen, leicht zu erkennen

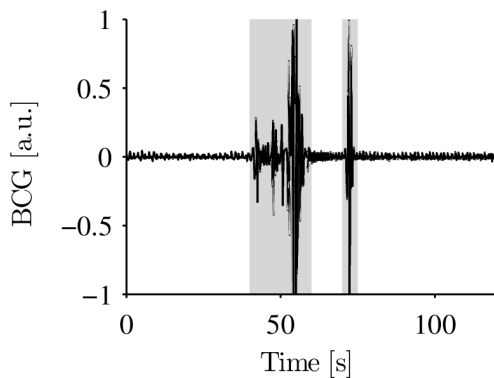


Abbildung 3.3: Artefakte mit hoher Energie

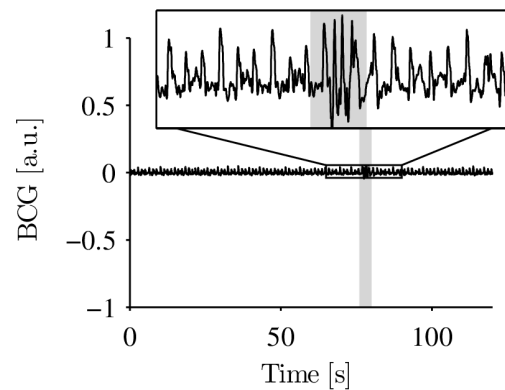


Abbildung 3.4: Artefakte mit niedriger Energie

- oft genutzt, dass Bewegungen stärkere Krafteinwirkungen verursachen, als Atmung und Herzschlag, das sind dann Artefakte mit hoher Energie
- bei Betrachtung verschiedener Ansätze zur Signalverarbeitung: Proband*innen oft angewiesen, sich möglichst wenig zu bewegen -> nicht realistisch für BKG-Aufnahmen in Betten
- ⁷ festgestellt, dass bei Messsystemen in Betten besonders die Messung während des Tages große Signaleile von schlechter Qualität aufweisen, mehr als nachts, sichtbar ist z.B. auch Zeit des Mittagessens
- teils EKG-Referenz zur Artefakterkennung genutzt -> im *unobtrusive* Kontext nicht zielführend
- im Folgenden ausgewählte Ansätze vorgestellt

3.3.3 Schwellwertbasierte Artefakterkennung

In „Noninvasive ambulatory measurement system of cardiac activity“ präsentieren Pino et al. einen Ansatz für die Erkennung von Körperbewegungen für ein in einen Stuhl eingebettetes BKG-Messsystem. Dafür werden über ein *moving window* Maximum, Minimum, Standardabweichung und Mittelwert ermittelt und daraus 2 Schwellwerte berechnet:

⁷Hoog Antink et al. 2020.

$$T_1 = \frac{\max + \min}{2},$$
$$T_2 = \text{mean} + 1,1 * \text{std.}$$

Die Länge des *moving window* ist mit 200 Samples bei einer Abtastrate von 200 Hz benannt. Untersucht wurden sowohl Freiwillige im Labor, als auch im Krankenhauswartzimmer für eine sehr kurze Messdauer von ein bis zwei Minuten. Mit diesem Ansatz wurde bei mehr als 50 % der Laborgruppe eine *Coverage* zwischen 87 % und 95 % erreicht. Die *Coverage* der im Krankenhaus aufgenommenen Gruppe war bedeutend niedriger; hier lagen 50 % der Messungen zwischen 48 % und 95 % *Coverage* erreicht. Zu der Genauigkeit der Herzschlagdetektion auf den akzeptierten Signalteilen wird keine Aussage in Zahlen getroffen sondern nur der folgende Bland-Altman Graph gezeigt.

