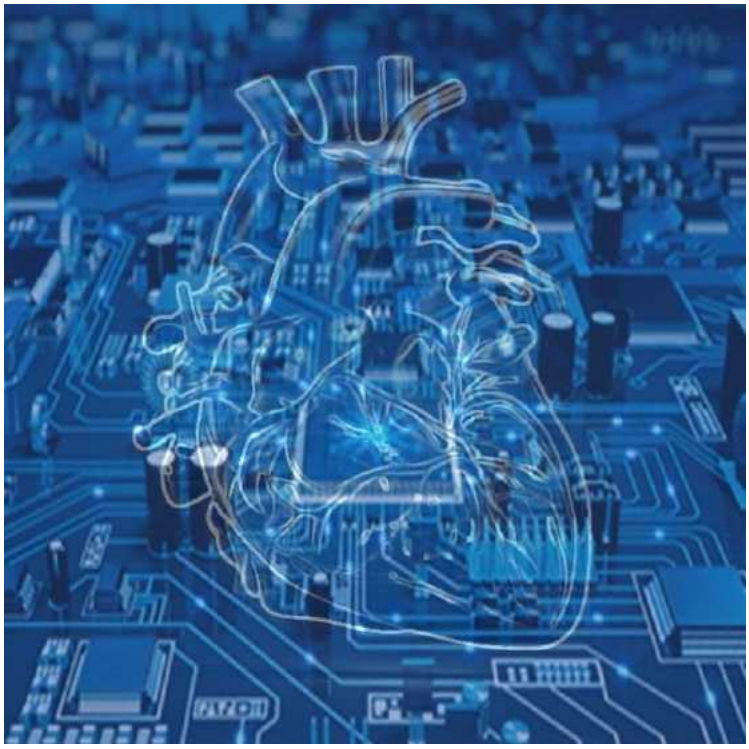


Models and Applications for Embedded Systems

Edited by
Massimo Conti, Simone Orcioni

Dipartimento di Ingegneria dell'Informazione
Università Politecnica delle Marche

ISBN: 978-88-87548-00-6

Models and Applications for Embedded Systems

Edited by
Massimo Conti, Simone Orcioni

Dipartimento di Ingegneria dell'Informazione
Università Politecnica delle Marche

ISBN: 978-88-87548-00-6

Models and Applications for Embedded Systems

Printed in Italy

ISBN: 978-88-87548-00-6 (printed and electronic)

Scientific Cooperation Partners

- **Massimo Conti** *(Università. Politecnica delle Marche, Italy)*
- **Simone Orcioni** *(Università Politecnica delle Marche, Italy)*
- **Natividad Martínez Madrid** *(Reutlingen University, Germany)*
- **Ann Nosseir** *(The British University of Cairo, Egypt)*
- **Ralf Seepold** *(HTWG Konstanz, Germany)*

Editorial Board

- **Massimo Conti** *(Università. Politecnica delle Marche, Italy)*
- **Simone Orcioni** *(Università Politecnica delle Marche, Italy)*

Objective

This book will present the work of the research results developed by the research cooperation between the Università Politecnica delle Marche (Italy), HTWG Konstanz (Germany), Reutlingen University (Germany) and the British University of Cairo (Egypt). The topics of the research work are signal detection, acquisition and processing, platforms and applications processing vital data, physiological measurement, low power wearable sensors, algorithm design for pattern classification, real-time data mining.

Contents

Monitoring Drivers' Heart Rates and Predicting Heart Attacks

Ann Nosseir, *British University in Egypt and Institute of National Planning*

***Abstract*— Driving for long hours can cause heart attack. Especially, track drivers spend more time driving to get more money and they take less rests even they are quite tired.This work develops a model that predicts a heart attack before happening. Support Vector Machine (SVM), Naïve- Bayses, Xgboost and KNN are used to develop the model. The accuracy reported are 91.73 Support Vector Machine, 85.12 Naïve Bayes, 98.7 Xgboost, and 88.42 KNN.**

I. INTRODUCTION

A heart attack, also known as a myocardial infraction, is a medical emergency situation that occurs when a heart muscle loses the access of blood that brings it oxygen because it becomes blocked. If the flow of blood does not get restored in a timely manner, the heart muscle will cease to function. A heart attack can also be referred to as a myocardial infraction.

With the traffic, the adrenalin pumping will increase heart rate and blood pressure and, in that time, increase the risk of heart attack.

This work develops a model using machine learning techniques to predict heart attacks. It uses. Support Vector Machine (SVM), Naïve-Bayses, Xgboost and KNN. the results show that accuracy reaches 98.7% using Xgboost. The following parts shows the details of the work.

* Ann Nosseir is at British University in Egypt and Institute of National (Email: ann.nosseir@bue.edu.eg; nosseir12@yahhoc.o.uk)

II. METHODOLOGY

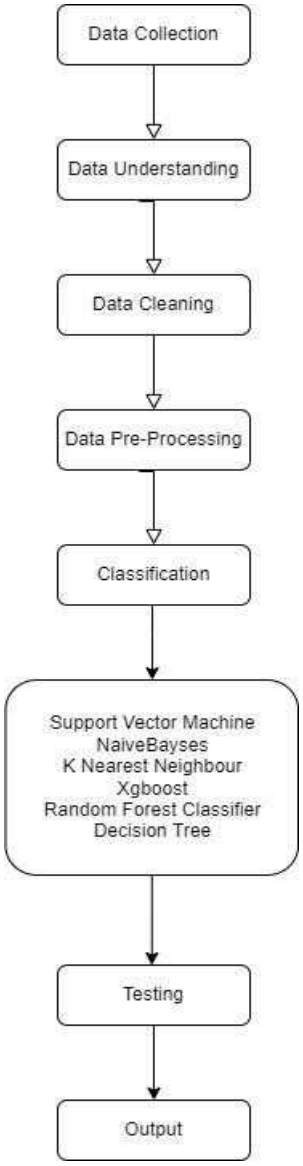


Figure 1. Steps of developing the model.

A. Data Collection and Preprocessing

First, the data is collected from Kaggel[11]. The data has the following attributes:

1. Age: Age of the patient
2. Sex: Sex of the patient
3. Exang: exercise induced angina (1=yes, 0=no)
4. ca: number of major vessels (0-3)
5. cp : Chest Pain type chest pain type; 1= typical angina, 2=atypical angina, 3= non-anginal pain, 4: asymptomatic
6. trtbps: resting blood pressure (in mm Hg)
7. chol : cholesterol in mg/dl fetched via BMI sensor
8. fbs : (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
9. rest_ecg : resting electrocardiographic results; 0= normal, 1= having ST-T wave abnormality, 2= showing probable ordefinite left ventricular hypertrophy by Estes' criteria
10. thalach: maximum heart rate achieved

Target: 0= less chance of a heart attack, 1= high chance of heart attack

The data has 700 record of different patients. Data was normalised with Min-Max. It is one of the most common ways to normalize data. it maps the minimum and maximum valuesof a feature to 0 and 1 respectively. So, all the other values are transformed to a value between 0 and 1 linearly.

B. Model Development

The researcher tested mainly four algorithms namely Support Vector Machine (SVM), Naïve-Bayses, Xgboost and KNN.

"Support Vector Machine" (SVM) is a supervised learning machine learning algorithm. SVM is able to perform linear classification which is optimal for this dataset to separate between the likely to get a heart attack or not. Finding a hyperplane with the maximum margin (margin is basically a protected space around hyperplane equation) and algorithm tries to have maximum margin with the closest points (known as support vectors) [12].

$$wT(\Phi(x)) + b < 0 \quad (1)$$

Figure 2 shows the hyperplane and how the classification works.

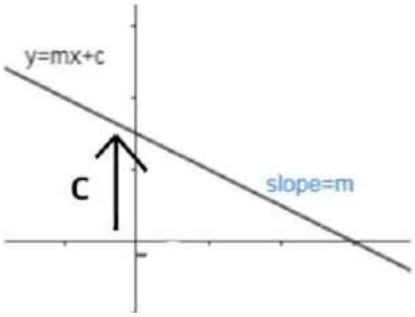
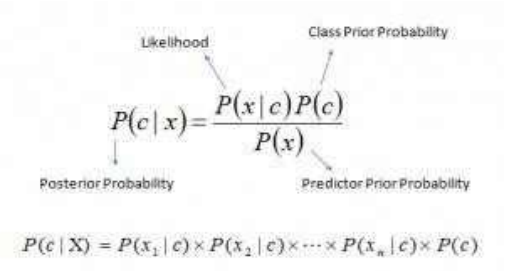


Figure 2. Support Vector Machine SVM

The second algorithm is Naïve-Bayses. Bayes Algorithm is famous for being both easy to use and efficient. Using this approach, model building and prediction are completed more quickly. The Naïve Byes works well on a small training data set on relatively small dimensions, it will perform better than the other models which applies in our case for our dataset.



(2)

- P(c|x) is the posterior probability of class (c, target) given predictor (x, attributes).
- P(c) is the prior probability of class.
- P(x|c) is the likelihood which is the probability ofthe predictor given class.
- P(x) is the prior probability of the predictor.

eXtreme Gradient Boosting (XGBoost) Is an Optimized Enhancement of Gradient Boosting. In Boosting, Weights are added to the model based on the residuals. However, In gradient boosted the loss function is optimized to correct errors made by previous models. XGBoost introduces new features to gradient boosting like regularization, tree pruning, and parallel processing.

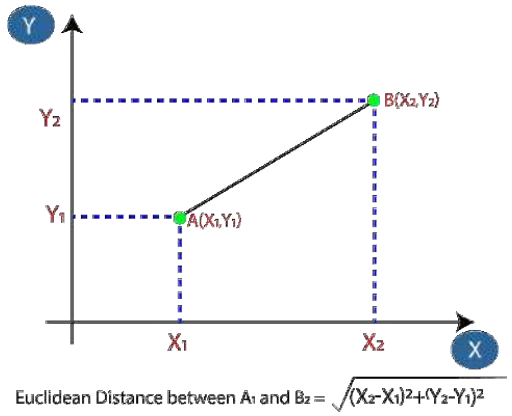
To Understand The XGboost Algorithm in classification, let's try to predict if students pass or fail basedon the number of hours studied. for binary classification, the probability for the base model is 0.5, meaning there is a fifty percent chance that a student will either pass or fail. consequently, the residuals for the first five readings are -0.5,0.5, -0.5, 0.5 and 0.5, respectively. XGBoost looks at which feature and split-point maximizes the gain. The maximum gain

is found where the sum of the loss from the child nodes most reduces the loss in the parent node. In math this is given by:

$$\text{Similarity Score} = \frac{\left(\sum \text{Residuals}\right)^2}{\sum_{N}[P\,(1 - P)] + \lambda} \quad (3)$$

P = Probability

KNN stands for K-nearest neighbour, it's one of the Supervised learning algorithm mostly used for classification of data on the basis how it's neighbour are classified. KNN stores all available cases and classifies new cases based on a similarity measure. K in KNN is a parameter that refers to the number of the nearest neighbours to include in the majority voting process.



(4)

We trained the algorithms on 80% of the data from the dataset, we tested the algorithms on 20% of the rest of the data to determine the accuracy of each algorithm model used.

III. Results

Table 1 shows the results of the four algorithms. The Xgboost reported the highest accuracy of 98.7%. that is followed by theSVM of 91.73% and then KNN of 88.42% andthe least accuracy is achieved by the Naïve Bayes which is 85.12%.

TABLE I. RESULTS

Machine Learning	SVM	Naïve Bayes	Xgboost	KNN
Accuracy	91.73	85.12	98.7	88.42

IV. Conclusion

Driving for long hours can cause heart attack. Additionally, Heart attacks are a very important and critical illness, it cannot be easily detected and can sometimes become active at very weird and unpredictable times which makes the ability to rightfully predetermine and predict if a patient has a chance of likely getting a heart attack or not by using a program that checks on the parameters and stats that surround what affects and causes heart attacks is greatly beneficial to the medical field as it can potentially save millions of lives.This work has used a dataset provided by kaggle to developa model this can predict this disease. We have compared the results for the following algorithms: Support Vector Machine (SVM), Naïve-Bayses, Xgboost and KNN. and, Xgboos provides the best accuracy which is 98.7.

References

[1] "Heart attack - what is a heart attack?," National Heart Lung and Blood Institute. [Online]. Available: https://www.nhlbi.nih.gov/health/heart-attack. [Accessed: 18-Jun-2022].

[2] NHS choices. [Online]. Available: https://www.nhs.uk/conditions/heart-attack/#:~:text=Symptoms%20of%20a%20heart%20attack%20can%20include%3A,jaw%2C%20neck%2C%20back%20and%20tummy. [Accessed: 18-Jun-2022].

[3] "Modern machine learning algorithms: Strengths and weaknesses," EliteDataScience, 09-Jun-2020. [Online]. Available: https://elitedatascience.com/machine-learning-algorithms. [Accessed: 18-Jun-2022].

[4] R. Rahman, "Heart attack analysis & prediction dataset," Kaggle, 22-Mar-2021. [Online]. Available: https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset?resource=download. [Accessed: 18-Jun-2022].

[5] "Big Data Analytics in heart attack prediction - researchgate.net." [Online]. Available: https://www.researchgate.net/profile/Lidong-Wang/publication/316851031_Big_Data_Analytics_in_Heart_Attack_Prediction/links/595fd6c4a6fdccc9b1c47e90/Big-Data-Analytics-in-Heart-Attack-Prediction.pdf. [Accessed: 18-Jun-2022].

[6] A. Widodo and B.-S. Yang, "Support Vector Machine in machine condition monitoring and fault diagnosis," Mechanical Systems and Signal Processing, 12-Jan-2007. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0888327007000027. [Accessed: 18-Jun-2022].

[7] "Implementation of Breiman's random forest machine ... - datajobstest.com." [Online]. Available: http://datajobstest.com/data-science-repo/Random-Forest-[Frederick-Livingston].pdf. [Accessed: 18-Jun-2022].

[8] T. C. U. of Washington, T. Chen, U. of Washington, C. G. U. of Washington, C. Guestrin, Ibm, Bosch, Amazon, Baidu, and O. M. V. A. Metrics, "XGBoost: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and data mining," ACM Conferences, 01-Aug-2016. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2939672.2939785. [Accessed: 18-Jun-2022].

[9] J. Cervantes, F. G. Lamont, A. López-Chau, L. R. Mazahua, and J. S. Ruíz, "Data selection based on decision tree for SVM classification on large data sets," Applied Soft Computing, 25-Sep-2015. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1568494615005591. [Accessed: 18-Jun-2022].

[10] "K-nearest neighbors | springerlink." [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-38652-7_2. [Accessed: 18-Jun-2022].

[11] https://www.kaggle.com/rashikrahmanpritom/heart-attack-analysis-prediction-dataset
[12] https://www.analyticsvidhya.com/blog/2020/10/the-mathematics-behind-svm/

Software Scripts for Sensor Data Extraction in Rasberry Pi: user-space and kernel-space comparison

Daniel Vélez, Natividad Martínez Madrid and Ralf Seepold

***Abstract*— This paper compares two popular scripting implementations for hardware prototyping: Python scripts execut from User-Space and C-based Linux-Driver processes executed from Kernel-Space, which can provide information to researchers when considering one or another in their implementations. Conclusions exhibit that deploying software scripts in the kernel space makes it possible to grant a certain quality of sensor information using a Raspberry Pi without the need for advanced real-time operational systems.**

I. INTRODUCTION

Fast prototyping over open-access Hardware platforms has become a valuable teaching tool for spreading technological knowledge among students and the general population. Since the introduction of cheap open Hardware platforms like Ar-Arduino and Raspberry Pi in 2010 and 2012, general enthusiasts have gained easy access to electronic components related to computing, robotics, IoT technologies, and smart devices, which until then were only accessible to high education institutes or specialized tech companies. Similarly, several research initiatives have benefited from these open-access hardware platforms, easing the implementation and testing of novel hardware arrangements for different purposes. This has settled down a set of preferred technologies in the last years in academia: around ten years ago, languages like Java and C++ used to lead the surveys [1], although python was settling down as preferred in academic computing [2]; Nowadays, Python has overcome the academic spheres andis positioned as a top trend for industry as well [3]. Support communities and collaborative work have been crucial for establishing an extensive collection of standard libraries that make Python-Language easy to use and adapt to domains such as Big Data, Artificial Intelligence, and Machine Learning. The popularity of interpreted languages over structured compiled languages has induced the misleading perception that scripting and Software development share the same process. Nonetheless, regarding hardware prototyping, scriptprocedures that automate system tasks are crucial; thus, interpreted languages like Python are important. This paper compares two popular scripting implementations for hardware prototyping: Python scripts executed from User-Space and C-based Linux-Driver processes executed from Kernel-Space. These comparisons could provide valuableinformation to researchers considering one or the other technology in their implementations.

* This research was partially funded by Carl Zeiss Foundation and the MORPHEUS-Project "Non-invasive system for measuring parameters relevant to sleep quality" (project number: P2019-03-003)

D. Vélez is with the Ubiquitous Computing Lab at HTWG Konstanz, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (phone: +49 7531 206-703; Email: jvelezg@htwg-konstanz.de).

II. CURRENT WIDESPREAD OPEN-ACCESS TECHNOLOGIES FOR HARDWARE PROTOTYPING

A. Arduino

The Arduino Board (UNO) was commercially launched in 2010 after several years of research [4]. The availability of this family of boards was a huge step forward in simplifying the process of microcontroller development for educational purposes at a low price, making it available for hobbyists, teachers, and students alike. Their regular modules include microcontrollers such as Atmel ATmega328p or ATmega168, which can provide enough capability to support a wide range of projects. Similarly, the Arduino-IDE platform simplifies the coding challenges when developing in programming languages like C and C++. Nowadays, it has become a popular choice in the research of electronics and prototypes in different domains, including hardware communication, software prototyping, home general automation, agriculture, healthcare, mining industry, energy, defense, IoT, and education, increasing affordability for both profit and non-profit institutions [5]. Some of the limitations of Arduino hardwareand Software identified in the literature are listed below [5]:

- Limited processing power
- Small storage space and memory
- Requires efforts to accomplish tasks such as scheduling and database storage.
- Cannot handle large complexities of advanced projects.
- Kits are not suitable for high-performance Hardware.

B. Raspberry-Pi

The Raspberry Pi became commercially available in 2012. It was intended to get people interested in computing and coding. Since its origins, it has provided support for coding in Python as the primary programming language, althoughit also included support for other languages like BASIC, C/C++, JAVA, Perl, and Ruby. The general availability of this achievable tool encouraged people to get in touch with the guts of the electronic devices, which until then was demotivated for the sake of the durability and secrecy of regular commercial electronics [6].

N. Martínez Madrid is with the IoT Lab at Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany (Email: natividad.martinez@reutlingen-university.de).

R. Seepold is with the Ubiquitous Computing Lab at HTWG Konstanz, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (Email: ralf.seepold@htwg-konstanz.de)

The Raspberry Pi 4 Model B hardware specifications include:

- Quad-core 64-bit ARM-Cortex A72 running at 1.5GHz
- 1, 2 and 4 Gigabyte LPDDR4 RAM options
- H.265 (HEVC) hardware decode (up to 4Kp60)
- H.264 hardware decode (up to 1080p60)
- VideoCore VI 3D Graphics
- Supports dual HDMI display output up to 4Kp60

Easy interface with peripherals is the main feature of the Raspberry products. Their GPIO pins interface can be used as straightforward software-controlled input and output. However, more recent models can also be switched into I2C, UART and SPI, which gives users much more flexibility when attaching add-on hardware [7]. Software capabilities of Raspberry Pi include a mature Linux software stack, recent Linux kernel support, and availability of GPU functions using standard APIs. The Operative Systems recommended to support these development boards come in different flavors; nonetheless, the most popular and recommended choices are based on Linux OS. Some of the weaknesses of theRaspberry Pi identified in the literature are listed below [5]:

- Raspberry Pi boards are not always available for long.
- The Operating System (OS) runs on an SD Card
- Absence of USB header connectors
- Absence of Real-time Clock with Battery Backup
- Absence of Onboard ADC
- Absence of EEPROM/FRAM/SPI Flash
- Limited Universal Asynchronous Receiver Transmitters(UARTs)
- Raspberry cannot handle External Power Supply
- Poor thermal Management
- Raspberry Pi boards were not designed for commercial products or large-scale projects but for small projects.

