Machine Learning
Data Analytics

Friedrich-Alexander-Universität
Technische Fakultät

Improving the robustness of heart rate estimation from continuous-wave radar data using a wavelet-based approach

Master's Thesis in Computer Science

submitted
by

Simon Meske

born 27.10.1994 in Fürth

Written at

Machine Learning and Data Analytics Lab
Department Artificial Intelligence in Biomedical Engineering
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

in Cooperation with

Institute of High-Frequency Technology, Hamburg University of Technology (TUHH)

Advisors: Robert Richer, M. Sc., Thomas Altstidl, M. Sc., Luca Abel, M. Sc., Daniel Krauß, M. Sc., Prof. Dr. Bjoern Eskofier (FAU), Nils Albrecht M. Sc., Kristina Heß, M. Sc., (TUHH)

Start: 15.02.2024

Ende: 15.08.2024

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.
Die Richtlinien des Lehrstuhls für Bachelor- und Masterarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Erlangen, den 15. August 2024

Übersicht

Herz-Kreislauf-Erkrankungen (CVD) sind weltweit die häufigste Todesursache. Das unterstreicht die Notwendigkeit einer effektiven Erhebung herzbezogener Vitalparameter wie Herzschlag und Herzfrequenz zur Früherkennung. Die derzeitige Goldstandard-Methode, das Elektrokardiogramm (EKG), erfordert einen direkten Kontakt mit Patienten, was in Fällen wie beispielsweise infektiösen Patienten oder Brandopfern nicht praktikabel ist. Daher arbeitet die Wissenschaft an Alternativen, die ohne direkten Kontakt zum Körper auskommen. Ein vielversprechender Ansatz stellt hierbei die Verwendung von RADAR-Sensoren zur kontaktlosen Erkennung von Herzschlägen dar. Der RADAR-Sensor misst die Veränderungen im empfangenen RADAR-Signal, diese werden unter anderem durch die Bewegungen des Brustkorbs aufgrund des Herzschlags verursacht. Um die einzelnen Herzschläge aus dem RADAR-Signal zu extrahieren, können verschiedene Ansätze angewandt werden, darunter Methoden der Zeit-Frequenz-Analyse wie Kurzzeit Fourier Transformation und in jüngerer Zeit auch Modelle des maschinellen Lernens. Diese Methoden verwenden jedoch häufig Datensätze mit geringer Bewegung der Probanden. Das erleichtert zwar die Erkennung von Herzschlägen, macht die Ergebnisse durch Minimierung von Artefakten allerdings weniger vergleichbar mit reelen Szenarien in denen Patienten Körperbewegungen aufweisen. Um eine brauchbare Alternative zu Methoden wie dem EKG zu sein, müssen weitere Forschungen durchgeführt werden, sodass bewertet werden kann, wie RADAR-Daten verarbeitet und die Modelle trainiert werden müssen, damit Herzschläge auch bei Vorhandensein von größeren Körperbewegungen erkannt werden. In der hier vorliegenden Arbeit werden zwei Modellarchitekturen zur Verbesserung der Nutzung von RADAR-Daten zur Erkennung von Herzschlägen verglichen. Die erste Architektur verwendet long short-term memory cells (LSTM), die die gefilterten inphase (I) und quadrature (Q)-Komponente sowie den Angle, die Power und das Envelope des RADAR-Signals verarbeiten. Beim zweiten Ansatz wird das RADAR-Signal zunächst mithilfe von kontinuierlicher Wavelet Transformation (CWT) verarbeitet, um ein Scalogramm zu erstellen, das direkt an das UNet-Modell übergeben werden kann. Beide Modelle wurden mit jeweils drei Datensätzen trainiert und getestet. Der erste Datensatz mit Aufnahmen von 22 Personen im Liegen zeigte kaum Artefakte durch große Körperbewegungen. Der zweite Datensatz mit 110 Probanden enthielt erhebliche Artefakte aufgrund von Bewegungen des Körpers. Studien bestätigten, dass die Kombination eines Datensatzes mit einer gemischten Menge an Artefakten die Leistung des Modells verbessern kann. Aus diesem Grund wurden beide Datensätze für den dritten Datensatz kombiniert, um die Auswirkungen gemischter Ausprägungen von Artefakten zu bewerten. Die Ergebnisse zeigen, dass die rechenintensiven continuous wavelet transform (CWT)-verarbeiteten RADAR-Daten mit der UNet-Architektur die LSTM-Modelle sowohl bei Daten mit geringer als

auch mit hoher Menge an Artefakten übertreffen. Der kombinierte Datensatz hat die Leistung des UNet-Modells nicht verbessert, und das LSTM-Modell konnte Herzschläge in Daten mit Körperbewegungen kaum erkennen. Insgesamt zeigt diese Arbeit das Potenzial der Kombination der CWT mit einer UNet-Architektur zur Verbesserung der Herzschlagerkennung aus RADAR-Daten. Trotz der hier vorgestellten Ergebnisse sind weitere Forschungen unerlässlich, um die Toleranz für größere Körperbewegungen oder Artefakte zu verbessern um diese kontaktlose Methode zu einer Alternative für EKG-Messungen zu machen.

Abstract

Cardiovascular diseases (CVD) are the leading cause of death worldwide, highlighting the need for effective tracking of heart-related vital signs, such as heart beats (HBs) and instantaneous heart rate (IHR), for early detection. The current gold standard method, electrocardiogram (ECG), requires direct contact with patients, which is sometimes impractical, like in cases of infectious patients or burn victims. As a result, researchers are exploring non-contact alternatives that do not require direct attachment of sensors to the body. One promising approach researchers have focused on is using radio detection and ranging (RADAR) sensors to detect HBs contactless. The RADAR sensor measures changes in received RADAR signals caused by chest movements due to the HB. For extracting the individual HBs from that RADAR signal, various approaches have been employed, including time-frequency analysis such as short-time fourier transform (STFT) and, more recently, machine learning (ML) models. However, these methods often use datasets with restricted subject movements, making it easier to detect HBs by minimizing artifacts caused by random large body movements (RLBM). In order to become a viable alternative to methods such as ECG, further research has to be conducted in order to evaluate how RADAR data needs to be processed and how models need to be trained to detect HBs even in the presence of RLBM. This thesis compares two model architectures to enhance the use of RADAR data to detect HBs. The first architecture utilizes LSTM cells, processing the filtered I, and Q component as well as the Angle, Power, and Envelope of the RADAR signal. The second approach first transforms the RADAR signal using CWT to create a scalogram that can be used as input for the second part of the approach - a UNet based model. Both models were trained and tested on three datasets. The first dataset, with recordings from 22 subjects lying down, had low noise levels. The second dataset with 110 subjects included a higher amount noise from RLBM. Studies confirmed that combining a dataset with less noisy and more noisy data can improve the model's performance. For this reason, both datasets were combined for the third dataset to evaluate the impact of mixed noise levels. The results indicate that the computationally intensive CWT-processed RADAR data with the UNet architecture outperforms the LSTM models in both low and high noise data. The combined dataset did not enhance the UNet model's performance, and the LSTM model failed to detect HBs in noisy data in most of the cases. Overall, this thesis shows the potential of combining CWT with a UNet architecture for improving HB detection from RADAR data. Despite the results presented here, further research is essential to enhance the tolerance for RLBM to make this contactless method a viable alternative to ECG measurements.

Contents

Chapter 1

Introduction

According to the World Heart Federation CVD are currently the leading cause of death globally, with a death rate of 51% in 2019 [Gai23]. This emphasizes the urgent need for innovative solutions to detect CVD at an early stage [Fed24; Gai23]. Vital signs such as HBs and blood pressure play an important role in diagnosing heart-related health issues [Pon19; Sch03; Yug23; Ele96]. While various methods exist to measure HBs, the current gold standard is the ECG [Sam16; Str22; De 12]. ECGs are a contact-based HB measurement that detects the heart's electrical activity produced during each heart cycle using multiple electrodes placed at the chest. Apart from the fact that constantly wearing the electrodes during the measurement can be perceived as uncomfortable and infeasible, the requirement for direct contact with the patient makes the ECG not an appropriate method for the use outside of a medical setting [Ele96].

RADAR-based methods offer an alternative to ECG for remote HB monitoring in various environments, such as at home, in vehicles, or clinical settings by eliminating the need for direct contact and sensor application [Mal20; Ric20; El 23; Yen22; Mur22]. RADAR-based technologies measure the HB by sending electromagnetic waves that hit the body and are reflected back to the RADAR device. When the heart beats, the chest moves slightly, causing small changes in the reflected signals which are measured by the RADAR and analyzed to detect the HB [Li13].

However, movements of the body, such as breathing or other activities, can lead to interference or artifacts in the measured signals, making it difficult to accurately record the HB [Che21]. For this reason, existing literature mainly focused on data where participants are in a stationary position, on one hand, decreasing the noise, but on the other hand providing limited exposure to real-world scenarios [Mog17; Shi22; Li17]. To become a viable alternative to ECG, RADAR-based methods must reliably detect the HB even in the presence of noise [Iwa21].

For this reason, the thesis presented here not only compares different ML approaches to examine the most suitable one for stationary subject data, but also includes one dataset containing noise, in order to evaluate the possibility of noise reduction and HB detection.
In detail, two approaches for detecting HBs from RADAR data in the two datasets and a combined version of both datasets were compared. Depending on the approach used, different preprocessing pipelines were utilized.
The first approach involves a LSTM artificial neural network (ANN) architecture for the used RADAR data. The second approach employs a time-frequency-based method, where the RADAR signal is additionally transformed using CWT and processed by a UNet-based architecture [Ron15] to detect HBs. Both architectures were used for both datasets separately, and also on a combined version of the datasets. By comparing the performance of these approaches under various conditions, this research aims to identify the best-performing method for HB detection using RADAR data. Additionally, this thesis investigates the generalizability of models trained on less noisy data and their ability to predict HBs from datasets with higher levels of RLBM. Therefore, this thesis seeks to enhance the understanding of RADAR-based HB detection methods and their potential applications in real-world settings. By addressing the challenges posed by noise and movement in RADAR data, the findings from this thesis may contribute to improve the robustness and accuracy of systems for non-contact HB monitoring, with implications for early detection and prevention of cardiovascular diseases.

In order to guide through the details of this thesis, it is structured as follows: Chapter 2 explains the cardiac physiology, the background of the used processing steps for the two different approaches, and the fundamentals of the used architectures of the ML models. Chapter 3 presents existing research relevant to this thesis in order to provide an overview of current advancements in RADAR based HB detection. Chapter 4 describes the used datasets, the data processing, and the evaluation methods for the used architectures as well as the computational settings in detail. The results and individual evaluations, as well as discussions of all used approaches, are presented in Chapter 5. Finally, Chapter 6 summarizes the findings and gives an outlook for future improvements.

Chapter 2

Background

2.1 Physiology of the cardiovascular system

The human cardiovascular system, consisting of the heart and blood vessels, is responsible for circulating blood throughout the body. It includes the heart, arteries, veins, and capillaries [Opi04]. The heart, located slightly to the left behind the sternum, serves as the system's central component. It is defined by two main loops managed by specialized cardiac muscle cells: (1) systemic circulation, which delivers oxygenated blood to all organs, and (2) pulmonary circulation, which oxygenates deoxygenated blood [Opi04; Kat10]. As depicted in Figure 2.1, the heart is divided in four separate chambers with four valves that ensure unidirectional blood flow. The right atrium receives deoxygenated blood from the body and passes it through the tricuspid valve to the right ventricle, which pumps it through the pulmonary valve and pulmonary arteries to the lungs for oxygenation. The left atrium then receives oxygenated blood from the lungs via the pulmonary veins and pumps it through the mitral valve to the left ventricle, which distributes it through the aortic valve into the aorta and throughout the body [Opi04; Kat10].

2.1.1 Heart beat and the cardiac electrical conduction system

The HB, is produced by the rhythmic contraction and relaxation of the heart muscle, which is essential for pumping blood and is regulated by the heart's electrical conduction system to ensure coordinated beating [Ele96]. The sinoatrial node, located at the top of the right atrium within the cardiac septum, acts as the heart's natural pacemaker, generating electrical impulses that initiate each heartbeat. The atrioventricular node, lying at the junction between the atria and ventricles in the cardiac septum, receives these impulses from the sinoatrial node and delays them slightly

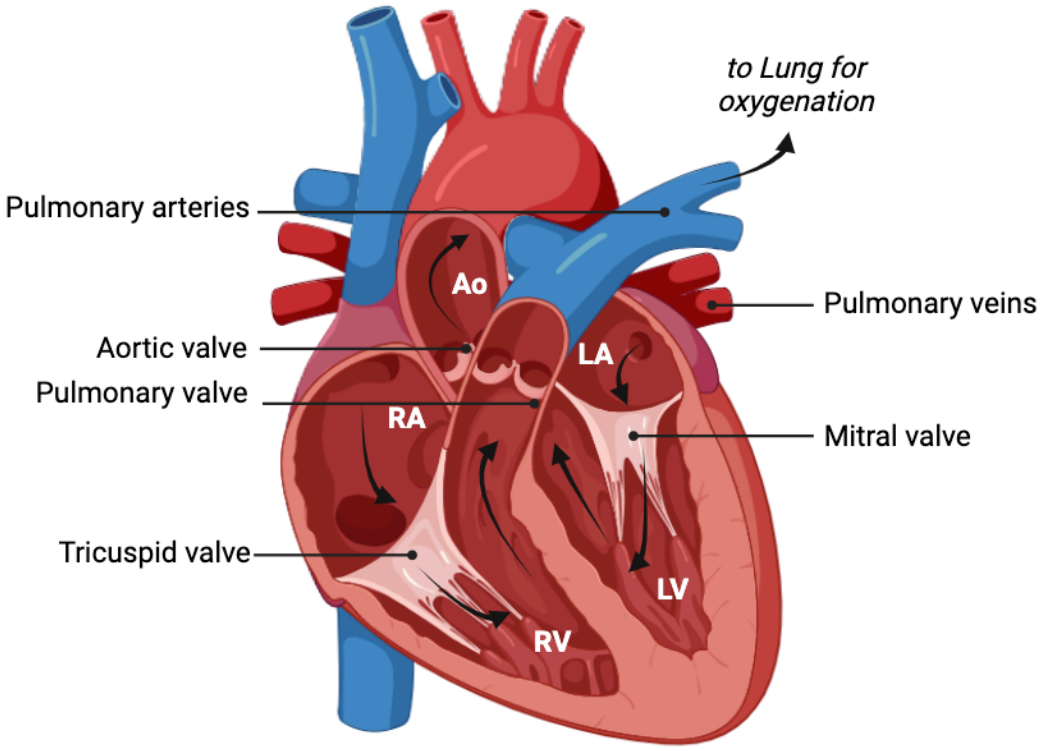

Figure 2.1: **Heart anatomy and physiology.** Representation of heart-related anatomical components for the enrichment of the blood with oxygen, the production of healthy heartbeat and the forwarding of the oxygen-rich blood for distribution throughout the body. Arrows represent the healthy blood flow through the heart. RA = Right Atrium, RV = Right Ventricle, LA = Left Atrium, LV = Left ventricle, Ao = Aorta. Graphic created in BioRender.com.

to allow the atria to contract and empty blood into the ventricles. The Bundle of His transmits the impulse from the atrioventricular node along the cardiac septum to the ventricles, and the Purkinje fibers, a network of fibers spread throughout the ventricles, ensure rapid and coordinated ventricular contraction [Ele96; Unu14; Maa23; Nik12].

2.1.2 Heart sounds

The HB is typically heard as two heart sounds (HSs), which are generated by the mechanical activities of the heart, specifically the opening and closing of the heart valves and the resultant blood flow [Ski22; Fel90]. The primary HS is typically differentiated into two phases, the first HS (S1), is often described as “lub” and is produced by the closure of the atrioventricular valves, which include the mitral and tricuspid valves (Figure 2.1). When the ventricles contract (begin-

ning of systole), the pressure rises, causing the mitral and tricuspid valves to close to prevent backflow into the atria. This sudden closure creates vibrations that produce the S1 sound. S1 is a low-pitched, slightly prolonged sound. The second HS (S2), often described as “dub”, is produced by the closure of the semilunar valves, which include the aortic and pulmonary valves (Figure 2.1). After the ventricles have ejected blood into the aorta and pulmonary artery, they begin to relax (beginning of diastole), the decrease in pressure causes the aortic and pulmonary valves to close, creating the S2 sound, which is a higher-pitched, shorter sound when compared to the S1 HS. Next to these two phases, there can also be a third and a fourth HS (S3 and S4), which are often pathologic. The third HS (S3) is a low-pitched sound that occurs just after S2 during the rapid filling phase of the ventricles in early diastole. S3 is produced by the rapid flow of blood from the atria into the ventricles, causing vibrations of the ventricular walls. S3 can be normal in children and young adults but may indicate heart failure or volume overload in older adults. The fourth HS (S4) is a low-pitched sound that occurs just before S1 during the atrial contraction phase of late diastole. S4 is produced by the atria contracting forcefully to push blood into a stiff or hypertrophic ventricle, causing vibrations. S4 is often associated with conditions that increase ventricular stiffness, such as hypertension, aortic stenosis, or ischemic heart disease [Ski22; Fel90].

2.2 Contact-based applications for HS and HB measurement

2.2.1 Electrocardiogram

In most clinical settings, physicians use the non-invasive ECG as a graphical representation of the electrical activity of the heart over a period of time [Ele96]. It is initiated by the sinoatrial node and leads to the contraction of the heart’s atrias and ventricles via a series of successive stimulus transmissions along the cardiac septum (see Section 2.1.1) [Unu14; Maa23; Nik12]. In order to measure these impulses, electrodes are placed at the chest and the limbs. While two electrodes can determine the heart’s rhythm which can then be used to derive the HBs and heart rate (HR), twelve leads are typically used since they provide more comprehensive information ensuring for example diagnosis of cardiac arrhythmias and heart failure [Khu14]. These leads include three bipolar limb leads (I, II, III) introduced by Einthoven, three augmented unipolar limb leads (aVR, aVL, and aVF) developed by Goldberger, and six unipolar chest leads (V1-V6) established by Wilson [Lut19; Hou19; Ger08]. Proper electrode placement is crucial for accurate and detailed data collection and is done by trained medical personnel [Ele96]. For less detailed applications, such as measuring the HBs the accuracy of the positioning is less relevant. Once

in place, the electrodes record the heart's electrical signals leading to an ECG signal (Figure 2.2). The interpretation of the ECG is separated into waves, intervals, and complexes. The P-

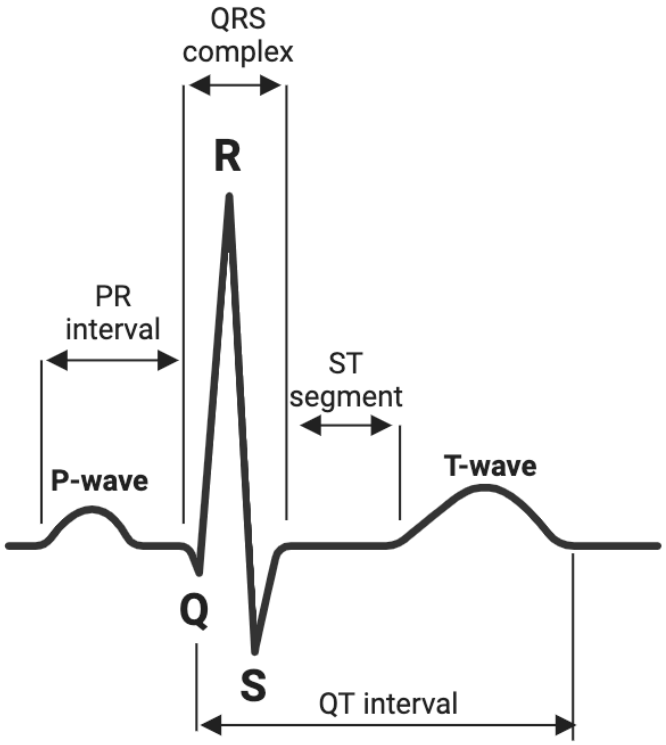

Figure 2.2: **ECG Signal illustration.** Depiction of the main components of an ECG including P-wave (atrial depolarization), QRS complex (Q-wave, R-wave and S-wave - ventricular depolarization) and T-wave (ventricular repolarization)[El 16].

wave represents the atrial depolarization (Figure 2.2), whereas the QRS complex illustrates the ventricular depolarization (consisting of the Q-wave, R-wave, and S-wave). The HR is calculated by measuring the intervals between R-waves (R-R interval). A normal resting HR lies between 60-80 beats per minute (BPM). The T-wave represents ventricular repolarization. The intervals refer to the time past from one wave or complex to another. The PR interval is the time between the start of atrial depolarization (P-wave) and the start of ventricular depolarization (QRS complex). The QT interval represents the total time for ventricular depolarization and repolarization and the ST segment is the period between ventricular depolarization and repolarization [Kun00; Abd21; Gar14] (Figure 2.2). The analysis of differing times or shapes of waves and intervals provides important information for diagnosing cardiovascular diseases or abnormalities in the conduction system [Ele96; Lut19; Hou19; Ger08].

2.2.2 Phonocardiography

Phonocardiography offers an alternative diagnostic tool to the ECG by focusing on the auditory aspects of cardiac activity. In contrast to ECG, which measures the cardiac electrical conduction, phonocardiography records the sounds and murmurs produced by the heart and transforms them into a biomedical signal known as phonocardiogram (PCG) [Gio19]. PCG captures HSs using a microphone attached to the precordium (the area of the chest above the heart) or a miniature sensor inserted into the heart chambers via blood vessels. It records the sounds that the heart makes during a cardiac cycle, including the closing of the heart valves and the associated sounds (see Section 2.1.2) [Ism18]. However, the PCG signal is affected by a variety of artifacts, making its interpretation a difficult task [Gio19]. The most challenging step is the segmentation of the HSs, because of that in most cases require the simultaneous recording of a ECG as a reference (Figure 2.3) [Mil22; Hui97].

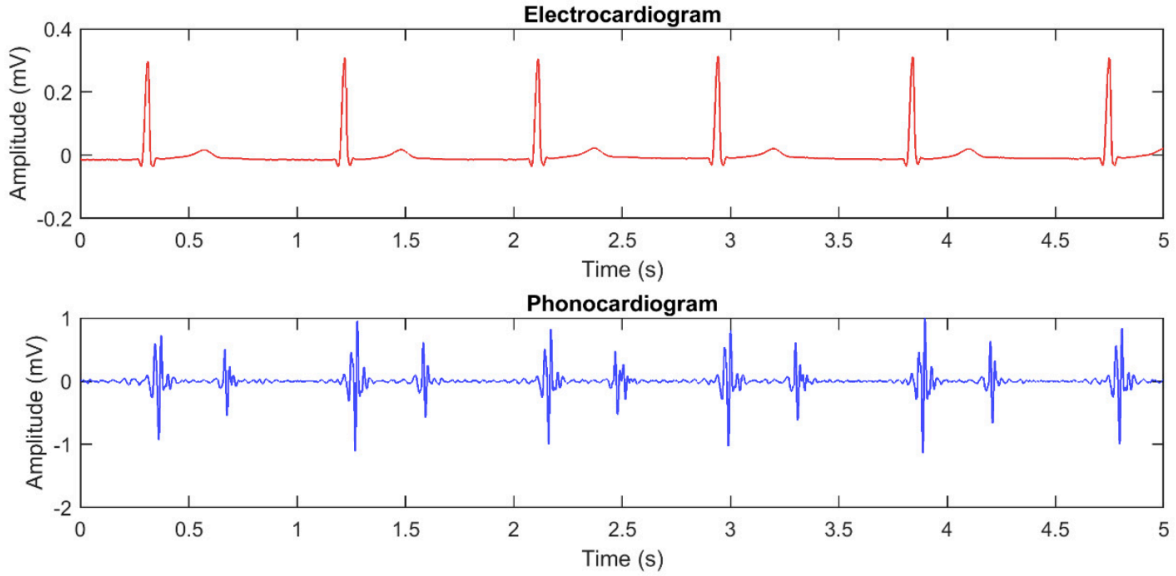

Figure 2.3: **Alignment of ECG and PCG signal.** Simultaneous recording of ECG (upper panel) and PCG (lower panel) aligned in time-domain, this alignment allows ECG to serve as a reference for detecting HSs [Gio19].