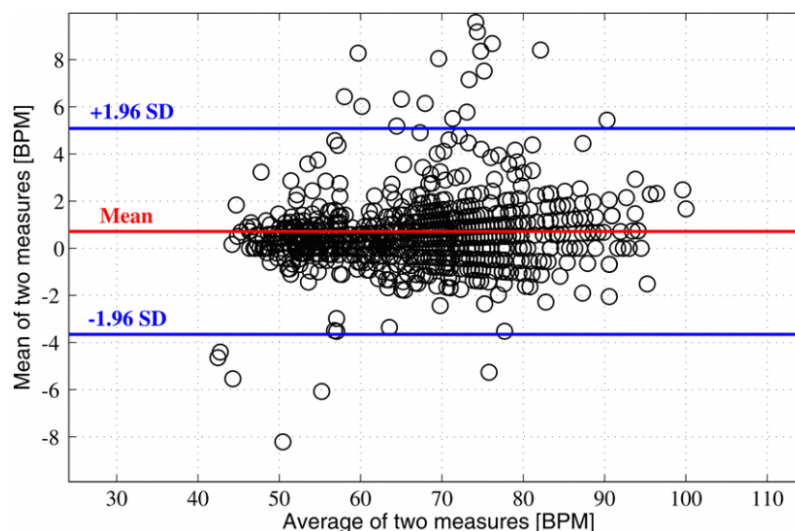


Abbildung 3.5: Bland-Altman Graph zwischen von EKG und BKG berechneter HR

Hier zeigt sich, dass beim Großteil des hier betrachteten Signals die HR größtenteils mit einer Genauigkeit von \pm Schläge pro Minute bestimmt werden konnte. Allerdings handelt es sich hier um im Sitzen aufgenommenes Signal, bei dem die Variabilität geringer ist als bei in Betten aufgenommenem BKG.

3.3.4 Maschinelles Lernen mit statistischen Merkmalen

Ein Algorithmus zur Beurteilung der Signalqualität mittels maschinellen Lernens wird von Sadek, Biswas, Yongwei et al. im Paper „Sensor data quality processing for vital

signs with opportunistic ambient sensing“ beschrieben. Betrachtet werden BKG-Signale, die in einem Massagesessel aufgenommen werden, also ebenfalls im Sitzen aufgenommenes Signal, bei dem eine geringere Variabilität als in unserem Anwendungsfall erwartet wird.

Die vorliegenden Daten wurden manuell von Expert*innen als informativ oder nicht informativ klassifiziert und in 10-Sekunden-Segmente, die sich nicht überlappen, aufgeteilt, die sich nicht überlappen. Insgesamt waren 58 % der Daten als informativ und 42 % als nicht-informativ gelabelt. Von diesen Segmenten wurden nach einer Bandpass-Filterung auf 1 bis 12 Hz 13 statistische Merkmale berechnet:

- Minimum
- Maximum
- Mittelwert
- Standardabweichung
- Schiefe
- Kurtosis
- Spannweite
- Interquartilspannweite
- mittlere absolute Abweichung
- Anzahl der Nulldurchgänge
- Varianz der lokalen Minima
- Varianz der lokalen Maxima
- Mittelwerte der Signalhüllkurve⁸

Für fünf verschiedene Modelle des maschinellen Lernens wurden jeweils die besten Hyperparameter über Kreuzvalidierung auf den Trainingsdaten ermittelt und die Modelle anschließend mit diesen Hyperparametern trainiert. Anschließend wurden die Modelle auf unbekannten Daten getestet. Das Training und Testen wurde mit getauschten Gruppen wiederholt. Das beste Ergebnis wurde mit einem Random Forest erreicht: Die durchschnittliche Genauigkeit der Kreuzvalidierung betrug 98,13 % bzw. 100 % bei getauschten Gruppen. Auf dem Testset wurde eine Genauigkeit von 92,3 % bzw. 97,99 % erreicht. Weitere Evaluationsmetriken außer eine *Confusion Matrix* für den besten Klassifikator sind nicht gegeben.

⁸Die Signalhüllkurve ist eine glatte Kurve, die die Extrema des Signals umreißt.

Diese Ergebnisse sind sehr gut, allerdings muss bei der Einordnung beachtet werden, dass bei für die Kreuzvalidierung die Segmente zufällig verteilt wurden und nicht beachtet wurde, dass der Algorithmus für aussagekräftige Validierung einzelne Personen nicht darf. Dadurch ist die Performance auf gänzlich unbekannten Daten weiterhin nicht bekannt und vermutlich schlechter, als die Zahlen es hier vermuten lassen.