III. FAST HARDWARE PROTOTYPE PROGRAMMING LANGUAGES

The programming and configuration languages used to operate the electronic components are at the top of the previously mentioned open-access hardware development platforms. Regular Hardware prototyping is usually based on simulation platforms, implemented over languages like SPICE (Simulation Program for Integrated Circuit Emphasis) and focused on circuit design with Hardware Description Languages (HDLs) like VHDL, Verilog, LabVIEW, Scilab Xcos. In comparison, prototyping Hardware using micro-computers like Arduino and Raspberry Pi focuses on enhancing the potential and functionality of commercially available electronic actuators by tailoring the processing capabilities of microprocessors to specific pre-established project objectives using Software scripts and procedures. Under this specific approach, three different code implementation approaches can be applied.

For the Arduino, the approach consists of deploying the execution instructions (Script) directly on the chip without any OS (bare metal design). The advantages of this typeof implementation include having complete control over the components provided on the board.

When using a Raspberry Pi, it is possible to run Scripts ontop of the operating system, either in user space (as a regularthird-party program) or kernel space (as drivers integratedinto the main OS procedures). A primary disadvantage is that running the applications on top of the OS works slower. This can be a challenging aspect to consider when the data extracted needs to comply with a certain level of quality to ensure the reliability or reproducibility of specific experiments. Although in counter position, this strategy allows the use of available community-supported libraries and produces code not tied to a particular hardware configuration, which is a valuable advantage when testing novel approaches with commercial Hardware and micro-controllers.

Regarding the Raspberry Pi approach, technologies to implement the procedures differ when focusing on user spaceor kernel space:

A. Python

Python is the de facto language for user-space configurations (as designed by the creators of the development board);it counts with a plethora of free-access and open-source libraries available and supported for different actuators. A tailored version of the language (MicroPython) is also available for "bare metal" or on-chip implementations optimized to run on microcontrollers and in constrained environments. Python software and documentation are licensed under the PSF License Agreement [8]. Some of this language's generally accepted advantages and disadvantages are listedbelow.

Advantages:

- Fast Learning Curve.
- Portability.
- Open-source.
- Efficient for Rapid Development.
- Automatic Memory Allocation.
- Large Built-In Objects and Libraries.
- Extensive Third-Party Library Availability.

Disadvantages:

- Not Very Fast.
- Memory Intensive.
- Not Optimized for Database Access.
- No Multi-threading Support.
- Prone to Overuse or Misuse.

B. C language

Despite the popularity and rise of other general-purpose languages, C and C++ are still considered top-class professional development languages for creating modern applications. Historically, the implementation of C-language markeda new era for Software development. By early 1973, C-based Unix OS became the most used environment at research-oriented academic and government organizations. Nowadays, C programming language is the base software technology behind most operating and embedded systems.

It is still the preferred language for working with system kernels and device drivers (including kernels of Linux, macOS, iOS, Microsoft Windows, and Android operating systems). Additionally, many compilers and interpreters used in other languages like Python are developed in C. Similarly, it is a common standard development tool for drivers in everyday electronics, the automotive industry, and smart home appliances [9]. The advantages and disadvantages of this language are listed below.

Advantages:

- Large Built-In Objects and Libraries across multiple OS.
- Multi-threading Support.
- It can be used for low-level programming, such as scripting for drivers and kernels.
- Its dynamic memory allocation capabilities make it the preferred choice for scripting drivers of embedded systems.
- It can efficiently work on enterprise applications.
- Facilitates the implementation of algorithms and data structures, enabling faster program computations.

Disadvantages:

- Extended learning curve.
- Low portability.
- Low security given its low level of abstraction.
- Originally, C does not support Object-Oriented programming, although this functionality is supported in C++.

C. JavaScript

JavaScript is another popular language for user-space execution. It is a multi-paradigm scripting language that has prototype-based object-oriented features. It was implemented by Brendan Eich in 1995 as a programming language for web pages as part of the Netscape browser using C-like syntax. By 1998, it became a standard through ECMA International (currently in the 6th revision). Despite being a recognized standard, it was usually implemented as a complementary feature in different browsers and included in frameworks for PHP and Java.

However, the lack of general support dampened its implementation. That was the case until 2009, when Ryan Dahl introduced Node.Js, a novel tool to build efficient web servers purely in JavaScript using Google's V8 engine. JavaScript is now widely used in web pages, web servers, and mobile applications, and multiple initiatives are looking to bring this language into the embedded systems [10] [11].

The main advantages of JavaScript are those of the interpreted languages; it is quite suitable for the rapid development of prototype programs, and its event-driven programming style is considered an exploitable advantage for data processing in embedded systems [11]. Disadvantagesare also shared, including their high memory usage and worst performance compared to low-level languages like C.

JavaScript differentiates from Python because it lacks specialized libraries and frameworks to interfacewith hardware devices. The approaches to bringing this technology into the hardware environment are still early.

IV. NON-INVASIVE CARDIORESPIRATION MEASUREMENT USIN G FORCE RESISTIVE SENSOR

The experiment was conducted as part of the Morpheus project trials. The MORPHEUS Box (MoBo) device aimsto monitor patients' body signals over long periods during sleep using exclusively non-invasive technologies[12, Chapter 6].

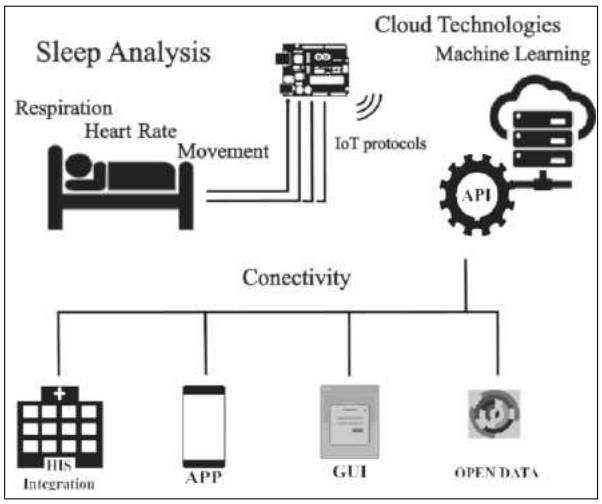


Fig. 1. Morpheus project layout

A. Hardware configuration

The system includes pressure sensors installed under a bed mattress and a computational unit, as described by Asadov et al. in [13]. Three Force Sensing Resistors (FSR) are connected to an I2C compatible, analog-to-digital converter, and the signals are collected using a Raspberry Pi4. The FSR sensors are attached to the bed's frame, below the mattress, but in contact with it.

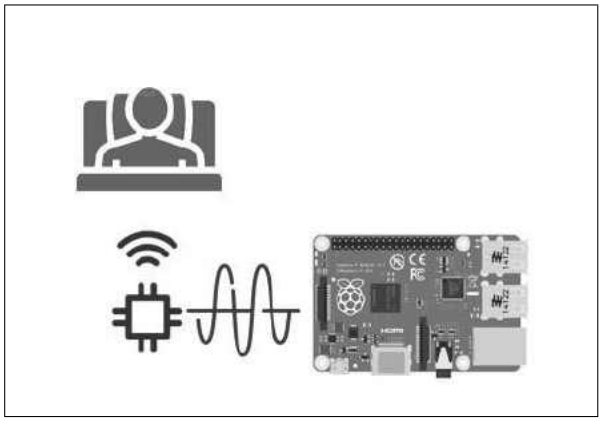


Fig. 2. Experimental system diagram

B. Experiment

During the experiment, the subjects participating in the study lay on the bed in different positions (prone, right side, supine, and left side). The data extracted from the sensors is processed to extract respiration and Heart rate measurementsfrom the subject.

C. Software configuration

Software Scripts were coded to extract and pre-process the signals from the FSR sensors, store the datalocally in the Raspberry Pi and transmit it to a backend server. Additional synchronization procedures are included to verify the availability of the backend server and secure the transmission of information. The application startup consistsof the following steps:

1) Sync device properties with the backend server (ID, IP, Assigned Name).
2) If absent, check patient ID parameters or assign a random uniquecross-platform ID (UUID).
3) Updates patient data ID data with the backend server.
4) Start recording signals from FSR sensors in JSON datastructures.
5) Store information locally.
6) Transmit information to the backend-server

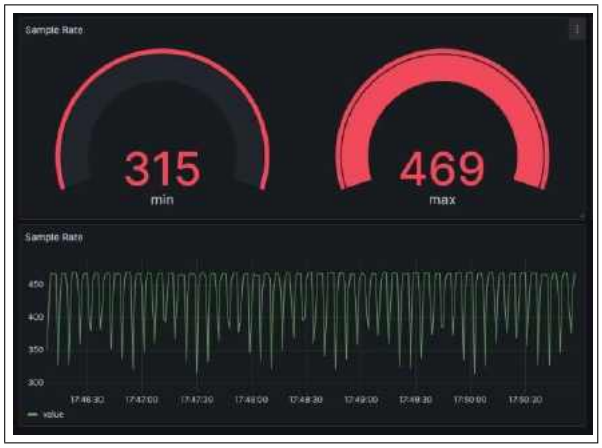


Fig. 3. Samples per second obtained from user-space

D. User-Space: Python Implementation

The first implementation was made for user space using Python on the top of a Raspberry Pi OS. The script uses the library Adafruit_ADS1015 to read the signals from the converter. Using a loop to control each iteration, the procedure reads each sensor once and stores the measurements locally using an embedded SQLite database. Unfortunately, when testing this version, the Software overloaded the memory of the Raspberry Pi device, slowing down the OS and making the transmission of information non-viable; after 10 minutes, only 1 minute of recorded information was transmitted to the server. Additionally, the SD card of Raspberry Pi got ”bricked out” due to the increased write cycles performed in the embedded database.

E. User-Space: Simplified Python Script

A second implementation was developed to run simplified Python Scripts. It consists of an orchestrator developed in nodeJS that runs the Python Scripts recorders to store the information in plain CSV files. An enhanced PythonScript can also transmit the information via the MQTT protocol. The sample rate obtained with this implementation (described in Picture 3) was variable, with a maximum of 469 and a minimum of 315 samples per second (SPS). A pattern of high and low SPS can be identified in the plotted visualization.

F. Kernel-Space: C Implementation

For the kernel-space approach, the implementation was coded using C language. The Script was deployed as a kernel module in a Raspberry Pi OS. This configuration also implies the execution of a complementary user-space script that would collect the information sent from the kernel via Procfs [14]. This Script extracts the sensor data from the conversion register of the ADS1115 [**?**]. The sample rate obtained from this Script is shown in Figure 4. In this case, the maximum sample rate obtained is 337, though the memory space of the Procfs file limits the output. The ADS1115 component can provide a maximum of 860 SPS for a single sensor. The plotted visualization (lower part of Figure 4) shows that the sample rate obtained with this implementation approach is stable across time.

Fig. 4. Samples per second obtained from kernel-space

V. RESULTS

Table I summarizes the results obtained in this comparison. Regarding the user-space approach, the OS process priority and allocation may interfere with the script's capacity to maintain a stable sample rate. In comparison, the kernel-space approach can maintain a stable sample rate across time, allowing setting a fixed samplerate without needing a more complex implementation or a specialized Real-Time Operative System.

TABLE I

SAMPLE RATES OBTAINED

Script	Space	Min SPS	Max SPS
Adafruit_ADS1015+MQTT	User-Space	315	469
Kernel+MQTT	Kernel-Space	337	337

VI. CONCLUSIONS

Fast prototyping over open-access Hardware platforms is helpful in scenarios with a pragmatic approach. The features provided by these development platforms, like Arduino and Raspberry Pi, are powerful but limited, and these limitations can affect the quality of data obtained during experimentation. The comparison described in this paper exhibits that by exploiting the benefits given by the Linux OS and deploying Software Scripts in the kernel space, it is possible to grant certain quality of information from the data extracted from sensors using a Raspberry Pi without the need to get involved in specialized implementations or using advanced Real-Time Operative Systems.

ACKNOWLEDGMENT

Carl-Zeiss-Foundation funded this research with the MORPHEUS-Project “Non-invasive system for measuring parameters relevant to sleep quality” (project number: P2019- 03-003).

REFERENCES

[1] L. BEN ARFA RABAI, B. COHEN, and A. MILI, “Programming language use in us academia and industry,” 2014. [Online]. Available: https://files.eric.ed.gov/fulltext/EJ1079004.pdf

[2] R. S. King, “The Top 10 Programming Languages - IEEE Spectrum,” 2011. [Online]. Available: https://spectrum.ieee.org/the- top-10-programming-languages

[3] S. Cass, “The Top Programming Languages 2023 - IEEE Spectrum,” 2023. [Online]. Available: https://spectrum.ieee.org/the- top-programming-languages-2023

[4] A. Team, “One board to rule them all: History of the Arduino UNO,” Dec. 2021. [Online]. Available: https://blog.arduino.cc/2021/12/09/one-board-to-rule-them-all- history-of-the-arduino-uno/

[5] H. K. Kondaveeti, N. K. Kumaravelu, S. D. Vanambathina, S. E. Mathe, and S. Vappangi, “A systematic literature review on prototyping with Arduino: Applications, challenges, advantages, and limitations,” Computer Science Review, vol. 40, p. 100364, May 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574013721000046

[6] P. Sachdeva and S. Katchii, “A review paper on raspberry pi,” 2014. [Online]. Available: http://www.inpressco.com/wp-content/uploads/2014/11/Paper53818-3819.pdf

[7] R. P. T. Ltd, “raspberry-pi-4-datasheet.” [Online]. Available: https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-datasheet.pdf

[8] “History and License.” [Online]. Available: https://docs.python.org/3/license.html

[9] T. Taylor, “A quick glance at the history of C programming languages | TechTarget,” 2022. [Online]. Available: https://www.techtarget.com/searchapparchitecture/tip/A-quick- glance-at-the-history-of-C-programming-languages

[10] T. Ugawa, H. Iwasaki, and T. Kataoka, “eJSTK: Building JavaScript virtual machines with customized datatypes for embedded systems,” Journal of Computer Languages, vol. 51, pp. 261–279, Apr. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1045926X18302416

[11] F. L. Oliveira and J. C. B. Mattos, “JSGuide: A Tool to Improve JavaScript Algorithms Focusing on IoT Devices,” in 2022 Symposium on Internet of Things (SIoT), Oct. 2022, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/document/10070155

[12] Non-Invasive Health Systems based on Advanced Biomedical Signal and Image Processing. [Online]. Available: https://www.routledge.com/Non-Invasive-Health-Systems-based-on-Advanced-Biomedical-Signal-and-Image/Al-Jumaily-Crippa-Mansour-Turchetti/p/book/9781032386942

[13] A. Asadov, J. Ortega, R. Seepold, and N. Martínez Madrid, “Hardware and Software supporting Physiological Measurement (HSPM-2022),” Oct. 2022.

[14] B. Wang, B. Wang, and Q. Xiong, “The comparison of communication methods between user and Kernel space in embedded Linux,” in International Conference on Computational Problem-Solving, Dec. 2010, pp. 234–237. [Online]. Available: https://ieeexplore.ieee.org/document/5696027

Heart Rate Estimation based on in-bed Accelerometer Sensor Measurement*

Maksym Gaiduk, *Member, IEEE*, Andrei Boiko, Massimo Conti, *Member, IEEE*, Simone Orcioni, *Senior Member, IEEE*, Natividad Martínez Madrid, Senior *Member, IEEE*, and Ralf Seepold, *Senior Member,IEEE*

***Abstract*— Accurate monitoring of a patient's heart rate is a key element in the medical observation and health monitoring. In particular, its importance extends to the identification of sleep-related disorders. Various methods have been established that involve sensor-based recording of physiological signals followed by automated examination and analysis. This study attempts to evaluate the efficacy of a non-invasive HR monitoring framework based on an accelerometer sensor specifically during sleep. To achieve this goal, the motion induced by thoracic movements during cardiac contractions is captured by a device installed under the mattress. Signal filtering techniques and heart rate estimation using the symlets6 wavelet are part of the implemented computational framework described in this article. Subsequent analysis indicates the potential applicability of this system in the prognostic domain, with an average error margin of approximately 3 beats per minute. The results obtained represent a promising advancement in non-invasive heart rate monitoring during sleep, with potential implications for improved diagnosis and management of cardiovascular and sleep-related disorders.**

***Clinical Relevance*—This research provides a foundation for contactless monitoring of health signs in domain of sleep analysis.**

I. Introduction

Continuous and comprehensive monitoring of a patient's vital signs remains a critical aspect of chronic disease management in various healthcare settings [1]. The traditional approach to monitoring, mainly in clinical settings, poses significant challenges due to its intrusive nature and limited accessibility, thereby affecting the patient's comfort [2].

The emergence of non-intrusive monitoring systems offers a promising solution to these challenges, enabling unobtrusive and remote patient monitoring [3]. The integration of these systems directly into a patient's bed, whether in a hospital or home setting, holds immense potential for continuous and discreet monitoring of vital signs [2].

Among these vital signs, heart rate (HR) plays a central role as an indicator of several serious health conditions such as arrhythmia or bleeding [4]. Its importance extends to sleep-related assessments due to its close relationship with cardiovascular health and its utility in automated sleep stage classification [5, 6].

However, the gold standard for sleep pattern assessment, polysomnography (PSG), has limitations in terms of cost, complexity, and accessibility that limit its widespread use beyond clinical settings [7]. This gap has sparked interest in the development of cost-effective, user-friendly ambulatory monitoring systems tailored for home care [8].

Numerous methodologies exist for the acquisition of cardiac activity data [9]. A comprehensive review of non-contact cardiorespiratory monitoring systems can be found in [10]. Various approaches have been developed to measure ballistocardiography (BCG) signals during sleep, including different sensor types, numbers, and placement strategies. Some studies have used load cells [11], piezoelectric transducers [12] or pressure sensors [13], among other variations.

The exploration of various measurement techniques, including ballistocardiography (BCG) signals, demonstrates advances in sensors such as accelerometers powered by micro-electro-mechanical systems (MEMS) technology [14]. Despite these advances, the potential of accelerometers to capture BCG signals remains relatively underexplored and warrants further investigation [15].

This work attempts to investigate non-invasive HR monitoring using a contactless accelerometer-based system. The primary objective is to evaluate the effectiveness of the system in different sleeping positions and to validate its performance against PSG, with the goal of innovating unobtrusive vital signs monitoring in healthcare settings.

The following sections explain the components of the proposed system, including autonomy, data acquisition methodology, signal processing algorithm, and experimental design. Each facet has been carefully developed to ensure robustness, accuracy, and practical applicability in real-world scenarios, offering potential advancements in patient care and chronic disease management. After that, the obtained results

*This research was partially funded by Carl Zeiss Foundation and the MORPHEUS-Project "Non-invasive system for measuring parameters relevant to sleep quality" (project number: P2019-03-003).

M. Gaiduk is with the Ubiquitous Computing Lab at HTWG Konstanz, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (phone: +49 7531 206-703; Email: maksym.gaiduk@htwg-konstanz.de).

A. Boiko was with the Ubiquitous Computing Lab at HTWG Konstanz, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany.

M. Conti is with the Dip. di Ingegneria dell'Informazione Università Politecnica delle Marche, Ancona, Italy (Email: m.conti@univpm.it).

S. Orcioni is with the Dip. di Ingegneria dell'Informazione Università Politecnica delle Marche, Ancona, Italy (Email: s.orcioni@univpm.it).

N. Martínez Madrid is with the IoT Lab at Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany (Email: natividad.martinez@reutlingen-university.de).

R. Seepold is with the Ubiquitous Computing Lab at HTWG Konstanz, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (Email: ralf.seepold@htwg-konstanz.de).

are being discussed, and conclusions, as well as ideas for futurework, are presented.

II. Methodology

The autonomous functionality of our system stands as a fundamental prerequisite, demanding self-sufficiency in both hardware and software components. This independence facilitates seamless data collection and processing without reliance on external equipment, ensuring affordability, high accuracy, and suitability for home use [16].

A. Data Acquisition

The data acquisition segment includes two primary components: the mechanical structure (sensor mount and suspension) and the direct data acquisition. This mechanical unit consists of a bed frame mount and a spring steel plate (hanger) that houses the ADXL355z 3-axis accelerometer sensor [17]. This sensor, which has been validated for vital sign detection [18], provides a sampling rate of 62 Hz, which is considered sufficient for the objectives of this study. In addition, the ESP32 module was chosen for its compact size, cost-effectiveness, and storage capabilities using a micro-SD module along with Wi-Fi data transfer capabilities [19].

Strategic placement of the sensor is critical, and based on previous research on accelerometer applications in cardiorespiratory measurements, chest level emerged as the optimal position [20].