2.3 RADAR-based applications for contactless HS and HB detection

RADAR technology provides an alternative approach for monitoring and analyzing cardiac activity. It allows for contactless measurements of HSs, HBs, and respiration, being specifically advantageous in situations where contact-based methods are impractical or uncomfortable. RADAR systems typically use either frequency modulated continuous wave (FMCW) RADAR or continuous wave (CW) RADAR technologies [Kra24].

FMCW RADAR continuously transmits a signal that sweeps through a range of frequencies over time [Bjo98]. The instantaneous frequency $f(t)$ of the transmitted signal can be described as:

$$f(t) = f_0 + \Delta f \cdot \sin(2\pi f_m t) \tag{2.1}$$

where f_0 is the carrier frequency, Δf is the frequency deviation, and f_m is the modulation frequency [Wan15]. The transmitted signal is reflected by the subject's chest and is received by the RADAR system and contains information about the distance and motion of the chest. The RADAR system calculates the beat frequency, by calculating the difference between the transmitted and received signals [Bjo98; Tur20; Ani09]. This beat frequency f_b is directly related to the time delay Δt and the velocity v of the chest motion:

$$f_b = \Delta f \cdot \Delta t \tag{2.2}$$

$$v = \frac{f_b \cdot \lambda}{2} \tag{2.3}$$

where λ is the wavelength of the transmitted signal [Wan15]. By continuously monitoring the distance and velocity of the chest wall, the RADAR system captures the movements associated with the HBs and respiration which is further processed to extract the HR and to detect any abnormalities in cardiac activity. However, the technology also faces challenges such as artifacts from other body movements and therefore needs signal processing algorithms to accurately extract HSs and HBs from the RADAR signal [Bjo98; Tur20; Ani09].

CW radar, in contrast, is a radar system that emits continuous high-frequency radio waves and uses the Doppler effect to detect moving objects. CW radar can be utilized in a non-contact sensing mode, when radio waves are directed towards a subject, and the reflected waves contain

information about the motion of the chest due to the HB. The Doppler frequency shift (Δf) can be related to the motion of the chest, the dominant frequency of this shift then can be identified by fast fourier transform (FFT) which in return can be converted into BPM resulting in the value for the HR [Mal20]. Mathematically, the Doppler frequency shift (Δf) can be expressed as follows:

$$\Delta f = \frac{2fv\cos(\theta)}{c} \tag{2.4}$$

where Δf refers to the Doppler frequency shift, f is representing the frequency of the emitted wave, v represents the velocity of the target (motion of the chest), θ the angle between the radar line of sight and the target's velocity vector and c refers to the speed of light [Sko91].
Nevertheless, several challenges and potential difficulties are associated with using RADAR data for HB detection. These include interference from external motion, such as breathing or body movements, which can affect the accuracy of HB measurements [Keb20; Pra24]. Therefore, precise signal processing techniques to extract the HB signal from RADAR data are crucial and require sophisticated algorithms to filter out noise and artifacts.

2.4 Signal processing techniques to improve HS and HB detection

Several methods have been developed to increase the accuracy of HB detection for multiple biomedical measurement applications. Initial methods for processing and interpretation of HS or HB signals from different sources have been summarized in works by Mansier et al. [Man96] and Hedge et al. [Heg13]. However, HS or HB signals are non-stationary because their frequency content and amplitude vary throughout cardiac cycles, which makes it challenging to analyze them with the proposed traditional analyses methods. To address the challenges occurring for the analysis of non-stationary biomedical data, more recent processing techniques like Hilbert transform, STFT or CWT have been utilized [Tha20; Ism18; Zin03; Deb21; Mei18; Cha15].

2.4.1 Hilbert transform

The Hilbert transform identifies HBs by detecting peaks in the envelope curve, distinguishing systoles and diastoles based on the spacing between the peaks [Tha20]. This method is well-suited for non-stationary signals, as it provides a clear representation of amplitude variations over time.

2.4.2 Short-time fourier transform

The STFT, on the other hand, provides a time-frequency representation of the signal through a three-step process: segmentation (dividing the signal into shorter, overlapping segments), windowing with functions like Hamming or Hanning (reducing artifacts arising through discontinuous analysis of segment boundaries) [Pod14], and applying the FFT to decompose each segment into a frequency-domain representation of each segment [Far16; Dje00]. The result is a spectrogram that visually represents how frequency components of the signal evolve over time, enabling the identification of patterns, such as the frequency shifts associated with different disease-related heart sounds or murmurs [Dje00].

2.4.3 Continuous wavelet transform

The CWT provides an alternative approach that is highly effective for analyzing non-stationary signals [Zin03; Deb21; Mei18]. Unlike the STFT, that uses windows with a fixed size to analyze the signal, the CWT utilizes a set of functions known as wavelets. At the core of this approach is the mother wavelet. The CWT is defined as:

$$W(a,b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \, dt$$

where:

- $W(a,b)$ represents the wavelet coefficient at scale a and position b,
- $f(t)$ is the original signal,
- ψ is the mother wavelet function,
- a is the scale parameter, and
- b is the translation or shift parameter.

The mother wavelet $\psi(t)$ acts as a prototype to generate a set of wavelets through scaling (adjusting size) and translation (shifting in time). This process allows the CWT to analyze the signal at various resolutions, capturing both high- and low-frequency details [Zin03; Deb21; Mei18]. This property of the wavelets enables the CWT to provide both time and frequency resolution, unlike the FFT, which only offers frequency information [Yog76]. During the CWT calculation, the mother wavelets are convolved with the original signal at different points in time and across various scales. This convolution process involves sliding the wavelet across the signal, starting at the beginning of the input signal, moving along the time axis to the end and repeating the process

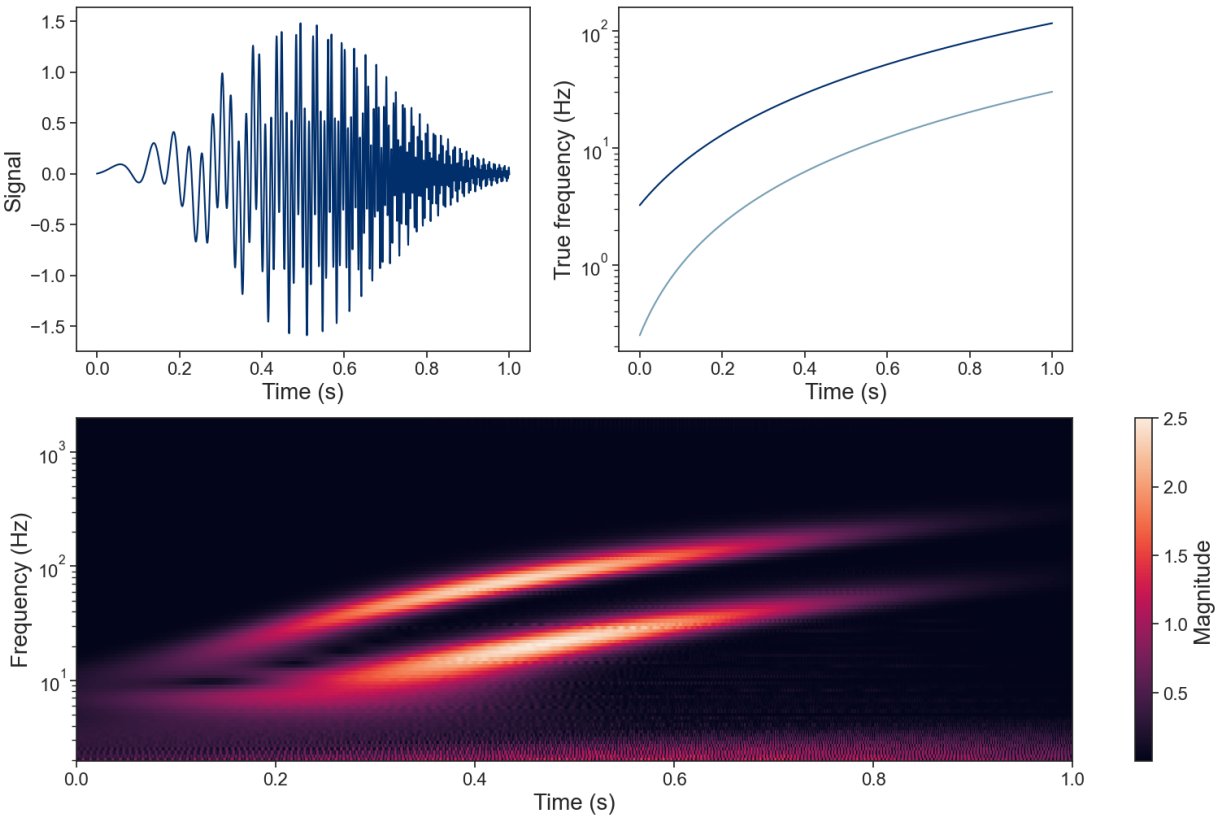

Figure 2.4: **CWT illustration**. Upper panel depicts an example signal (left) and the two time-dependent frequency components it contains (right). Lower panel shows the scaleogram containing signal strength vs. time and frequency after CWT. Graphic adapted from [Lee19].

with different scaled versions of the original mother wavelet to transforms the input signal from a one-dimensional signal into a two-dimensional one, containing the time and the so-called scale information [Zin03; Deb21; Mei18], that can be visualized in a scalogram (Figure 2.4), which displays how the signal's frequency content changes over time. The scales used in the CWT can be converted into frequencies using the formula:

$$f = \frac{f_c}{a}$$

where:

- f_c is the central frequency of the used mother wavelet, and
- a is the scale parameter of the CWT.

By using wavelets of different sizes, the CWT provides a scalogram, which shows how the signal's

frequency content changes over time. This flexibility makes the CWT particularly effective for identifying complex cardiac signals.

2.5 Filters to reduce noise and artifacts in RADAR data

In RADAR signal processing, especially for applications such as non-contact HB measurement, the presence of noise introduced by RLBMs, artifacts or technical interferences can drastically affect the accuracy and reliability of the data. To overcome these issues, various filtering techniques have been developed.

2.5.1 High-pass filter

A high-pass filter is designed to allow frequencies higher than a specified cutoff frequency while attenuating lower frequencies. This procedure is particularly effective in eliminating low-frequency noise from radar signals [Jol08; She05; Paa01]. The frequency response $H(f)$ of an ideal high-pass filter is expressed as:

$$H(f) = \begin{cases} 1 & \text{if } |f| > f_c \\ 0 & \text{if } |f| \leq f_c \end{cases}$$

where f_c denotes the cutoff frequency [Paa01].

2.5.2 Low-pass filter

Conversely, a low-pass filter allows frequencies lower than a specified cutoff frequency while attenuating higher frequencies. This enables effective removal of high-frequency noise from radar signals [Jol08; She05; Paa01]. The frequency response $H(f)$ of an ideal low-pass filter is given by:

$$H(f) = \begin{cases} 1 & \text{if } |f| \leq f_c \\ 0 & \text{if } |f| > f_c \end{cases}$$

where f_c represents the cutoff frequency [Paa01].

2.5.3 Band-pass filter

A band-pass filter permits frequencies within a specified range while attenuating frequencies outside this range. This filter is useful for isolating specific frequency components of interest in radar

signals [Jol08; She05; Paa01], such as those corresponding to the HB. The frequency response $H(f)$ of an ideal band-pass filter is given by:

$$H(f) = \begin{cases} 1 & \text{if } f_l < |f| < f_h \\ 0 & \text{if } |f| \leq f_l \text{ or } |f| \geq f_h \end{cases}$$

where f_l and f_h are the lower and upper cutoff frequencies, respectively [Paa01].

2.5.4 Butterworth filter

The Butterworth filter is characterized by a maximally flat frequency response in the passband, making it a popular choice for applications requiring minimal signal distortion. It can be implemented as a high-pass, low-pass, or band-pass filter [Jol08; Paa01]. The magnitude response $|H(f)|$ of a Butterworth filter is given by:

$$|H(f)| = \frac{1}{\sqrt{1 + \left(\frac{f}{f_c}\right)^{2n}}}$$

where: f_c is the cutoff frequency and n is the order of the filter. A higher order results in a steeper roll-off [Paa01].

2.6 AI-based applications to improve HS or HB detection

Signal processing techniques like Hilbert Transform, STFT, and CWT have made significant strides in improving the accuracy of HB detection. When used in combination with filters — such as butterworth or band-pass filters — these methods can effectively reduce noise and artifacts, while enhancing the signal clarity and improving detection performance. Despite these advancements, these techniques still face limitations when dealing with non-stationary data and complex noise patterns. Therefore, artificial intelligence (AI)-based applications have emerged as a powerful alternative for HB detection under such challenging conditions [Kai20; Sha23]. AI models based on LSTM units or the UNet architecture can help in adapting to differing signal conditions and still detect HBs in the RADAR signal.

2.6.1 Long short-term memory networks

LSTM networks, are a type of recurrent neural network (RNN), which are characterized by managing long-term dependencies within time-series data, making it well-suited for analyzing HB signals and being comprised of several LSTM cells.
The core component of an LSTM cell is the memory cell, which holds long-term information and helps the network remember context over extended sequences. This memory cell is managed by several gates that control the flow of information (Figure 2.5) [She20; Gre17]. The forget gate decides what information from the previous memory cell state (blue x_{t-1} in Figure 2.5) should be discarded. It uses a sigmoid function to output values between 0 and 1, where 0 means "forget" and 1 means "keep". The input gate determines what new information should be added to the memory cell. It includes a sigmoid function (green σ in Figure 2.5) to decide the importance of the new data and a $tanh$ function (green $tanh$ in Figure 2.5) to generate candidate values for updating the cell state. The output gate controls what part of the memory cell state should be used to form the hidden state (yellow h_t in Figure 2.5). It uses a sigmoid function to filter the cell

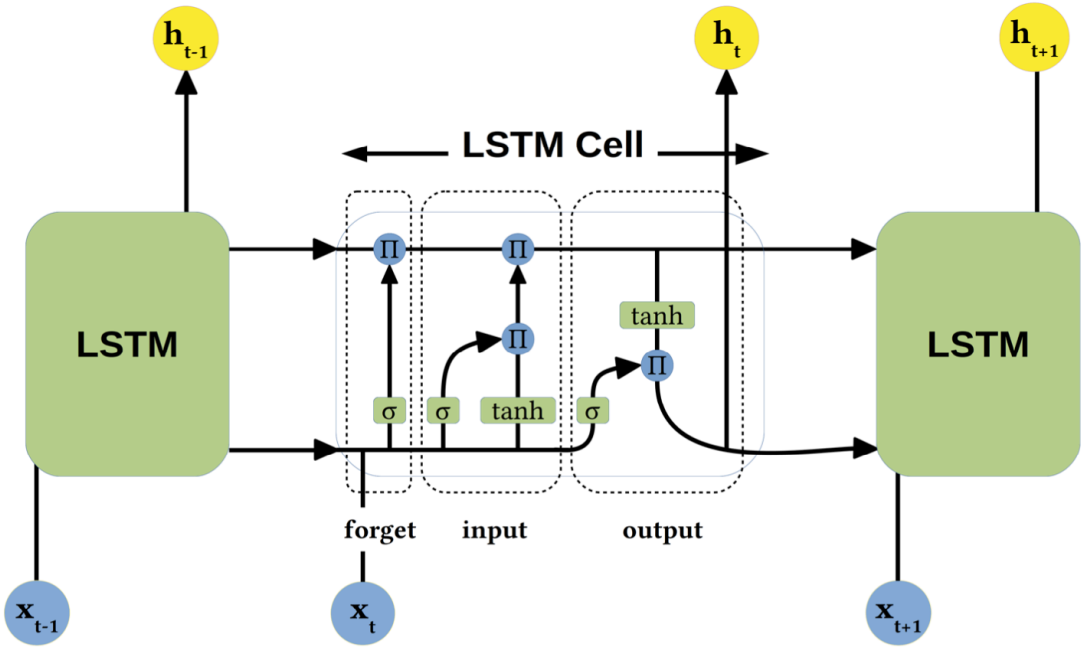

Figure 2.5: **LSTM cell architecture.** Middle part of the graph depicts the different gates (forget, input, output) of the LSTM cell with all relevant components (σ, $tanh$, h_t, x_t)[She20]. For a detailed architectural and functional description see main text.

state and a $tanh$ function to scale it, ensuring that only the relevant information is passed forward. The hidden state (yellow h_t in Figure 2.5 is the primary output of the LSTM cell at each time step which contains the information from the memory cell. This hidden state can be used for different tasks, like detecting HSs, classifying HB patterns, or detecting anomalies [She20; Gre17].

2.6.2 UNet

In comparison to LSTM networks, the UNet architecture is a convolutional neural network (CNN) designed primarily to segment images, like scalograms created by CWT, or biomedical images [Du20]. It features a distinctive U-shaped structure, that consists of three parts: the contracting path (encoder), the expanding path (decoder) and skip connections (Figure 2.6) [Ron15]. The

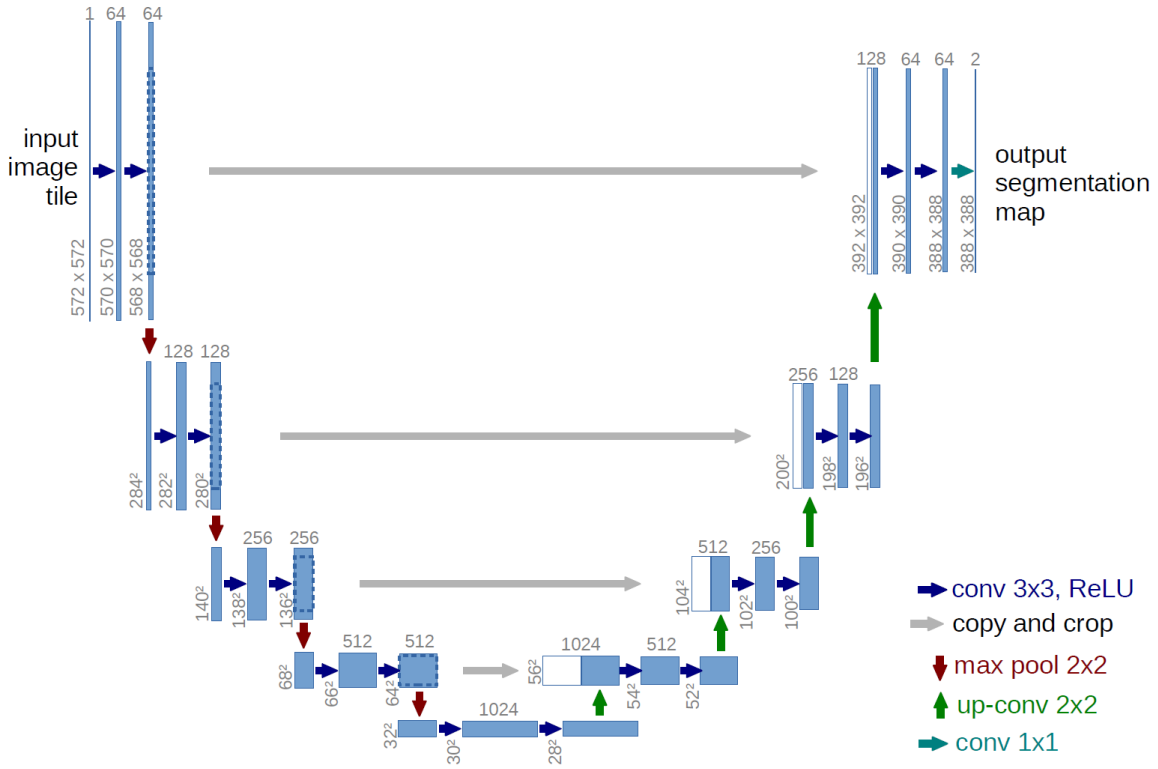

Figure 2.6: **UNet architecture**. Blue boxes are referring to a multi-channel feature map with the number of features on top and x-y-size on the left. Boxes in white depict copied feature maps. Different operations are represented by arrows [Ron15]. For detailed architectural and functional description see main text.

contracting path progressively reduces the spatial dimensions of the input image while capturing its essential features. This process involves a series of convolutional and pooling layers. Convo-

lutional layers apply multiple filters to the input image to detect various features such as edges, textures, and patterns. Each convolutional layer is followed by a rectified linear unit (ReLu) activation function, which introduces non-linearity into the model, enabling it to learn more complex patterns (blue arrows in Figure 2.6). Max pooling layers perform downsampling by selecting the maximum value from each region of the feature map (red arrows in Figure 2.6). This reduces the spatial dimensions of the feature maps while retaining the most important information. Typically, a 2x2 max pooling operation is used, which halves the dimensions of the feature maps. As the image passes through the encoder path, the number of feature maps (channels) increases, allowing the model to capture more complex and abstract features. At the bottleneck, the image is represented in its most abstract and compressed form. This stage captures the essence of the input image, with the spatial dimensions drastically reduced but the depth (number of channels) increased. The bottleneck serves as the transition point between the encoder and decoder paths [Ron15]. The expanding path reverses the process of the encoder to reconstruct the high-resolution segmentation map by progressively increasing the spatial dimensions of the image while adding finer details. This involves a series of up-convolutional (transposed convolution) layers and skip connections. Up-convolution performs upsampling, increasing the spatial dimensions of the feature maps (green arrows in Figure 2.6). Skip connections link corresponding layers in the encoder and decoder paths, enabling the model to combine high-resolution features from the encoder with the upsampled features in the decoder (grey lines illustrated in Figure 2.6). This mechanism ensures that the detailed spatial information lost during downsampling is retained and integrated into the reconstruction process. Each step in the decoder path combines the upsampled image with the corresponding feature maps from the encoder path, refining the details and improving the accuracy of the segmentation. The feature map is processed by various convolutional layers, and after that, batch normalization and ReLu activation are applied. Through that, more intricate patterns can be learned, and the precision of the segmentation map can be improved. At the final level of the expanding path, a final convolution is applied (turquoise arrow in Figure 2.6) to ensure that the number of channels matches the desired output channels. The final output is a segmented image where each pixel is labeled as part of a specific object [Ron15].

Chapter 3

Related Work

In recent years, various methodologies have been explored to enhance the detection and analysis of HBs.

3.1 Spectrum-based methods

Spectrum-based techniques for HB detection focus on analyzing the frequency components of cardiac signals. These methods include techniques such as the STFT and CWT, which use time-frequency representations to capture non-stationary HSs and HBs [Car03]. Chen et al. [Che09] characterized time and frequency of HSs by using STFT and used this method for differentiation of sample signals of coronary heart disease and non-coronary heart disease. However, the authors did not mention the RADAR system used or the experimental setup for measuring the heart-related vital signs. Mogi et al. [Mog17] used Doppler RADAR data to measure heart rate variability by R-R-intervals (changes in the intervals between successive HBs) from participants sitting still. They used a band-pass filter to reduce the noise caused by respiration and small body motion together with STFT to calculate the spectrogram. The authors could verify that the root mean square errors (RMSEs) of R-R intervals were improved for all subjects compared to conventional algorithm using adaptive scale factor selection on learning [Mog15]. In a recent study from Shi et al. [Shi22] an approach to extract the HB from CW RADAR data was utilized. The recording was performed on five subjects sitting down and instructed to move as little as possible. After the recording, the RADAR data was first filtered with a band-pass filter, so only frequencies in the range of 0.05 to 15 Hz remained. After this preprocessing step, the data was transformed utilizing STFT. The authors showed, with simulated as well as recorded data, that their approach could successfully extract the HR from RADAR data with the advantage of being able to determine the

HR in real time. Vikhe et al. [Vik09] compared the usage of STFT and CWT for heart sounds and could show that STFT was unable to determine the time between the closure of the aortic (A2) and pulmonary valve closure (P2) in S2 of the HSs, which play a crucial role in various cardiac diseases. However, they could also show that CWT, on the other hand, was not only able to distinguish between A2 and P2 but also to measure the delay in time. However, the authors used PCG signals recorded using electronic stethoscopes as input data for their analyses instead of RADAR data [Vik09]. Li et al. [Li17] used CWT on RADAR data of participants sitting in front of the radar, breathing normally, and compared the performance with a Fourier transform-based algorithm. With that approach, the authors could achieve an average error reduction from >26% to 3.5%.