3.3.5 Ähnlichkeit der Intervallschätzer des CLIE-Algorithmus

Ein weiteres Maß für die Signalqualität basiert auf dem in 3.2 vorgestellten Algorithmus zur Intervallschätzung. Dieser Signal Quality Index misst, wie enig sich die drei Intervallschätzer sich. Wenn diese sich uneinig sind, ist der Signal Quality Index (SQI) bei 0, je ähnlicher sich die Schätzungen sind, desto höher ist er. Für jedes Fenster i wird er wie folgt berechnet:

$$q = \frac{E_f[n_{opt}, i]}{\sum E_f[n, i]}$$

3.4 Messdaten

3.4.1 Erfassung

- aufgenommen in der Gefäßstation des Universitätskrankenhauses in Tampere in Finnland
- 14 Patient*innen wurden bis zu 24 h überwacht
- 2 weiblich, 12 männlich
- Durchschnittsalter: 69,57 Jahre
- nach verschiedenen gefäßchirurgischen Eingriffen
- durchschnittliche Messdauer: 17.7 h, range 4,46 bis 22,96 h
- EMFit QS Bettsensor, zwischen Matratze des Krankenhausbettes und Bettgestehl positioniert
- Samplingrate des EMFit QS Systems: 100 Hz, Bandpass-limitiert auf 1 bis 5 Hz
- Referenz EKG: Faros 360 5 lead Holter monitor, 1 kHz Abtastrate
- variabler Drift zwischen beiden Signalen

3.4.2 Vorliegende Form

- unbearbeitetes BKG-Signal, abgesehen von der Bandpasslimitierung auf 1 bis 5 Hz
- unbearbeitetes 3-Kanal EKG Signal
- mit CLIE-Algorithmus detektierte Herzschläge, schon nach Qualität gefiltert mit samt Brüser SQL, Länge und Länge des Herzschlages der EKG Referenz
- Vektoren, die den Drift der beiden Signale beschreiben, Form Sekunde BKG-Signal und entsprechende Sekunde in EKG-Referenz

3.4.3 Verarbeitung und Datenstruktur

- Datensatz von einem Patient besteht aus BKG Signal und EKG Referenz
- beides wird eingelesen, geprüft ob schon Detektion von Herzschlägen (Erkennen von R-Peaks bzw. CLIE Algorithmus schon durchgeführt wurde und als csv-Datei existiert

3.4.4 Annotation der Daten

Die vorliegenden Daten sind nicht annotiert. Es ist im Rahmen dieser Arbeit nicht möglich, die Annotation durch Expert*innen durchführen zu lassen, weshalb auf das parallel aufgenommene EKG zurückgegriffen wird.

- aufgrund des nicht-linearen Drifts der Daten herzs Schlaggenaue Synchronisierung schwierig
- Entscheidung Annotation von Bereichen von mehreren Sekunden möglich zu machen
- Ablauf: existiert EKG Signal zu diesem Zeitpunkt, bei dem eine Herzfrequenz ermittelt werden konnte
- Berechnung dieser -> Anzahl der R-Peaks in diesem Bereich -1 geteilt durch den Abstand des letzten und des ersten Peaks
- Berechnung BKG-Herzfrequenz: Durchschnitt der geschätzten Längen der erkannten Peaks im Bereich
- Annotation anhand von relativer oder absoluter Abweichung der beiden

4 Analyse

4.1 Anwendung existierender Verfahren

- gibt einige Verfahren
- Vermutung, dass nicht hinreichend für unsere Daten aus Gründen