B. Signal Processing

The signal processing algorithm is an integral part of the contactless HR monitoring system and includes several steps for HR estimation. First, motion artifacts associated with sleep position changes are minimized by excluding the first and last 10 seconds of data for each position measurement. Next, offset reduction via mean signal amplitude subtraction and cross-axis signal magnitude calculation precedes BCG signal derivation through a Butterworth bandpass filter (6th order, range: 0.7 to 3.25 Hz) for HR detection. Further enhancement with a symlets6 wavelet improves the quality of the processed signal, facilitating HR estimation through peak detection. In general, wavelet symlets are widely used in biosignal processing to effectively handle complex biological signals with varying frequencies and transient characteristics, including ECG analysis and feature extraction in medical data [21].

B. Experiment Design

A standard wooden single bed was used in the experiment to ensure uniformity and reproducibility as well as standard foam mattress. Subjects were instructed to stay in four different sleeping positions - prone, right side, supine, and left side - designated P1 to P4, respectively. The experiment began in position P1 and ended in position P4, with subjects having a relaxation period of at least three minutes prior to data collection.

Data collection in each position lasted 140 seconds, during which subjects were instructed to minimize movement while behaving naturally. A total of 10 participants were included in the analysis. The experiment would be stopped immediately if the subjects felt uncomfortable, prioritizing their well-being, which fortunately did not happen during the experiment.

III. Results

A total of 10 subjects participated in the data collection phase, consisting of 5 males and 5 females, with a mean age of 31.8 ± 8.6 years, a mean height of 175.0 ± 6.0 cm, and a mean weight of 79.1 ± 12.9 kg. Prior to participation in the experiment, all subjects provided informed consent by signing the required documents and were thoroughly informed of the specifics of the research. None of the subjects reported any known medical conditions, diseases, ongoing treatments, or medication use.

Evaluation of the contactless HR monitoring system was performed by comparing the obtained HR values with reference data measured with a ECG system. Mean absolute error (MAE) was used as the primary evaluation metric. To facilitate HR analysis and MAE calculation, each signal epoch was segmented into 20-second intervals. Approximately 240 segments, totaling 4800 seconds, were considered for further analysis, consistent with the experimental design described in the preceding sections and subsections (60 segments for each sleep position). Table 1 shows the MAE values, which are essential for assessing the performance of the HR monitoring system.

TABLE I. MAE for Heart Rate Estimation

	Sleep Position			
	Prone (P1)	***Right lateral (P2)***	***Supine (P3)***	***Left lateral (P4)***
MAE, bpm	2.95	2.78	3.34	3.07

To assess the performance of the contactless system and to provide a comparative analysis with the ground truth, individual Bland-Altman analyses were performed for each sleeping posture. The results of these visualisations are presented visually in Figures 1-4 and allow detailed analysis.

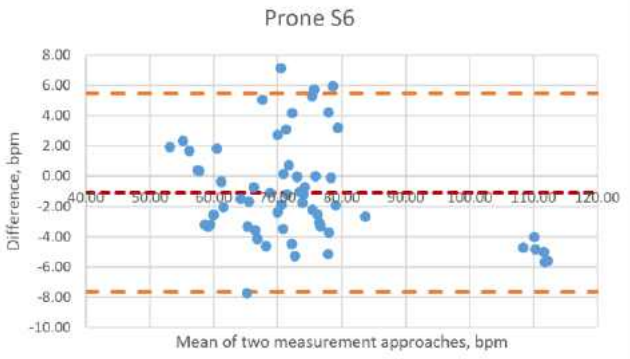


Figure 1. Bland-Altman plot for HR measurement in prone position applying symlets6 wavelet.

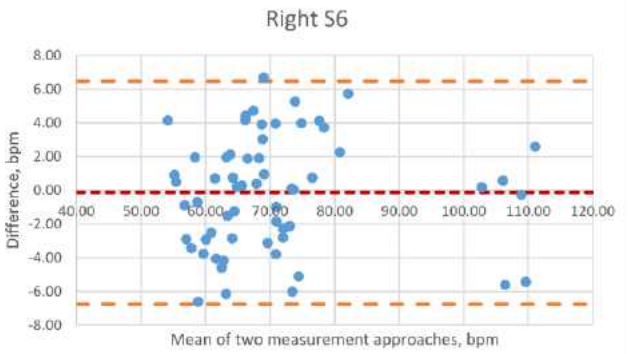


Figure 2. Bland-Altman plot for HR measurement in right lateral position applying symlets6 wavelet.

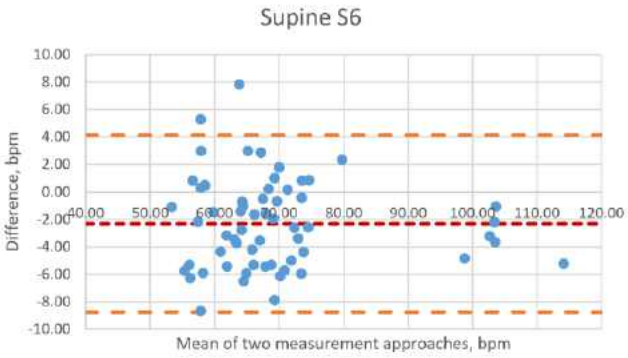


Figure 3. Bland-Altman plot for HR measurement in supine position applying symlets6 wavelet.

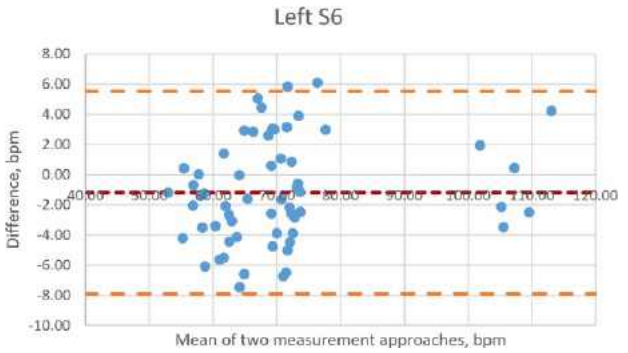


Figure 4. Bland-Altman plot for HR measurement in left lateral position applying symlets6 wavelet.

The Bland-Altman analysis revealed remarkably accurate HR measurements across all observed sleep positions. However, this commendable accuracy was mainly observed within a specific HR range of [50; 90] bpm. Cases in which the HR exceeded 90 bpm indicated a slightly higher deviation in the performance of the algorithm between the reference and contactless systems. However, it's important to note that only limited data was available within this higher HR range, precluding a definitive statement regarding this observed trend. Further data collection within this HR range is warranted to conclusively support these initial findings.

IV. Conclusion and Outlook

The findings suggest that the non-contact system using accelerometer sensors is promising for heart rate (HR) monitoring achieving the MAE of about 3 bpm. Comparable accuracy to existing methods was observed [22]. This system, which uses both hardware and software components, has the potential to be a modern, user-friendly approach with advantages over traditional technologies. It operates autonomously under a bed, taking measurements unobtrusively, suggesting its potential applications in hospital settings, rehabilitation centres and even home environments. Coupled with its processing algorithm, the system offers advantages such as simplicity, cost-effectiveness, and non-contact measurement capabilities.

Nevertheless, the study carried out has several limitations, which are listed below:

- The sample size includes a limited number of subjects, which may limit the diversity and representation of different demographic and physiological variations within the population.
- Relatively short measurement periods were used, which may not capture long-term trends or variations in physiological parameters.
- The study was conducted in a controlled laboratory setting, which may not fully replicate real-life conditions, thereby limiting the generalisability of the findings to real-life settings.
- The age range of the subjects was limited to 31.8 ± 8.6 years, which limits the applicability of the findings to broader age ranges and different demographic groups.

Efforts to address these limitations in future research include expanding the pool of subjects to include different age groups, heights, and various other characteristics. In addition, hardware improvements are being pursued to improve signal quality and signal-to-noise ratio. For example, proposed improvements include upgrading the mechanical support to mitigate hanger drift, thereby refining signal quality and data accuracy during patient movement.

In addition, future work aims to integrate this system and methodology into a comprehensive sleep monitoring framework, potentially in conjunction with algorithms proposed in [23]. This integration could provide a more comprehensive approach to sleep monitoring, potentially addressing some of the limitations observed in this study and broadening its applicability in real-world scenarios. This would allow for an improvement in sleep quality through the timely detection and treatment of disorders, thus leading to an enhancement in the overall health of the population [24].

Acknowledgment

This research was partially funded by Carl Zeiss Foundation and the MORPHEUS-Project “Non-invasive system for measuring parameters relevant to sleep quality” (project number: P2019-03-003).

References

[1] A. Rahaman, M. Islam, M. Sadi, and S. Nooruddin, "Developing IoT Based Smart Health Monitoring Systems: A Review," RIA, vol. 33, no. 6, pp. 435–440, 2019, doi: 10.18280/ria.330605.

[2] M. Conti, S. Orcioni, N. M. Madrid, M. Gaiduk, and R. Seepold, "A Review of Health Monitoring Systems Using Sensors on Bed or

Cushion," in Lecture Notes in Computer Science, Bioinformatics and Biomedical Engineering, I. Rojas and F. Ortuño, Eds., Cham: Springer International Publishing, 2018, pp. 347–358.

[3] S. M. Lee and D. Lee, "Healthcare wearable devices: an analysis of key factors for continuous use intention," Serv Bus, vol. 14, no. 4, pp. 503–531, 2020, doi: 10.1007/s11628-020-00428-3.

[4] P. Palatini and S. Julius, "Heart rate and the cardiovascular risk," Journal of hypertension, vol. 15, no. 1, pp. 3–17, 1997, doi: 10.1097/00004872-199715010-00001.

[5] E. Stamatakis, L. F. M. de Rezende, and J. P. Rey-López, "Sedentary Behaviour and Cardiovascular Disease," in Springer Series on Epidemiology and Public Health, Sedentary Behaviour Epidemiology, M. F. Leitzmann, C. Jochem, and D. Schmid, Eds., Cham: Springer International Publishing, 2018, pp. 215–243.

[6] M. Gaiduk, Á. Serrano Alarcón, R. Seepold, and N. Martínez Madrid, "Current status and prospects of automatic sleep stages scoring: Review," Biomedical engineering letters, vol. 13, no. 3, pp. 247–272, 2023, doi: 10.1007/s13534-023-00299-3.

[7] S. M. Caples, W. M. Anderson, K. Calero, M. Howell, and S. D. Hashmi, "Use of polysomnography and home sleep apnea tests for the longitudinal management of obstructive sleep apnea in adults: an American Academy of Sleep Medicine clinical guidance statement," Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, vol. 17, no. 6, pp. 1287–1293, 2021, doi: 10.5664/jcsm.9240.

[8] J. Wang, N. Spicher, J. M. Warnecke, M. Haghi, J. Schwartze, and T. M. Deserno, "Unobtrusive Health Monitoring in Private Spaces: The Smart Home," Sensors (Basel, Switzerland), vol. 21, no. 3, 2021, doi: 10.3390/s21030864.

[9] D. Fan, A. Ren, N. Zhao, D. Haider, X. Yang, and J. Tian, "Small-Scale Perception in Medical Body Area Networks," IEEE journal of translational engineering in health and medicine, vol. 7, p. 2700211, 2019, doi: 10.1109/JTEHM.2019.2951670.

[10] A. Boiko, N. Martínez Madrid, and R. Seepold, "Contactless Technologies, Sensors, and Systems for Cardiac and Respiratory Measurement during Sleep: A Systematic Review," Sensors (Basel, Switzerland), vol. 23, no. 11, 2023, doi: 10.3390/s23115038.

[11] A. R. Malik and J. Boger, "Zero-Effort Ambient Heart Rate Monitoring Using Ballistocardiography Detected Through a Seat Cushion: Prototype Development and Preliminary Study," JMIR rehabilitation and assistive technologies, vol. 8, no. 2, e25996, 2021, doi: 10.2196/25996.

[12] W. Xu, C. Yu, B. Dong, Y. Wang, and W. Zhao, "Thin Piezoelectric Sheet Assisted PGC Demodulation of Fiber-Optic Integrated MZI and its Application in Under Mattress Vital Signs Monitoring," IEEE Sensors J., vol. 22, no. 3, pp. 2151–2159, 2022, doi: 10.1109/JSEN.2021.3128601.

[13] J. Gomez-Clapers, A. Serra-Rocamora, R. Casanella, and R. Pallas-Areny, "Towards the standardization of ballistocardiography systems for J-peak timing measurement," Measurement, vol. 58, pp. 310–316, 2014, doi: 10.1016/j.measurement.2014.09.003.

[14] Y. D'Mello et al., "Real-Time Cardiac Beat Detection and Heart Rate Monitoring from Combined Seismocardiography and Gyrocardiography," Sensors (Basel, Switzerland), vol. 19, no. 16, 2019, doi: 10.3390/s19163472.

[15] F. Sana, E. M. Isselbacher, J. P. Singh, E. K. Heist, B. Pathik, and A. A. Armoundas, "Wearable Devices for Ambulatory Cardiac Monitoring: JACC State-of-the-Art Review," Journal of the American College of Cardiology, vol. 75, no. 13, pp. 1582–1592, 2020, doi: 10.1016/j.jacc.2020.01.046.

[16] M. Gaiduk, R. Seepold, N. Martínez Madrid, and J. Ortega, "Digital Health and Care Study on Elderly Monitoring," Sustainability, vol. 13, no. 23, p. 13376, 2021, doi: 10.3390/su132313376.

[17] A. Boiko et al., "Monitoring of Cardiorespiratory Parameters during Sleep Using a Special Holder for the Accelerometer Sensor," Sensors (Basel, Switzerland), vol. 23, no. 11, 2023, doi: 10.3390/s23115351.

[18] A. Boiko, M. Gaiduk, R. Seepold, and N. M. Madrid, "Accelerometer based system for unobtrusive sleep apnea detection," Procedia Computer Science, vol. 225, pp. 1592–1600, 2023, doi: 10.1016/j.procs.2023.10.148.

[19] S. Hersek, B. Semiz, M. M. H. Shandhi, L. Orlandic, and O. T. Inan, "A Globalized Model for Mapping Wearable Seismocardiogram Signals to Whole-Body Ballistocardiogram Signals Based on Deep Learning," IEEE journal of biomedical and health informatics, vol. 24, no. 5, pp. 1296–1309, 2020, doi: 10.1109/JBHI.2019.2931872.

[20] J. Skoric et al., "Respiratory Modulation of Sternal Motion in the Context of Seismocardiography," IEEE Sensors J., vol. 22, no. 13, pp. 13055–13066, 2022, doi: 10.1109/JSEN.2022.3173205.

[21] M. R. Kaloop, C. O. Yigit, A. El-Mowafy, A. A. Dindar, M. Bezcioglu, and J. W. Hu, "Hybrid Wavelet and Principal Component Analyses Approach for Extracting Dynamic Motion Characteristics from Displacement Series Derived from Multipath-Affected High-Rate GNSS Observations," Remote Sensing, vol. 12, no. 1, p. 79, 2020, doi: 10.3390/rs12010079.

[22] A. Ullal et al., "Non-invasive monitoring of vital signs for older adults using recliner chairs," Health Technol., vol. 11, no. 1, pp. 169–184, 2021, doi: 10.1007/s12553-020-00503-9.

[23] M. Gaiduk et al., "Estimation of Sleep Stages Analyzing Respiratory and Movement Signals," IEEE journal of biomedical and health informatics, vol. 26, no. 2, pp. 505–514, 2022, doi: 10.1109/JBHI.2021.3099295.

[24] J. Spiesshoefer et al., "Sleep - the yet underappreciated player in cardiovascular diseases: A clinical review from the German Cardiac Society Working Group on Sleep Disordered Breathing," European journal of preventive cardiology, vol. 28, no. 2, pp. 189–200, 2021, doi: 10.1177/2047487319879526.

Influence of gender and age distinction on patient data for sleep apnea detection using artificial intelligence models

Angel Serrano Alarcon, *Reutlingen University* and Natividad Martinez Madrid*, Reutlingen University* and Ralf Seepold, *HTWG Konstanz*

***Abstract*— The massive use of patient data for the training of artificial intelligence algorithms is common nowadays in medicine. In this scientific work, a statistical analysis of one of the most used datasets for the training of artificial intelligence models for the detection of sleep disorders is performed: sleep health heart study 2. This study focuses on determining whether the gender and ageof the patients have a relevant influence to consider working with differentiated datasets based on these variables for the training of artificial intelligence models.**

I. INTRODUCTION

Nowadays, the development of artificial intelligence algorithms has invariably influenced many fields of knowledge in industry and science [1]. In medicine, the application of artificial intelligence models for detecting and monitoring physiological events is becoming increasingly common [2-4]. These algorithms can handle both tabular data and complex data, such as time series. They are not onlyused to detect and monitor diseases but also medicalsolutions aimed at predicting these events to improvethe results of treatment, surgery, or simply the prevention of more serious diseases [5].

Among the problems faced by using artificial intelligence in medicine is the well-known problem of blackboxes, that is, the lack of transparency that the model providesafter making the prediction [6]. On the other hand, a large number of devices apply artificial intelligence after collecting data from people who may or may not be sick people, along with the small number that these devices represent as certified medical devices [7]. Apart from these major drawbacks, if we go deeper into how these artificial intelligence models work, we will see that not always the amount of data used is sufficient, the quality of the data is farfrom ideal, and on the other hand, the null distinction between genders and age ranges when artificial intelligence algorithmsare used for the detection of physiological events of interest [8-9]. This is even more worrisome when it comes to detecting and monitoring events related to sleep disorders [10].

Sleep disorders are highly correlated with age and the psychological and physical changes that people undergo as they get older. If we look at the scientific literature, there is practically no distinction between men and women when it comes to the application of artificial intelligence algorithms in detecting sleep-related diseases [11-14].

Based on the above, this scientific work aims to analyze patients suffering from sleep apnea to study whether there are major differences between men and womenalong with age. In this way, it can be verified if it is necessaryto adapt the artificial intelligence algorithms to work with gender and age ranges. For this purpose, 2651 patients from study 2 of the global sleep health heart study have been used. However, artificial intelligence is not applied throughout thisscientific work.

The study of the influence of age and gender on the severity of sleep apnea that the patient may have is something very relevant and should be clarified prior to the training of the artificial intelligence models. Therefore, at this initial stage of the study of this influence, no results are shown with deep learning models trained with gender bias in the data.

II. MATERIALS AND METHODS

For data analysis, data were obtained from the National Sleep Research Resource (NSRR), specifically from the repository Sleep Heart Health Study (SHHS). The Sleep Heart Health Study (SHHS) is a multicenter cohort study initiated by the National Heart Lung & Blood Institute to determine sleep-related breathing disorders' cardiovascular and other consequences. In total, 6,441 men and women aged 40 years or older were enrolled between November 1, 1995, and January 31, 1998, to participate in SHHS Visit 1. During examination cycle 3 (January 2001 to June 2003), a second polysomnogram (SHHS Visit 2) was obtained in 3,295 of the participants [15].

Once the data were obtained, we worked with a Python programming language environment in Jupyter Notebook. Python libraries such as Numpy, Pandas, Scipy, matplotlib, seaborn, etc. were used for data analysis. Prior to the visualization of the results, the data was cleaned and processed. For this work, the features of age, gender, and apnea-hypopnea index were selected.

* This research was funded by ZiM project “Sleep Lab at Home” (SLaH) grant: ZF4825301AW9 and Carl Zeiss Foundation and the MORPHEUS-Project “Non-invasive system for measuring parameters relevant to sleep quality” (project number: P2019-03-003).

Ángel Serrano Alarcón is with the IoT Lab at Reutlingen University, Alteburgstr, 150, 72762 Reutlingen, Germany (Email: angel.serrano_alarcon@reutlingen-university.de)

N. Martínez Madrid is with the IoT Lab at Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany (Email: natividad.martinez@reutlingen-university.de).

R. Seepold is with the Ubiquitous Computing Lab at HTWG Konstanz, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (Email: ralf.seepold@htwg-konstanz.de)

In order to conclude this study, the inclusionof graphs with statistical information is essential. This is because the large number of patients used is too large to simply use descriptive statistics without providing any visualization technique. As this study does not intend to go further with the study of model training outcomes with data divided between males and females, this study aims to understand how the sex and age of patients influence AHI. Therefore, graphs that include populations and AHI indices may yield deterministic information about the influence on sleep apnea detection.