3.2 Deep learning-based techniques

With the advent of AI, deep learning methods have emerged as powerful tools for cardiac signal analysis. Techniques such as LSTM networks and UNet architectures have been shown to efficiently handle complex and noisy data. For example, Shi et al. [Shi19] focused on detecting the S1 and S2 HSs using a dataset of 30 test persons and different LSTM architectures in order to accurately extract the HSs from RADAR data. After different model variations, the authors were able to train a model that could predict S1, Systole, S2 and diastole with an accuracy of 93.4%. Han-Trong et al. [Han22] used CW RADAR and ECG signals of 30 participants breathing normally. Motion artifacts of the participants were extracted using a Butterworth filter in the range of 0.83 to 2.33 Hz and and the filtered data was fed into a LSTM model. With this setup, the authors could achieve a correlation coefficient of 96.64% between the CW RADAR and the ECG signal. Chowdhury et al. [Cho24] compared four different AI based methods: UNet, FPN, LinkNet und MultiResLinkNet. The authors used a publicly available dataset from [Lu23], including RADAR and ECG data from 30 subjects in five different lying down scenarios: resting, valsalva (forceful exhalation against a closed airway), apnea (holding the breath), tilt-up, and tilt-down. The authors could show that the MultiResLinkNet outperformed all other tested architectures (F1 score > 88% in all measurement scenarios), even though all of them achieved correlation > 55% between the ECG and RADAR signal.

3.3 Noise-based approaches

Noise-based approaches involve advanced signal processing techniques to filter unwanted artifacts and improve signal clarity. Yang et al. [Yan18], for example, suggested a tool for HB measurements using a band-pass filter and CW Doppler RADAR sensor to develop a quarantine system at airports for dengue fever. However, almost all data they used (>90%) needed to be excluded due to RLBM or movements while breathing. Iwata et al. [Iwa21] on the other hand, combined a matched filter and singular value decomposition to detect HR information from chest surface vibrations with high accuracy. Chen et al. [Che21] presented a software system intending to extract vital signs even in the presence of motion artifacts. This system is based on deep contrastive learning, so signals coming from body movements can be distinguished from heart-related vital signs. To test their system, they compared the predictions from data recorded in varying amounts of body movements: playing on the phone, typewriting, swaying the body, leg shaking, walking on the treadmill, sitting down, and standing up. They could show that even in scenarios with larger amounts of motion, the system could still predict the inter beat interval (IBI) with an error of under 5%. However, this system was only tested using data recorded with impulse-radio ultra-wideband radar (IR-UWB) and FMCW RADAR. Tang et al. [Tan23] utilized a dataset containing less noisy data and data with motion artifacts. In this case, the I and the Q components of the RADAR data were preprocessed using STFT and afterwards separately fed into a dual autoencoder network. The output was then combined to a single spectrogram from which the HBs were extracted. They could show that this type of network could improve the signal to noise ratio (SNR) by combining less noisy data with data containing motion artifacts. But, the authors did not provide an explanation how the waveform was recovered from the resulting spectrogram

Chapter 4

Methods

4.1 Data acquisition

In order to analyze which models and preprocessing steps can handle or tolerate motion artifacts, two different data sets with varying levels of patient movements were used in the work presented here. Both datasets, Radarcardia (EmpkinS D05; in the following referred to as D05 dataset) and EKSpression (EmpkinS D02; hereafter called D02 dataset) were recorded as part of studies conducted within the Collaborative Research Center "Empatho-Kinaesthetic Sensor Technology – Sensor Techniques and Data Analysis Methods for Empatho-Kinaesthetic Modeling and Condition monitoring (EmpkinS)".

4.1.1 D05 dataset

The first dataset was recorded in November 2023, containing 22 participants, evenly split with 50% females. This data was collected as part of a study conducted by [Oes24]. The ECG data was recorded using the BIOPAC MP36 device (BIOPAC Systems Inc., Goleta, CA, USA) in combination with the SS2LB module, as described by Oesten [Oes24] and Albrecht et al. [Alb24]. The SS2LB module captures a single-channel ECG following Lead II as described by Einthoven, with three electrodes placed as illustrated in Figure 4.1A [Jin12]. The recordings were saved in an unprocessed format without applying any filter and were subsequently exported as an .acq file. For the RADAR measurements, radar sensors were placed approximately 10 to 15 cm above the measurement location (Figures 4.1B and 4.2). The used measurement setup for the study is depicted in Figure 4.1B [Oes24]. The RADAR sensor measured at four anatomical locations: the arteria carotis communis, and the proximal, medial, and distal aorta (Figure 4.2). Two measure-

ment modalities were considered: (1) Normal - which indicates a normal breath and (2) Hold - which refers to a participant holding their breath. The arteria carotis communis was only measured during normal breathing, resulting in seven distinct measurement modalities for each participant. The single RADAR sensor recorded the I and Q components at a frequency of 1953.125 Hz and transmitted them to a synchronization node, which also transmits a synchronization signal to other systems. All data were stored in a database, extracted using Python, and saved as .hdf5 files. In all of the recordings the participants were lying down, resulting in a minimum amount of artifacts caused by RLBM [Oes24].

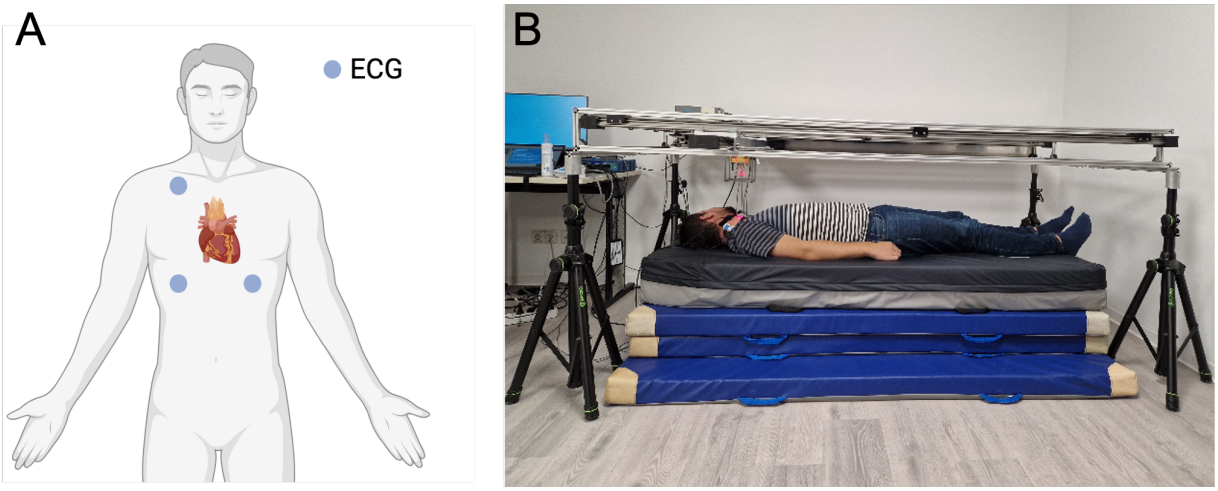

Figure 4.1: **Experimental setup for HB measurement.** (A) Electrode placement for ECG according to Einthoven. As represented with blue dots, electrodes are placed under the right clavicle, on the lower left and the lower right lib creating an imaginary triangle around the heart. (B) Measurement setup for RADAR and reference signal (ECG), graphic modified from [Oes24].

4.1.2 D02 dataset

The second dataset contained a total of 110 participants (balanced age, gender and equal parts of participants with and without depression based on ICD-10 classification, interviews, visual analogue scale (VAS) and questionnaires) with corresponding HB measurements [KeiND]. This data was collected as part of a study conducted by Keinert et al. [KeiND]. For the purpose of the work presented here, solely ECG and CW RADAR sensor recordings were used for all subsequent analyses. The ECG signal was recorded using the BIOPAC MP160 device (BIOPAC Systems Inc., Goleta, CA 93117, USA). The RADAR sensor recorded the I and Q components at 61 GHz. The RADAR signals were transmitted to a synchronization node, which also transmits

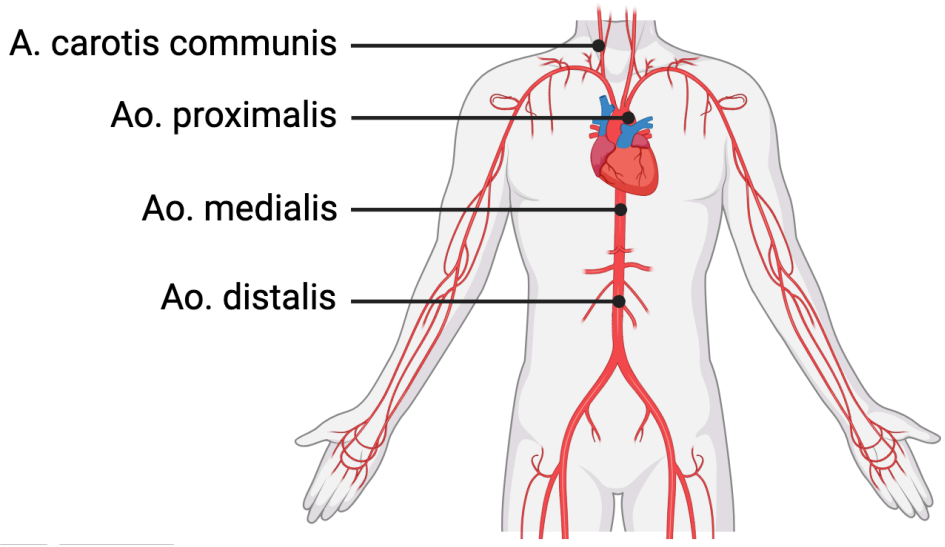

Figure 4.2: **Locations of HB measurements.** Representation of the different anatomical locations used for HB measurement which were depicted in the thesis presented here. As shown, four different locations along the aorta and arteria carotis communis were used.

a synchronization signal to other systems [KeiND]. Four measurement phases were considered: emotion induction, coping, training and latency, resulting in an increased amount of movement and therefore artifact caused by RLBM.

4.2 Data processing

To evaluate the effectiveness of different models and preprocessing methods in handling or tolerating motion artifacts, both datasets were analyzed with two separate approaches - LSTM and UNet, each requiring a separate preprocessing pipeline. Both of these pipelines were implemented using the tpcp Python package [Küd23]. If not explicitly mentioned, all parameters were used with the default value.

4.2.1 ECG as ground truth

While the PCG is valuable for detecting HSs and mechanical functions, the timing and intensity of HSs can be influenced by factors that do not affect the ECG, leading to potential discrep-

ancies when used as a reference. For this reason, the ECG was chosen as the reference measurement in this thesis. Data were collected from two sensors, ECG and CW RADAR, each utilizing different internal clocks. For accurate comparison, both signals need to be perfectly synchronized, which was not the case, necessitating synchronization. First, the actual sampling rate of both signals was determined using FFT, since the actual sampling rate of the sensors can deviate from the set sampling rate due to clock drift. Subsequently, both signals were resampled to 1000 Hz to ensure equal temporal resolution. Cross-correlation was then computed by cutting both signals to equal length and identifying the peak of the cross-correlation function to determine the positions with the best alignment. This information was used to calculate the time offset, which was then applied to the RADAR signal for final alignment with the ECG signal in the time domain. Following synchronization, the signals were divided into 5-second segments with a 40% overlap. The ECG signal was then downsampled to 200 Hz.

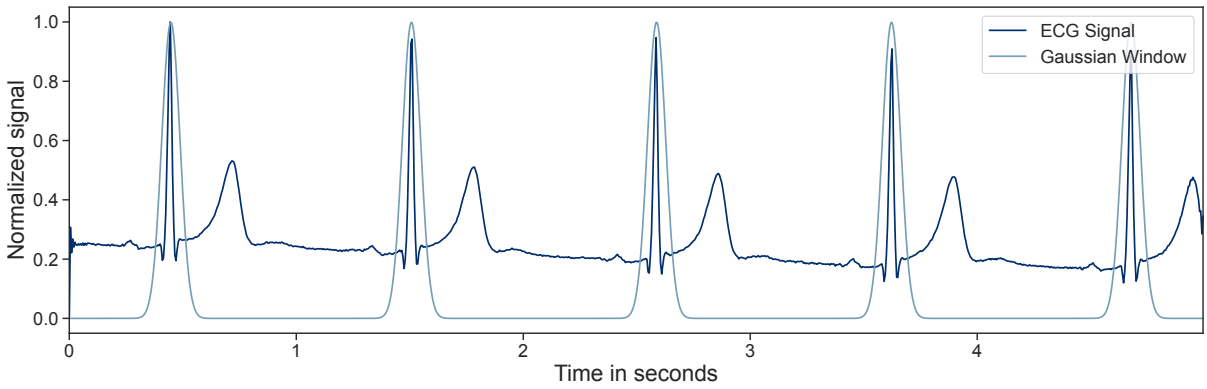

Figure 4.3: **Label representation.** Illustration of the cleaned ECG signal and the transformed signal used as labels in the thesis presented here. Dark blue signal is referring to the cleaned ECG signal, whereas the light blue line is representing the transformed signal.

Using the downsampled ECG data, the R-peak locations (of the QRS complex) were identified using the ecg_process method of the Neurokit2 package in Python. At the R-peak locations, Gaussian windows with a window length of 400 were placed and the rest of the signal remained 0 to use the Gaussian window in following analyses as sole label for identifying the HB (Figure 4.3). This method showed good results in previous research from Lu et al. [Lu23] and was chosen as label for all trained models in this thesis. Using the cleaned ECG signal (instead of the Gaussian windows) as labels was also tested, but this did not lead to a sufficient performance.

4.2.2 LSTM approach

The preprocessing pipeline for the LSTM approach is illustrated in Figure 4.4 and involves six steps. The first steps — synchronizing the measurements and segmenting the data — are con-

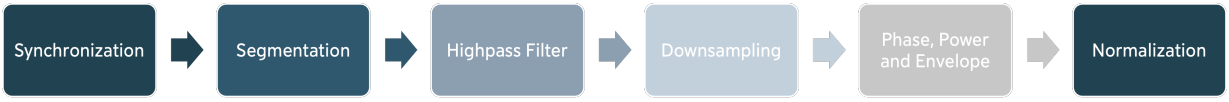

Figure 4.4: **Preprocessing pipeline for the LSTM approach.** Outlines the different preprocessing steps performed for the LSTM approach.

ducted in the same way as described for the preprocessing of the ECG data (see Section 4.2.1). Subsequent preprocessing steps include applying a Butterworth high-pass filter separately to the I and Q component of the RADAR data in order to filter out frequencies outside of a possible HB frequency. The filter used has a 5th-order and a cutoff frequency of 0.4 Hz. Following filtering, both the I and Q components were downsampled to 200 Hz. The angle and power of the signals are then calculated using Equations 4.1 and 4.2, respectively. A 4th-order Butterworth band-pass filter is applied to the power signal, with a frequency range of 18 to 80 Hz.

$$\theta = \tan^{-1}\left(\frac{Q}{I}\right) \tag{4.1}$$

$$P = \frac{1}{T}\int_0^T |x(t)|^2\, dt \tag{4.2}$$

$$E(t) = \sqrt{x(t)^2 + (\mathcal{H}\{x(t)\})^2} \tag{4.3}$$

The envelope of this filtered signal is computed using Equation 4.3, where the Hilbert Transform is used as in Equation 4.4. These steps were used, because a similar preprocessing showed promising results in Albrecht et al. [Alb24]. After completing these preprocessing steps, the last step involved standardizing and normalizing the five calculated components: I, Q, angle, power, and envelope. The different results are standardized using the z-score (Equation 4.5) and then normalized to a range of [0,1] using unity-based normalization (Equation 4.6), as this showed good performance in Chowdhury et al. [Cho24].

$$\mathcal{H}\{x(t)\} = \frac{1}{\pi}\,\text{P.V.}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}\,d\tau \tag{4.4}$$

where:

- $x(t)$ is the original signal,
- τ represents the time delay or shift,
- t is the current time point,
- $P.V.$ stands for the Cauchy principal value.

$$z = \frac{x-\mu}{\sigma} \tag{4.5}$$

where:

- x refers to the value of the data point,
- μ is the mean of the dataset,
- σ represents the standard deviation (SD) of the dataset.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{4.6}$$

A log transformation of the input values was also tested, but this resulted in worse performances and was therefore not used for further experiments.

The resulting preprocessing output is exemplary illustrated in Figure 4.5. All segments were saved as individual files which could not fit into the systems memory. To address this, segments were loaded individually using TensorFlow datasets [Mar15], which can load segments as needed and prefetch additional segments for upcoming training steps. As architecture for the LSTM approach, a sequential model in Keras[1] was applied (Figure A.1). The input shape of the first layer has a shape of (5, 1000), where the five arrays correspond to the I and Q component as well as the power, angle and the envelope of the RADAR signal, respectively. After the following layers of this sequential model, a dense layer was chosen, ensuring that the output of the model

[1] https://keras.io/guides/sequential_model/

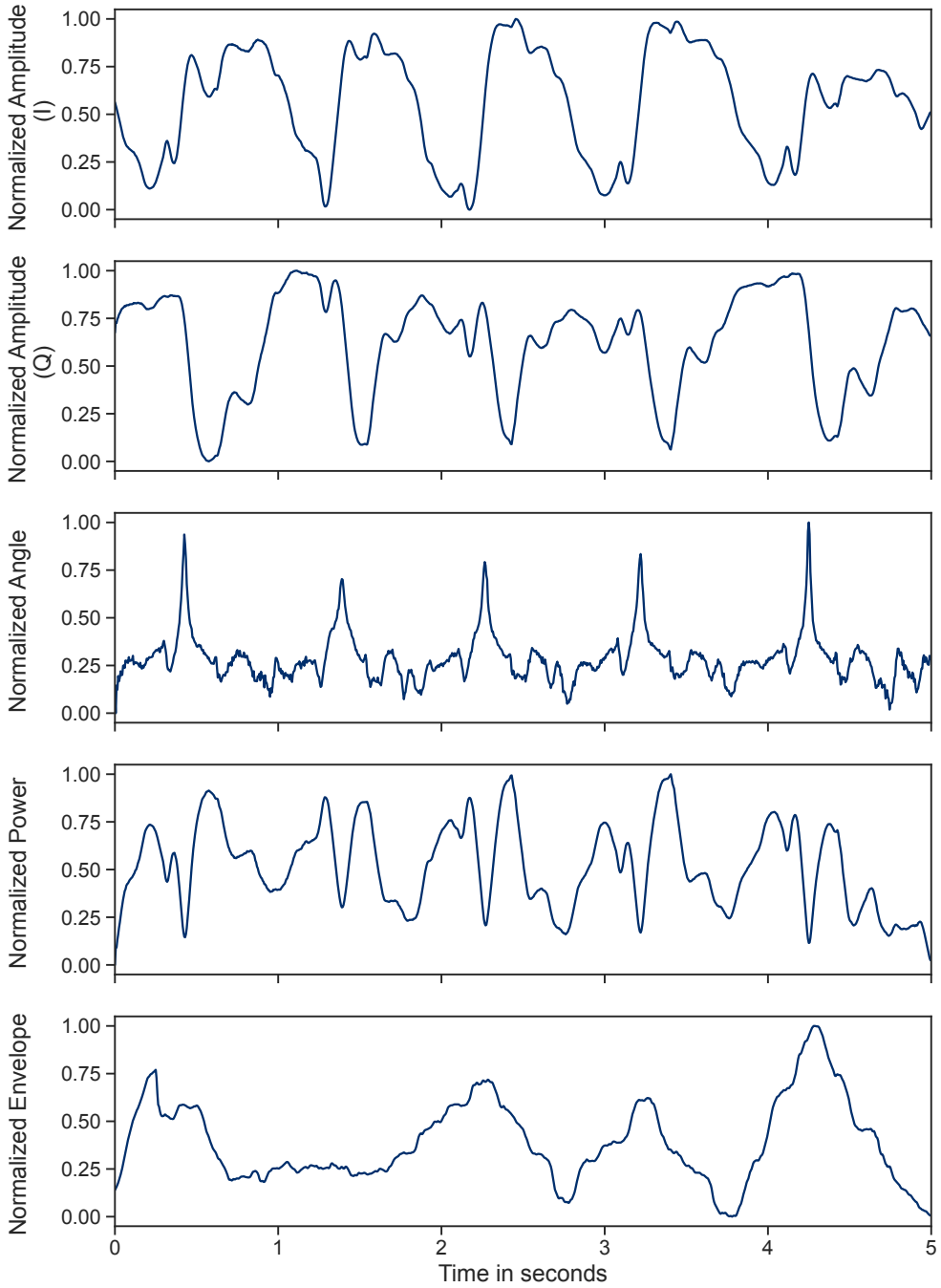

Figure 4.5: **Representation of the preprocessing output for the LSTM approach.** Line plots depicting the different value changes over time for I, Q, angle, power, and envelope aligned in the time domain. The X and Y axes refer to normalized values and time. White "X" represents the corresponding mean performance metric.

has the same dimensions as the label vectors. For this model the Binary Cross entropy function as the loss function and the Adam algorithm as an optimizer, both from Keras[2], were selected. This architecture was tested with varying amounts of LSTM units (128, 256, 512) in the corresponding layers. Additionally, different learning Rates (LRs) ($1e^{-3}$ and $1e^{-4}$), which adjust the weights during the training, were tested to find the best fitting one. Furthermore 40, 100 or 300 epochs were used for training.

4.2.3 UNet approach

The data preprocessing for the UNet approach includes several subsequent steps (Figure 4.6). Synchronizing and segmenting the RADAR data was performed the same way as described for

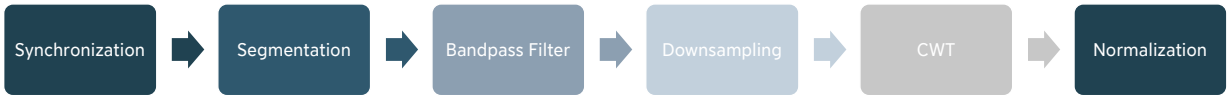

Figure 4.6: **Preprocessing pipeline for the UNet approach.** Outlines all used steps for preprocessing performed for the UNet approach.

the ECG data (see Section 4.2.1). Each segment is then filtered using a 4th-order Butterworth band-pass filter with a frequency range of 18 to 80 Hz. This frequency range is selected based on previous research [Wil18] that demonstrated its effectiveness in enhancing HB detection in RADAR data. The filtered data are then downsampled to 200 Hz, resulting in 1000 data points per segment. For the Continuous Wavelet Transform (CWT), various mother wavelets are tested to identify the most suitable one for subsequent analysis using the PyWavelets package in Python [Lee19]. As illustrated in Figure 4.7, four different mother wavelets are compared: Morlet, Shannon, Mexican Hat, and Gaussian wavelet, in the following referred to as morl, shan1-1, mexh, and gaus1, respectively. Scales for the CWT are selected within a range of 1 to 257, as recommended by the PyWavelets package documentation [Lee19], and are transformed into a geometric progression, where each number is a constant multiple of the previous number [Har20]. Visual inspection of the CWT results confirms that this range and transformation provided the best representation of the HB signal, since the HB were clearly visible in the spectrogram of the CWT signal (see Figure 5.3). Following standardization and normalization was conducted the same way as described for the LSTM approach (see Section 4.2.2).

[2]https://keras.io/api/optimizers/adam/, https://keras.io/api/losses/probabilistic_losses/

The computed segments were also saved individually and loaded during the training and test-

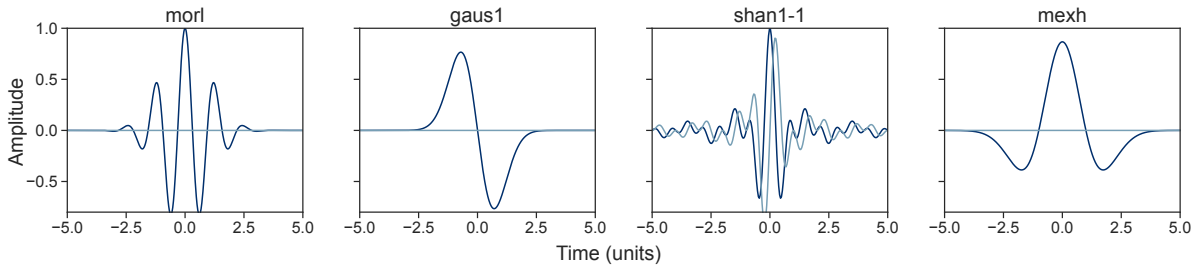

Figure 4.7: **CWT mother wavelets.** Illustration of the different mother wavelets used for the CWT in the work presented here. As shown, morl, gaus1, shan1-1 and mexh were included. The light blue lines show the imaginary part of the mother wavelet function, and the dark blue lines the real part.

ing process using a TensorFlow dataset. The standardized and normalized CWT results were used as TensorFlow datasets to provide the input for a sequential model using Keras[3] (Supplementary Figure A.2). A UNet model is used as first layer with an input size of (256, 1000, 1). For each up- and downsampling level filters with the sizes 16, 32 and 64 were used. The output activation for this layer was set to a linear activation function. Initial weights are randomly set, and the number of labels the model has to differentiate from is set to 1. Following the UNet layer, a 2D-convolutional layer was applied with a kernel size of (256, 1) to reduce the dimensions from (256, 1000, 1) to (1, 1000, 1). A flattening layer then adjusts the output to match the label vector size (1000). The Binary Cross Entropy from Keras[4] was also used as a loss function and the Adam algorithm (from Keras[5]) as the optimizer. For the UNet models, also different LRs ($1e^{-3}$ and $1e^{-4}$), were tested to find the best fitting one.

