# Impact of Varying Collision Avoidance Strategies on Human Stress Level in Human-Robot Interaction

**Master Thesis**

submitted to

Institute of Control Theory and Systems Engineering

Faculty of Electrical Engineering and Information Technology

Technische Universität Dortmund

by

Mohammed Faizan

Chennai, India

Date of Submission: 4. Februar 2024

*Responsible Professor*:
Univ.-Prof. Dr.-Ing. Prof. h.c. Dr. h.c. Torsten Bertram

*Academic Supervisors*:
M.Sc. Heiko Renz
M.Sc. Khazar Dargahi Nobari

*„Das ist die Widmung / This is the dedication (optional)“*

# Acknowledgement

Das ist die Danksagung / This is the acknowledgement (optional)

# Abstract

Das ist die Kurzfassung (siehe Abschnitt **??**) / This is the abstact (see section **??**).

# Inhaltsverzeichnis

# Nomenclature

**Greek symbols**

**Abbreviations and Acronyms**

| | |
|---|---|
| EDA | Electrodermal Activity |
| GSR | Galvanic Skin Respons |
| HR | Heart Rate |
| NASA-TLX | National Aeronautics Space Administration–Task Load Index |
| SAM | Self-Assessment Manikin |

**Usage of generative AI models**

| | |
|---|---|
| $\ddagger_{\text{MO}}$ | Media optimization: Correction, optimization, or restructuring of entire passages |

**Explanations for the usage of generative AI models and its notation:**
The bottommost level at which the identification is presented regarding the possible uses of generative AI models are subchapters of the 2nd order (e.g., 1.1.1, which may also appear without numbering), as otherwise, the identification would disrupt the reading flow due to frequent occurrences. Algorithms used for implementing generative AI models are mentioned at least in the text or provided as pseudo-code to facilitate appropriate identification.

# 1 Introduction

‡MO

## 1.1 Motivation

‡MO Industry 4.0, also known as the Fourth Industrial Revolution, has brought about significant transformations within the industry, particularly in the manufacturing sector. This revolution has been characterized by the introduction of intelligent technologies such as the Internet of Things (IoT), cloud connectivity, big data, and human-robot collaboration, among others. These advancements have led to notable improvements and innovations, with the core principle and driving force of innovation in Industry 4.0 being the enhancement of efficiency and productivity. Human-robot collaboration, a key component of Industry 4.0, has played a huge role in this advancement by bringing humans closer together and facilitating more efficient and cooperative workflows.

Looking towards the future of the emerging Industry 5.0, the focus shifts towards a more human-centric approach. Industry 5.0 aims to strike a balance between technological advancements and human needs and interests, placing a strong emphasis on sustainable and resilient industrial practices. The goal is to merge the technological efficiency of Industry 4.0 with a greater emphasis on enhancing human well-being and personalizing the production process. Industry 5.0 brings back the human workforce to the factory, where humans and machines are paired to increase process efficiency by utilizing human brainpower and creativity through the integration of workflows with intelligent system (Nahavandi 2019). This shows a significant shift from purely efficiency-driven operations to those that also prioritize human factors and environmental sustainability.

Traditionally, industrial robots like manipulator arms, autonomous mobile robots, and gantry models have been kept separate from human workers primarily due to concerns regarding safety. These robots are typically characterized by their large size, substantial weight, and high speed, attributes that pose potential hazards when in close proximity to humans. Consequently, their design is predominantly focused on fulfilling specific tasks such as drilling, welding, or loading and unloading where their size and speed are necessary for efficiency but also necessitate isolation from human workers to ensure safety. This traditional approach prioritized the physical separation of robots and humans in industrial settings. However, advancements in Industry 4.0 have significantly increased the use of

collaborative robots (cobots) , bringing them closer together to jointly accomplish tasks. This evolutionary progression has witnessed the transformation of robots from being secluded behind safety barriers to now operating side-by-side with their human counterparts, effectively capitalizing on their unique capabilities which combine human adaptability and decision-making skills with the precision and consistency offered by robots.
While technological advancements aim to optimize production, the comfort and well-being of human workers have not always been prioritized. This thesis aims to delve into the human aspect of human-robot interaction, considering how proximity to robots might affect the operator's physiological state. It aims to investigate how continuous interaction with robots impacts the stress levels experienced by humans and emphasizes the importance of monitoring and accurately assessing stress levels in human-robot collaborative environments. Sauppé und Mutlu (2015) have previously indicated that co-bots have the ability to influence the mental states of human workers, as they are often perceived as social entities. The close proximity of humans to robots in the workplace can lead to heightened levels of mental stress, particularly if the movements of the robot appear to be potentially harmful (Lasota und Shah 2015). For instance, if a co-robot swiftly moves towards a worker or follows an unpredictable path, it may induce feelings of anxiety or fear due to the perceived risk of sustaining an injury. This, in turn, can negatively impact both productivity and the efficacy of human-robot collaboration. Furthermore, it can impede the complete utilization of the advanced capabilities offered by collaborative robots. Identifying and addressing these stress factors is key to optimizing the human-robot collaboration for enhanced productivity and make the working environment effective, efficient and safe.