III. Results

As can be seen in Figure 1, the age range for men between 55 and 65 years constitutes the largest population of men suffering from sleep apnea of any severity. However, for women, this is not the case, and the severity of apnea is more pronounced for women between 65 and almost 80 years of age. This fact does not imply that sleep apnea is accentuated for women as they age. However, it can be stated that womenolder than 70 who suffer from apnea are more likely to havemore apnea events.

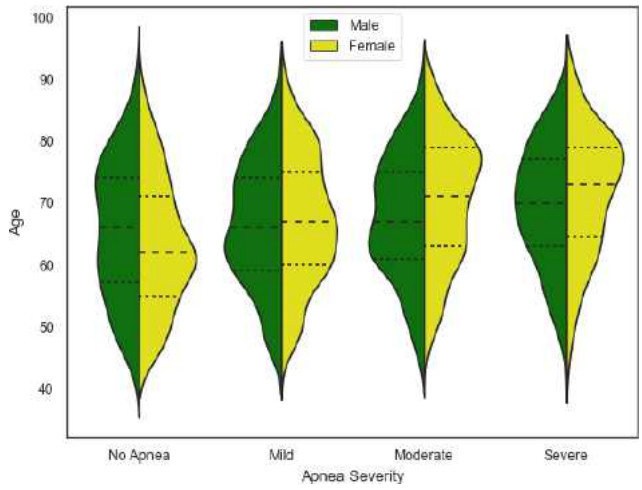


Figure 1. Violin plot of the age and gender distribution of shhs2. The dashed line indicates the mean, and the dotted line is the standard deviation.

In order to visualize the influence of age and gender on the detection of sleep apnea, a basic descriptive statistical analysis was performed, followed by several visualizations to study the data better. One aspect to remember is that the patient data used are age, apnea-hypopnea-index (AHI), and gender. In this sense, the age of the patients is 40 years and older, so there is a large bias in the data. Since we do not have data for patients under 40, this analysis does not represent a large part ofthe population. Therefore, despite not finding evidence to justify an increase in AHI in patients over the years, it cannot beaffirmed that this is not the case due to the drawback above.

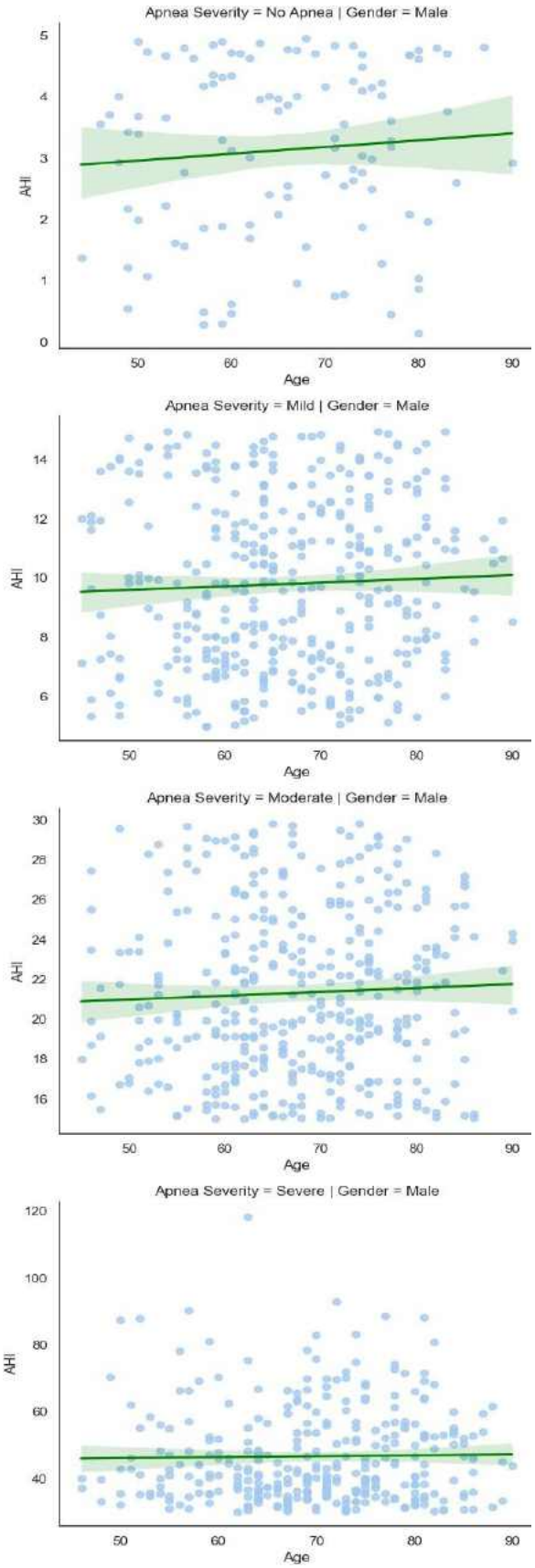


Figure 2. Scatter plot and Regression of the age and gender distribution of SHHS-2. The dashed line indicates the mean and the dotted line the standard deviation.

As seen in Figures 2 and 3, there is no evidence of gender predominance as a determinant in the development ofsleep apnea. Apart from this, no scientific evidence shows that the older the person is, the greater the number of apnea events he/she achieves.

Figures 1 and 2 show a certain tendency for men and women to have a higher AHI as age increases. However, it cannot be affirmed that this is always the case, and the positive correlation is not considered significant enough to confirm it. Apart from this, something that has already beenmentioned is that SHHS-2 contains data from patients older than 40, which implies a large range of populations aged between 18 and 40 years that have not been included in this study. Patients younger than 18 years are considered children. Artificial intelligence models shoulddiscriminate sleep apnea events regardless of patient age andgender based on the information included in this manuscript. Age and gender should not influence the decision-making of artificial intelligence models. Another aspect to consider is that some of the patients had other pathologies, which may act as a bias for decision-making based on the study proposed here.

In this matter, it seems that other decisions to be made before feeding the artificial intelligence algorithm for sleep apnea detection are more relevant than gender and age: the duration of the apnea events, whether the algorithms are trained with time windows of a certain duration or with complete time series, whether the patients suffer from other pathologies, the quality of the signals, number of artifacts, etc.

A priori, the information extracted from this study may be relevant in developing artificial intelligence models for the detection of sleep apnea and its subsequent implementation in a portable monitor or clinical detection system. In this way, the inconveniences that exist during detection could be reduced. The fact that gender and age do not have a great influence when detecting sleep apnea events facilitates the development of these artificial intelligence models, which reduces the complexity of the process and subsequent decision-making by the doctors.

Finally, it is worth highlighting that in order to clarify the fact that age and gender do not have a great impact on the development of artificial intelligence models for the detection of sleep apnea, more studies and a greater number of data with broader age ranges.

IV. I. CONCLUSION

Although the development of models is becoming increasingly common in the development of devices that detect sleep apnea; there is still insufficient evidence to demonstrate that it is unnecessary to classify training data by gender and age. The analysis shown here should be supported by future research. However, the large number of patient data used is a large enough sample to take the results into account. It cannot be stated that sleep apnea worsens with age and that there are large differences between men and women. Therefore, no gender and age distinction should be made in training artificial intelligence models.

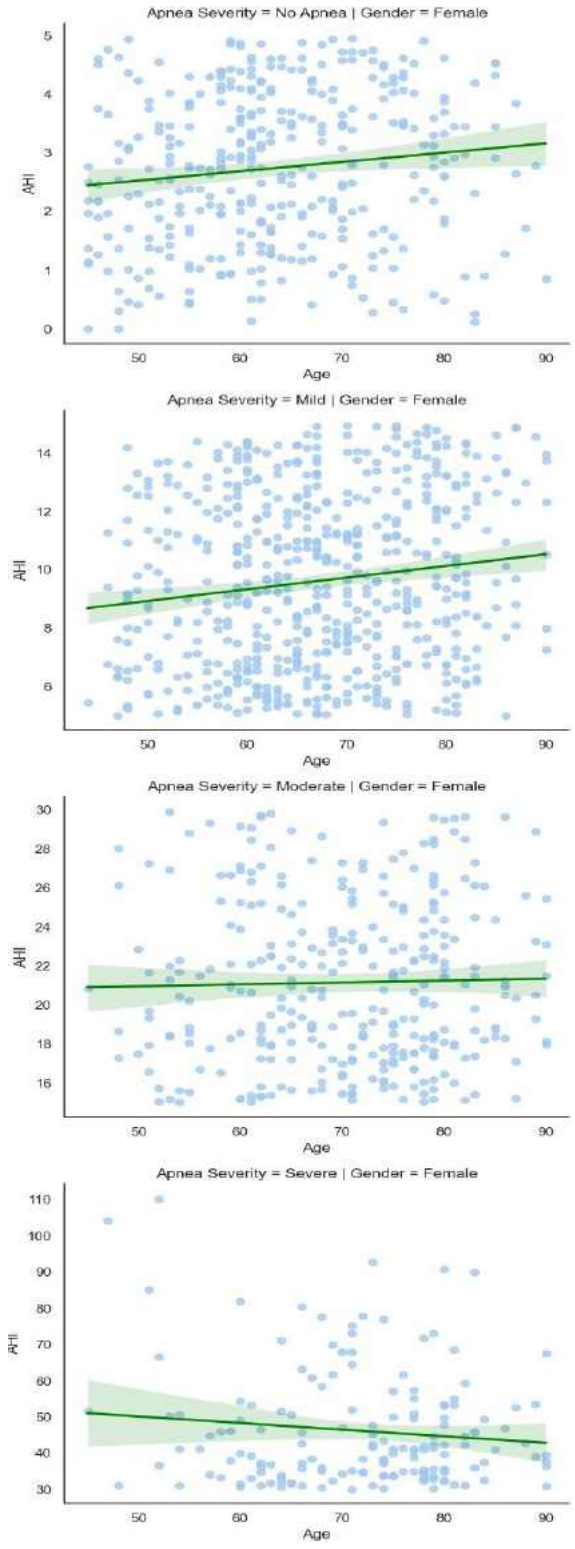


Figure 3. Scatter plot and Regression of the age and gender distribution of SHHS-2. The dashed line indicates the mean and the dotted line the standard deviation.

REFERENCES

[1] Biswal, S., Sun, H., Balaji Goparaju, M., Westover, B., Sun, J., and Bianchi, M. T. (2018). Expert-level sleep scoring with deep neural networks. J. Am. Med. Inform. Assoc. 25, 1643–1650. doi: 10.1093/jamia/ocy131

[2] Thorey, Valentin, Hernandez, Albert Bou, Arnal, Pierrick J., and During, Emmanuel H. (2019). "AI vs humans for the diagnosis of sleep apnea", Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

[3] Gaube, S., Suresh, H., Raue, M., Merritt, A., Berkowitz, S. J., Lermer, E., et al. (2021). Do as AI say: susceptibility in deployment of clinical decision-aids. Npj Digital Medicine 2021 4:1 4, 1–8. doi: 10.1038/s41746-021-00385-9

[4] Goldstein, C. A., Berry, R. B., Kent, D. T., Kristo, D. A., Seixas, A. A., Redline, S., et al. (2020). Artificial intelligence in sleep medicine: background and implications for clinicians. J. Clin. Sleep Med. 16, 609–618. doi: 10.5664/JCSM.8388

[5] Mendonça, F., Mostafa, S. S., Ravelo-García, A. G., Morgado-Dias, F., and Penzel, T. (2019). A review of obstructive sleep apnea detection approaches. IEEE J. Biomed. Health Inform. 23, 825–837. doi: 10.1109/JBHI.2018.2823265

[6] Wang, F., Kaushal, R., and Khullar, D. (2020). Should health care demand interpretable artificial intelligence or accept 'black box' medicine? Ann. Intern. Med. 172, 59–61. doi: 10.7326/M19-2548

[7] Alarcón, S., Ángel, N. M. M., and Seepold, R. (2021). A minimum set of physiological parameters to diagnose obstructive sleep apnea syndrome using non-invasive portable monitors. A systematic review. Life 11:1249. doi: 10.3390/LIFE11111249

[8] Zemouri, R., Zerhouni, N., and Racoceanu, D. (2019). Deep learning in the biomedical applications: recent and future status. Appl. Sci. 9:1526. doi: 10.3390/APP9081526

[9] Chollet, F. (2021). Deep learning with Python. 2nd Edn. Manning

[10] Chaw, H. T., Kamolphiwong, S., and Wongsritrang, K. (2019). Sleep apnea detection using deep learning. Tehnički Glasnik 13, 261–266. doi: 10.31803/tg-20191104191722

[11] Chen, X., Wang, R., Zee, P., Lutsey, P. L., Javaheri, S., Alcántara, C., et al. (2015). Racial/ethnic differences in sleep disturbances: the multi-ethnic study of atherosclerosis (MESA). Sleep 38, 877–888. doi: 10.5665/sleep.4732

[12] Fawaz, I., Hassan, G. F., Weber, J., Idoumghar, L., and Muller, P. A. (2019). Deep learning for time series classification: a review. Data Min. Knowl. Disc. 33, 917–963. doi: 10.1007/s10618-019-00619-1

[13] Gjevre, J. A., Taylor-Gjevre, R. M., Skomro, R., Frcpc, M. D., Reid, J., Fenton, M., et al. (2011). Comparison of polysomnographic and portable home monitoring assessments of obstructive sleep apnea in Saskatchewan women. Can. Respir. J. 18, 271–274. doi: 10.1155/2011/408091

[14] Kim, D., Lee, J., Woo, Y., Jeong, J., Kim, C., and Kim, D. K. (2022). Deep learning application to clinical decision support system in sleep stage classification. J. Person. Med. 12:136. doi: 10.3390/JPM12020136

[15] Zhang, G. Q., Cui, L., Mueller, R., Tao, S., Kim, M., Rueschman, M., et al. (2018). The National Sleep Research Resource: towards a sleep data commons. J. Am. Med. Inform. Assoc. 25, 1351–1358. doi: 10.1093/jamia/ocy064

Assessing Body Position During Sleep Using FSR Sensors and Machine Learning Algorithms

Akhmadbek Asadov, Hochschule Konstanz and Juan Antonio Ortega, Seville University and
Natividad Martínez Madrid, Reutlingen University and Ralf Seepold, HTWG Konstanz

***Abstract*— This study investigates the application of Force Sensing Resistor (FSR) sensors and machine learning algorithms for non-invasive body position monitoring during sleep. Although reliable, traditional methods like Polysomnography (PSG) are invasive and unsuited for extended home-based monitoring. Our approach utilizes FSR sensors placed beneath the mattress to detect body positions effectively. We employed machine learning techniques, specifically Random Forest (RF), K-Nearest Neighbors (KNN), and XGBoost algorithms, to analyze the sensor data. The models were trained and tested using data from a controlled study with 15 subjects assuming various sleep positions. The performance of these models was evaluated based on accuracy and confusion matrices. The results indicate XGBoost as the most effective model for this application, followed by RF and KNN, offering promising avenues for home-based sleep monitoring systems.**

I. Introduction

Cardiorespiratory monitoring is essential for evaluating sleep quality and diagnosing and managing sleep disorders. Polysomnography (PSG) [1,2] has traditionally been considered the most reliable assessment method, providing comprehensive insights into sleep-related cardiorespiratory functions. However, its invasive nature, cost, and complexity pose significant limitations, particularlyfor extended home-based monitoring.

Recent research has increasingly focused on using pressure sensors placed beneath the mattress as a non-invasive, cost-effective, and user-friendly method for monitoring heartbeat and breathing rates [3, 4]. These sensors are pivotal in the evolving domain of home health monitoring. The potential of wearable sensors as a viable alternative to Polysomnography (PSG) in evaluating sleep quality, particularly in conditions such as obstructive sleep apnea, is underscored.

This study further explores the application of Force Sensing Resistor (FSR) sensors [5], arranged under the bed in a triangular configuration [6, 7] to detect major body positions optimally. The research focuses on predicting sleeping positions using machine learning techniques, specifically the Random Forest (RF), K-Nearest Neighbors (KNN), and XGBoost algorithms, based on sensor data. The performance of these models is evaluated using accuracy as the key metric. The investigation involves meticulous stages of data preprocessing, processing, and analysis to make a significant contribution to the field of non-invasive sleep monitoring technologies and methods.

RF is an ensemble learning method primarily used for classification and regression [8]. It operates by constructing a multitude of decision trees at training time and outputting the class, that is, the mode of the classes (classification) or mean prediction (regression) of the individual trees. Each decision tree in a Random Forest splits the data into branches to make predictions, which are then aggregated to produce a more accurate and stable prediction. Random Forest is particularly effective due to its ability to handle large datasets with higher dimensionality and can estimate missing data while maintaining accuracy when a large proportion of the data is missing.

KNN is a simple, non-parametric algorithm used for classification and regression [9, 10]. In KNN, the input consists of the k closest training examples in the feature space, where k is a positive integer, typically small. The output is a class membership for classification problems: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors. KNN is inherently a lazy learning algorithm, meaning it does not build a model explicitly but stores training data instances. Classification is computed from a simple majority vote of the k nearest neighbors of each point.

XGBoost, short for Extreme Gradient Boosting, is an efficient and scalable implementation of a gradient boosting framework [11]. It utilizes gradient-boosted decision trees designed for speed and performance. XGBoost provides parallel tree boosting that solves many data science problems quickly and accurately. The core principle behind XGBoost is to build a series of weak learners (decision trees) sequentially, where each tree attempts to correct its predecessor's mistakes. The model adjusts for the errors of the previous trees in the series through a process called boosting,and the combination of these weak learners results in improved accuracy and robustness.

II. Materials and Methods

A. Hardware setup

Our research utilized a hardware setup designed for precise data collection using FSR sensors for sleep monitoring. The

Akhmadbek Asadov – Corresponding author, HTWG Konstanz – University of Applied Sciences, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (Akhmadbek.Asadov@htwg-konstanz.de).

Juan Antonio Ortega, University of Sevilla, Av. de la Reina Mercedes, s/n, 41012 Sevilla, Spain (jortega@us.es).

Natividad Martínez Madrid, Reutlingen University, Alteburgstr. 150,72762 Reutlingen, Germany (Natividad.Martinez@Reutlingen- University.DE)

Ralf Seepold, HTWG Konstanz – University of Applied Sciences, Alfred-Wachtel-Str. 8, 78462 Konstanz, Germany (ralf.seepold@htwg-konstanz.de)

core of this system consisted of three FSR 406 sensors. These sensors were strategically placed under the mattress in three positions -right, left, and center- to capturemovements and pressure variations during sleep effectively.The system also included electronic components for signalamplification and conversion [12]. This involved twoamplification boards and an Analog-to-Digital Converter(ADC) board, which converted analog signals from the sensorsto digital format for easier processing. An IICinterfacing converter also connected the ADC to a Raspberry Pi 4B,chosen for its efficient processing capabilities [13]. The circuitry featured power supply stabilization, a voltage dividercircuit with amplifier gain, and an active low-pass filter, essential for the system's stable operation [14]. An embeddedsystem within the hardware setup was responsible for data collection and storage, ensuring systematic recording for subsequent analysis.

B. Experiment design

Our study aimed to gather 15 subjects to assess the accuracy of FSR sensors in sleep monitoring. During the sessions, participants assumed four common sleep positions: prone, right-lateral, supine, and left-lateral. Data was recorded for 20 minutes, ensuring comprehensive datasets. Participant comfort was prioritized, and sessions were terminated immediately if any discomfort was reported, ensuring participant safety and the integrity of the collected data.

Data was collected using FSR sensors at a sampling rate of 160 Hz.

C. Data Preparation and Model Evaluation

In the preprocessing stage, the collected data was cleansed of motion artifacts. As part of data preprocessing, a windowing technique was employed to structure the dataset effectively for the machine learning models. A window length of 100 data points was chosen. This approach involved segmenting the data into overlapping windows, each encapsulating 100 sequential datapoints.

For each window, we removed columns irrelevant to our analysis, such as time stamps, and appended the target variable indicating the sleeping position. The resulting data frame comprised columns named to reflect the value (v) and time point (t) within each window, leading to a format conducive for machine learning analysis. This method ensured that each window was treated as a distinct observation, aligning with the target position, facilitating more accurate modeling and prediction.

In the next data preparation stage, the datasets were divided into training and testing sets to evaluate the machinelearning models effectively. The training set included data from 12 individuals, providing a robust basis for the models to learn and adapt to various sleep position patterns. The testing set, composed of data from 3 individuals, was used to assess the model's generalization capabilities and accuracy in unseen data scenarios.

Before model training, the data underwent standardization using the StandardScaler. This step normalized the features within the dataset, ensuring that each feature contributed equally to the model training and reducing potential biasfrom varying scales of data values.