4.3 Evaluation

To evaluate the performance of the ML models, performance metrics for the continuous data obtained by reconstructing the original time series from the segmented data was calculated.

[3]https://keras.io/guides/sequential_model/
[4]https://keras.io/api/losses/probabilistic_losses/
[5]https://keras.io/api/optimizers/adam/

As main performance metric the F1 score was used, which balances precision and recall. Precision (Equation 4.7) measures the proportion of true positives (TPs) among predicted positives (TPs and false positives (FPs)), while recall (Equation 4.8) measures the proportion of TPs among actual positives (TPs and false negatives (FNs)).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.7}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.8}$$

A prediction is classified as TP if it was detected within a tolerance window of 50ms or 100ms around the peak of the Gaussian window of the label data (with 50% of the tolerance window in either direction). The amount of detected TPs was subsequently divided by the sum of all detected peaks in the predictions to calculate the Precision. For the Recall, the amount of TPs was divided by the number of peaks in the label data. The F1 score (Equation 4.9) combines these metrics into a single value, where 1 represents a perfect detection and 0 represents the worst.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.9}$$

Additionally, the Pearson correlation coefficient (Equation 4.10) was calculated to measure the linear relationship between predicted and actual values as follows:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{4.10}$$

where X_i and Y_i are individual data points, and $\bar{X}$ and $\bar{Y}$ are the means of the two variables, respectively. The Pearson coefficient ranges from -1 (perfect negative correlation) to 1 (perfect positive correlation), with 0 indicating no correlation. For all approaches discussed in this thesis the data was split into a training and testing dataset. 80% of the subjects were randomly assigned to the training and validation dataset and the remaining 20% to the testing dataset, leading to the distribution of segments illustrated in Table 4.1.

4.4 Computational settings

Python version 3.9 was used for all preprocessing procedures, whereas Python version 3.10 was utilized for training and analyses. All models were trained using a single NVIDIA GeForce RTX2080Ti or a single NVIDIA GeForce RTX3080, Tensorflow version 2.15.0, and version

Table 4.1: **Number of segments per dataset.** Segments with an overlap of 40% and a length of 5 seconds for the D05 dataset, D02 dataset and in total.

	D05	**D02**
Train and Validation Set	4,411	105,235
Test Set	1,385	24,610
Total	5,796	129,845

2.15.0 of Keras.

Chapter 5

Results and Discussion

In this chapter, the results of the different approaches used to detect the HB from the obtained data are presented and discussed in detail. Each method's performance is evaluated and compared to each other.

5.1 Individual approaches

5.1.1 Long short-term memory approach

For this approach, the preprocessed data (see Section 4.2.3) was used to train a LSTM model for both datasets separately and in an combined dataset, with the architecture illustrated in Supplementary Figure A.1.

D05

To determine the best-performing LSTM model, various hyperparameters, including the LR, and the number of epochs, as well as different amounts of LSTM units before and after the first dropout layer were used.
After training the model with a LR of $1e^{-4}$ and $1e^{-3}$ for 40 epochs, it was examined that the lower LR of $1e^{-4}$ performed better (Table 5.1). This not only becomes apparent when comparing the two mean values of the F1 score with a tolerance window of 100ms, where the model with a LR of $1e^{-4}$ achieves a 25% higher score than with a LR of $1e^{-3}$ but also since the model with a LR of $1e^{-3}$ fails to detect any HBs for one modality. The same behavior was observed for the F1 score (50ms) and the correlation coefficient, as well as their corresponding SD. To explore the full potential of this model, several other LSTM models with different numbers of training epochs

Table 5.1: **Performance metrics separated by LR.** Differentiation of the LSTM model performance trained with data from the D05 study with 40 Epochs and a LR of $1e^{-4}$ and $1e^{-3}$, respectively.

	LR of $1e^{-4}$			**LR of $1e^{-3}$**		
	F1 score (100ms)	**F1 score (50ms)**	**Correlation**	**F1 score (100ms)**	**F1 score (50ms)**	**Correlation**
Mean	0.49	0.28	0.39	0.39	0.22	0.35
SD	0.23	0.17	0.09	0.26	0.19	0.10

and LSTM units were trained. However, increasing the number of LSTM units to 512 for both layers did not improve the model's overall performance. As shown in Table 5.2, the measured F1 score with a tolerance window of 100ms is at the maximum when 256 LSTM units were utilized as opposed to 128, or 512 units. It could also be observed that the correlation stays stable when 128 or 256 LSTM units were used. Since using 256 LSTM units led to the best performance, two separate LSTM models were trained with more training epochs. This increase in training epochs improved the models' performances (Figure 5.1). The largest improvement occurred between 40 and 100 epochs, with the mean F1 score (100ms) increasing by 32.51%. Between 40 and 300 epochs, the F1 score (100ms) improved by 36.77%, with a modest increase of 1.91% between 100 and 300 epochs. The correlation also rose with higher epochs, increasing by 8.34% from 100 to 300 epochs. Therefore, the best-performing model for this dataset uses 256 LSTM units, is trained for 300 epochs, and has a LR of $1e^{-4}$. When comparing the different measurement locations of the RADAR data (Figure 5.2), it becomes evident that the measurement location influences the accuracy of detecting HBs. The model performed best at the Arteria carotis communis, with the mean of every metric being the highest, and the interquartile range (IQR) among the lowest. Furthermore, the performance was generally better for modalities at which the subject held their breath, reducing motion artifacts (Figure 5.2). The performance metrics for different subjects show that the subject impacts the model's performance as well (Table 5.3). Subjects VP 01 and

Table 5.2: **Performance metrics for different LSTM units.** F1 score with tolerance window of 50 and 100ms and correlation coefficient between observed and predicted HRs for models with differing number of LSTM units for the D05 dataset.

	128 LSTM units	**256 LSTM units**	**512 LSTM units**
Mean F1 score (100ms)	0.40	0.49	0.34
Mean F1 score (50ms)	0.24	0.22	0.25
Correlation	0.40	0.40	0.31

VP 15 performed worse across all three metrics compared to the other participants.

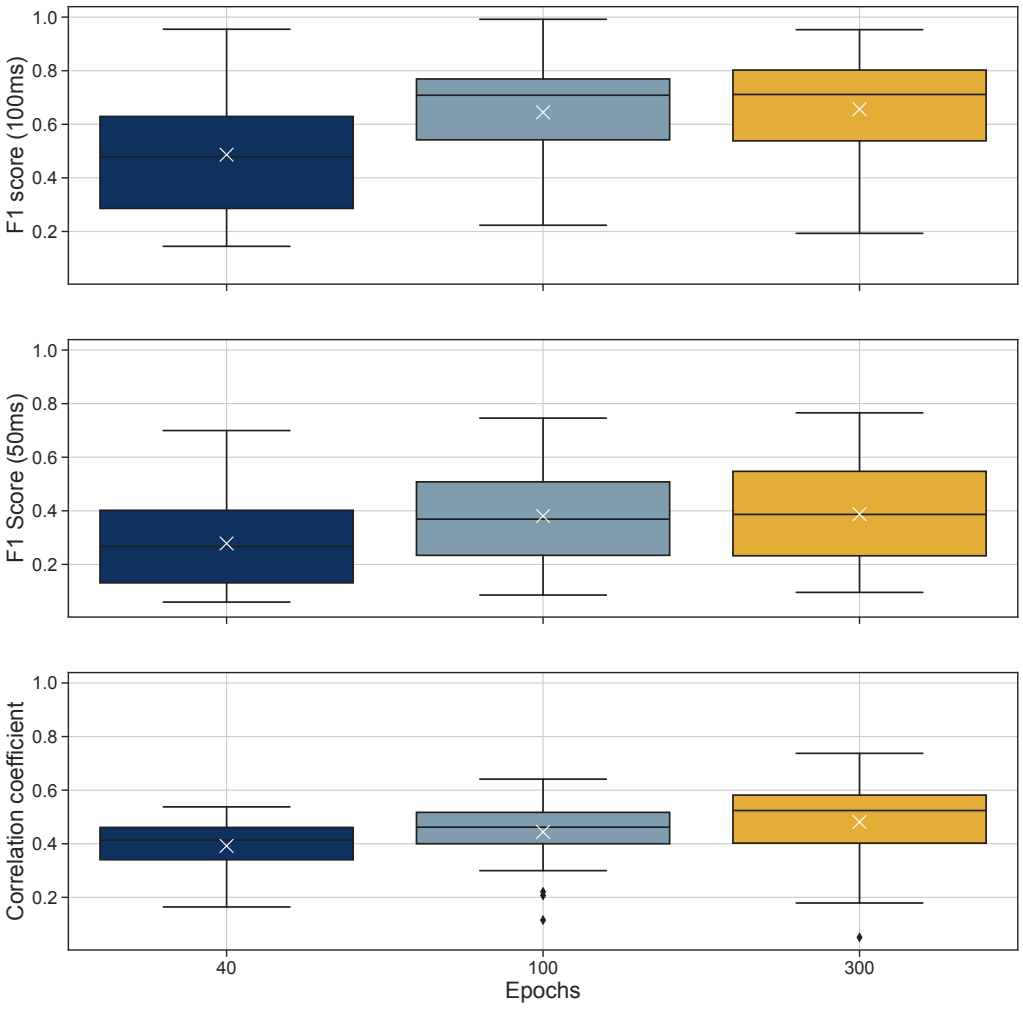

Figure 5.1: **Performance metrics for different epochs for the D05 dataset.** F1 Score with tolerance window of 50 and 100ms and correlation coefficient between observed and predicted HBs for models with differing number of epochs.

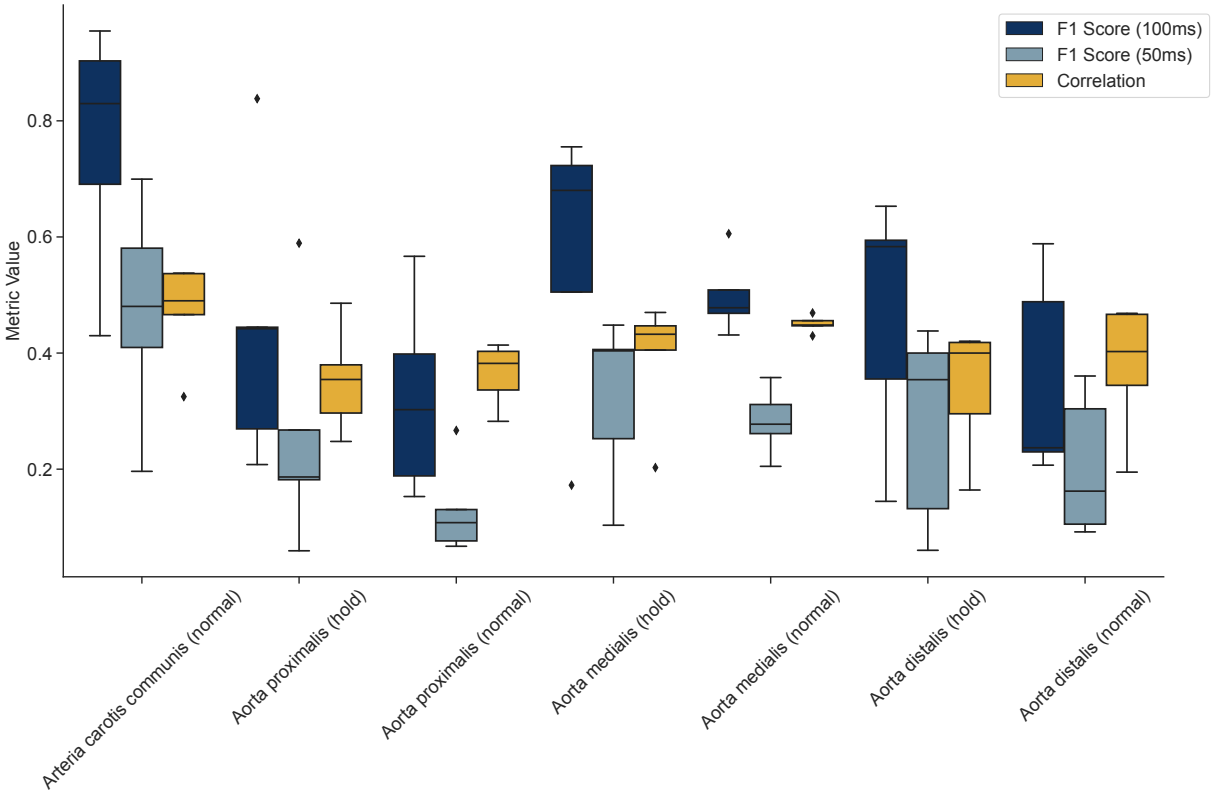

Figure 5.2: **Location-dependent performance metrics for LSTM approach of the D05 dataset.** F1 score with tolerance window of 50 and 100ms and correlation coefficient between observed and predicted HRs for differing measurement locations. *hold* = Participant is holding the breath while measuring; *normal* = Participant is breathing normally.

Table 5.3: **Subject-dependent performance metrics of the D05 dataset.** F1 score with tolerance window of 50 and 100ms and correlation coefficient between observed and predicted HRs for differing subjects.

	F1 score (100ms)				F1 score (50ms)				Correlation			
Subject	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
VP 01	0.57	0.18	0.29	0.74	0.31	0.12	0.14	0.43	0.42	0.14	0.18	0.56
VP 03	0.75	0.09	0.61	0.85	0.49	0.15	0.32	0.67	0.55	0.08	0.44	0.66
VP 11	0.74	0.12	0.59	0.95	0.44	0.20	0.12	0.75	0.55	0.10	0.44	0.74
VP 15	0.51	0.24	0.19	0.88	0.29	0.19	0.10	0.56	0.35	0.21	0.05	0.67
VP 18	0.71	0.20	0.38	0.93	0.41	0.22	0.10	0.63	0.53	0.15	0.27	0.69

D02

Despite the extensive data available in the D02 dataset for training ML models, it is noisier than the D05 dataset (see Section 4.1). A notable challenge was that using a learning rate of $1e^{-3}$ hindered the model from learning effectively, resulting in no detectable HBs. To get around that, it was necessary to increase the number of training epochs to 300 to ensure sufficient learning. Nonetheless, even with these adjustments, the predictions from the D02 dataset were worse across all performance metrics compared to those from the D05 dataset, as shown in Table 5.4.

Table 5.4: **Performance metrics for the LSTM model of the D02 dataset.** Evaluation of prediction with 100ms and 50ms tolerance windows, trained with 300 epochs, 256 units, and a LR of $1e^{-4}$.

	Mean	SD	Min	Max
F1 score (100ms)	0.12	0.05	0.02	0.24
F1 score (50ms)	0.06	0.03	0.00	0.14
Correlation	0.02	0.04	-0.05	0.15

Combined data

To enhance predictions for datasets with higher noise levels, both the D05 and D02 datasets were utilized together to test whether the model could improve its learning when exposed to data with high and low SNR at the same time.

Multiple LSTM models were trained with different numbers of epochs and LRs with the combined dataset. Since the models with 128 or 512 LSTM units performed poorly in the DO5 dataset, all models discussed in this chapter feature 256 units in both LSTM layers. Across both LRs, the LSTM models trained with only 40 epochs consistently underperformed, often failing to detect any HBs. Increasing the number of epochs led to improved detection of the HB for the D02 part of the combined dataset. Additionally, a higher LR positively impacted the performance metrics. Consequently, the best results were achieved with a model trained for 100 epochs, 256 LSTM units and a LR of $1e^{-3}$. When analyzing the performance of the best LSTM model for the combined data, it could be shown, that the model exhibited the same results for all subjects of the D02 part of the combined dataset. On the contrary, the measurement locations of the D05 dataset differed in performance, where the modalities Arteria carotis communis (normal) and Aorta proximalis (normal) performed worse than all other modalities. However, over all modalities, no model could achieve a mean F1 score (100ms) higher than 0.1, indicating that this model architecture is not

suited for the combined dataset.

5.1.2 UNet approach

To improve the aforementioned prediction of HBs further, a more extensive approach to enhance robustness against respiratory movements and RLBM was used. For this approach, the two datasets were preprocessed as discussed in Section 4.2.1 and 4.2.3. As depicted in Figure 5.3, transforming the power of the RADAR data using the CWT aligns well with the measurements taken by the ECG. A visual inspection of the CWT signal from the D05 dataset reveals that different HSs can be detected within the transformed signal. Notably, the strongest signal corresponds to the ST segment, corresponding to the time between the ventricular de- and repolarization. All

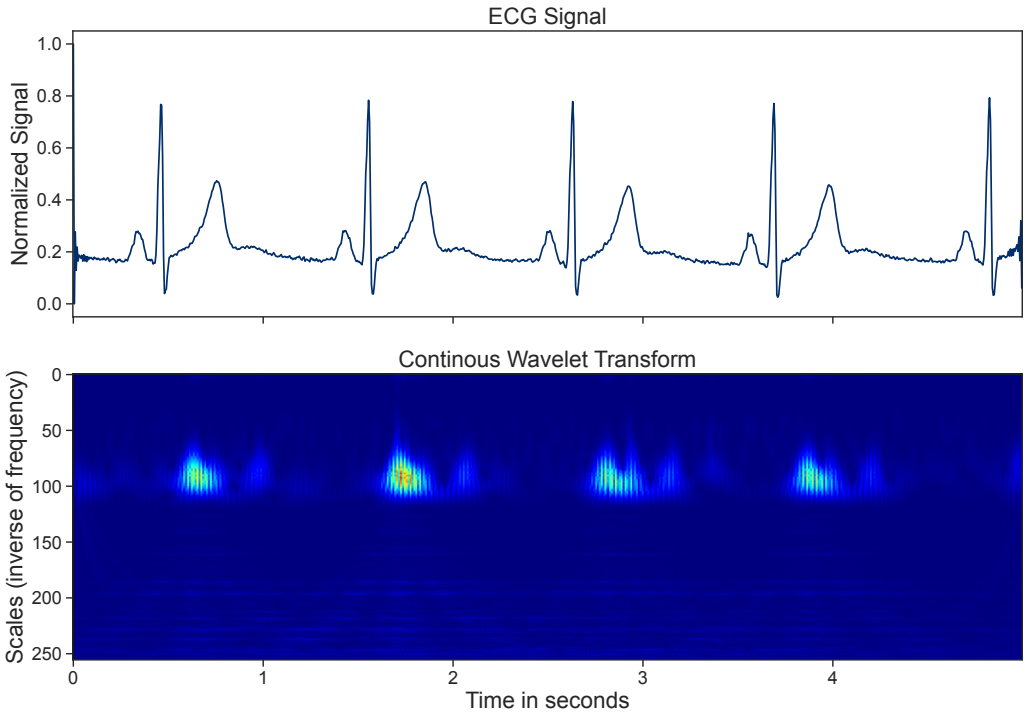

Figure 5.3: **Alignment of ECG Measurements with CWT Transformed Data.** Upper Panel depicts ECG signal. Lower Panel represents the Spectrogram obtained after CWT. Alignment demonstrates the effectiveness of the CWT preprocessing method in accurately detecting HBs. In this figure, only one measuring location for one subject is shown as an example in order to illustrate the preprocessing results.

in all, this indicates that the preprocessing method for the UNet approach is suitable for the accurate detection of HBs. The preprocessed data was used to train a UNet model for both datasets separately and in an combined approach, with the architecture illustrated in Supplementary Figure A.2. To determine the most suitable mother wavelet for detecting HBs, four different mother wavelets were tested. For each wavelet type, two models were trained with LR of $1e^{-4}$ and $1e^{-3}$, respectively, each model underwent 40 epochs of training.

D05

After training the UNet models, the performance metrics were calculated for each subject and location. The comparison of the F1 scores with a tolerance window of 100ms and 50ms is illustrated in Figure 5.4 and shows that the median F1 scores (100ms) for all models are above 0.80 when using a LR of $1e^{-4}$. Furthermore, the mother wavelet morl performed better than the other functions in terms of the mean F1 score (100ms) and exhibited the smallest IQR, indicating its high potential for HB detection. Although the gaus1 wavelet also yielded good results with a median F1 score (100ms) higher than that of morl, its mean F1 score (100ms) was lower, and the IQR was larger, indicating more variation in that score. Additionally, the minimum F1 score (100ms) for gaus1 was lower than that for morl. The shan1-1 wavelet showed promising results with a distribution of F1 scores (100ms) similar to morl, albeit with more outliers. The mexh wavelet had a smaller IQR than gaus1 but the lowest minimum score among all four functions.

In contrast to the LSTM model, an increased LR of $1e^{-3}$ using CWT generally led to improved scores for this dataset, except for the models using the mother wavelet functions gaus1 and mexh, where both the mean and median F1 scores (100ms and 50ms) dropped. However, the predictions of the model using mexh for preprocessing became more stable, as evidenced by a reduced IQR and an increased minimum F1 score (100ms and 50ms) due to the higher LR. The models trained with data preprocessed using morl and shan1-1 wavelets benefited from the reduced LR for some performance metrics. While the mean and median F1 scores (100ms) for the morl-based model did not change, the minimum F1 score (100ms) increased by 24.57%.

When analyzing the more stringent F1 score with a tolerance window of 50ms, the mean scores were lower across all types of mother wavelets used for preprocessing. However, an increased LR positively influenced the performance of models trained with morl or shan1-1 wavelets. For these models, both the median and mean F1 scores (50ms) were higher compared to those trained with a LR of $1e^{-4}$. Additionally, these models exhibited an increased IQR, which resulted in fewer

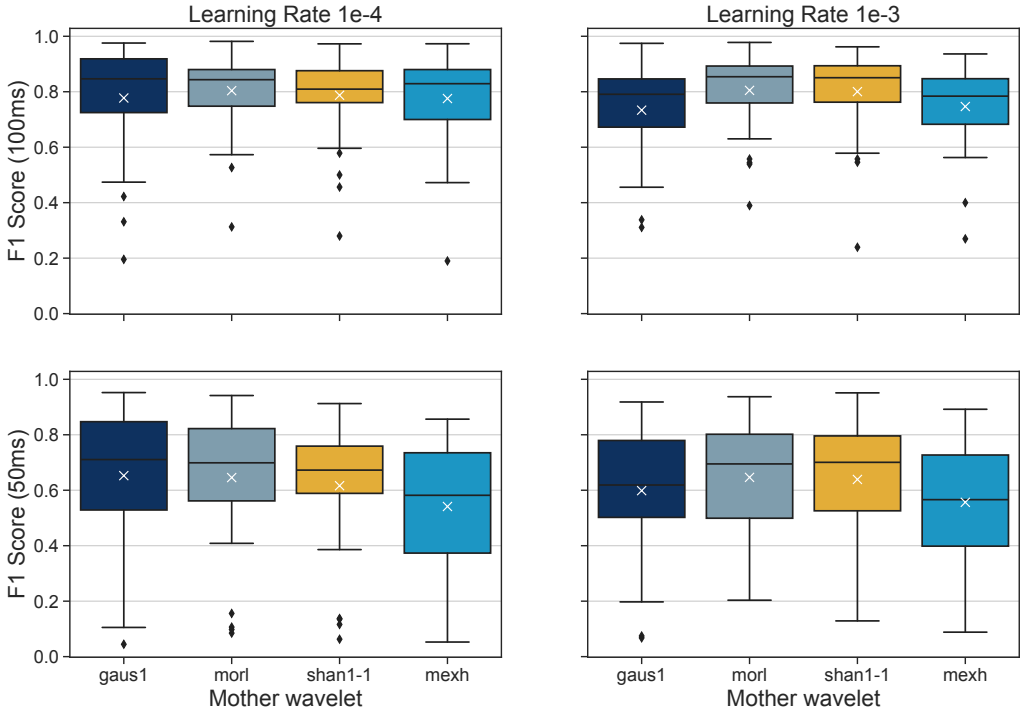

Figure 5.4: **Performance comparison of models across different LRs and tolerance windows for the D05 dataset.** Left Panel: Boxplots of F1 scores with a 100ms tolerance window (top) and 50ms tolerance window (bottom) for models trained with a LR of $1e^{-4}$. Right Panel: Boxplots of F1 scores with a 100ms tolerance window (top) and 50ms tolerance window (bottom) for models trained with a LR of $1e^{-3}$. Each graph includes data for the four mother wavelets used, with white "X" indicating the mean F1 score.

scores being considered outliers, as the lowest F1 score (50ms) also increased. Thus, the higher LR of $1e^{-3}$ positively affected these two models.