4.1.1 Ähnlichkeit der Intervallschätzer nach Brüser

4.1.2 Schwellwerte für Standardabweichung, Minimum und Maximum

4.1.3 Maschinelles Lernen mittels statistischer Merkmale

4.2 Datenanalyse und Merkmalskonstruktion

5 Synthese

5.1 Eingabeform der Daten

5.2 Evaluationsmetriken

5.3 Modelle

6 Evaluierung der Ergebnisse

7 Zusammenfassung und Ausblick

7.1 Zusammenfassung

7.2 Ausblick

Literatur

- Albukhari, Almothana, Frederico Lima und Ulrich Mescheder (2019). „Bed-embedded heart and respiration rates detection by longitudinal ballistocardiography and pattern recognition“. In: *Sensors (Switzerland)* 19.6.
- Brüser, C., S. Winter und S. Leonhardt (2013). „Robust inter-beat interval estimation in cardiac vibration signals“. In: *Physiological Measurement* 34.2, S. 123–138.
- Brüser, Christoph et al. (2011). „Adaptive beat-to-beat heart rate estimation in ballistocardiograms“. In: *IEEE Transactions on Information Technology in Biomedicine* 15.5, S. 778–786.
- de Lalla, V., M. A. Epstein und H. R. Brown (1950). „Analysis of H wave of ballistocardiogram.“ In: *Circulation* 2.5, S. 765–769.
- Gordon, J W (1877). „Certain Molar Movements of the Human Body produced by the Circulation of the Blood.“ In: *Journal of anatomy and physiology* 11.Pt 3, S. 533–6.
- Hoog Antink, Christoph et al. (Aug. 2020). „Ballistocardiography can estimate beat-to-beat heart rate accurately at night in patients after vascular intervention“. In: *IEEE Journal of Biomedical and Health Informatics* 24.8, S. 2230–2237. URL: <https://ieeexplore.ieee.org/document/8974196/>.
- Inan, Omer T. et al. (2015). „Ballistocardiography and Seismocardiography: A Review of Recent Advances“. In: *IEEE Journal of Biomedical and Health Informatics* 19.4, S. 1414–1427.
- Nizami, Shermeen, James R. Green und Carolyn McGregor (2013). *Implementation of artifact detection in critical care: A methodological review*.
- Orphanidou, Christina et al. (Mai 2015). „Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring“. In: *IEEE Journal of Biomedical and Health Informatics* 19.3, S. 832–838.
- Paalasmaa, Joonas, Hannu Toivonen und Markku Partinen (2015). „Adaptive heartbeat modeling for beat-to-beat heart rate measurement in ballistocardiograms“. In: *IEEE Journal of Biomedical and Health Informatics* 19.6, S. 1945–1952.

- Pinheiro, Eduardo, Octavian Postolache und Pedro Girão (2010). „Theory and Developments in an Unobtrusive Cardiovascular System Representation: Ballistocardiography“. In: *The Open Biomedical Engineering Journal* 4.1, S. 201–216.
- Pino, Esteban J., Javier A.P. Chavez und Pablo Aqueveque (2015). „Noninvasive ambulatory measurement system of cardiac activity“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. Bd. 2015-November. Institute of Electrical und Electronics Engineers Inc., S. 7622–7625.
- Rosales, Licet et al. (2012). „Heartbeat detection from a hydraulic bed sensor using a clustering approach“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, S. 2383–2387.
- Sadek, Ibrahim, Jit Biswas und Bessam Abdulrazak (2019). „Ballistocardiogram signal processing: a review“. In: *Health Information Science and Systems* 7.1. eprint: 1807.00951.
- Sadek, Ibrahim, Jit Biswas, Zhu Yongwei et al. (2016). „Sensor data quality processing for vital signs with opportunistic ambient sensing“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. Bd. 2016-October. Institute of Electrical und Electronics Engineers Inc., S. 2484–2487.
- Starr, Isaac et al. (1939). „Studies on the Estimation of Cardiac Output in Man, and of Abnormalities in Cardiac Function, From the Heart's Recoil and the Blood's Impacts; the Ballistocardiogram“. In: *American Journal of Physiology-Legacy Content* 127.1, S. 1–28.
- Zink, Matthias Daniel et al. (2017). „Unobtrusive Nocturnal Heartbeat Monitoring by a Ballistocardiographic Sensor in Patients with Sleep Disordered Breathing“. In: *Scientific Reports* 7.1.