For the machine learning models, we employed Random Forest with 50 estimators, K-Nearest Neighbors with 30 neighbors, and XGBoost. Each model was trained on the designated training dataset, encompassing data from 12 participants. The training process involved adjusting the models to recognize patterns and correlations pertinent to sleep position detection. Upon completing the training phase, the models were tested using a separate testing dataset from 3 individuals.

In the final stage of our analysis, we evaluated the performance of the machine learning models using two key metrics: accuracy and the confusion matrix. Accuracy provided a straightforward measure of the model's overall ability to correctly predict the sleep positions, calculated as the proportion of correct predictions to the total number of predictions made. On the other hand, the confusion matrix offered a more nuanced view by illustrating the specific types of errors made by the models. This matrix displayed the number of true positive, true negative, false positive, and false negative predictions for each sleep position, thereby giving insight into the models' strengths and weaknesses in different scenarios.

III. RESULTS

In the results section of this study, we critically analyze the performance metrics of three machine learning models: RF, KNN, and XGBoost. The evaluation focuses on the accuracy of each model in correctly classifying sleep positions, a fundamental aspect of our objective to advance sleep monitoring technologies.

An accuracy comparison table (Table 1) summarizes the percentage of correct predictions made by each model. This quantitative analysis clearly indicates which model most effectively discerns between different sleeping positions, a key consideration in the practical application of these models.

Table 1. Comparison table for models with their accuracy.

Model	**Accuracy**
RF	57.74
KNN	50.75
XGBoost	60.05

The Random Forest model, with an accuracy of 57.74%, shows a moderate level of effectiveness in predicting sleep positions. While it offers better accuracy than KNN in our study, it is not the most accurate model among the three. This performance suggests that while the ensemble method of Random Forest adds some robustness to the model, there may be complexities or nuances in the data it is not fully capturing.

KNN's performance, with an accuracy of 50.75%, is the lowest among the three models.

This might indicate that the simplistic approach of KNN, which relies on proximity in the feature space for classification, may not be sufficiently sophisticated for the complexities of sleep position data.

The choice of k (number of neighbors) and the distance metric used might also affect its performance. XGBoost shows the highest accuracy at 60.05%, indicating that its gradient-boosting approach is more effective. The ability of XGBoost to build upon the errors of previous trees and its sophisticated handling of various data conditions might contribute to its superior performance in our study.Furthermore, we delve into the detailed outcomes of the models through confusion matrices (Fig 1). These matrices offer a granular view of the predictive performance, demonstrating the overall accuracy and how eachmodel differentiates between the four classified sleeppositions: prone, right-lateral, supine, and left-lateral. By presenting a confusion matrix for each model, we gain insights into the specific areas where each model excels or falters, providing a comprehensive understanding of their capabilities and limitations in sleep position classification.

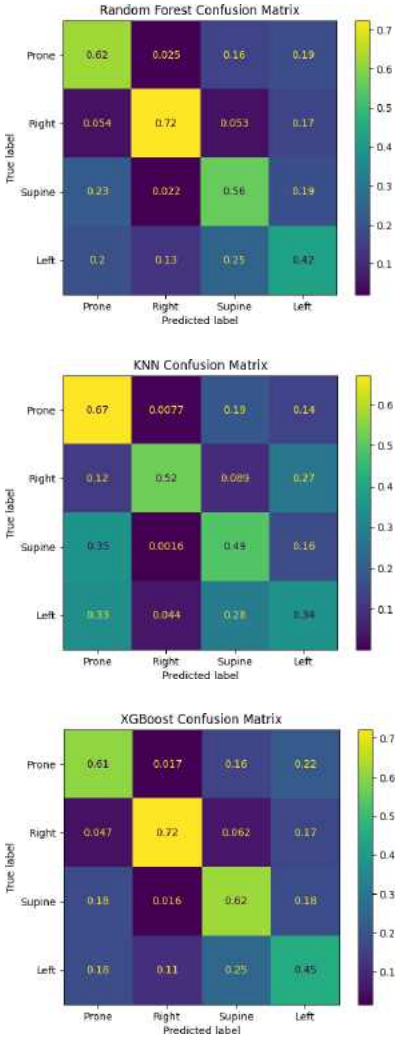


Figure 1. Confusion Matrix for RF, KNN, XGBoost

The confusion matrices provided represent the performance of three machine learning models – Random Forest, K-Nearest Neighbors, and XGBoost – for classifying sleep positions into four categories: prone, right (right-lateral), supine, and left (left-lateral).

The confusion matrix for the Random Forest model shows the highest correct classification rate for the right-lateral position, with notable weakness in distinguishing betweenthe prone and left-lateral positions. This suggests that the RF model may be confusing these positions more frequently.The KNN model displays a spread of classification accuracyacross all positions with notably lower performance than the RF model. The prone and left-lateral positions have thehighest misclassification rates, indicating that the KNNalgorithm struggles to differentiate between these positions accurately.

Lastly, the XGBoost model presents a more balanced classification with the highest accuracy in the right-lateral position classification. However, like the other models, XGBoost also shows some confusion between prone and left-lateral positions, though to a lesser extent.These results suggest that in the context of our study, XGBoost is the most suitable model for predicting sleep positions, followed by Random Forest, with KNN being the least effective. It is important to consider that model performance can also be influenced by factors like data quality, feature selection, and hyperparameter tuning, which might need to be explored further for potential improvements.

IV. Discussion

In this section, we reflect on the various aspects of our study, considering both the methodologies employed and the results obtained, and discuss potential avenues for future research.

Firstly, the participant pool size was limited. While this was sufficient for an initial investigation, a broader dataset encompassing a more diverse range of individuals would enhance the generalizability of the findings. Expanding the sample size could provide a more comprehensive understanding of the models' performance across a wider spectrum of physiological variations inherent in a larger population.

Secondly, the role of the window size in data preprocessing merits further exploration. The window length during the windowing process is a critical parameter that impacts the resolution of the time-series data. Adjusting the window size could potentially refine the models' ability to capture relevant features, and it would be valuable to investigate the effects of varying this parameter on model performance.

The potential of advanced feature engineering alsopresents an exciting opportunity for improving model accuracy. Incorporating statistical measures, frequency domain features, time-domain features, wavelet-based features, and signal-shape features could provide deeper insights into the data. Conducting further research with these additional features would allow us to evaluate the extent to which they enhance the predictive capability of the models.

Moreover, fine-tuning the hyperparameters of the machine learning algorithms is suggested. For the KNN model, experimenting with the number of neighbors (n_neighbors), and for the RF model, adjusting the number of trees (n_estimators) could lead to optimized performance.
Determining the optimal values for these parameters is crucial for improving accuracy and should be a focal pointof subsequent studies.

Lastly, transitioning towards deep learning models, such as convolutional neural networks (CNNs) and long short-term memory networks (LSTMs), could provide substantial benefits. These models have demonstrated exceptional capabilities in capturing complex patterns in high-dimensional data, which could be particularly beneficial for sleep position classification tasks. Future research should explore the application of these deep learning architectures and compare their performance against the current models.

V. Conclusion

This study aimed to improve the precision of non-invasive sleep monitoring by using machine learning models to predict sleep positions from pressure sensor data. The study conducted accurate experimentation and analysis, revealing valuable insights into various algorithms' capabilities and limitations.

The findings suggest that although models such as Random Forest, KNN, and XGBoost can provide acceptable accuracy, there is significant opportunity for improvement. Our investigation into the effects of data preprocessing, feature engineering, and model hyperparameters has established a basis for future research.

Expanding the dataset and incorporating more advanced machine learning and deep learning techniques has the potential further to enhance the accuracy and reliability of sleep position classification. Integrating advanced feature extraction methods and fine-tuning model parameters is a promising direction for future studies.

Acknowledgment

This research was funded by the Carl Zeiss Foundation as part of the MORPHEUS project "Non-invasive system for measuring parameters relevant to sleep quality" (project number: P2019-03-003).

References

[1] Hall, Andrew. "Sleep physiology and the perioperative care of patients with sleep disorders." BJA Education 15.4 (2015): 167-172.

[2] Electronic resource. Psychiatry Data Base. Polysomnography (PSG) Access:https://www.psychdb.com/neurology/polysomnography#polysomnography-psg. Date of access: 10.04.2023. Author name / Procedia Computer Science 00 (2019) 000–000 9

[3] Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C., & Vaughn, B. V. (2012). The AASM manual for the scoring of sleep and associated events. Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine, 176, 2012.

[4] Ibáñez, Vanessa, Josep Silva, and Omar Cauli. "A survey on sleep assessment methods." PeerJ 6 (2018): e4849.

[5] Gaiduk, M., Kuhn, I., Seepold, R., Ortega, J. A., & Martínez Madrid, N. (2017, April). A sensor grid for pressure and movement detection supporting sleep phase analysis. In International Conference on Bioinformatics and Biomedical Engineering (pp. 596-607). Springer, Cham.

[6] Seepold, Ralf, et al. "Identifying an Appropriate Area to Facilitate the Cardiorespiratory Measurement during Sleep Monitoring." 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2023.

[7] Haghi, Mostafa, et al. "A triangle-shape region of interest in cardiorespiratory estimation during sleep monitoring." Current Directions in Biomedical Engineering. Vol. 9. No. 1. De Gruyter,2023.

[8] Breiman, Leo. "Random forests." Machine learning 45 (2001): 5-32.

[9] Guo, Gongde, et al. "KNN model-based approach in classification." On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings. Springer Berlin Heidelberg, 2003.

[10] Zhang, Zhongheng. "Introduction to machine learning: k-nearest neighbors." Annals of translational medicine 4.11 (2016).

[11] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." Proceedings of the 22nd acm sigkdd international conferenceon knowledge discovery and data mining. 2016.

[12] Asadov, Akhmadbek, et al. "Performance improvement of cardiorespiratory measurements using pressure sensors with mechanical coupling techniques." Procedia Computer Science 225(2023): 1891-1899.

[13] Asadov, Akhmadbek, et al. Non-invasive cardiorespiration monitoring using force resistive sensor. Hochschule Reutlingen, 2022.

[14] Boiko, Andrei, et al. "Gamification system to improve the personalhealth of bedridden patients in long-term care." German-Italian Workshop Social Innovation in Long-Term Care through Digitalization. Cham: Springer International Publishing, 2021.

Comparative Study of Applying Signal Processing Techniques on Ballistocardiogram in Detecting J-Peak using Bi-LSTM Model*

Oluwaseun Awonuga, Priyanka Chaurasia, *Member, IEEE*, Maksym Gaiduk, *Member, IEEE,* Natividad Martínez Madrid, *Senior Member, IEEE*, Ralf Seepold, *Senior Member, IEEE*, and Mostafa Haghi, *Member, IEEE*

***Abstract*— Cardiovascular diseases (CVD) are leading contributors to global mortality, necessitating advanced methods for vital sign monitoring. Heart Rate Variability (HRV) and Respiratory Rate, key indicators of cardiovascular health, are traditionally monitored via Electrocardiogram (ECG). However, ECG's obtrusiveness limits its practicality, prompting the exploration of Ballistocardiography (BCG) as a non-invasive alternative. BCG records the mechanical activity of the body with each heartbeat, offering a contactless method for HRV monitoring. Despite its benefits, BCG signals are susceptible to external interference and present a challenge in accurately detecting J-Peaks. This research uses advanced signal processing and deep learning techniques to overcome these limitations. Our approach integrates accelerometers for long-term BCG data collection during sleep, applying Discrete Wavelet Transforms (DWT) and Ensemble Empirical Mode Decomposition (EEMD) for feature extraction. The Bi-LSTM model, leveraging these features, enhances heartbeat detection, offering improved reliability over traditional methods. The study's findings indicate that the combined use ofDWT, EEMD, and Bi-LSTM for J-Peak detection in BCG signals is effective, with potential applications in unobtrusive long-term cardiovascular monitoring. Our results suggest that this methodology could contribute to HRV monitoring, particularly in home settings, enhancing patient comfort and compliance.**

I. Introduction

An individual's health assessment significantly relies on the fundamental vital signs of cardio-respiratory activity, as established by notable sources. Cardiovascular diseases (CVD) currently stand as the leading cause of mortality worldwide, claiming approximately 17 million livesannually. To mitigate this mortality rate, clinical and researchstudies emphasize the importance of monitoring vital signs, particularly Heart Rate Variability (HRV) and Respiratory Rate, critical cardiovascular health indicators [1].

HRV is a way to measure the changes in the time between consecutive heartbeats. As the heart beats, a temporary interval exists before the next beat. HRV is a popular method extensively used by clinical researchers as a metric for noticing how these intervals change from one beat to the following [2]. HRV is a pivotal physiological phenomenon proven to be a discernible marker of the functional dynamics of the autonomic nervous system (ANS). The ANS exerts control over an array of involuntary physiological functions, encompassing heart rate modulation, respiration, and digestion [3]. Traditionally, this heart variability manifests as oscillations in the temporal intervals between sequential R-peaks (prominent peaks) observed in a cardiac cycle on an electrocardiogram (ECG) signal [4].

The R-R interval (changes in the time between consecutive R-peaks) in ECG is widely known as the gold standard for HRV monitoring, with monitoring experiments conducted either in the short term (5 mins) or in the long term (24 hrs). The ECG recording is an obtrusive method that involves multiple physiological sensors attached toa patient's body to record several body functions during sleep,such as breathing patterns, sleep stages, heartbeats, and body movements. Although ECG short-term HRV monitoring is the most commonly used HRV monitoring approach because it is relatively easy to obtain, long-term HRV monitoring offers the opportunity for predictive performance in detecting symptoms or categorizing patients with CVDs [5]. Despite the feasibility of the ECG in presenting real and accurate information about vital sign monitoring and associated physiological functions, it introduces limitations like discomfort (due to prolonged periods of direct contact of the sensor to the skin) and complexity in setting up the experiment. As a result, this has led to the demand for inexpensive and scalable non-obtrusive methods suited for both long-term and short-term cardiac activity monitoring experiments [6].

Recently, researchers have proposed ballistocardiograph (BCG) as an alternative means of unobtrusive and non-invasive HRV monitoring. BCG is a contactless method of recording the mechanical activity of the body with each heartbeat. The cardiac activity of the heart causes microforces to be generated as blood is pumped throughout the body. BCG is a recording of these micro-movements, and it is usually employed with other sensing methods that measure sensor readings of displacement, pressure, force, or acceleration [7]. A major merit of BCG is that the recording

*Research funded by Carl Zeiss Foundation (project number: P2019-03-003).

M. Haghi is with the HTWG Konstanz – University of Applied Sciences, Konstanz 78462 Germany (corresponding author, mostafa.haghi@htwg-konstanz.de).

O. Owonuga and P. Chaurasia are with the School of Computing, Engineering & Intelligent Systems, Ulster University, Derry, Londonderry BT487JL, UK ({awonuga-o,p.chaurasia}@ulster.ac.uk)

M. Gaiduk and R. Seepold are with the HTWG Konstanz – University of Applied Sciences, Konstanz 78462 Germany (e-mail: {maksym.gaiduk, ralf.seepold}@htwg-konstanz.de).

N. Martínez Madrid is with the Internet of Things Laboratory, School of Informatics, Reutlingen University, Reutlingen 72762 Germany (natividad.martinez@reutlingen-university.de).

system can be set up in users' homes without interrupting the user's privacy or lifestyle activities. Also, BCG can be scaled toward experiments related to sleep monitoring and sleep-disordered breathing [8]. In a BCG recording, the HRV manifests as oscillations in the intervals between sequential J-peaks (prominent peaks) observed in a cardiac cycle on a BCG signal. The changes in the time between consecutive J-peaks are known as the J-J interval. The J-J interval (JJ-I) is similar to the R-R interval (RR-I) in ECG.

While offering insights into mechanical heart activity, the BCG technique has certain limitations. Firstly, BCG signals can be affected by external movements or vibrations, which can introduce noise into the BCG signal, making it challenging to detect J-peak, which is essential in extracting cardiac signals for HRV evaluation. Also, the output of BCG signals is a composite signal consisting of cardiac activities, respiration, and muscle activity; therefore, advanced signal processing techniques are needed to separate the signals and extract meaningful information for HRV analysis and heartbeat detection [9].

Recent studies have shown the application of signal processing and machine learning algorithms for signal decomposition, feature extraction, and heartbeat detection using BCG signals. Chen et al. proposed using four piezoelectric sensors: one under the pillow and three under the mattress, and the data was gathered from five healthy participants in their twenties during a 2-hour nap. ECG and nasal thermistor signals were reference points for heart rate and respiration. Employing Cohen–Daubechies–Feauveau biorthogonal decomposition, heart rate was calculated by finding the 6th level approximation waveform, which matched the respiratory rhythm. Heart rate detection was effective using a combination of the 4th and 5th scale coefficients [10]. C.H. Antink investigated the viability of using BCG for post-surgical patient monitoring by comparing beat-to-beat intervals and ultra-short-term HRV with ECG references [11]. In another study, Katz et al. utilized a contact-free piezoelectric sensor beneath the mattress to measure cardiac interbeat intervals. Data from 25 home sleep recordings involving 14 healthy subjects in a two-in-bed setup were collected. The methodology included three algorithms: first, using empirical mode decomposition to locate candidate peaks for interbeat intervals; second, classifying intervals into three groups using binomial logistic regression based on ballistocardiogram signal properties; third, creating discrete interbeat interval distribution maps through overlapping 15-minute windows during night recording. These algorithms collectively affirmed the system's efficacy for heart rate variability analysis [12].

Rosales et al. utilized hydraulic transducer sensors to measure heart rate, employing a clustering-based approach for heartbeat computation. Data were gathered from four subjects in a supine position for six minutes. Filtering and feature extraction were applied to transducer signals. Extracted features were classified into two groups using k-means clustering. Heartbeat positions were compared to a reference signal from a finger-worn device. While the clustering approach showed promise, its applicability may be limited to specific situations and could require manual data labeling for practical implementation [13]. In 2021, researchers C. Jiao et al. introduced a bidirectional long short-term memory regression network to estimate non-invasive heart rates using BCG signal [14]. Akhbardeh et al. employed supported wavelet and biorthogonal wavelet transforms to extract key features from BCG signals. Thereafter, neural networks and a supervised fuzzy adaptive resonance theory were applied to classify healthy and heart disease subjects based on these features [15 – 16]. In a study by Yu et al., wavelet multi-resolution analysis was applied to extract wavelet coefficients from the BCG signal. Subsets of these coefficients were employed as features for neural network-based classification of normal and cardiovascular disease subjects [17].

As a result of the difficulties in detecting the J-peak in long-term monitoring, most of these methods have focused on HRV estimation based on BCG feature extraction. Few methods currently involve using advanced signal processing techniques and deep learning to determine heartbeat or beat-to-beat normality in patients for a given long and short-term HRV monitoring experiment during sleep. This study employs accelerometers to collect BCG data for long-term HRV monitoring of 20 patients during sleep. A deep learning approach is introduced utilizing bidirectional long short-term memory (Bi-LSTM). The method enhances heartbeat detection in comparison with earlier studies through:

- Combining discrete wavelet transforms (DWT) for converting BCG time series into the time-frequency domain for feature extraction.
- Implementing ensemble empirical mode decomposition (EEMD) parallel with DWT for HRV feature detection from smoothed BCG cardiac signals.
- Utilizing the Bi-LSTM model that benefits from both DWT and EEMD, using time-scale segmented smoothed BCG signals and HRV features to detect J-peaks
- Employing J-J Intervals variation from DWT and EEMD as target labels, comparing prediction results across different sleeping positions to assess model efficiency.

II. MATERIALS AND METHODS

A. BCG Data Preprocessing

The preprocessing phase of BCG data entailed an intricate series of steps to ensure the accuracy and reliability of subsequent HRV analysis. This phase lays the foundation for accurately interpreting physiological patterns during sleep. In the pursuit of unobtrusive monitoring of HRV during sleep using BCG signals, the initial step involves the meticulous preprocessing of the acquired accelerometer readings captured as BCG data using three sensors (sampled at 50 Hz), are synchronously recorded alongside ECG measurements (sampled at 256 Hz) during sleep sessions. The raw BCG and ECG data are stored in discrete 3-minute segments to facilitate accurate analysis, forming the foundation for subsequent processing stages.