Conversely, the increased LR had negative effects on the model trained with gaus1 as the mother wavelet. Here, both the mean and median F1 scores (50ms) decreased, along with a reduced IQR, indicating opposite behavior compared to morl and shan1-1. Lastly, the model using mexh-based CWT was least affected by the change in LR, with only slight increases in the mean F1 score (50ms) and the corresponding IQR. When further analyzing the predictions of the different models, it becomes apparent why the two mother wavelets, morl and shan1-1, performed better overall than the other models. An example prediction from the different models with a LR of $1e^{-4}$ is illustrated in Figure 5.5. This figure shows the variations in HB detection among the models.

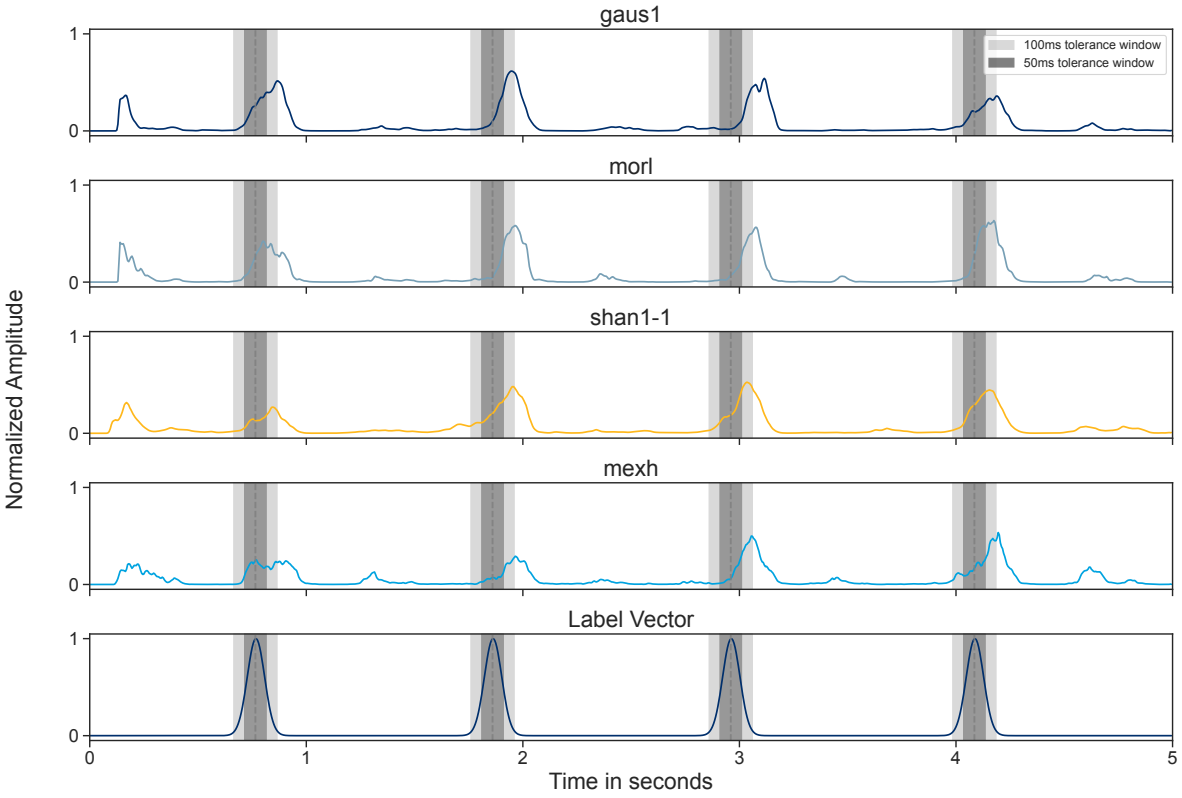

Figure 5.5: **Comparison of predictions and label vector across different mother wavelets of the D05 dataset.** Depiction of the predictions from the four different mother wavelets used. The last plot in the series shows the ground truth label vector, providing a reference for evaluating the performance of the predictions generated by each wavelet. Dotted lines depict the ground truth R-peak at the corresponding time point, dark grey bars refer to the 50ms tolerance window, light grey bars depict the 100ms tolerance window. For simplicity, only one measuring location for one subject is shown.

The mexh-based model not only produces the lowest peaks but also spreads them out, making it difficult for the peak detection algorithm to identify clear peaks. The remaining three model types produce similar predictions, with shan1-1 generating peaks that best match the label vector. However, this example also highlights a common issue across all four model types, where a FP occurs at the beginning of the predictions. This specific type of FP does not impact the overall model performance, as it occurs in the overlapping part of the predicted segment and is excluded when calculating scores.

To compare the results of the D05 dataset between the LSTM approach and the UNet approach

— only the best-performing model from each approach is selected for further comparison. For the LSTM model, this is the one trained for 300 epochs with a LR of $1e^{-4}$ and 256 LSTM units. For the UNet approach, the best-performing model was trained using the morl function as mother wavelet with a LR of $1e^{-3}$, due to its high mean and low IQR in the F1 score (100ms).

When comparing the IHR predictions with the actual measured IHR, as illustrated in Figure 5.6, it is evident that both approaches (Unet and LSTM) yield good results. However, the UNet model tends to underestimate the IHR, particularly in the range of 57 to 65 BPM. Conversely, the LSTM approach struggles more with estimating IHR values in the range of 65 to 68 BPM. Although Figure 5.6 suggests similar performance in terms of the mean IHR, the comparison of F1 scores and correlation metrics reveals differences between the two models.

Performance metrics are detailed in Table 5.3 and Figure 5.2 for the LSTM approach, and in Tables 5.5 and 5.6 for the UNet approach. The comparison reveals that the UNet model achieves higher overall scores, indicating better performance. Unlike the LSTM model, which shows a marked performance difference between measurements taken while subjects held their breath versus breathed normally, the UNet model demonstrates no performance disparity under these conditions. Additionally, despite a performance drop in the stricter F1 score metric, the UNet model shows a smaller variation between subjects. This suggests that the UNet approach is more robust against variations in measurement conditions and individual subjects.

Table 5.5: **Location dependent performance metrics for the UNet approach of the D05 dataset.** F1 score with tolerance window of 50 and 100ms and correlation coefficient between observed and predicted HRs for differing measurement locations. hold = Participant is holding the breath while measuring; normal = Participant is breathing normally.

	F1 score (100ms)				**F1 score (50ms)**				**Correlation**			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Aorta distalis (hold)	0.74	0.10	0.63	0.88	0.41	0.18	0.22	0.65	0.55	0.13	0.35	0.68
Aorta distalis (normal)	0.71	0.23	0.39	0.92	0.42	0.22	0.20	0.68	0.52	0.23	0.22	0.77
Aorta medialis (hold)	0.81	0.07	0.71	0.89	0.74	0.06	0.65	0.83	0.72	0.05	0.67	0.80
Aorta medialis (normal)	0.83	0.18	0.54	0.98	0.74	0.23	0.42	0.93	0.73	0.18	0.45	0.91
Aorta proximalis (hold)	0.84	0.04	0.80	0.88	0.76	0.11	0.61	0.85	0.74	0.06	0.67	0.80
Aorta proximalis (normal)	0.86	0.10	0.75	0.98	0.71	0.17	0.45	0.91	0.73	0.12	0.56	0.84
Arteria carotis communis (normal)	0.83	0.15	0.56	0.92	0.73	0.16	0.46	0.89	0.73	0.13	0.53	0.87

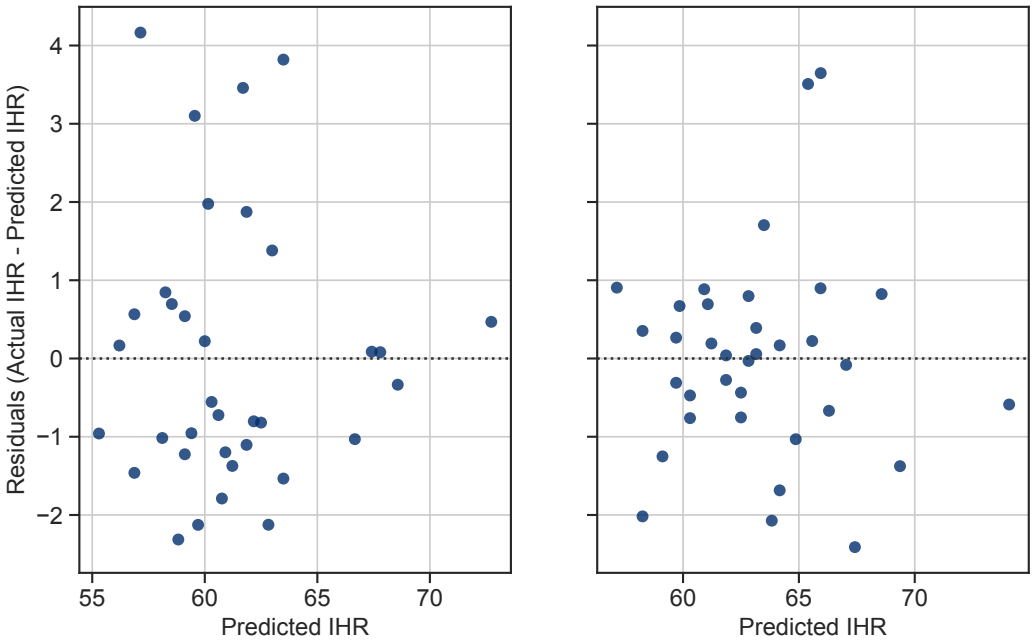

Figure 5.6: **Comparison of predicted IHR and corresponding residuals for both approaches examined in the D05 dataset.** Left panel represents scatterplot comparing predicted IHR and corresponding residuals for the UNet approach, right panel for the LSTM approach.

Table 5.6: **Subject-dependent performance metrics for the UNet approach of the D05 dataset.** F1 score with tolerance window of 50 and 100ms and correlation coefficient between observed and predicted HBs for differing measurement locations. hold = Participant is holding the breath while measuring; normal = Participant is breathing normally.

	F1 score (100ms)				**F1 score (50ms)**				**Correlation**			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
VP 01	0.79	0.19	0.39	0.98	0.63	0.29	0.20	0.89	0.67	0.23	0.22	0.87
VP 03	0.81	0.13	0.54	0.94	0.71	0.17	0.42	0.91	0.70	0.13	0.45	0.83
VP 11	0.77	0.13	0.56	0.88	0.62	0.16	0.36	0.84	0.67	0.10	0.53	0.79
VP 15	0.77	0.14	0.54	0.97	0.61	0.27	0.21	0.94	0.62	0.21	0.33	0.91
VP 18	0.89	0.07	0.75	0.98	0.67	0.19	0.43	0.91	0.72	0.10	0.56	0.83

Table 5.7: **F1 score (100ms) statistics for the two best performing mother wavelets and all phases of the D02 dataset.** Descriptive parameters separated by mother wavelet and phase. ei = emotion induction, lat = latency.

Mother Wavelet	Phase	Mean	SD	Min	Max
mexh	coping	0.33	0.23	0.03	0.83
	ei	0.31	0.18	0.06	0.78
	lat	0.32	0.25	0.04	0.89
	training	0.31	0.18	0.09	0.83
morl	coping	0.36	0.16	0.10	0.74
	ei	0.35	0.13	0.13	0.77
	lat	0.38	0.20	0.06	0.81
	training	0.36	0.14	0.12	0.72

D02

Multiple models were trained on preprocessed data from the D02 dataset. Each model utilized one of four mother wavelets and different LRs of either $1e^{-3}$ or $1e^{-4}$. It was quickly evident that a LR of $1e^{-3}$ was too high for this dataset, as all models failed to predict any HBs regardless of the chosen mother wavelet. When the LR was adjusted to $1e^{-4}$ for the D02 dataset, the model performance improved, and the models were able to detect HBs (Figure 5.7). Due to the increased noise in the D02 dataset, the metrics showed worse performance compared to the D05 dataset. When analyzing the performance of different mother wavelets, morl and mexh demonstrated the best performance in terms of the F1 score (100ms) and F1 score (50ms). This contrasts with the D05 dataset, where the mexh yielded the poorest results for most metrics. Specifically, mexh outperformed other wavelets in predicting the IHR in the D02 dataset, highlighting the impact of the selected mother wavelet on model predictions and performance. All four models exhibited several outliers in metrics such as the F1 score (100ms), F1 score (50ms), and correlation. However, these outliers were towards increased performance, differing from the behavior observed in the D05 models. To further investigate the source of these outliers, predictions were analyzed by location and phase. Table 5.7 shows the performance differences across the four phases used in this dataset, revealing that, for the two best performing models, phase differences were smaller than those between the used mother wavelets for the two best-performing models. This suggests that the wavelet function influences performance more than the phase. Moreover, the influence of the subject on performance was found to be greater than that of the phase, as illustrated in Figure 5.8. These large differences between subjects could be attributed to the measurement settings or the participants themselves, as some may move less or have more easily detectable HBs.

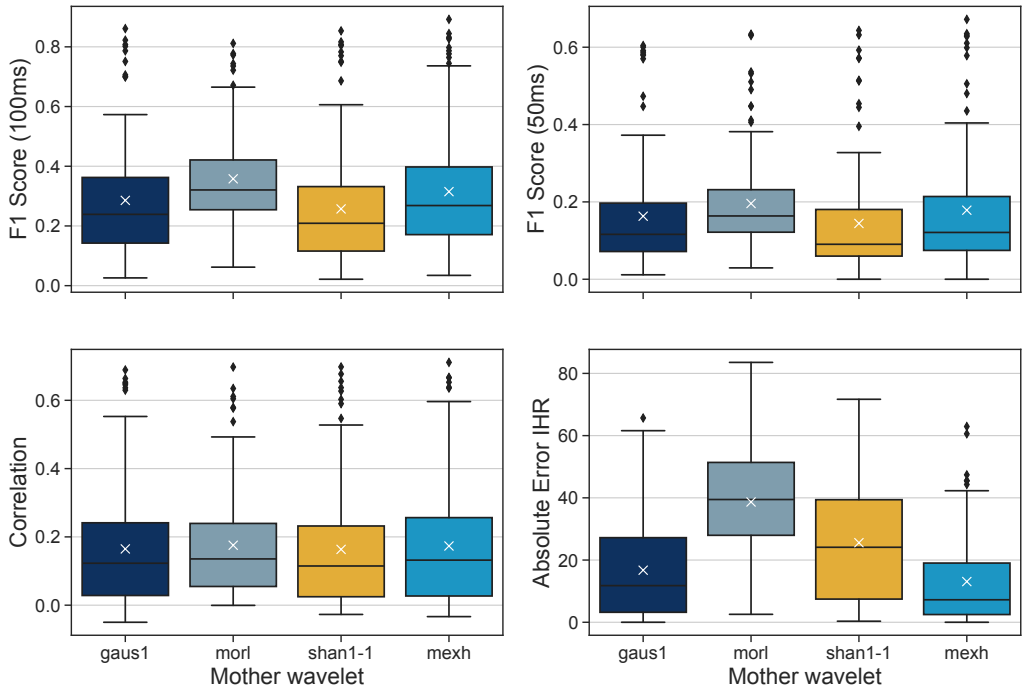

Figure 5.7: **Performance metrics for the D02 dataset with LR 1e^{-4}.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for IHR. White "X" indicates the mean for the corresponding metric.

Combined data

Similar to the LSTM approach, the combined dataset was also used to train the UNet models. Eight models were trained, each using one of four different mother wavelets and a LR of either 1e^{-4} or 1e^{-3}. Results of these eight models indicated that a LR of 1e^{-4} consistently yielded better performance metrics, so this chapter will focus on these.

Among the models, the one trained with the morl mother wavelet performed best overall. However, the performance was dependent on the origin of the test data. For data from the D05 dataset of the combined dataset, the morl model had the worst performance regarding the F1 score (100ms), with a mean of 0.56. In contrast, the gaus1 model, which performed best for this dataset, achieved a mean F1 score (100ms) of 0.72. This pattern was similar for the F1 score

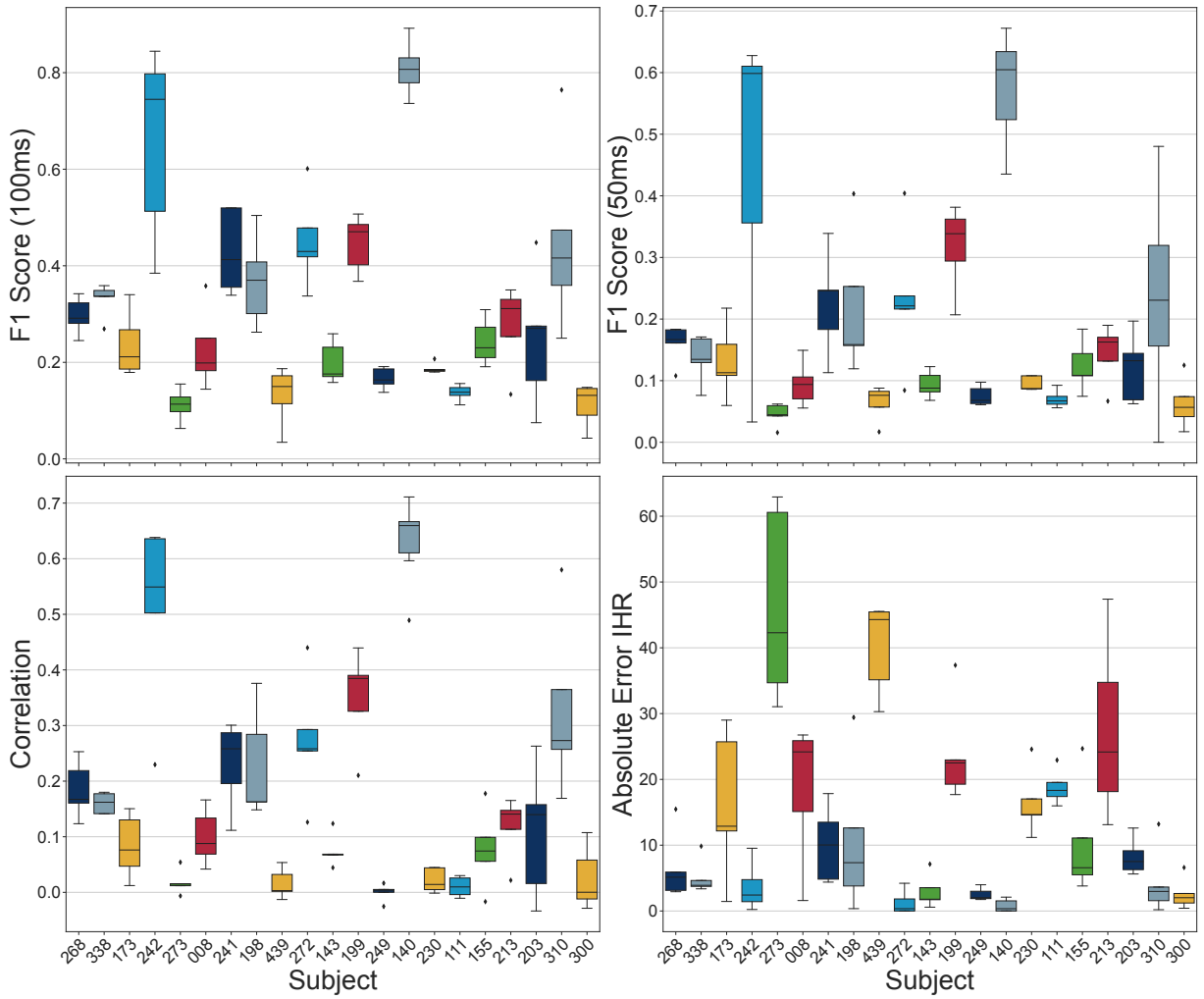

Figure 5.8: **Subject dependent performance metrics for the UNet approach of the D02 dataset.** F1 score with tolerance window of 50 and 100ms, correlation coefficient and absolute error between observed and predicted HBs for all subjects.

(50ms) and the absolute error of the IHR. For correlation, the gaus1 and mexh models outperformed the others, with mean correlations of 0.60 and 0.61 respectively, while other models scored 0.56. Figure 5.9 compares the best performing model trained solely on the D05 dataset with the model trained on the combined dataset. This Figure illustrates that, for the F1 score (100ms) and correlation, the model did not benefit from the increased dataset size. The only improvement was seen in the F1 score (50ms), with the mean, median, and first quartile being higher, although the maximum was lower. Figure 5.10 illustrates the distribution of performance metrics for the two best-performing models: one trained with the combined dataset and one trained exclusively on

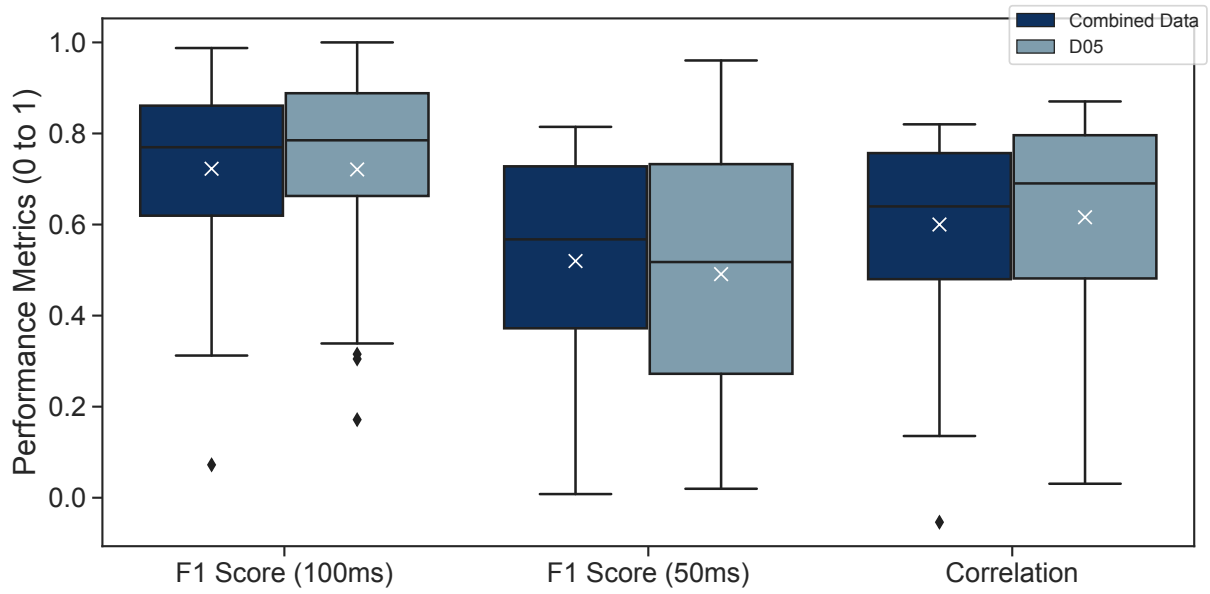

Figure 5.9: **Performance comparison of best-performing models for the D05 and combined dataset.** Boxplots depicting the F1 score for both tolerance windows (100ms and 50ms) as well as correlation between predicted values and ground truth labels for the D05 and combined dataset. Best performing model for D05: CWT with UNet, mother wavelet morl and LR of $1e^{-3}$. Best-performing model combined data: CWT with UNet, morl as mother wavelet and LR of $1e^{-4}$. White "X" represents the corresponding mean performance metric.

the D02 dataset. This figure shows that the model trained on the combined dataset performed worse than the one trained solely on the D02 dataset. Since this overall approach did not yield improvements, performance based on location and subject will not be discussed.

Table 5.8: **Combined performance metrics** for Correlation, F1 score (100ms), and F1 score (50ms).

Metric	Origin	Mean	SD	Min	Max
F1 score (100ms)	D02	0.2847	0.1945	0.0346	0.9483
	D05	0.5608	0.2634	0.0567	0.9412
F1 score (50ms)	D02	0.1638	0.1547	0.0000	0.8534
	D05	0.4216	0.2680	0.0253	0.8863
Correlation	D02	0.1771	0.1912	-0.0476	0.8111
	D05	0.6064	0.1610	0.2265	0.8497

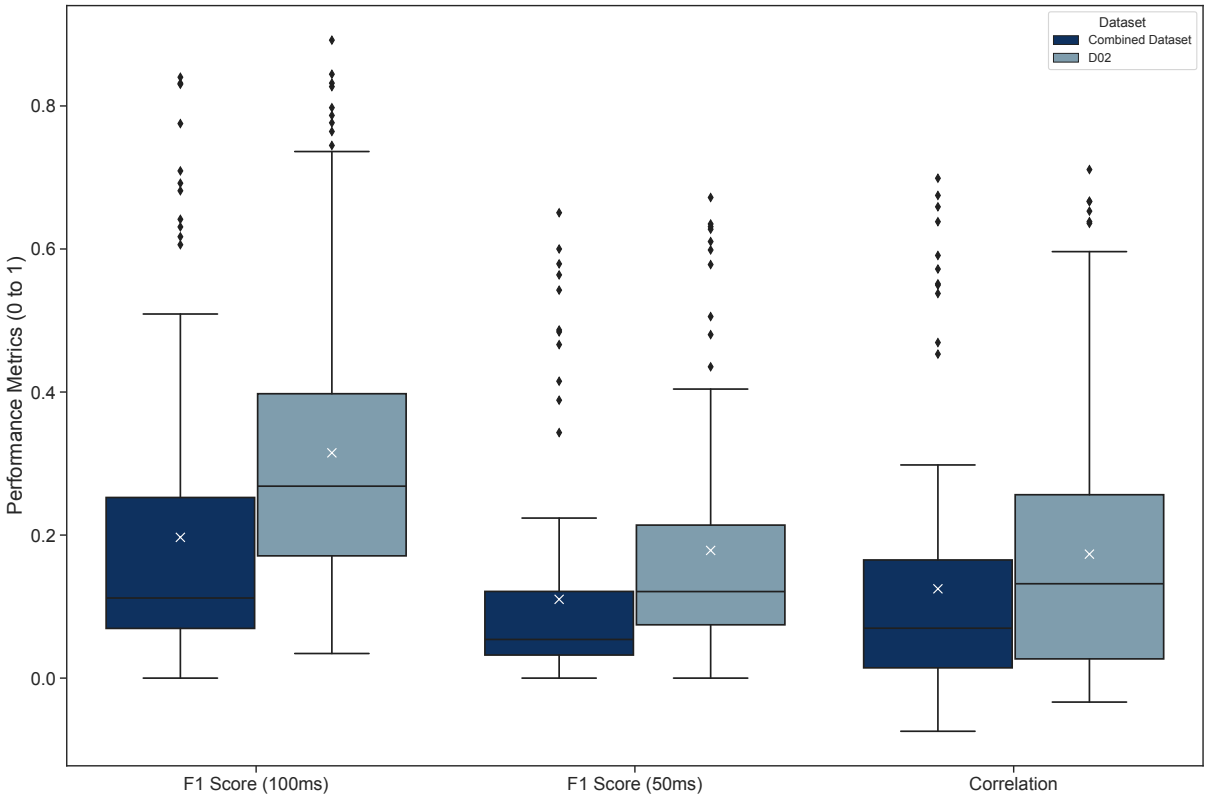

Figure 5.10: **Performance comparison of best-performing models for the D02 and combined dataset.** Boxplots depicting the F1 score for both tolerance windows (100ms and 50ms) as well as correlation between predicted values and ground truth labels for the D02 and combined dataset. Best performing model for D02: CWT with UNet, mother wavelet morl and LR of $1e^{-4}$. Best-performing model combined data: CWT with UNet, morl as mother wavelet and LR of $1e^{-4}$. White "X" represents the corresponding mean performance metric.

5.2 General discussion and Limitations

In the thesis presented here, the performance of processing data using either a LSTM model or CWT in combination with a UNet approach was investigated. The datasets used were D02 and D05, each exhibiting differing levels of noise and artifacts and a combined version of both.

The UNet model demonstrated better performance compared to the LSTM approach, especially when using noisy data. While the LSTM model failed to detect HBs in most of the cases, the UNet model was still able to detect the HBs, highlighting the worth of a more computational intensive preprocessing and training. Even though the UNet model handled the detection of the HBs in the presence of noisier data better than the LSTM models, it still showed poor performance (Figure 5.7).

When compared to existing literature, the UNet model presented here underperformed. Lu et al. [Lu23] achieved an F1 score (with a tolerance window of 75ms) of 0.98 for HR detection, using a bidirectional gated recurrent unit network on a similar dataset as the D05 dataset, where the participants were lying down or in an tilt-up position. With the approach presented here, only a mean F1 score of 0.80 (100ms) and 0.64 (50ms) could be achieved.

Chowdhury et al. [Cho24] compared different model architectures, including UNet, for the same dataset as Lu et al. [Lu23]. In their approach, a MultiResLinkNet architecture performed better than UNet, achieving an F1 score of 0.94. The MultiResLinkNet architecture has a similar structure as the UNet architecture, as it contains an encoding and expanding path. However, it replaces the convolutional blocks in the UNet with Multi-Residual Blocks, containing several aligned CNN blocks that work in parallel. Each block processes the input, passes its output to the next block, and to a concatenation block. The concatenation block combines the outputs from all the CNN blocks and adds them with the output of a 1x1 CNN block to increase details and information. Additionally, in the decoding path, MultiResLinkNet uses Residual Paths instead of direct skip connections. These Residual Paths involve various CNN blocks that further process the output, refining the reconstruction of the image. Even though Chowdhury et al. [Cho24] proposed a better performance for the MultiResLinkNet model compared to the UNet model, the authors did not mention the size of their tolerance window used, complicating the comparison to the results presented here. Nevertheless, Chowdhury et al. [Cho24] achieved a lower F1 score for the modality where subjects holding there breath. The opposite trend was observed for the LSTM approach used for the D05 dataset presented here, exhibiting better performance if the subjects

held their breath (Figure 5.2). On the contrary, the UNet approach shown here does not reveal strong difference between the modalities of breathing normally vs. holding the breath (Table 5.5), indicating an increased robustness against artifacts caused by breathing.

Tang et al. [Tan23] used a combined approach, using a less noisy dataset (data collected using a digital stethoscope) with a more noisy RADAR dataset. The authors used a Dual-UNet architecture, where both UNets used the spectrogram computed using STFT, but one processed the I and the other one the Q component of the RADAR signal. The authors could show, that this setup led to an overall increased performance and robustness against background noise. Comparing this to the combined approach in the thesis presented here, the LSTM model did not profit from the combined dataset, and the metrics still show a low performance. The UNet models in this thesis, on the other hand, showed a contrary behavior to the one observed by Tang et al. [Tan23], as their performance became worse when a combined dataset was used (Figures 5.9 5.10 and Table 5.8).

Chen et al. [Che21] used a different approach for the detection of HBs in the presence of RLBMs. The authors utilized deep contrastive learning using FMCW RADAR data from different modalities such as walking on a treadmill, leg shaking or exercising. With their approach, the authors could achieve a consistent error < 4% for all modalities tested. In addition, Chen et al. [Che21] also detected a differences in performance between subjects. A comparison to the results presented here might not be perfectly accurate due to differing RADAR data and models used. However, the UNet model used in the thesis here had several outliers throughout the different phases with relative errors of the IHR of 10% in 50% of the predictions, suggesting that the UNet model used on CW RADAR data here performed worse than the one used by Chen et al. [Che21]. In addition, also a variation in performance depending on the subject could be observed in both datasets (Table 5.6 and Figure 5.8) presented here, comparable to the results of Chen et al. [Che21].

Taking all of the discussed results and comparisons into account, there are several limitations that should be considered when interpreting the aforementioned findings. The scope of the thesis presented here was to focus on analyzing CW RADAR data using a LSTM and a UNet based architecture. Evaluating other models might have identified a better suitable architecture for the datasets analyzed here. A more detailed hyperparameter tuning could also improve the overall performance. Furthermore, only four mother wavelets have been tested in the thesis here. A more expanded collection of mother wavelets might further increase the models' performance, due to the shown influence of the mother wavelet type. Lastly, the two datasets, D02 and D05, were col-

lected under different experimental conditions and for different subjects. These setup differences limit the comparability of the results between the two datasets and suggests a standardized data acquisition protocol.

Chapter 6

Conclusion and outlook

In the scope of this thesis, two datasets with varying levels of noise and artifacts were processed and analyzed in order to detect HBs from CW RADAR data. The best-performing approach for the D02 dataset could be identified with CWT in combination with UNet, morl as mother wavelet, and a LR of $1e^{-4}$ with a mean F1 score (100ms) of 0.36. In contrast, the D05 dataset performed best under the following conditions: CWT in conjunction with UNet, morl as mother wavelet and a LR of $1e^{-3}$ showing a mean F1 score (100ms) of 0.80. For the combined dataset, again CWT together with UNet, morl as mother wavelet and a LR of $1e^{-4}$ was determined as the best-performing approach with a mean F1 score (100ms) of 0.35. All in all, CWT together with UNet and morl as mother wavelet always performed best, independent of the dataset used. Only the LR seems to be dataset dependent, suggesting a lower learning rate is more suitable for noisy data. Even though the highest performance across all combinations and datasets did not reach an F1 score >0.85 or a correlation of >0.7 between the prediction and the ground truth, these findings provide insights into the performance of the described models for the analyzed datasets and highlight areas for future projects in order to improve contactless HR detection and analyses methods.

Future perspectives concerning the used model would be beneficial. For example, architectures like Dual-UNet or MultiResLinkNet would be a good alternative since they have already been reported to outperform the UNet architecture used [Tan23; Cho24]. Finally, in order to improve the comparability between noisy and non-noisy datasets, it might be helpful to collect CW RADAR data from the same subjects for both conditions (with body movement and without). This might also enhance the evaluation of a model's noise reduction capability. In summary, further investigation of HB detection using CW RADAR data and their analyses are needed in order to make it a viable contactless alternative to the current ECG standard.

List of Figures

List of Tables

Bibliography

[Abd21] ABM Abdullah. *ECG in medical practice*. Jaypee Brothers Medical Publishers, 2021. ISBN: 978-93-89587-25-8. URL: https://books.google.com/books?id=GMV0EAAAQBAJ.

[Alb24] Nils C. Albrecht, Dominik Langer, Daniel Krauss, Robert Richer, Luca Abel, Bjoern M. Eskofier, Nicolas Rohleder, and Alexander Koelpin. "EmRad: Ubiquitous Vital Sign Sensing using Compact Continuous-Wave Radars". In: *IEEE Open Journal of Engineering in Medicine and Biology* (2024), pp. 1–10. DOI: 10.1109/OJEMB.2024.3420241.

[Ani09] Laura Anitori, Ardjan de Jong, and Frans Nennie. "FMCW radar for life-sign detection". In: *2009 IEEE Radar Conference*. 2009, pp. 1–6. DOI: 10.1109/RADAR.2009.4976934.

[Bjo98] John E Bjornholt and Terry B Wilson. "FM-CW radar for imaging applications". In: *Radar Processing, Technology, and Applications III*. Vol. 3462. SPIE. 1998, pp. 89–98. DOI: https://doi.org/10.1117/12.326736.

[Car03] J.L.A. Carvalho, A.F. Rocha, L.F. Junqueira, J.S. Neto, I. Santos, and F.A.O. Nascimento. "A tool for time-frequency analysis of heart rate variability". In: *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*. Vol. 3. 2003, 2574–2577 Vol.3. DOI: 10.1109/IEMBS.2003.1280441.

[Cha15] Tamal Chakrabarti, Sourav Saha, Sathi Roy, and Ishita Chel. "Phonocardiogram signal analysis - practices, trends and challenges: A critical review". In: *2015 International Conference and Workshop on Computing and Communication (IEMCON)*. 2015, pp. 1–4. DOI: 10.1109/IEMCON.2015.7344426.

[Che09] Tianhua Chen, Li-qun Han, Su-xia Xing, Hai-ran Wang, and Kun-peng Wang. "STFT-Based Comparative Studies of Heart Sound Signals". In: *2009 3rd International Con-*

ference on Bioinformatics and Biomedical Engineering. 2009, pp. 1–4. DOI: 10.1109/ICBBE.2009.5163425.

[Che21] Zhe Chen, Tianyue Zheng, Chao Cai, and Jun Luo. "MoVi-Fi: Motion-robust vital signs waveform recovery via deep interpreted RF sensing". In: *Proceedings of the 27th annual international conference on mobile computing and networking*. 2021, pp. 392–405. DOI: 10.1145/3447993.3483251.

[Cho24] Farhana Ahmed Chowdhury, Md Kamal Hosain, Md Sakib Bin Islam, Md Shafayet Hossain, Promit Basak, Sakib Mahmud, M Murugappan, and Muhammad EH Chowdhury. "ECG waveform generation from radar signals: A deep learning perspective". In: *Computers in Biology and Medicine* 176 (2024), p. 108555. DOI: 10.1016/j.compbiomed.2024.108555.

[De 12] Antoni Bayes De Luna. *Clinical electrocardiography: a textbook*. John Wiley & Sons, 2012. ISBN: 9781118392010. URL: https://books.google.com/books?id=oEo4IYa_tBIC&oi.

[Deb21] SM Debbal and C Hamza. "Heart sounds analysis using the three wavelet transform versions the continuous wavelet transform (CWT), the discrete wavelet transform (DWT) and the wavelet packet transforms (PWT)". In: *Journal of Cardiology Interventions* 1.1 (2021). DOI: 03.2020/1.1003.

[Dje00] A. Djebbari and F. Bereksi Reguig. "Short-time Fourier transform analysis of the phonocardiogram signal". In: *ICECS 2000. 7th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.00EX445)*. Vol. 2. 2000, 844–847 vol.2. DOI: 10.1109/ICECS.2000.913008.

[Du20] Getao Du, Xu Cao, Jimin Liang, Xueli Chen, and Yonghua Zhan. "Medical Image Segmentation based on U-Net: A Review." In: *Journal of Imaging Science & Technology* 64.2 (2020). DOI: 10.2352/J.ImagingSci.Technol.2020.64.2.020508.

[El 16] Oussama El B'charri, Rachid Latif, Wissam Jenkal, and Abdenbi Abenaou. "The ECG signal compression using an efficient algorithm based on the DWT". In: *International Journal of Advanced Computer Science and Applications* 7.3 (2016). DOI: 10.14569/IJACSA.2016.070325.

[El 23] Amal El Abbaoui, David Sodoyer, and Fouzia Elbahhar. "Contactless heart and respiration rates estimation and classification of driver physiological states using CW radar and temporal neural networks". In: *Sensors* 23.23 (2023), p. 9457. DOI: 10.3390/s23239457.

[Ele96] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. "Heart rate variability: standards of measurement, physiological interpretation, and clinical use". In: *Circulation* 93.5 (1996), pp. 1043–1065. DOI: 10.1161/01.CIR.93.5.1043.

[Far16] Irmalia Suryani Faradisa, Dimas Okky Anggriawan, Tri Arief Sardjono, and Mauridhi Heri Purnomo. "Identification of phonocardiogram signal based on STFT and Marquart Lavenberg Backpropagation". In: *2016 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. 2016, pp. 25–30. DOI: 10.1109/ISITIA.2016.7828628.

[Fed24] World Heart Federation. *Deaths from Cardiovascular Disease Surged 60% Globally Over the Last 30 Years: Report*. Accessed: 2024-08-03. 2024. URL: https://world-heart-federation.org/news/deaths-from-cardiovascular-disease-surged-60-globally-over-the-last-30-years-report/.

[Fel90] John M. Felner. "The First Heart Sound". In: *Clinical Methods: The History, Physical, and Laboratory Examinations*. Ed. by H. Kenneth Walker, W. Dallas Hall, and J. Willis Hurst. 3rd. Boston: Butterworths, 1990. Chap. 22. URL: https://www.ncbi.nlm.nih.gov/books/NBK333/.

[Gai23] Oleg Gaidai, Yu Cao, and Stas Loginov. "Global cardiovascular diseases death rate prediction". In: *Current problems in cardiology* 48.5 (2023), p. 101622. DOI: 10.1016/j.cpcardiol.2023.101622.

[Gar14] Tomas B Garcia. *Introduction to 12-lead ECG: the art of interpretation*. Jones & Bartlett Publishers, 2014. ISBN: 978-1-284-04088-3. URL: https://books.google.com/books?id=wbXpAwAAQBAJ.

[Ger08] Marc Gertsch. *Das EKG: auf einen Blick und im Detail*. Springer Science & Business Media, 2008. ISBN: 978-3-540-79121-8. URL: https://books.google.com/books?id=kJjCWQqUVqAC.

[Gio19] Noemi Giordano and Marco Knaflitz. "A novel method for measuring the timing of heart sound components through digital phonocardiography". In: *Sensors* 19.8 (2019), p. 1868. DOI: 10.3390/s19081868.

[Gre17] Klaus Greff, Rupesh K. Srivastava, Jan Koutník, Bas R. Steunebrink, and Jürgen Schmidhuber. "LSTM: A Search Space Odyssey". In: *IEEE Transactions on Neural Networks and Learning Systems* 28.10 (2017), pp. 2222–2232. DOI: 10.1109/TNNLS.2016.2582924.

[Han22] Thanh Han-Trong and Hoang Nguyen Viet. "An efficient heart rate measurement system using medical radar and LSTM neural Network". In: *Journal of Electrical and Computer Engineering* 2022.1 (2022), p. 4696163. DOI: 10.1155/2022/4696163.

[Har20] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

[Heg13] Veena N Hegde, Ravishankar Deekshit, and PS Satyanarayana. "A review on ECG signal processing and HRV analysis". In: *Journal of Medical Imaging and Health Informatics* 3.2 (2013), pp. 270–279. DOI: 10.1166/jmihi.2013.1163.

[Hou19] Andrew Houghton. *Making sense of the ECG: a hands-on guide*. CRC press, 2019. ISBN: 978-0-429-19908-0. DOI: 10.1201/9780429199080.

[Hui97] Liang Huiying, L. Sakari, and H. Iiro. "A heart sound segmentation algorithm using wavelet decomposition and reconstruction". In: *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No.97CH36136)*. Vol. 4. 1997, 1630–1633 vol.4. DOI: 10.1109/IEMBS.1997.757028.

[Ism18] Shahid Ismail, Imran Siddiqi, and Usman Akram. "Localization and classification of heart beats in phonocardiography signals—a comprehensive review". In: *EURASIP Journal on Advances in Signal Processing* 2018.1 (2018), pp. 1–27. DOI: 10.1186/s13634-018-0545-9.

[Iwa21] Yuki Iwata, Han Trong Thanh, Guanghao Sun, and Koichiro Ishibashi. "High accuracy heartbeat detection from CW-Doppler radar using singular value decomposition and matched filter". In: *Sensors* 21.11 (2021), p. 3588. DOI: 10.3390/s21113588.

[Jin12] Benjamin Jin, Heike Wulff, Jonathan Widdicombe, Jie Zheng, Donald M Bers, and Jose L Puglisi. "A simple device to illustrate the Einthoven triangle". In: *Biophysical Journal* 102.3 (2012), 211a. DOI: 10.1016/j.bpj.2011.11.1153.

[Jol08] Harry M Jol. *Ground penetrating radar theory and applications*. elsevier, 2008. ISBN: 978-0-444-53348-7. URL: https://books.google.de/books?id=y__uIi-5RvgC.

[Kai20] Naira Kaieski, Cristiano Andre Da Costa, Rodrigo da Rosa Righi, Priscila Schmidt Lora, and Bjoern Eskofier. "Application of artificial intelligence methods in vital signs analysis of hospitalized patients: A systematic literature review". In: *Applied Soft Computing* 96 (2020), p. 106612. DOI: 10.1016/j.asoc.2020.106612.

[Kat10] Arnold M Katz. *Physiology of the Heart*. Lippincott Williams & Wilkins, 2010. ISBN: 978-1-60831-171-2. URL: https://books.google.de/books?id=24CcilHdzC4C.

[Keb20] Mamady Kebe, Rida Gadhafi, Baker Mohammad, Mihai Sanduleanu, Hani Saleh, and Mahmoud Al-Qutayri. "Human vital signs detection methods and potential using radars: A review". In: *Sensors* 20.5 (2020), p. 1454. DOI: 10.3390/s20051454.

[KeiND] M. Keinert, L. Schindler-Gmelch, L. H. Rupp, M. Sadeghi, K. Capito, M. Hager, F. Rahimi, R. Richer, B. Egger, B. M. Eskofier, and M. Berking. "Facing Depression: Evaluating the Efficacy of the EmpkinS-EKSpression Reappraisal Training augmented with Facial Expressions – Protocol of a Randomized Controlled Trial". unpublished manuscript. N.D.

[Khu14] Kirti Khunti. "Accurate interpretation of the 12-lead ECG electrode placement: A systematic review". In: *Health Education Journal* 73.5 (2014), pp. 610–623. DOI: 10.1177/0017896912472328.

[Kra24] Daniel Krauss, Lukas Engel, Tabea Ott, Johanna Bräunig, Robert Richer, Markus Gambietz, Nils Albrecht, Eva M. Hille, Ingrid Ullmann, Matthias Braun, Peter Dabrock, Alexander Kölpin, Anne D. Koelewijn, Bjoern M. Eskofier, and Martin Vossiek. "A Review and Tutorial on Machine Learning-Enabled Radar-Based Biomedical Monitoring". In: *IEEE Open Journal of Engineering in Medicine and Biology* (2024), pp. 1–22. DOI: 10.1109/OJEMB.2024.3397208.

[Küd23] Arne Küderle, Robert Richer, Raul C. Sîmpetru, and Bjoern M. Eskofier. "tpcp: Tiny Pipelines for Complex Problems - A set of framework independent helpers for algorithms development and evaluation". In: *Journal of Open Source Software* 8.82 (2023), p. 4953. DOI: 10.21105/joss.04953. URL: https://doi.org/10.21105/joss.04953.

[Kun00] Mahantapas Kundu, Mita Nasipuri, and Dipak Kumar Basu. "Knowledge-based ECG interpretation: a critical review". In: *Pattern Recognition* 33.3 (2000), pp. 351–373. DOI: 10.1016/S0031-3203(99)00065-5.

[Lee19] Gregory R. Lee, Ralf Gommers, Filip Waselewski, Kai Wohlfahrt, and Aaron O'Leary. "PyWavelets: A Python package for wavelet analysis". In: *Journal of Open Source Software* 4.36 (2019), p. 1237. DOI: 10.21105/joss.01237.

[Li13] Changzhi Li, Victor M. Lubecke, Olga Boric-Lubecke, and Jenshan Lin. "A Review on Recent Advances in Doppler Radar Sensors for Noncontact Healthcare Monitoring". In: *IEEE Transactions on Microwave Theory and Techniques* 61.5 (2013), pp. 2046–2060. DOI: 10.1109/TMTT.2013.2256924.

[Li17] Meiyu Li and Jenshan Lin. "Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8-GHz CW Doppler radar". In: *IEEE Transactions on Microwave Theory and Techniques* 66.1 (2017), pp. 568–576.

[Lu23] Hui Lu, Markus Heyder, Marvin Wenzel, Nils C. Albrecht, Dominik Langer, and Alexander Koelpin. "Accurate Heart Beat Detection with Doppler Radar using Bidirectional GRU Network". In: *2023 IEEE Radio and Wireless Symposium (RWS)*. 2023, pp. 52–54. DOI: 10.1109/RWS55624.2023.10046202.

[Lut19] Atul Luthra. *ECG made easy*. Jaypee Brothers Medical Publishers, 2019. ISBN: 978-93-89188-72-1. URL: https://books.google.de/books?id=dnkxEAAAQBAJ.

[Maa23] Lieve E van der Maarel, Alex V Postma, and Vincent M Christoffels. "Genetics of sinoatrial node function and heart rate disorders". In: *Disease models & mechanisms* 16.5 (2023). DOI: 10.1242/dmm.050101.

[Mal20] Nebojša Malešević, Vladimir Petrović, Minja Belić, Christian Antfolk, Veljko Mihajlović, and Milica Janković. "Contactless real-time heartbeat detection via 24 GHz continuous-wave Doppler radar using artificial neural networks". In: *Sensors* 20.8 (2020), p. 2351. DOI: 10.3390/s20082351.