From the raw data, three distinct signals emerge: body movement, heartbeat, and respiration. However, the presence

of body movement presents a significant challenge to the quality of physiological signals; therefore, extracting respiratory signals becomes a pivotal step for the successful removal of body movement artifacts. This is accomplished by applying a 3rd-order Butterworth bandpass filter, effectively isolating the relevant frequency range for respiration. The lower and upper cutoff frequencies of the filters are carefully set at 0.5 Hz and 10 Hz to capture the respiratory components while eliminating undesired frequencies. Second, the median filter is used to suppress thebaseline drift of the BCG signal. Finally, a cross-correlation function is used to solve the problem of the signal's peak notbeing prominent after the median filter to enhance the qualityof the signals further. The introduction of an amplitude threshold becomes imperative. This threshold serves the purpose of identifying and addressing the saturated signal segments resulting from abrupt motion artifacts. By strategically setting the threshold, the signal's integrity is preserved, minimizing distortion caused by extreme motion-induced fluctuations. In the pursuit of comprehensive signal analysis, a normalization process is introduced. This step involves subtracting the mean value of the signals and subsequently normalizing each channel by dividing it by its standard deviation. This normalization process aims to accentuate the signal's shape and characteristics, shifting the focus from absolute values to the underlying signal dynamics.

B. BCG Data Decomposition

For an efficient, unobtrusive monitoring of HRV using BCG signals, the process of BCG data decomposition forms a critical component of the methodology. This section focuses on extracting vital respiratory and heart activity signals through the DWT to unveil the underlying physiological dynamics.

The accelerometer signals, captured in three dimensions (xyz axis), assume prominence as the z-axis comes into focus. This axis, corresponding to the heart's pumping motion, serves as the key determinant for signal analysis [17]. After identifying and eliminating artifacts, the subsequent phase involves the extraction of respiratory and heart activity signals, achieved through two distinct wavelength decomposition techniques—Wavelet Decomposition and EEMD. The foundation of the BCG data decomposition rests on the DWT, functioning as a digital filter bank comprising high-pass (HPF) and low-pass (LPF) filters. These filters operate in tandem, cascading from high to low frequencies, effectively decomposing the BCG signal into detail coefficients (dj) and approximation coefficients (aj) across different scales (j levels). For the context of this research, a meticulous decomposition at j = 8 levels has been undertaken [12]. In this endeavor, the selection of the wavelet basis function holds paramount significance. While the daubechies, sym4, and biorthogonal wavelets find common application, the biorthogonal (bior3.9) wavelet is harnessed for its exceptional symmetrical signal properties. Despite a potential energy trade-off, adopting this wavelet ensures precise signal reconstruction and positional accuracy, thus enhancing the overall fidelity [18]. The values of wavelet coefficients, comprising both approximation (aj) and detail (dj) coefficients, play a pivotal role in expressing the outcome of BCG signal decomposition. This representation emerges through the calculation of inner products represented by the equation

$$dj,(k) = [\ (x(l),\ \ \Psi j,k\ (l)\ ] \quad (1)$$

$$aj.(k) = [\ (x(l),\ \Phi j,k\ (l)\ ] \quad (2)$$

This inner product $\Psi_{j.k}(l)$ and $\Phi_{j.k}(l)$ are scaled and dilated versions of the basis functions (Daubechies type) associated with HPF and LPF impulse response, effectively encapsulating the underlying physiological patterns within the signal structure. The extracted respiratory (Resp(n)) and (Heart(n)) signals, emerging as the outcome of the DWT-based decomposition, find their digital form representation.

$$\Psi j.k(l) = 2^{-j/2}\ \ \Psi(2^{-j}l - k) \quad (3)$$

$$\Phi j.k(l) = 2^{-j/2}\ \ \Phi(2^{-j}l - k) \quad (4)$$

The acquired Resp(n) and Heart(n) waveforms will undergo processing to facilitate the extraction of crucial respiration rate and heart rate values, subsequently subjected to a rigorous comparison with measurements obtained through reference measurement equipment.

The BCG signal for subject 1 in the supine position was used as the case study for the DWT decomposition. DC-3 contains noise, making it hard to accurately distinguish all the J-J intervals. Also, DC-5 loses a part of the contour and peak point information. Meanwhile, the DC-4 contains greater signal contour information but does not lose peak point information. Therefore, DC4 is better than other decomposition components to obtain the J-J intervals. However, it is recommended that in obtaining J-J intervals, utilizing simplyone stationary component (e.g., DC-5) of the decompositioncomponents is not advisable [19]. Therefore, the heart rate signal will be obtained from a combination of different decomposition components (DC-3 and DC-4) within the usual frequency spectrum for a heart rate signal of 0.6 – 3.5 HZ.

The reconstructed signal (DC3 + DC4) was used as a heart rate signal for estimating HR respectively with the BCG signal. It was observed that the reconstructed signal for each sleeping position contained noise and loss of peak information. Therefore, the moving average smoothing filter of window = 10s using the Savitzky-Golay filter is applied to the cardiac signal to filter out the noise and distinguish the J-peak. Careful examination was taken to ensure the window size did not blur rapid changes or small details in the signal. The window size was set to 10 to ensure we consider an equal number of data points with peak information on each side of the current point within the 12 sec time domains. In Fig. 2 above, the reconstructed signal of the subject in lateral positions shows a better cardiac waveform and a smoother level coefficient with J-peak information in contrast to the results in the other 3 sleeping positions, however, it can be observed that the J-peak in the lateral position BCG cycle may be submerged in the nearby peak, especially in lateral which would bring difficulties to the labelling the data using the J-J intervals.

In applying the alternative method using EEMD, the goal is to deal with noise present in BCG signals and provide a solution to limitations inherent in the conventional ensemble mode decomposition (EMD). The EEMD is an innovative

technique representing an advancement over the conventional EMD method. EMD breaks down complex and changing signals, like BCG signals, into intrinsic mode functions (IMFs). These IMFs serve as building blocks that capture theunderlying dynamics of a signal in its simplest forms within the time domain [20].

However, EMD faces challenges such as "end effects," which distort results at the signal's endpoints, and "mode mixing," where oscillations of varying amplitudes interfere with one another. To tackle these challenges, EEMD combines the EMD process with a multitude of signals augmented by white Gaussian noise (WGN). This ensemble approach utilizes the way EMD filters signals to overcome mode mixing issues effectively. The process involves:

- Adding controlled noise to the signal, breaking downthe signal with EMD,
- Repeating these steps with different noise series leads to a set of IMF averages.

This ensemble technique mitigates the mixing phenomenon inherent in EMD, ensuring the accurate extraction of key metrics like JJ intervals, which play a crucial role in heart rate variability analysis. It is observed that J-peaks become prominent from IMF4 to IMF7. It can be seen that the decomposition levels above IMF7 have a low-frequency spectrum and generally lose peak information, which makes them unusable for estimating heart rate and respiration rate values. On the otherhand, IMF4 and IMF7 contain a part of peak point information, making it hard to distinguish all the J-J intervals accurately. Meanwhile, the IMF5 and IMF6 contain greater signal contour information and don't lose peak point information. Therefore, both IMF5 and IMF6 are better thanother decomposition components to obtain the J-J intervals. Similarly, in obtaining J-J intervals, utilizing simply one stationary component (e.g., IMF5) of the decomposition components is not advisable [19]. Therefore, the heart rate signal will be obtained by combining IMF4 with IMF5 and IMF6 because it has a strong weight of 3Hz, which is withinthe usual frequency spectrum for heart rate signals.

C. J-Peak Detection

A critical aspect of our methodology is the local peak detection algorithm. We introduce a distinctive approach for accurate J-peak identification. First, a moving average smoothing filter underwent a parameter tuning of window = 3s and 25s using the Savitzky-Golay filter for the DWT and EEMD reconstructed cardiac signal to filter out noise and distinguish the J-peak. Next, we define a recommended minimum peak-to-peak distance of 30 milliseconds, which is derived from visual observations of the peak-to-peak distances between using reference ECG R-peaks. This distance is tuned so that detected J-peaks are suitably spaced apart, enhancing the reliability of the analysis. To implement this method, a local peak detector is employed to analyze the derivative of the reconstructed cardiac signal. This detector identifies explicitly local peaks, which are characteristic ofJ-peaks. The reliability of this methodology is further confirmed through the synchronization.

D. Data Labelling and Bi-LSTM

This section outlines the methodology of data labeling and the implementation of a Bidirectional Long Short-Term Memory (Bi-LSTM) model. The goal is to leverage DWT and EEMD to extract HRV features from BCG signals and subsequently employ a Bi-LSTM model for heartbeat extraction and HRV analysis. This section entails several key steps that culminate in creating and utilizing an effective predictive model.

- Time-Scale segmentation and candidate J-Peak labeling:

From Fig. 2 above, there is no exact one-to-one correspondence between the R peak and the J peak in the time domain, and this could cause difficulties when labeling the data for data modeling and heartbeat extraction. Therefore, the signal of interest (ROI) is introduced for labeling the J-peak. The sampling points surrounding the reference R-peaks are labeled as J-peak. A label of '0' is set if the variation between consecutive BCG heartbeats is less than the width of the region of interest. However, if the width interval exceeds the signal of interest, the data is labeled as '1'. The threshold of probability is set at 0.5.

- Feature Extraction:

In modeling the Bi-LSTM network, the forward and backward features of a moving 3s sliding period, the BCG signals are fed as input data into the model. The forward and backward features are extracted from a signal length of 12s and 30s using the DWT and EEMD cardiac signals respectively. As a result of shift variant property limitation in DWT, which downsamples the raw signal, causing a reduction in sample size of the coefficient by 2, the extracted sample size for the heart rate signal is decomposed to a 12s window. However, since long-term HRV monitoring offers the opportunity for predictive performance in heartbeat detection, a thirty-second time scale was considered using EEMD cardiac signal for model evaluation. The target data is the variation (seconds) of consecutive beat-to-beat J-J intervals within the entire duration of the cardiac signal. A central timestamp of the window is determined as the midpoint between the start and end timestamps of the cardiac signal, which is crucial for synchronization. As data within the defined window is extracted, the model creates a windowed synchronized cardiac signal as input and corresponding output, which is the variation between consecutive J-J intervals at the start and end timestamp of the 3s sliding window.

- Bi-LSTM model architecture:

The Bi-LSTM architecture is pivotal in extracting temporal patterns from BCG signals and predicting heartbeatintervals. The Bi-LSTM model, a variant of the Long Short-Term Memory (LSTM) network, possesses the unique ability to capture both forward and backward temporal dependencies within sequential data. In the context of HRV analysis, the Bi-LSTM architecture is a powerful tool to predict heart rate abnormality by comparing BCG signal-derived J-J intervals (JJIs) with reference R-R intervals (RRIs) obtained from ECG measurements. Theproposed Bi-LSTM model leverage

TABLE I. COMPARISON OF AVERAGE BPM USING DWT AND EEM

Subject	Left Lateral (average bpm)		Right Lateral (average bpm)		Supine (average bpm)	
	ECG	BCG	ECG	BCG	ECG	BCG
1	66.6	76.1	74.3	66.1	68.4	76.6
2	63.5	63.9	68.6	69.3	63.2	60.0
3	60.2	73.4	66.7	63.9	61.6	64.2
4	58.0	64.3	61.8	59.1	57.7	62.0
5	60	65.0	58.5	69.3	61.7	77.5
6	69.6	66.1	67.3	64.0	72.1	59.2
7	64.7	70.0	68.5	66.5	68.0	69.3
8	65.2	60.4	67.2	60.6	64.3	63.8
9	68.5	73.0	77.3	71.5	78.4	70.0
10	68.5	53.8	77.3	67.9	78.4	74.0

TABLE II. BBI VARIATION COMPARISON BETWEEN ECG AND BCG

Right Lateral DWT BBI		Right Lateral EEMD BBI	
ECG (RRI)	BCG (JJI)	ECG (RRI)	BCG (JJI)
0.76	0.72	0.76	0.92
0.74	1.02	0.74	0.9
0.76	0.74	0.76	0.92
0.76	1.32	0.76	0.88
0.75	0.68	0.75	1.34
0.73	0.86	0.74	1.0
0.73	1.08	0.73	0.9
0.76	0.76	0.76	0.84
0.76	0.6	0.76	1.04
0.76	0.6	0.76	0.94

large time-scale segments of BCG signals. These segments encapsulate multiple heartbeat intervals and serve as the input to the network. By considering both forward and backward temporalcharacteristics, the Bi-LSTM effectively captures the rhythmic features present within BCG signals [13]. Thisenables the model to learn complex patterns and relationshipswithin the data. The Bi-LSTM architecture comprises two main components: a sequential class model and a Bi-LSTM model. The sequential class model processes the large time-scale segment to extract spatial characteristics, particularly the rhythm features inherent in BCG segments across multiple heartbeat intervals. The output of this model is a feature matrix that captures important spatial patterns. The Bi-LSTMmodel, on the other hand, focuses on capturing temporal dependencies of feature sequences extracted by the sequential class model. It involves a single layer of Bi-LSTM units, with each unit containing 128 neurons. This layer is instrumental in memorizing the context of the input time signal, enhancingthe network's ability to understand temporal relationships within the data. The Bi-LSTM model's mathematicalformulation involves a series of transformations:

$$y = f(W * x + b) \quad (5)$$

Where:

- y is the output of the model
- f() represents the activation function, commonly the rectified linear unit (ReLU) function
- W denotes the weight matrix of the model
- x signifies the input data
- b refers to the bias term of the model

The ReLU activation function introduces non-linearity into the model. It maps positive inputs to themselves and negative inputs to zero, allowing the network to capture complex relationships. The weight matrix, W, encodes the connections between neurons, determining the flow of information through the network. The bias term, b, adjusts the output of the model.

To train the Bi-LSTM model, 10 k fold cross-validation strategy is adopted because it provides a balance between the bias of a single train-test split (which might not be representative of the entire dataset) and the high computational cost of leave-one-out cross-validation (where each data point is treated as a test set) [21]. The Mean Square Error (MSE loss) serves as the optimization objective, and the Adam algorithm functions as the optimizer. A learning rate of 0.0003 facilitates efficient convergence and accurate prediction [19].

The Bi-LSTM model is trained using the labeled DWT and EEMD datasets associated with subjects 1,2,3,4,5,6 and 7. The efficiency of the trained Bi-LSTM model is evaluated using subjects 8,9 and 10. The model's predictive capabilities are tested on these separate datasets to assess its accuracy in detecting actual J-peaks from DWT and EEMD cardiac signals.

III. RESULTS

The proposed heartbeat detection scheme was conducted by conducting an experiment using 3 minutes and 30 seconds of BCG signals measured from 14 subjects. BIOPAC MP160 simultaneously recorded ECG signals as a reference. In this research context, The BCG signals of 10 subjects were used for HR analysis to validate the measure of agreement of BCG and ECG heart rate in line with clinical research ground truth HR values varying between 60-100 beats per minute [22].

A. HRV Time Domain

The table above shows each subject's average beats per minute using cardiac signals. The extraction of these features was implemented under a 12s window. Table 1 above shows that in the "Right Lateral" sleeping position,

the DWT average beats-per-minutes (bpm) generally showgood agreement with the ground truth ECG, with relatively small differences. This suggests that the DWT approach performs well in this right lateral position and can provide accurate bpm estimates.

In contrast, there is more variability in the "Left Lateral" and "Supine" sleeping positions. In some cases, the DWT bpm values are close to the ECG values (e.g., Subject 6 in all sleeping positions), while in others, there are larger values (e.g., Subject 5 in all sleeping positions). The data shows that bpm values can vary significantly based on body position. For example, in Subject 7, there is a notable difference in bpm between the Right Lateral and Supine positions for both ECG and BCG HR values. Generally, the bpm estimates using DWT show a good degree of consistency across different sleeping positions for each subject. In addition, the bpm values tend to follow a similar increasing or decreasing trend as the ground truth ECG-average bpm values.

Table 1 shows that the EEMD-derived bpm values exhibit subject-specific variability - each subject has a unique response to different sleeping positions, resulting in varying HR values. The EEMD-derived bpm estimates are sensitive to sleeping positions. Several subjects show variations in HR values when transitioning between sleeping positions. For example, Subjects 4 and 8 have relatively stable HR across positions, while Subjects 5 and 9 display significant variability. Also, the EEMD-derived HR values often deviate from reference values, with varying degrees of difference across subjects and positions.

In contrast to the DWT, the EEMD-derived bpm estimates tend to be lower than the corresponding ground truth ECG heart rate values for most subjects and positions. This suggests that EEMD may underestimate Heart Rate values compared to DWT. Although the DWT-derived HR values also exhibit variability, similar to EEMD, DWT tends to exhibit greater consistency and agreement with reference HR values across subjects and positions. This might result from DWT robustness against noise and baseline drift, reflectedin the smoother bpm estimates and smoother and more stablebpm trends.

The inter-beat intervals using BCG were derived for each subject using the simple peak detector algorithm. The threshold for the peak distance was set by observing the periodicity trend of beat-to-beat intervals with reference ECG. Table 2 above shows ten (10) consecutive beat-to-beat intervals for Subject1 derived using DWT and EEMD Heart Rate signals. From the table above, it was observed that the inter-beat interval shows a consistent time interval of 0.76s with reference ECG. In addition, the variation between consecutive heartbeats was 30 milliseconds (0.03s). In contrast, the variation in consecutive heartbeat intervals with DWT is within the range of 0.1 to 0.6s (100 – 600ms), while the time interval for consecutive heartbeat is between 0.6 and 1.3 seconds. For EEMD, the variation between consecutive heartbeats was observed to be within 0.1 to 0.3s (100 – 300ms). It can also be observed that in beat-to-beat intervals with EEMD and DWT, the heart rate signal had a deviation interval of 1.3sec between consecutive heartbeats, which, according to clinical researchers [19], is an anomaly because it exceeds the recommended heartbeat interval of 0.8s – 1.2s. Therefore, during time domain analysis, heartbeat intervals exceeding this recommended range will be omitted. HRV time domain features extracted will be evaluated using a short-term (12-second) window. The time-domain variability to be measured is the Mean of a normal heartbeat interval (Mean_NN) and standard deviation of normal heartbeat intervals (SDNN). Three subjects (subjects 1, 4, and 8) were selected for the purpose of analyzing the time domain variability across different sleeping positions (Supine, Left Lateral, and Right Lateral).

TABLE III. HEART RATE TIME DOMAIN ANALYSIS FOR SUBJECTS IN SLEEPING POSITIONS

Subject	Sleeping Position	Technique	BCG SDNN	ECG SDNN	BCG meanNN	ECG meanNN
1	Right	DWT	0.20	0.02	0.84	0.76
1	Right	EEMD	0.18	0.02	0.92	0.76
4	Left	DWT	0.19	0.05	0.85	0.87
4	Left	EEMD	0.21	0.05	0.96	0.87
8	Supine	DWT	0.13	0.06	0.81	0.66
8	Supine	EEMD	0.22	0.06	0.83	0.66

Table III above shows that the ECG-derived SDNN values are consistently lower than BCG-derived SDNN values for all subjects, techniques, and sleeping positions with a variation of 30 milliseconds. This suggests that the average peak-to-peak distance between consecutive R-peaks in the ECG signal falls within the 20–60 milliseconds range, depending on the sleeping position. On the other hand, the BCG SDNN tends to be higher than ECG-reference values in most cases for all subjects, indicating that BCG may result in longer beat-to-beat intervals on average with a minimum beat-to-beat variation of 130 milliseconds. Sometimes, there is variation between consecutive heartbeats reaching 500 – 600 ms.

Overall, the average heartbeat interval of both ECG and BCG seems to have a close agreement regarding the actual beat-to-beat time interval; however, in most cases, BCG captures wide variability in interval periods while ECG tends to have a consistent time interval. The Coverage(CR) performance indicator was analyzed to derive the percentage of J-J intervals, which was successfully aligned with corresponding ground truth RR intervals. The coverage formula is defined below

$$CR = (Number\ of\ Matched\ JJI\ /\ Total\ Number\ of\ RRI) * 100\% \quad (6)$$

Coverage is an important metric in assessing the accuracy and reliability of a DWT and EEMD method in deriving beat-to-beat intervals, as it provides insights into how well the derived intervals correspond to the ground truth intervals. The average beat-to-beat interval error is used to measure how accurately both signal processing methods capture the timing of heartbeats between the J-J interval and the ground truth R-R interval. Since beat-to-beat interval measurement is essential for diagnosing heart conditions like arrhythmias, which involve irregular heartbeats, understanding the error is critical in clinical applications.