[Man96] Pascale Mansier, Jean Clairambault, Nathalie Charlotte, Claire Médigue, Christophe Vermeiren, Gilles LePape, François Carré, Athanassia Gounaropoulou, and Bernard Swynghedauw. "Linear and non-linear analyses of heart rate variability: a minireview". In: *Cardiovascular research* 31.3 (1996), pp. 371–379. DOI: 10.1016/S0008-6363(96)00009-0.

[Mar15] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. *TensorFlow: Large-Scale Machine Learning*

on Heterogeneous Systems. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[Mei18] Andries Meintjes, Andrew Lowe, and Malcolm Legget. "Fundamental Heart Sound Classification using the Continuous Wavelet Transform and Convolutional Neural Networks". In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2018, pp. 409–412. DOI: 10.1109/EMBC.2018.8512284.

[Mil22] MG Manisha Milani, Pg Emeroylariffion Abas, and Liyanage C De Silva. "A critical review of heart sound signal segmentation algorithms". In: *Smart Health* 24 (2022), p. 100283. DOI: 10.1016/j.smhl.2022.100283.

[Mog15] Eriko Mogi and Tomoaki Ohtsuki. "Heartbeat detection with Doppler sensor using adaptive scale factor selection on learning". In: *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. 2015, pp. 2166–2170. DOI: 10.1109/PIMRC.2015.7343656.

[Mog17] Eriko Mogi and Tomoaki Ohtsuki. "Heartbeat detection with Doppler radar based on spectrogram". In: *2017 IEEE International Conference on Communications (ICC)*. 2017, pp. 1–6. DOI: 10.1109/ICC.2017.7996378.

[Mur22] Keito Murata, Daichi Kitamura, Ryo Saito, and Daichi Ueki. "Heart Rate Estimation of Car Driver Using Radar Sensors and Blind Source Separation". In: *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 2022, pp. 1156–1163. DOI: 10.23919/APSIPAASC55919.2022.9980190.

[Nik12] T Nikolaidou, OV Aslanidi, H Zhang, and IR Efimov. "Structure–function relationship in the sinus and atrioventricular nodes". In: *Pediatric cardiology* 33 (2012), pp. 890–899. DOI: 10.1007/s00246-012-0249-0.

[Oes24] Marie Oesten. "Radar-Based Investigation of Pulse Wave and Heart Sound Propagation". Master's thesis. Erlangen, Germany: Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2024.

[Opi04] Lionel H Opie. *Heart physiology: from cell to circulation*. Lippincott Williams & Wilkins, 2004. ISBN: 978-0-7817-4278-8. URL: https://books.google.de/books?id=CPVSg69CPMsC.

[Paa01] Larry D Paarmann. *Design and analysis of analog filters: a signal processing perspective*. Vol. 617. Springer Science & Business Media, 2001. ISBN: 0-306-48012-3. DOI: https://books.google.de/books?id=vGWN49squBAC.

[Pod14] Prajoy Podder, Tanvir Zaman Khan, Mamdudul Haque Khan, and M Muktadir Rahman. "Comparative performance analysis of hamming, hanning and blackman window". In: *International Journal of Computer Applications* 96.18 (2014), pp. 1–7. DOI: 10.5120/16891-6927.

[Pon19] Piotr Ponikowski, Ilaria Spoletini, Andrew JS Coats, Massimo F Piepoli, and Giuseppe MC Rosano. "Heart rate and blood pressure monitoring in heart failure". In: *European Heart Journal Supplements* 21.Supplement_M (2019), pp. M13–M16. DOI: 10.1093/eurheartj/suz217.

[Pra24] Sourav Kumar Pramanik and Shekh Md Mahmudul Islam. "Through the wall human heart beat detection using single channel CW radar". In: *Frontiers in Physiology* 15 (2024), p. 1344221. DOI: 10.3389/fphys.2024.1344221.

[Ric20] Manola Ricciuti, Gianluca Ciattaglia, Adelmo De Santis, Ennio Gambi, and Linda Senigagliesi. "Contactless Heart Rate Measurements using RGB-camera and Radar." In: *Proceedings of the 6th International Conference on Information and Communication Technologies for Ageing Well and e-Health*. 2020, pp. 121–129. DOI: 10.5220/0009793201210129.

[Ron15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *Medical image computing and computer-assisted intervention – MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Ed. by Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi. Springer International Publishing. 2015, pp. 234–241. ISBN: 978-3-319-24574-4. DOI: 10.1007/978-3-319-24574-4_28.

[Sam16] Stefan Sammito and Irina Böckelmann. "Options and limitations of heart rate measurement and analysis of heart rate variability by mobile devices: A systematic review". In: *Herzschrittmachertherapie+ Elektrophysiologie* 27 (2016), pp. 38–45. DOI: 10.1007/s00399-016-0419-5.

[Sch03] Emily B Schroeder, Duanping Liao, Lloyd E Chambless, Ronald J Prineas, Gregory W Evans, and Gerardo Heiss. "Hypertension, blood pressure, and heart rate variabil-

ity: the Atherosclerosis Risk in Communities (ARIC) study". In: *Hypertension* 42.6 (2003), pp. 1106–1111. DOI: 10.1161/01.HYP.0000100444.71069.73.

[Sha23] Thanveer Shaik, Xiaohui Tao, Niall Higgins, Lin Li, Raj Gururajan, Xujuan Zhou, and U Rajendra Acharya. "Remote patient monitoring using artificial intelligence: Current state, applications, and challenges". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 13.2 (2023), e1485. DOI: 10.1002/widm.1485.

[She05] Belle A Shenoi. *Introduction to digital signal processing and filter design*. John Wiley & Sons, 2005. ISBN: 978-0-471-46482-2. URL: https://books.google.de/books?id=37g8oUqaS_AC.

[She20] Alex Shenfield and Martin Howarth. "A novel deep learning model for the detection and identification of rolling element-bearing faults". In: *Sensors* 20.18 (2020), p. 5112. DOI: 10.3390/s20185112.

[Shi19] Kilin Shi, Sven Schellenberger, Leon Weber, Jan Philipp Wiedemann, Fabian Michler, Tobias Steigleder, Anke Malessa, Fabian Lurz, Christoph Ostgathe, Robert Weigel, and Alexander Koelpin. "Segmentation of Radar-Recorded Heart Sound Signals Using Bidirectional LSTM Networks". In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019, pp. 6677–6680. DOI: 10.1109/EMBC.2019.8857863.

[Shi22] Qiuxia Shi, Bin Hu, Fuze Tian, and Qinglin Zhao. "Noncontact Doppler Radar-based Heart Rate Detection on the SVD and ANC". In: *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2022, pp. 1710–1713. DOI: 10.1109/BIBM55620.2022.9994893.

[Ski22] Practical Clinical Skills. *Fundamentals of Heart Sounds*. Accessed: 2024-08-03. 2022. URL: https://www.practicalclinicalskills.com/fundamentals-heart-sounds/60.

[Sko91] Merrill I. Skolnik. *Introduction to Radar Systems*. 2nd. New York: McGraw-Hill, 1991. ISBN: 0-07-066572-9.

[Str22] Tibor Stracina, Marina Ronzhina, Richard Redina, and Marie Novakova. "Golden standard or obsolete method? Review of ECG applications in clinical and experimental context". In: *Frontiers in Physiology* 13 (2022), p. 867033. DOI: 10.3389/fphys.2022.867033.

[Tan23] Hui Tang, Yu Rong, Li Chai, and Daniel W. Bliss. "Deep Learning Radar for High-Fidelity Heart Sound Recovery in Real-World Scenarios". In: *IEEE Sensors Journal* 23.15 (2023), pp. 17803–17814. DOI: 10.1109/JSEN.2023.3288764.

[Tha20] Angelika Thalmayer, Samuel Zeising, Georg Fischer, and Jens Kirchner. "A robust and real-time capable envelope-based algorithm for heart sound classification: Validation under different physiological conditions". In: *Sensors* 20.4 (2020), p. 972. DOI: 0.3390/s20040972.

[Tur20] Emmi Turppa, Juha M Kortelainen, Oleg Antropov, and Tero Kiuru. "Vital sign monitoring using FMCW radar in various sleeping scenarios". In: *Sensors* 20.22 (2020), p. 6505. DOI: 10.3390/s20226505.

[Unu14] Sathya D Unudurthi, Roseanne M Wolf, and Thomas J Hund. "Role of sinoatrial node architecture in maintaining a balanced source-sink relationship and synchronous cardiac pacemaking". In: *Frontiers in physiology* 5 (2014), p. 446. DOI: 10.3389/fphys.2014.00446.

[Vik09] P. S. Vikhe, N. S. Nehe, and V. R. Thool. "Heart Sound Abnormality Detection Using Short Time Fourier Transform and Continuous Wavelet Transform". In: *2009 Second International Conference on Emerging Trends in Engineering and Technology*. 2009, pp. 50–54. DOI: 10.1109/ICETET.2009.112.

[Wan15] Siying Wang, Antje Pohl, Timo Jaeschke, Michael Czaplik, Marcus Köny, Steffen Leonhardt, and Nils Pohl. "A novel ultra-wideband 80 GHz FMCW radar system for contactless monitoring of vital signs". In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2015, pp. 4978–4981. DOI: 10.1109/EMBC.2015.7319509.

[Wil18] Christoph Will, Kilin Shi, Sven Schellenberger, Tobias Steigleder, Fabian Michler, Jonas Fuchs, Robert Weigel, Christoph Ostgathe, and Alexander Koelpin. "Radar-based heart sound detection". In: *Scientific reports* 8.1 (2018), p. 11551. DOI: 10.1038/s41598-018-29984-5.

[Yan18] Xiaofeng Yang, Koichiro Ishibashi, Le Hoi, Trung Nguyen Vu, Kinh Nguyen Van, and Guanghao Sun. "Dengue Fever Detecting System Using Peak-detection of Data from Contactless Doppler Radar". In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2018, pp. 542–545. DOI: 10.1109/EMBC.2018.8512355.

[Yen22] Hoang Thi Yen, Masaki Kurosawa, Tetsuo Kirimoto, Yukiya Hakozaki, Takemi Matsui, and Guanghao Sun. "A medical radar system for non-contact vital sign monitoring and clinical performance evaluation in hospitalized older patients". In: *Biomedical Signal Processing and Control* 75 (2022), p. 103597. DOI: 10.1016/j.bspc.2022.103597.

[Yog76] Ajit P Yoganathan, Ramesh Gupta, Firdaus E Udwadia, J Wayen Miller, William H Corcoran, Radha Sarma, John L Johnson, and Richard J Bing. "Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man". In: *Medical and biological engineering* 14 (1976), pp. 69–73. DOI: 10.1007/BF02477093.

[Yug23] Louise Buonalumi Tacito Yugar, Juan Carlos Yugar-Toledo, Nelson Dinamarco, Luis Gustavo Sedenho-Prado, Beatriz Vaz Domingues Moreno, Tatiane de Azevedo Rubio, Andre Fattori, Bruno Rodrigues, Jose Fernando Vilela-Martin, and Heitor Moreno. "The role of heart rate variability (HRV) in different hypertensive syndromes". In: *Diagnostics* 13.4 (2023), p. 785. DOI: 10.3390/diagnostics13040785.

[Zin03] Z.M. Zin, S.H. Salleh, S. Daliman, and M.D. Sulaiman. "Analysis of heart sounds based on continuous wavelet transform". In: *Proceedings. Student Conference on Research and Development, 2003. SCORED 2003.* 2003, pp. 19–22. DOI: 10.1109/SCORED.2003.1459656.

Appendix A

Additional Figures

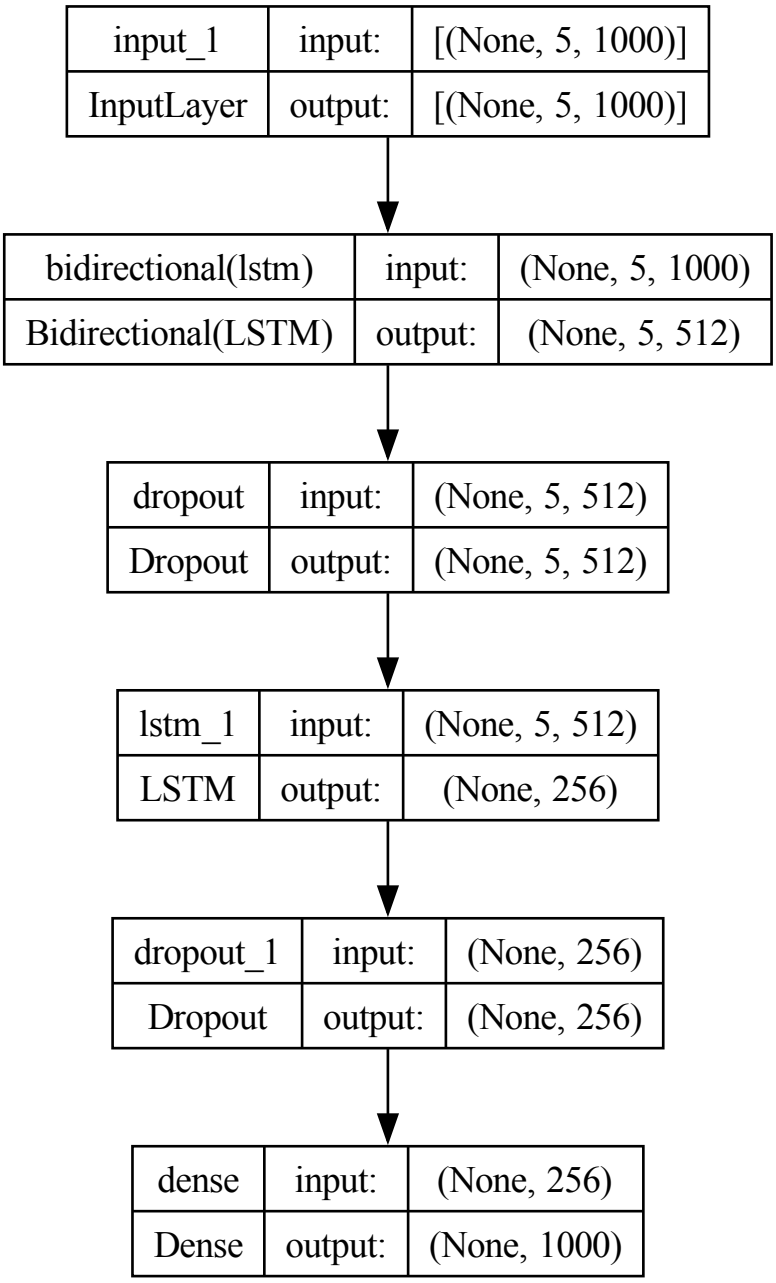

Figure A.1: **Model architecture used for the LSTM approach.** Outlines complete architecture used for the LSTM approach.

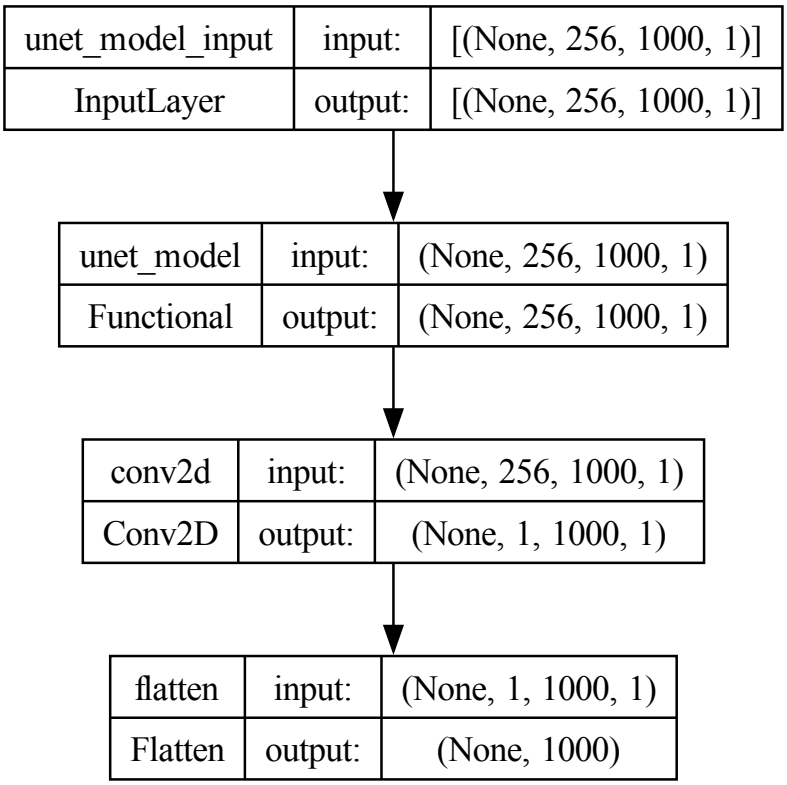

Figure A.2: **Model architecture used for the UNet approach.** Outlines reduced architecture used for the UNet approach.

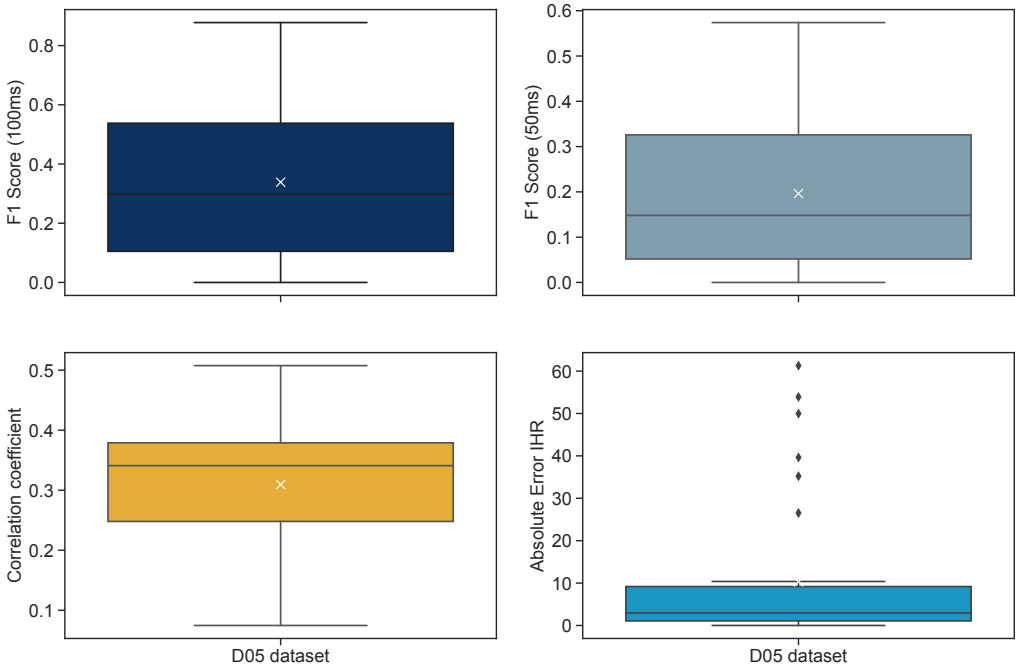

Figure A.3: **Performance metrics for the LSTM model with a LR of 1e^{-3}, 40 epochs and 256 LSTM units for the D05 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

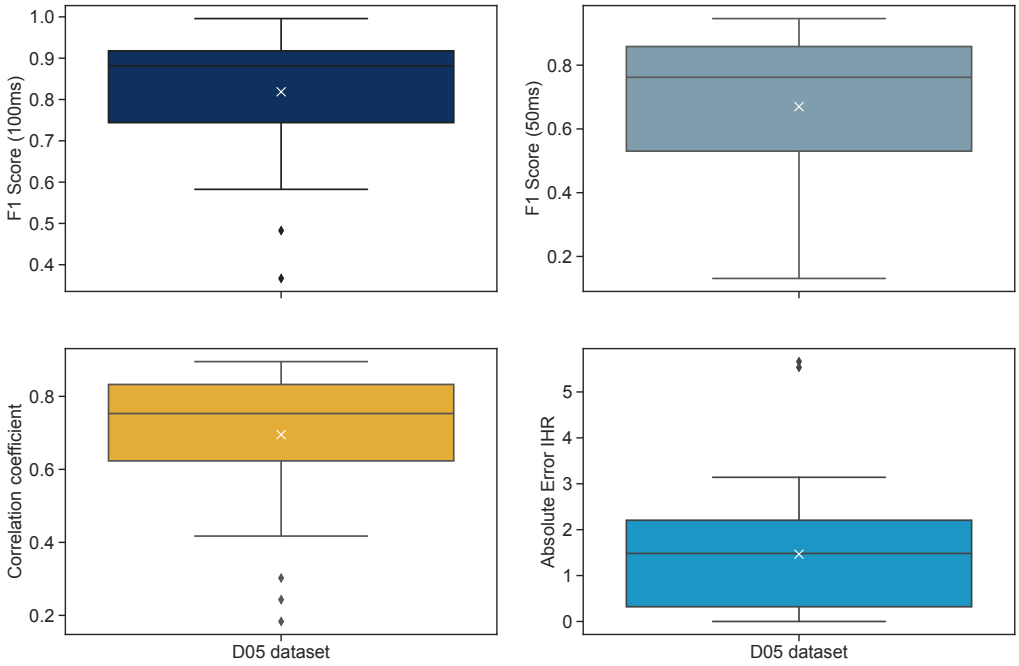

Figure A.4: **Performance metrics for the LSTM model with a LR of 1e^{-3}, 100 epochs and 256 LSTM units for the D05 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

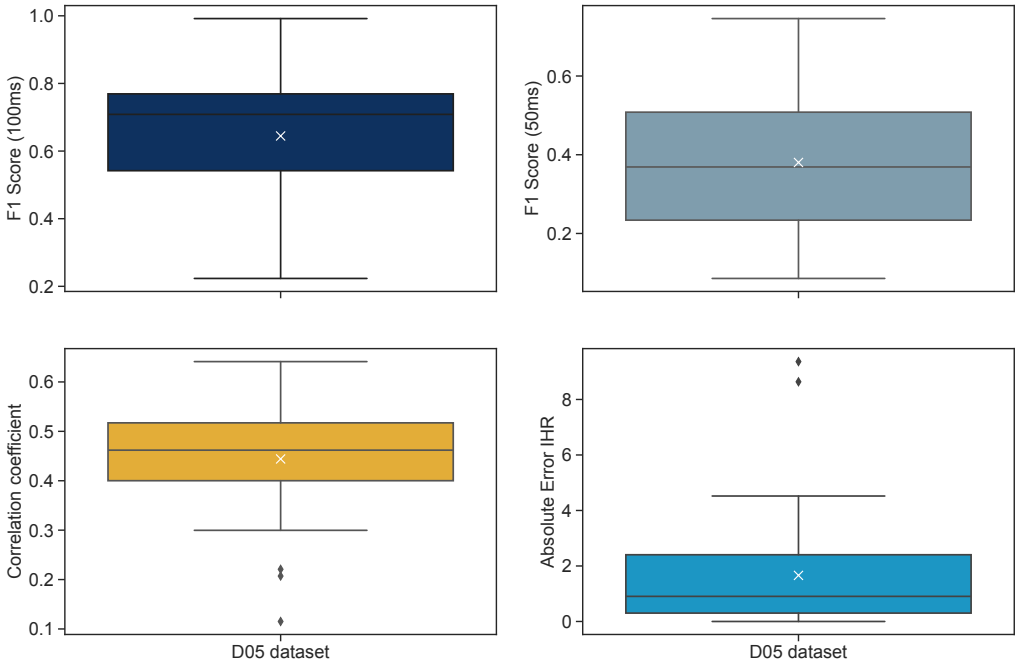

Figure A.5: **Performance metrics for the LSTM model with a LR of 1e^{-4}, 100 epochs and 256 LSTM units for the D05 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

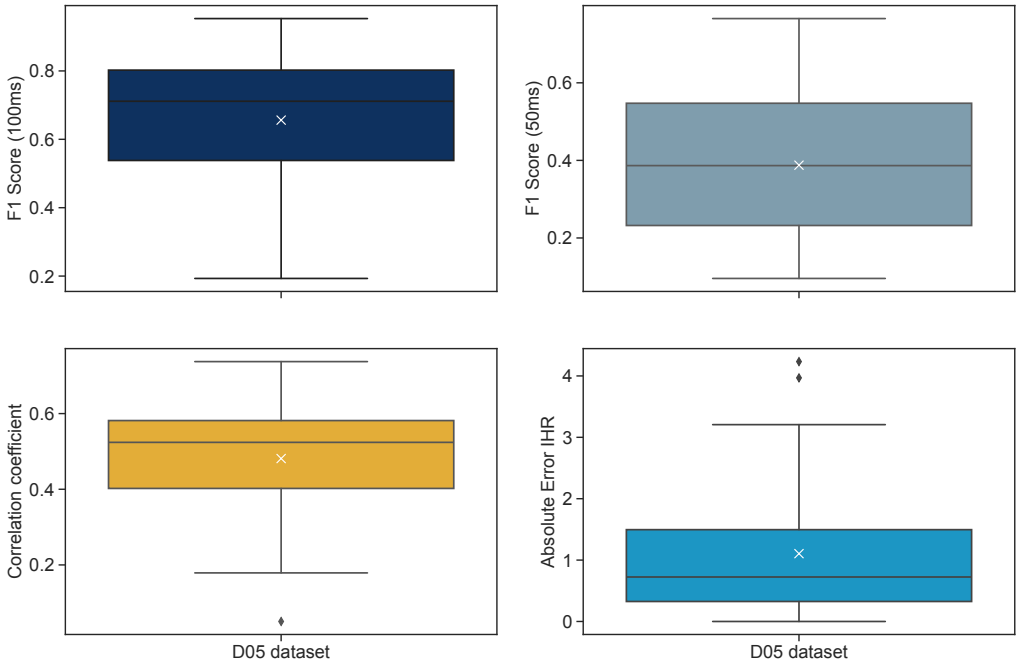

Figure A.6: **Performance metrics for the LSTM model with a LR of 1e^{-4}, 300 epochs and 256 LSTM units for the D05 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

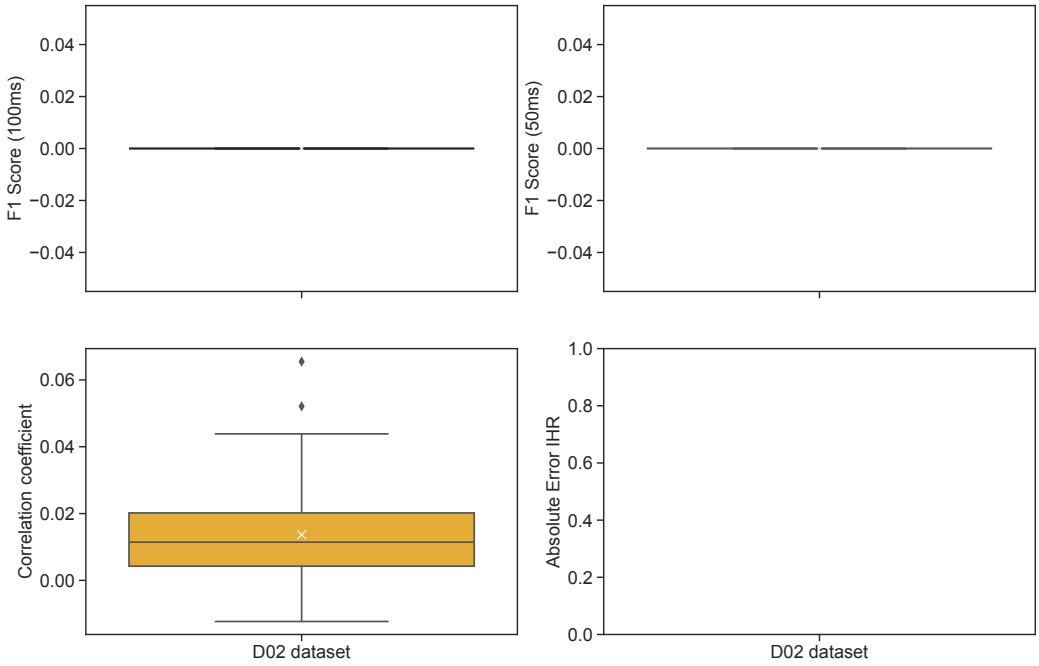

Figure A.7: **Performance metrics for the LSTM model with a LR of 1e^{-3}, 100 epochs and 256 LSTM units for the D02 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

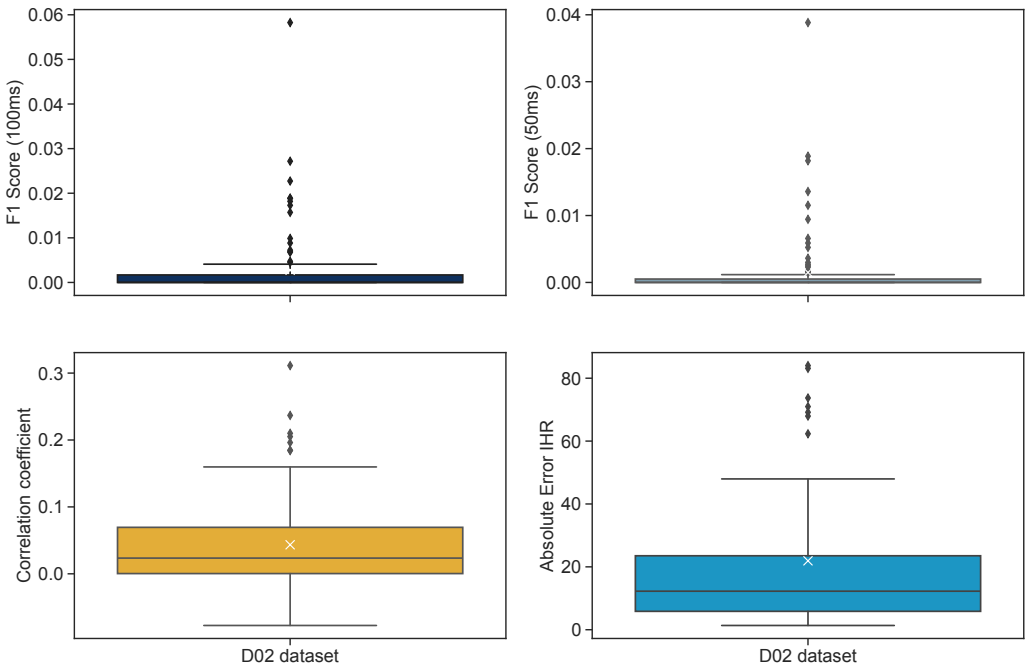

Figure A.8: **Performance metrics for the LSTM model with a LR of 1e^{-4}, 100 epochs and 256 LSTM units for the D02 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

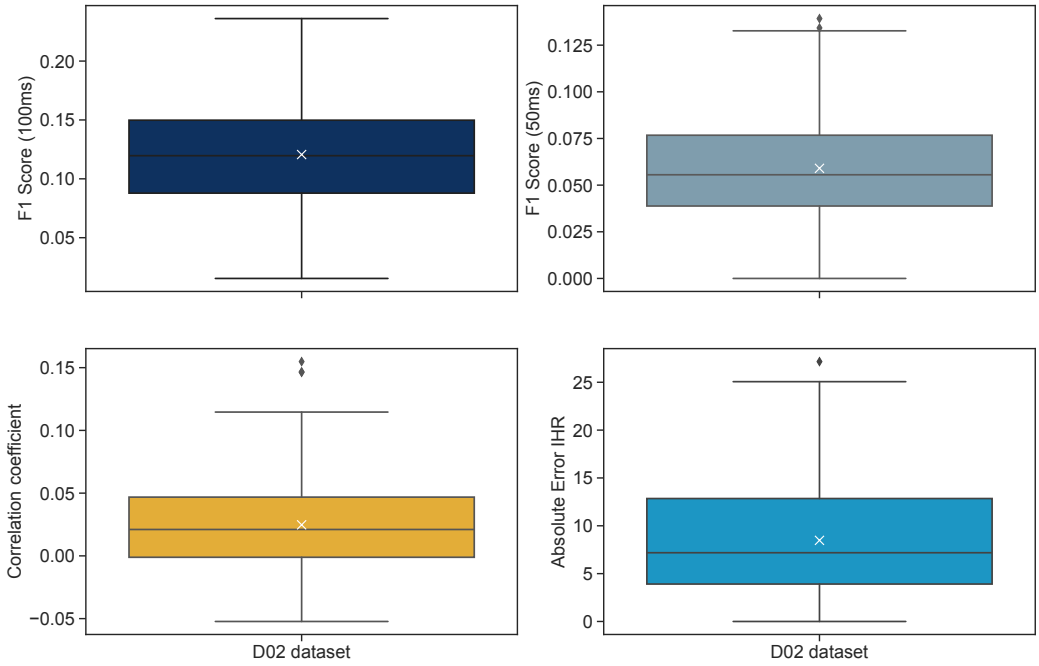

Figure A.9: **Performance metrics for the LSTM model with a LR of 1e^{-4}, 300 epochs and 256 LSTM units for the D02 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

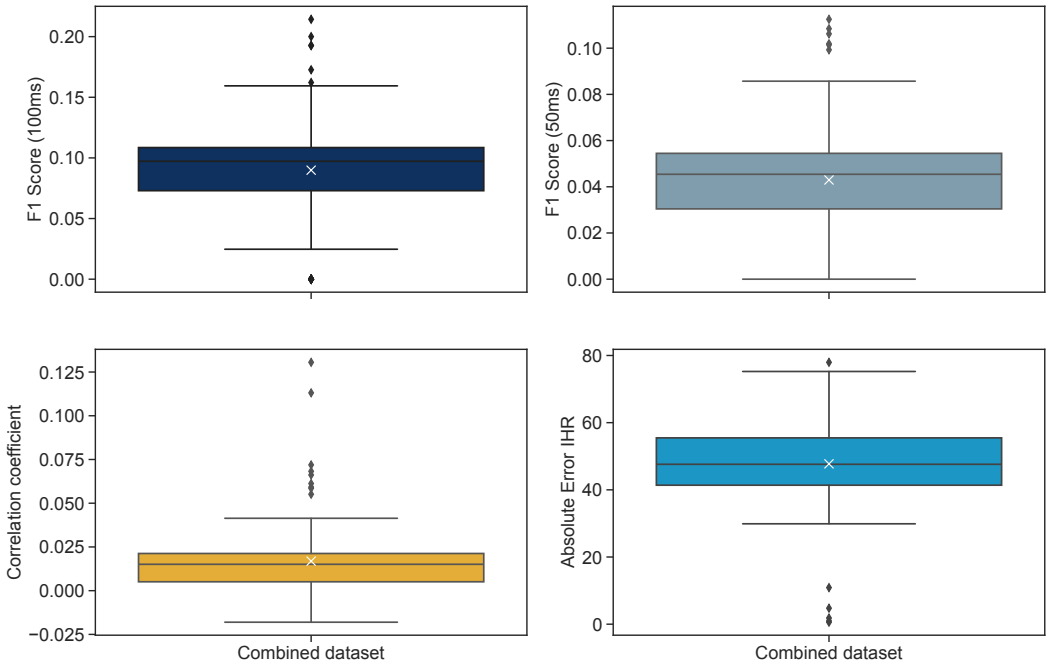

Figure A.10: **Performance metrics for the LSTM model with a LR of 1e^{-3}, 300 epochs and 256 LSTM units for the combined dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

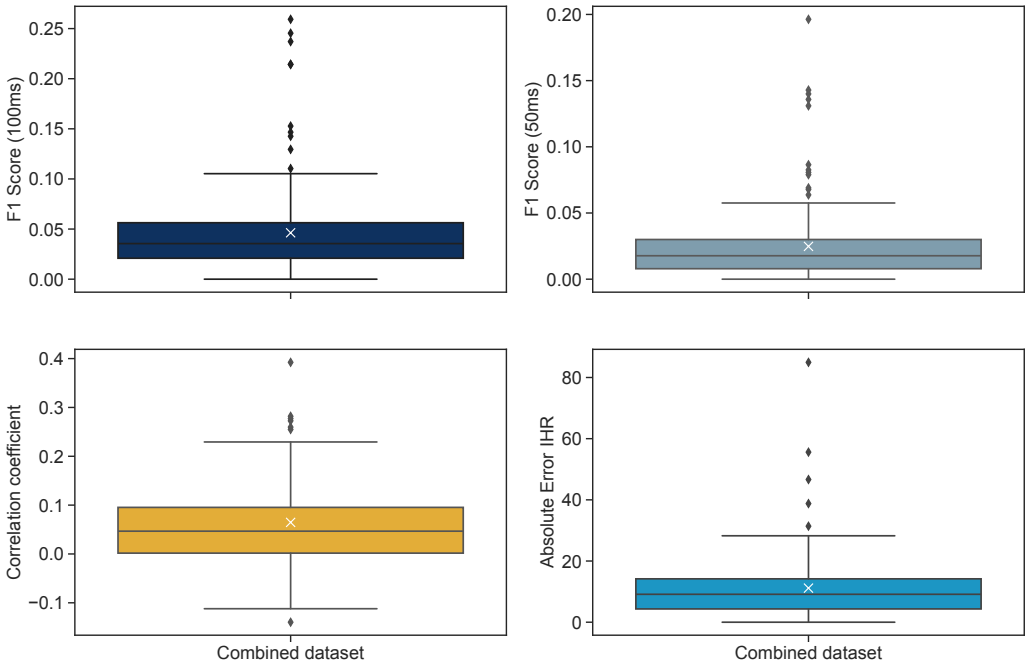

Figure A.11: **Performance metrics for the LSTM model with a LR of 1e^{-4}, 300 epochs and 256 LSTM units for the combined dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. White "X" depicts the mean of the corresponding metric.

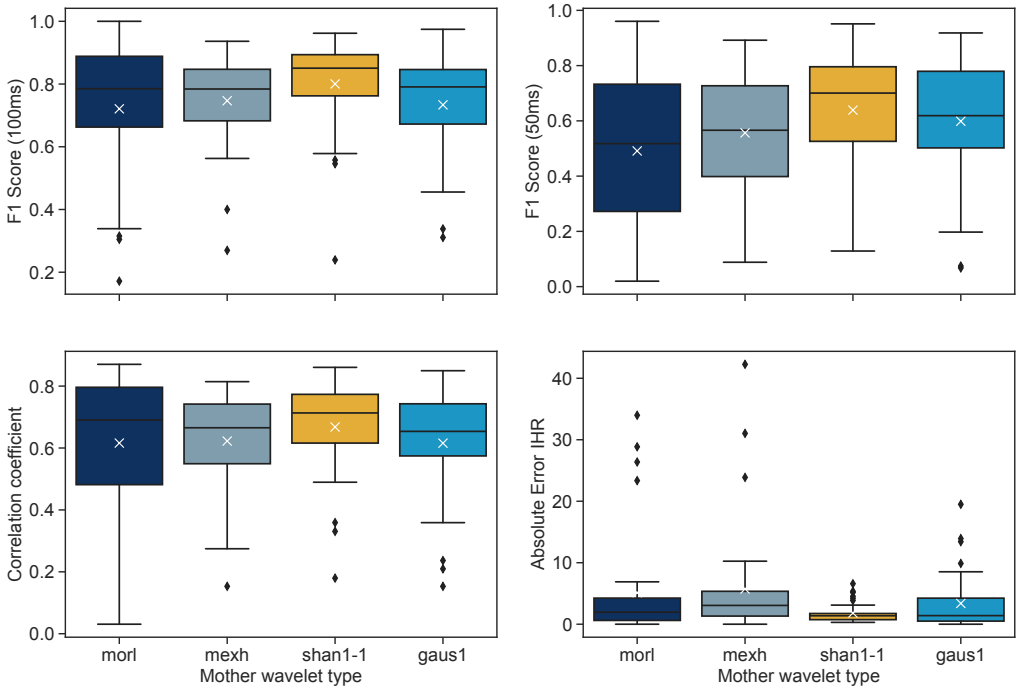

Figure A.12: **Performance metrics for the UNet model with a LR of 1e^{-3} and all mother wavelets tested for the D05 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. Each graph includes data for the four mother wavelets used, with white "X" indicating the mean for the corresponding metric.

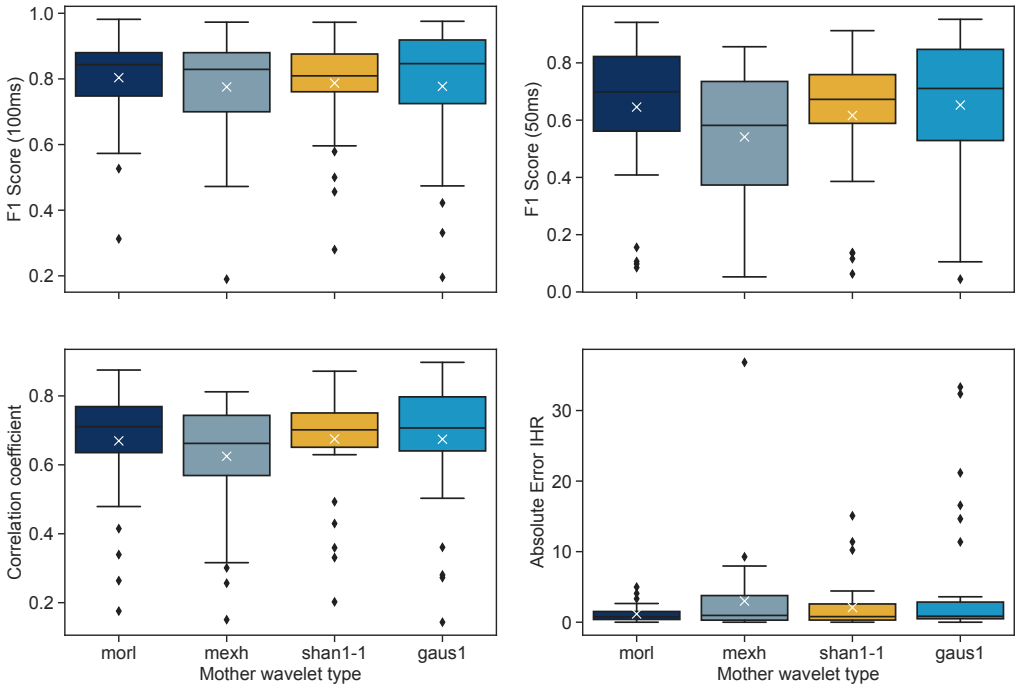

Figure A.13: **Performance metrics for the UNet model with a LR of 1e^{-4} and all mother wavelets tested for the D05 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. Each graph includes data for the four mother wavelets used, with white "X" indicating the mean for the corresponding metric.

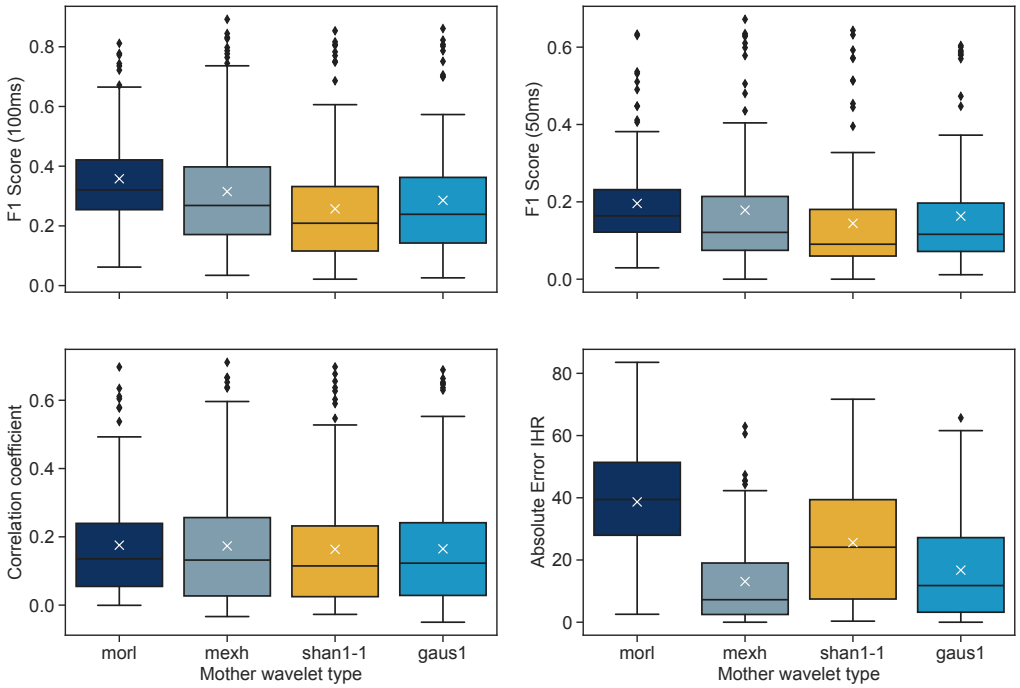

Figure A.14: **Performance metrics for the UNet model with a LR of 1e^{-4} and all mother wavelets tested for the D02 dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. Each graph includes data for the four mother wavelets used, with white "X" indicating the mean for the corresponding metric.

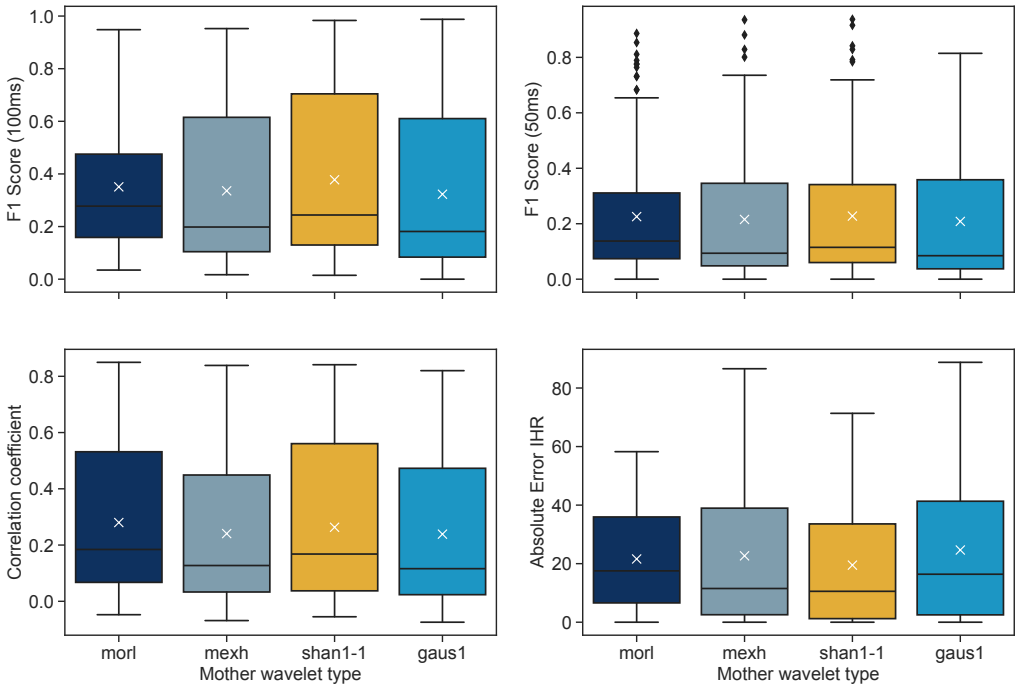

Figure A.15: **Performance metrics for the UNet model with a LR of 1e^{-4} and all mother wavelets tested for the combined dataset.** Boxplots depicting the of F1 scores with a 100ms tolerance window (top left) and 50ms tolerance window (top right) as well as the correlation coefficient and absolute error for the IHR. Each graph includes data for the four mother wavelets used, with white "X" indicating the mean for the corresponding metric.

Appendix B

Acronyms

AI	artificial intelligence
ANN	artificial neural network
BPM	beats per minute
CNN	convolutional neural network
CVD	cardiovascular diseases
CW	continuous wave
CWT	continuous wavelet transform
ECG	electrocardiogram
FFT	fast fourier transform
FMCW	frequency modulated continuous wave
FN	false negative
FP	false positive
HB	heart beat
HR	heart rate
HS	heart sound

I in-phase

IBI inter beat interval

IHR instantaneous heart rate

IQR interquartile range

IR-UWB impulse-radio ultra-wideband radar

LR learning Rate

LSTM long short-term memory

LSTM long short-term memory cells

ML machine learning

PCG phonocardiogram

Q quadrature

RADAR radio detection and ranging

ReLu rectified linear unit

RLBM random large body movements

RMSE root mean square error

RNN recurrent neural network

SD standard deviation

SNR signal to noise ratio

STFT short-time fourier transform

TP true positive