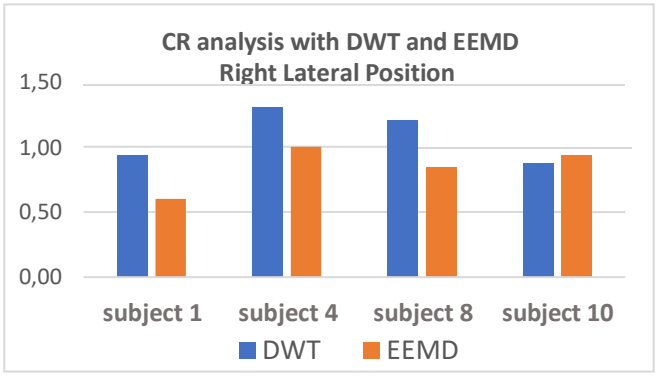


Figure 1. Coverage analysis of heartbeat interval using DWT and EEMD.

There is significant variability in coverage between DWT and EEMD for each subject in the right lateral position. DWT tends to have higher CR for subjects 1, 4, and 8, while EEMD performs slightly better than DWT for subject 10. Both methods are relatively close in CR subject 10. The trends in the coverageresults suggest that the EEMD method performs better at estimating the J-J interval from the R-R interval for subjectswith more regular heartbeats (subjects 1,8 and 10). Similarly, the DWT CR suggests that this technique could better estimate the J-J interval from the R-R interval for subjects with more irregular heartbeats (subject 4). Generally, DWT shows higher CR for most subjects in thisanalysis, suggesting it may be a more robust choice for subjects with different heart conditions.

TABLE IV. AVERAGE BEAT-TO-BEAT INTERVAL ERROR USING DWT AND EEMD.

Subject	DWT	EEMD
1	0.18s	0.21s
4	0.18s	0.20s
8	0.22s	0.25s
10	0.19s	0.16s

The average beat-to-beat interval error (E) is calculated using the formula

$$E = \text{Sum [absolute difference (JJI - RRI)] / n-intervals} \quad (7)$$

Where n is the number of heartbeat intervals to be considered for the analysis. For computing (E), Ten (10) consecutive beat-to-beat intervals were utilized as the input. According to scholars Yijun et al. [20], 5–8 BCG intervals that fall within the range of 5s are considered sufficient to understand the long-term rhythmic features of a heartbeat. Therefore, this suggests 10 BCG intervals prove sufficient to understand the synchronization of RR and JJ heart-beat intervals. In Table 3 above, Subjects 1, 4, and 10 show relatively consistent performance with DWT, with average beat-to-beat interval error ranging from 0.18 to 0.19 seconds. This suggests that DWT consistently aligns J-J intervals with RR intervals with a minimal lag time of 0.18s for these subjects. Like DWT, EEMD also shows relatively consistent performance across subjects, with error values ranging from 0.16 to 0.25 seconds.

In general, EEMD generally exhibits a slightly longer lag time compared to DWT for all subjects. However, the differences are relatively small, with a maximum difference of 0.05 seconds (Subject 8). This lag time implies that there is no exact one-to-one correspondence between the R peak and the J peak in the time domain, and this could bring difficulties when labeling the data for data modeling and heartbeat extraction. Therefore, the signal of interest (SOI) is introduced to label the J-peak. The sampling points within the range of -0.2 to +0.2 surrounding the reference R-peaks are labeled as J-peak. Since reference ECG has been observed to have an average variation of -0.06s to +0.06s (Table 3), a label of '0' is set as R-peak while width intervals exceeding average variation in ECG are labeled as J-peak (label = 1). threshold of probability is set at 0.5. A tolerance level of +0.

B. Data Modelling and Evaluation Results

Evaluation metrics such as recall, precision, accuracy, positive prediction value (PPV), and negative prediction value (NPV) were used to assess model performance. In this context, precision is the fraction of predicted J-peaks that are actually J-peaks. Recall is the fraction of actual J-peaks that are correctly predicted. Negative prediction value (NPV) is the fraction of predicted non-J-peaks that are actually non-J-peaks. Positive prediction value (PPV) is the fraction of actual J-peaks that are correctly predicted. It is calculated as follows. The number of False Negatives (where they are actually J-peaks but are predicted not to be J-peaks) will also be considered in evaluating the performance of J-peak prediction.

TABLE V. MODEL EVALUATION RESULTS ON VALIDATION DATA IN RIGHT AND SUPINE POSITION

Technique	Position	Precision	Recall	F1	NPV	PPV	Accuracy
DWT	Right	0.94	0.89	0.91	0.33	0.94	0.85
DWT	Supine	0.84	0.84	0.84	0.77	0.84	0.81
EEMD	Right	0.76	0.87	0.81	0.33	0.76	0.70
EEMD	Supine	0.79	0.84	0.81	0.44	0.79	0.72

The Table above shows the results evaluating the model's performance on validation data using DWT and EEMD Bi-LSTM models. The DWT-based models, both on the right Lateral and supine positions, outperform the EEMD-based model across all evaluation metrics. They exhibit higher precision, recall, F1-scores, NPV, PPV, and test accuracy. For instance, The DWT Right Lateral model shows higher precision (0.94) and recall (0.89). This indicates it has up to 90% ability to correctly identify J-peaks (with low false positives) while capturing a significant portion of the actual J-peaks. Also, the DWT models show balanced precision and recall, reflected in the F1-scores for the right and supine positions (0.91 and 0.84). This indicates consistent performance in J-peak detection, with relatively low false positives.

On the other hand, while the EEMD models show decent performance, they fall short in precision for both sleeping positions (0.76 and 0.79) and test accuracy compared to DWT models. Overall, the Bi-LSTM model is more effective at detecting J-peaks using DWT Right Lateral cardiac signal than in the other three conditions using the validation data. For instance, Subjects 1, 4, and 9 show notable differences in BPM across positions. However, in some instances, the BCG-derived bpm values are slightly higher than the ECG values (e.g., Subject 1 in the Right Lateral position), while in others, they are slightly lower (e.g., Subject 2 in the Left Lateral position). Also, the BCG-derived bpm values are relatively consistent across different sleeping positions. This indicates that the BCG method may be robust for monitoring heart rate regardless of sleeping posture. In several cases, the DWT-derived bpm values are more accurate in the right position than in other positions.

REFERENCES

[1] J. F. Thayer and R. D. Lane, "The role of vagal function in the risk for cardiovascular disease and mortality," Biological Psychology, vol. 74, no. 2, pp. 224–242, 2007. doi:10.1016/j.biopsycho.2005.11.013

[2] R. Mccraty and F. Shaffer, "Heart rate variability: New perspectives on physiological mechanisms, assessment of self-regulatory capacity, and Health Risk," Global Advances in Health and Medicine, vol. 4, no. 1, pp. 46–61, 2015. doi:10.7453/gahmj.2014.073

[3] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," Frontiers in Public Health, vol. 5, 2017. doi:10.3389/fpubh.2017.00258

[4] S. M. Koenig, D. Mack, and M. Alwan, "Sleep and sleep assessment technologies," Aging Medicine, pp. 77–120. doi:10.1007/978-1-59745-233-5_5

[5] Y.-L. Zheng et al., "Unobtrusive sensing and wearable devices for Health Informatics," IEEE Transactions on Biomedical Engineering, vol. 61, no. 5, pp. 1538–1554, 2014. doi:10.1109/tbme.2014.2309951

[6] O. T. Inan et al., "Novel wearable seismocardiography and machine learning algorithms can assess clinical status of heart failure patients," Circulation: Heart Failure, vol. 11, no. 1, 2018. doi:10.1161/circheartfailure.117.004313

[7] Sadek, I. and Biswas, J. (2018) 'Nonintrusive heart rate measurement using Ballistocardiogram Signals: A Comparative Study', Signal, Image and Video Processing, 13(3), pp. 475–482. doi:10.1007/s11760-018-1372-z.

[8] Chen, H. et al. (2008) 'Evaluation of a portable recording device (apnealinkTM) for case selection of obstructive sleep apnea', Sleep and Breathing, 13(3), pp. 213–219. doi:10.1007/s11325-008-0232-4.

[9] Hoog Antink, C. et al. (2020) 'Ballistocardiography can estimate beat-to-beat heart rate accurately at night in patients after Vascular intervention', IEEE Journal of Biomedical and Health Informatics, 24(8), pp. 2230–2237. doi:10.1109/jbhi.2020.2970298

[10] Katz, Y., Karasik, R. and Shinar, Z. (2016) 'Contact:free piezo:electric sensor used for real:time analysis of Inter beat interval series', 2016 Computing in Cardiology Conference (CinC) [Preprint]. doi:10.22489/cinc.2016.222-272.

[11] Rosales, L. et al. (2012) 'Heartbeat detection from a hydraulic bed sensor using a clustering approach', 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society [Preprint]. doi:10.1109/embc.2012.6346443.

[12] Jiao, C. et al. (2021) 'Non-invasive heart rate estimation from Ballistocardiograms using bidirectional LSTM regression', IEEE Journal of Biomedical and Health Informatics, 25(9), pp. 3396–3407. doi:10.1109/jbhi.2021.3077002

[13] Akhbardeh, A. et al. (2005) 'The heart disease diagnosing system based on force sensitive chair's measurement, biorthogonal wavelets and Neural Networks', Proceedings, 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. [Preprint]. doi:10.1109/aim.2005.1511060.

[14] Akhbardeh, A. et al. (2006) 'Applying novel supervised Fuzzy Adaptive Resonance theory (SFART) neural network and biorthogonal wavelets for ballistocardiogram diagnosis', 2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control [Preprint]. doi:10.1109/cacsd-cca-isic.2006.4776638.

[15] Xinsheng Yu, Dent, D. and Osborn, C. (1996) 'Classification of ballistocardiography using wavelet transform and Neural Networks', Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society [Preprint]. doi:10.1109/iembs.1996.652649.

[16] "High blood pressure ," www.heart.org, https://www.heart.org/en/health-topics/high-blood-pressure/ (accessed Sep. 9, 2023).

[17] Song, Y. et al. (2015) 'Extracting features for cardiovascular disease classification based on Ballistocardiography', 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom) [Preprint]. doi:10.1109/uic-atc-scalcom-cbdcom-iop.2015.223.

[18] Fontecha, J. et al. (2017) 'Ambient Intelligence for Health: Advances in vital signs and gait monitoring systems within mHealth Environments', Human Monitoring, Smart Health and Assisted Living: Techniques and technologies, pp. 183–202. doi:10.1049/pbhe009e_ch9.

[19] D. P. Kingma and J. Ba, "Adam: a method for stochasticoptimization," CoRR, vol. abs/1412, 2015.

[20] Fushiki, T. (2009) 'Estimation of prediction error by using k-fold cross-validation', Statistics and Computing, 21(2), pp. 137–146. doi:10.1007/s11222-009-9153-8.

[21] Landreani, F. et al. (2015) 'Feasibility Study for beat-to-beat heart rate detection by Smartphone's accelerometers', 2015 E-Health and Bioengineering Conference (EHB) [Preprint]. doi:10.1109/ehb.2015.7391493.

[22] Liu, Y. et al. (2022) 'ResNet-BiLSTM: A multiscale deep learning model for heartbeat detection using Ballistocardiogram signals', Journal of Healthcare Engineering, 2022, pp. 1–11. doi:10.1155/2022/6388445.

[23] G. Williamson and A. Carughi, "Polyphenol content and health benefits of raisins," Nutrition Research, vol. 30, no. 8, pp. 511–519, 2010.

[24] M. Khojastehnazhand and H. Ramezani, "Machine vision system for classification of bulk raisins using texture features," Journal of Food Engineering, vol. 271, p. 109864, 2020.

A. O. Raji and A. O. Alamutu, "Prospects of Cumputer Vision Automated Sorting Systems in Agricultural Process Operations in Nigeria," Agricultural Engineering International: CIGR Journal, 2005..

Data analysis of non-invasive ballistocardiographic sensors

Mattioli Sara, Bruschi Sara, Maksym Gayduk, Ralf Seepold, Natividad Martinez Madrid,
Simone Orcioni, Massimo Conti

***Abstract*— Unintrusive health monitoring systems is important when continuous monitoring of the patient vital signals is required. In this paper, signals obtained from accelerometers placed under a bed are processed with ballistocardiography algorithms and compared with synchronized electrocardiographic signals.**

I. Introduction

Ballistocardiography (BCG) is a non-invasive medical method based on the measurement of movement of the body generated by the expulsion of blood in each cardiac cycle. The ballistocardiogram is the reaction (displacement, velocity or acceleration) of the whole body to the ejection of blood [1]. The BCG signal comes from the fact that with each heartbeat blood travels through the veins produces a change in the body's center of mass. The BCG measures these movements which can be acquired in the form of displacement, velocity or acceleration along the three directional axes.

In recent decades, a renaissance has occurred in the field of evaluation non-intrusive cardiomechanics through ballistocardiographic or seismographic (SCG) signal measurement and interpretation methods. These two types of systems can be implemented on wearable devices, whose main advantage consists in the possibility of acquiring data continuously during the normal life of the subject being measured. Furthermore, another advantage is that they are able to carry out measurements in any environment and under any situation of stress. Typically, the most used sensor for this type of acquisition is the accelerometer. Recent review papers in this field have been published [2-6].

In some applications the BCG systems are applied on the bed structure with the aim of evaluating a subject's sleep and any related disorders. In this case, the system is non-invasive system and there is no need to use electrodes attached to the patient's body. Accelerometer sensors placed under the bed can often provide additional information about breathing and movement of the body, together with the measurement of heartbeats [7].

Since ballistocardiography is a technique that measures the heartbeat through vibrations body mechanics, it can be used to estimate the heart rate by exploiting sensors that can be mounted on the frame of a bed, under the mattress.

To this end, different types of sensors can be used, including hydraulic sensors, accelerometers, and force sensors. There selection of the most suitable sensor for your purpose is based on a comparison between costs and noise level and accuracy [8]. To derive the heartbeat waveform starting from the BCG signal, different algorithms are used, including FFT spectrum analysis and the autocorrelation function.

A data set of measurements to estimate heart rate and breathing rate with the person laying on a bed is reported in [9]. The vital signals have been acquired by two commercial medical devices and one low cost accelerometer prototype:

1. a commercial ECG Holter (M12R by Global Instrumentation) measuring the 12-lead ECG signals;
2. a commercial wearable sensor attached to a chest strap (BioHarness 3.0 by Zephyr Technology) for ECG and breathing measurements;
3. an accelerometer sensor prototype placed under the mattress.It consists of a low cost but accurate accelerometric sensor ADXL355Z of Analog Devices connected to a Raspberry with the SPI bus.

The aim of the work presented in this paper is the synchronization and preliminary signal processing of the data acquired in [9] by the different sensors, in order to compare the results and estimate the accuracy of the ballistocardiographic measurements.

The objective of the work is the development of non-invasive measurement methods to monitor useful medical parameters, also for the purpose of creating telemedicine systems. Starting from acquisitions carried out on a bed constructed in such a way as to accommodate the useful sensors an analysis is carried out aimed mainly at the synchronization of the signals. The synchronization is necessary to make the comparison between the signals in order to estimate the precision of the heart rate estimation obtained from the ballistocardiographic signals.

S. Mattioli and S. Bruschi are with the Università Politecnica delle Marche, Ancona, Italy .

M. Gaiduk is with the Ubiquitous Computing Lab at HTWG Konstanz, Konstanz, Germany (maksym.gaiduk@htwg-konstanz.de).

N. Martínez Madrid is with the IoT Lab at Reutlingen University, Reutlingen, Germany (natividad.martinez@reutlingen-university.de).

R. Seepold is with the Ubiquitous Computing Lab at HTWG Konstanz, Konstanz, Germany (Email: ralf.seepold@htwg-konstanz.de)

M. Conti is with the Dip. di Ingegneria dell'Informazione Università Politecnica delle Marche, Ancona, Italy (Email: m.conti@univpm.it).

S. Orcioni is with the Dip. di Ingegneria dell'Informazione Università Politecnica delle Marche, Ancona, Italy (Email: s.orcioni@univpm.it).

II. Sincronization Procedure

A. *Synchronization on acquisition time*

The experimental data have been acquired as described in [9]. The measurement procedure has been applied for each one of the 21 volunteers. The volunteer lies down on a bed in a supine position, the holter electrodes have been positioned on his chest, then the zephyr in an elasticated strap is applied to the chest, and acceleration sensors have been placed on the structure of the bed.

The protocol, for a total time of about 20 minutes, consists of: a deep breath (about 30 seconds) and a recovery phase (about 1 minute) repeated five times, followed by an apnea phase (about 30 seconds) and a recovery phase (about 1 minute) repeated five times. Finally, the person moves from supine to prone position and re-mains in prone position for about 5 minutes with normal breathing.

The Zephyr and Holter ECG signals acquisition have a duration of twenty minutes, while the acceleration signal lasts approximately in ten minutes. For this reason, only the first ten minutes of data acquisition are considered.

The data have been acquired with the time stamp of the clock of the different devices, apart from the holter that does not store the time stamp.

An initial synchronization between acceleration and Zephyr ECG signals is based on time stamp of the two devices.

The initial s synchronization between Zephyr and Holter ECG signals is obtained considering that a maximum of the Zephyr breathing signal and a minimum of the Zephyr and Holter ECG correspond to each deep breath. Therefore, the first maximum of the Zephyr breathing signal and the first a minimum of the Holter ECG signal have been aligned.

Figure 1 reports the results of this preliminary alignment for the Zephyr breathing and ECG signals, Holter ECG signal, accelerometer signal. The five deep breath and the first two apnea can be seen in Figure 1.

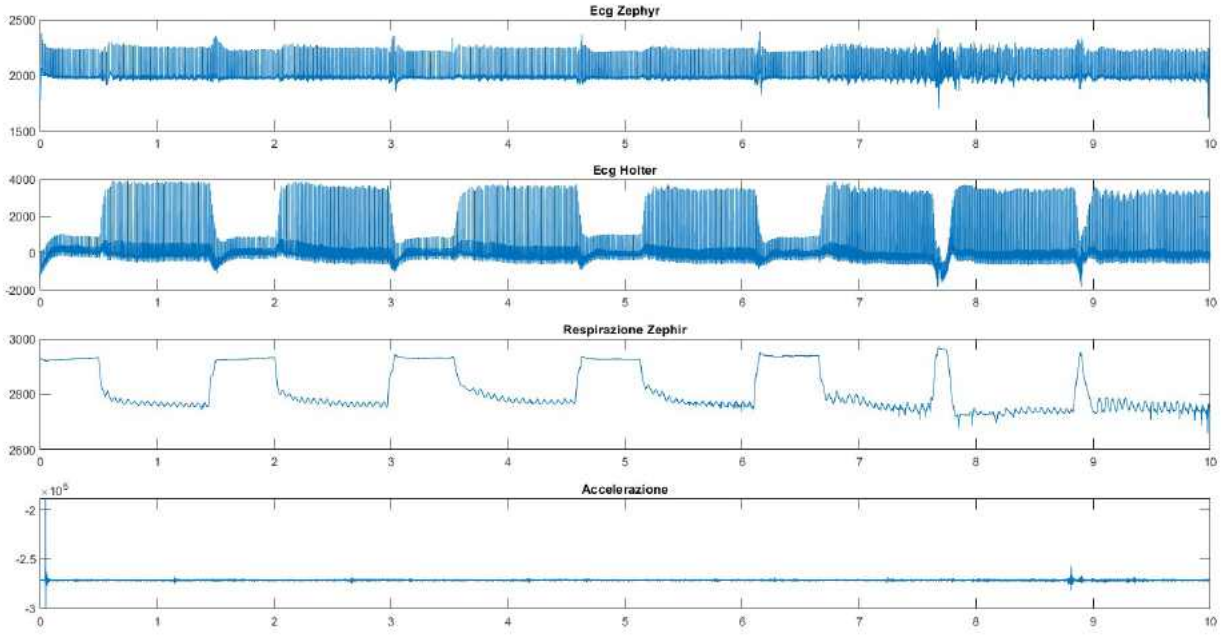


Figure 1. Preliminary alignment for the Zephyr breathing and ECG signals, Holter ECG signal, accelerometer signal

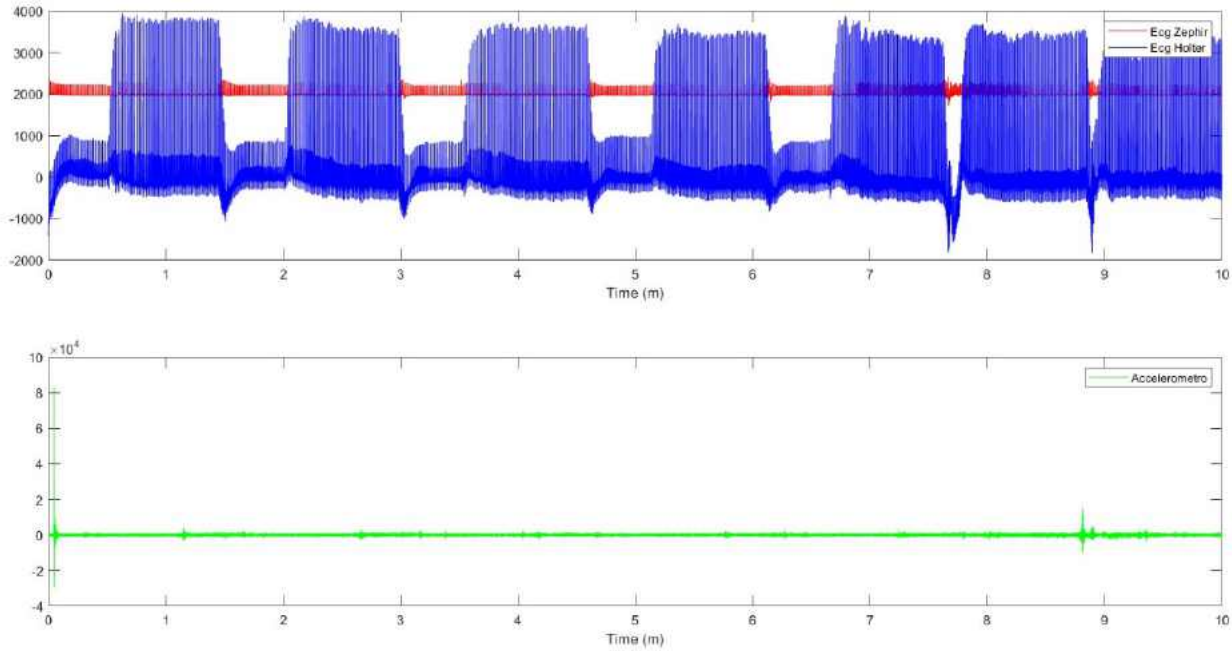


Figure 2. Zephyr ECG, Holter ECG and acceleration signals resynchronized. Zephyr ECG and Holter ECG signals has been synchronized on the basis of the peaks. Zephyr ECG and acceleration signals has been synchronized on the basis of the time stamp.

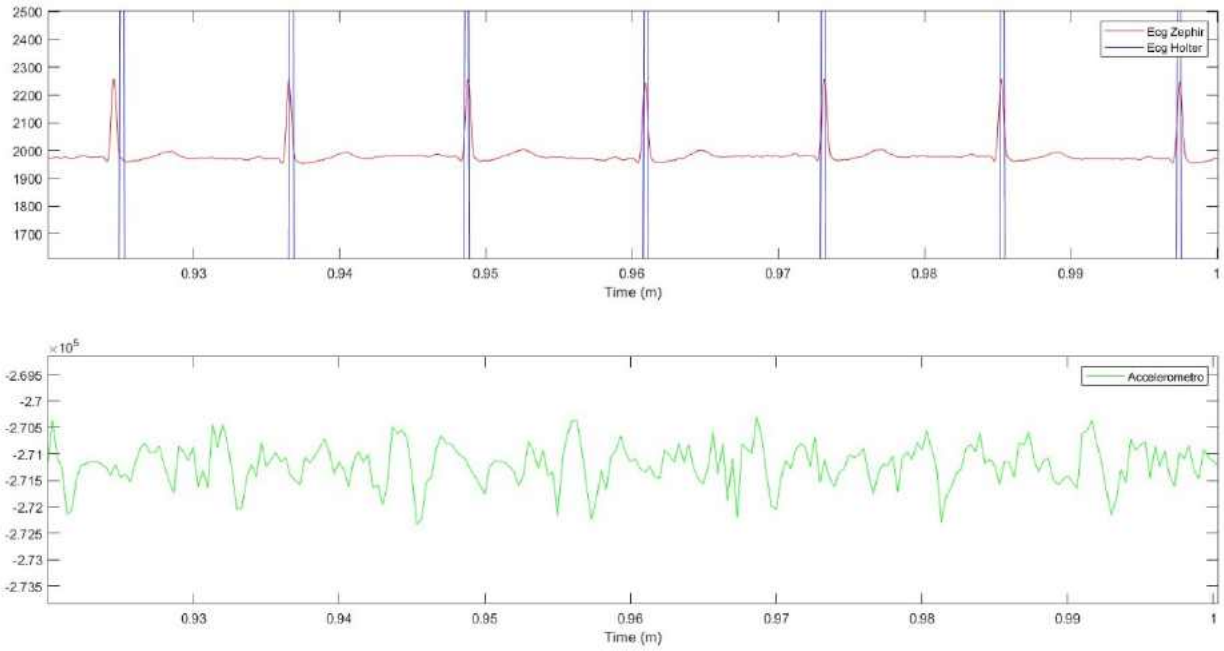


Figure 3. Zephyr ECG, Holter ECG and acceleration signals of Figure 2 for a short time interval during a deep breath.

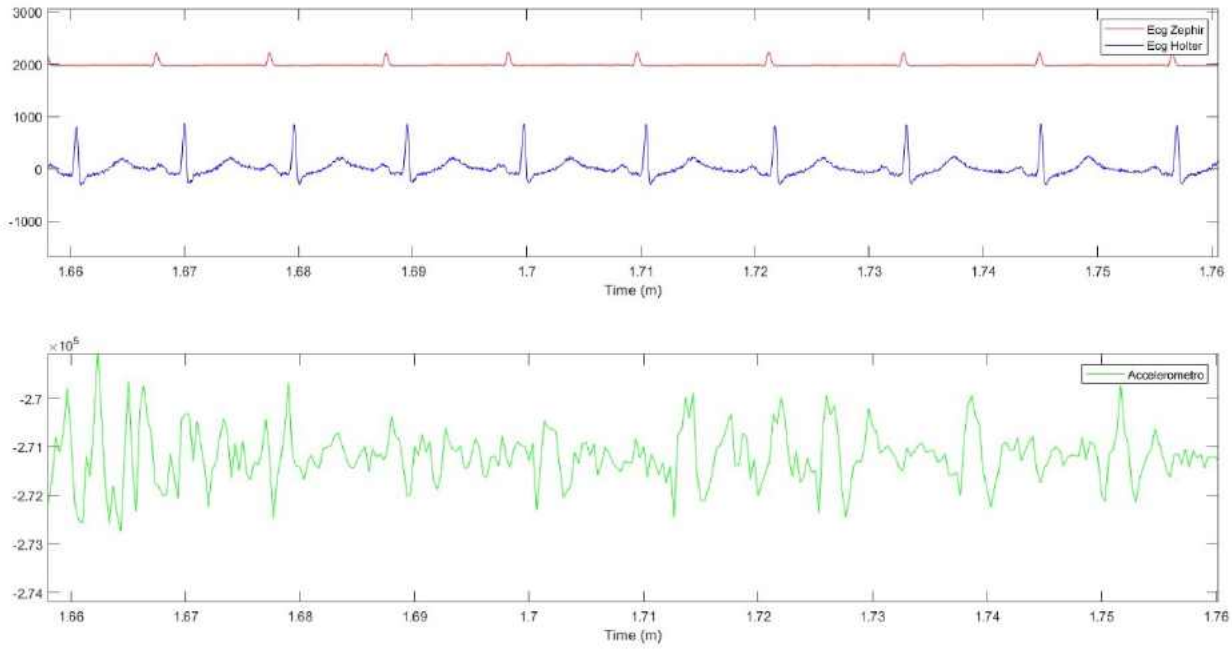


Figure 4. Zephyr ECG, Holter ECG and acceleration signals of Figure 2 for a short time interval during an apnea.

B. Synchronization on signals peaks

In a second phase we considered that the time stamp of the devices can be not accurate and that, even if the initial time stamp is accurate, during the acquisition this synchronization can be lost due to the drift of clock of the devices.

To make the synchronization more precise, we aligned the peaks of the Zephyr and Holter ECG signals. The peaks of the ECG signals have been obtained using the Matlab function called “Pantompkins” [10]. Zephyr ECG and Holter ECG signals has been synchronized on the basis of the peaks. Zephyr ECG and acceleration signals has been synchronized on the basis of the time stamp. Figure 2 reports the Zephyr ECG, Holter ECG and acceleration signals resynchronized in this way. Figure 3 reports the Zephyr ECG, Holter ECG and acceleration signals of Figure 2 for a short time interval during a deep breath. Figure 4 reports the Zephyr ECG, Holter ECG and acceleration signals of Figure 2 for a short time interval during an apnea.

The peaks of the accelerometer signal have been obtained through a signal processing using a convolution with a reference function reported in [8].

To confirm the result of this further synchronization operation, we calculated the time instants in minutes of the peaks of Zephyr ECG, Holter ECG and acceleration signals during the ten minutes of acquisition time, reported in Figure 5.

The perfect synchronization between Zephyr ECG and Holter ECG can be seen in Figure 5. An initial synchronization between Holter ECG an acceleration signal can be seen in Figure 5. The effect of drift of the clock of the Raspberry of the system with the accelerometer can be seen in Figure 5. At the end of the ten minutes of measurements the synchronization between the Zephyr ECG signals and acceleration signal is lost.

A final study is carried out by analysing of the distance between two adjacent peaks of the ECG and the acceleration signals. This operation allows you to observe the result of the synchronization from another point of view. Figure 6 reports the time interval in seconds between two adjacent peaks of the Zephyr ECG, Holter ECG and acceleration signals during the 10 minutes of acquisition time. The peak detection algorithm of the acceleration signal is not always accurate, due to the fact that the signal is not periodic as ca be seen in Figures 3 and 4. Therefore sometimes the period detected is double or half of the period of the ECG signal.

It is possible to make the ECG and acceleration signals more comparable by filtering the acceleration signal with the Matlab "medfilt" function, which filters with respect to the median of the signal. The results are reported in Figure 7.

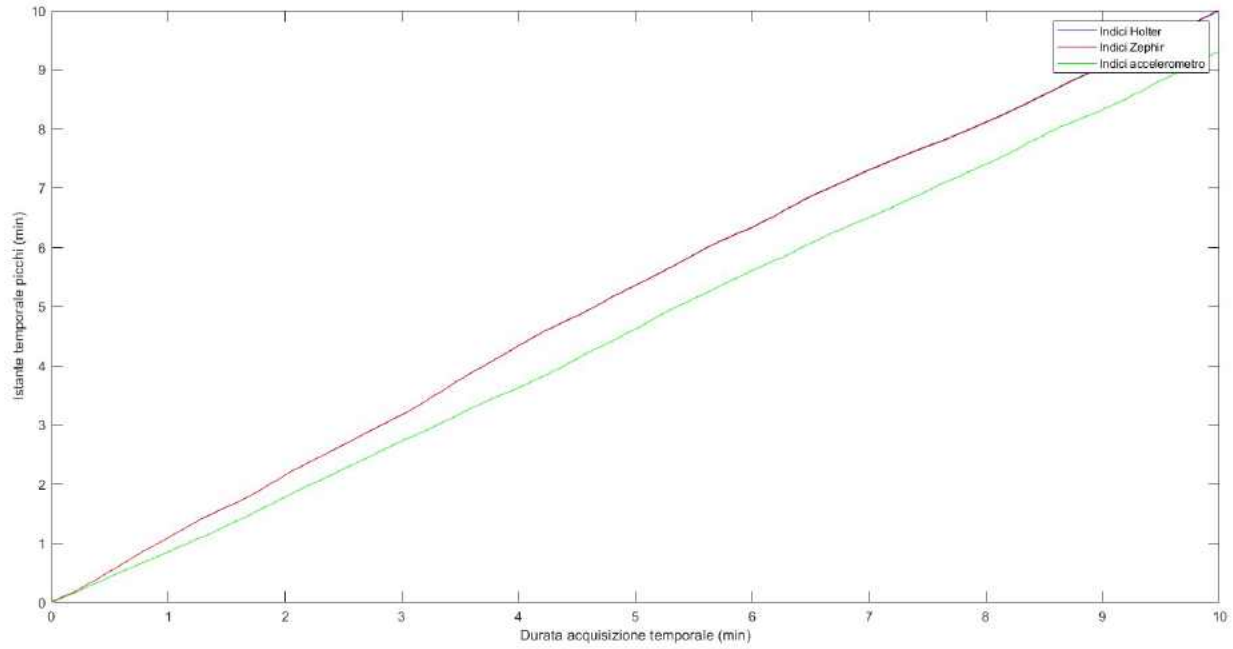


Figure 5. Verification of synchronization. The time instants in minutes of the peaks of Zephyr ECG, Holter ECG and acceleration signals during the acquisition time in minutes.

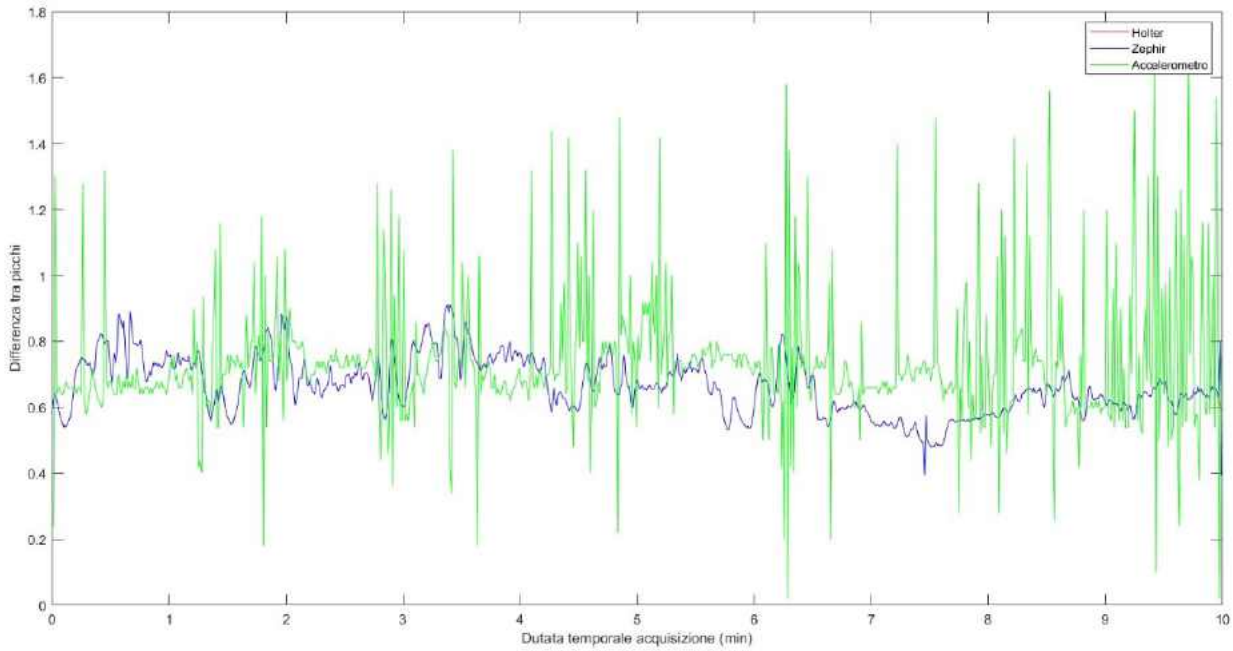


Figure 6. Distance in seconds between two adjacent peaks of the Zephyr ECG, Holter ECG and acceleration signals during the acquisition time in minutes.

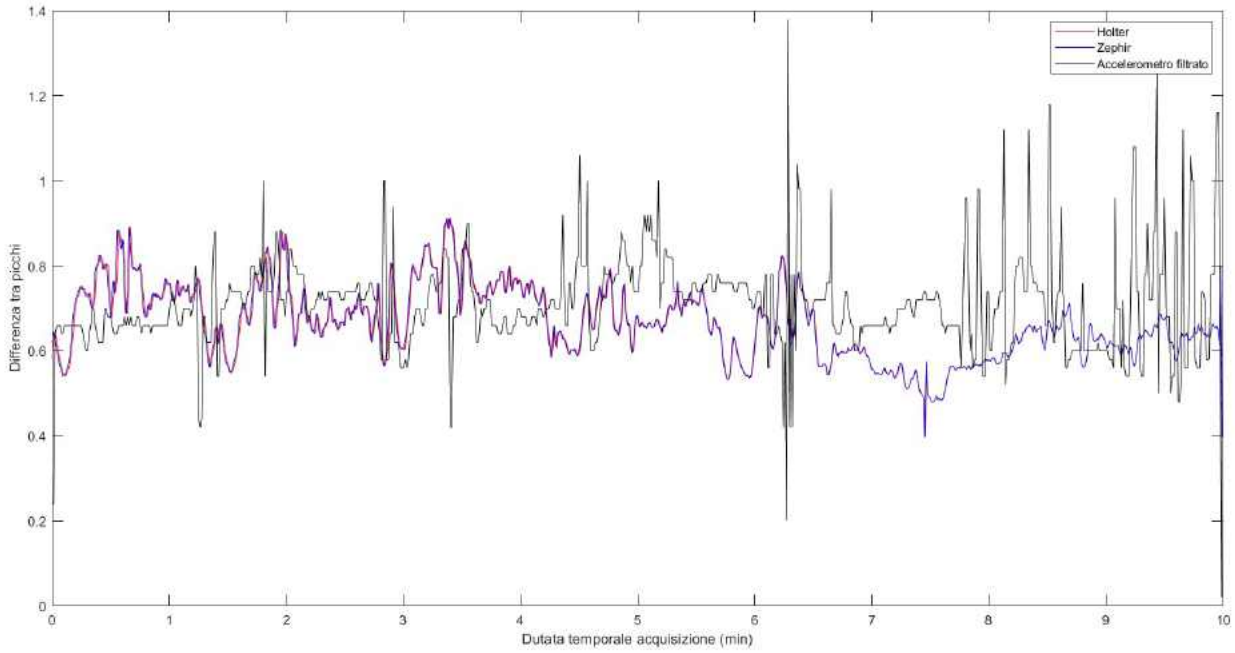


Figure 7. Distance in seconds between two adjacent peaks of the Zephyr ECG, Holter ECG signals and acceleration signal after filtering during the acquisition time in minutes.

III. Conclusions

The relevance of the proposed acquisition system is due to the possibility to acquire information on a patient's cardiac activity using of a low cost accelerometer in non-invasive way.

The work carried out in this work have the objective to verify the accuracy of the accelerometer system. To this aim, the time alignment of ECG and accelerometer signals have been obtained. Future development will be carried out applying the described procedure to the all the data of the 21 volunteers.

References

[1] Giovangrandi, Laurent, et al. "Ballistocardiography—a method worth revisiting." 2011 annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2011.

[2] Gaiduk M, Penzel T, Ortega JA, Seepold R, "Automatic sleep stages classification using respiratory, heart rate and movement signals," Physiol Meas 39(12). December 2018.

[3] Omer T. Inan, et al., "Ballistocardiography and Seismocardiography: A Review of Recent Advances," IEEE Jou of Biomedical and Health informatics, Vol. 19, No. 4, pp. 1414-1427, July 2015.

[4] A. Taebi, et al., "Recent Advances in Seismocardiography," Vibration 2019, 2, 64–86.

[5] I Sadek, J Biswas, "Nonintrusive heart rate measurement using ballistocardiogram signals: a comparative study," Signal, Image and Video Processing 13 (3), 475-482, 2019.

[6] Maksym Gaiduk, Simone Orcioni, Massimo Conti, Ralf Seepold, Thomas Penzel, Natividad Martínez Madrid, Juan A. Ortega, "Embedded system for non-obstrusive sleep apnea detection," Proc. of Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC'20), Montréal, Québec, Canada July 20-24, 2020

[7] Inan, Omer T., et al. "Ballistocardiography and seismocardiography: A review of recent advances." IEEE journal of biomedical and health informatics 19.4 (2014): 1414-1427.

[8] Conti, Massimo, et al. "Heart rate detection with accelerometric sensors under the mattress." 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2020.

[9] Gaiduk, M., et al. (2022, October). Heart and breathing rate measurement using low intrusive monitoring systems. In Social Innovation in Long-Term Care Through Digitalization: Proceedings of the German-Italian Workshop LTC-2021 (pp. 37-49)..

[10] https://it.mathworks.com/matlabcentral/fileexchange/45840-complete-pan-tompkinsimplementation-ecg-qrs-detector (16/04/2023)

Models and Applications for Embedded Systems

Edited by
Massimo Conti, Simone Orcioni
Dipartimento di Ingegneria dell'Informazione
Università Politecnica delle Marche

ISBN: 978-88-87548-00-6