## 1.2 Aim of the Thesis

[‡MO] The aim of this thesis is to evaluate the impact of varying collision avoidance strategies on human stress levels in the context of human-robot interaction. Our objective was to conduct a study to collect and analyze data to understand the different stress levels in relation to varying robot collision avoidance strategies. This is done taking different collaboration levels and robot control strategies into account. We then used this data to create a predictive model that can identify and address sources of stress during human robot collaboration.
Specifically, the objectives of the thesis are:

- **Assessment of Human Stress Levels**: Develop a holistic approach for evaluating stress levels in human-robot interactions, combining both objective physiological measures and subjective experiences. Objectively, the study will employ various physiological indicators such as Galvanic Skin Respons (GSR), Electrodermal Activity (EDA), Heart Rate (HR), and body posture analysis. These indicators will provide quantifiable data on the body's physiological response to robot interactions. Subjectively, the study aims to incorporate personal feedback from participants, gathered through questionnaires. This will offer insights into their personal feelings and perceptions regarding their interactions with robots. By blending these objective and subjective methods, the study aims to provide a comprehensive understanding of stress in human-robot interactions.

- **Development of Data Acquisition and Synchronization System:** Designing an acquisition system that successfully takes data from several sensors at different frequencies and synchronizes it. Devices such as the Empatica E4 wristband are utilized for gathering data on GSR, EDA, and other parameters, as well as a motion capture system to record human posture and movement. A vital aspect would be to synchronize these many data streams, ensuring accurate and consistent assessment of human physiological states across different robot interaction scenarios.
- **Data Collection and Evaluation**: Designing and conducting a subject study to collect data on participants' physiological responses while doing different assembly tasks under different robot-human interaction scenarios. These scenarios included three distinct levels of robot collision avoidance strategies: No Collision Avoidance, Dynamic Collision Avoidance, and Predictive Collision Avoidance as well as three different collaboration levels: Different Workspace with the cobot, Shared Workspace, and Shared Workspace with Direct Collaboration. The aim is to gather comprehensive data to analyze the impact of these varying robot control strategies on human stress levels.
- **Stress Prediction Model**: Developing a model for predicting and classifying stress levels during human robot collaboration. This model trained on the dataset of human physiological responses collected from the subject study.Various preprocessing techniques and feature engineering techniques are used to prepare the data for the model. Various machine learning models such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and others, are evaluated to determine the best model for predicting stress levels.

## 1.3 Related Work

Still not sure if to add realted work related to various stress detection strategies here or add it in the therotical foundations citing how people have used this for stress detection.Will a one page of related work suffice here or does it have to be detailed also looking at if any robot human collaboration experiments/studies have been done before.

## 1.4 Structure of the thesis

Chapter 1 presents the motivation behind the study and the aims of the thesis. In Chapter 2, we delve into the fundamental concepts of stress and examine the relationship between stress and key physiological signals such as EDA, HR and GSR etc. Chapter 3 describes the data collection process of the subject study that was conducted to collect data on participants' physiological responses. Chapter 4 describes the data analysis process of the subject study to analyze the collected data and the different pre-processing techniques and feature selection used to build the model. Chapter 5 describes the results of the subject study and the different machine learning models that were evaluated to determine the best model for predicting stress levels. Chapter 6 concludes the thesis and outlines the future work and research directions.

# 2

# Theoretical foundation

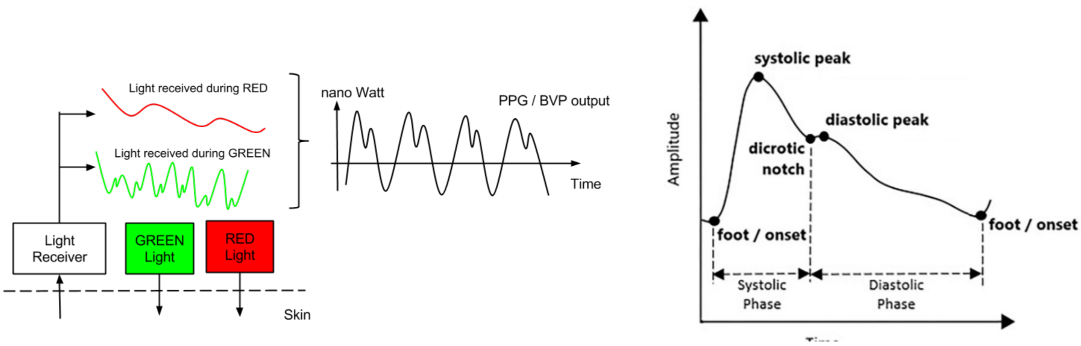
## 2.1 Stress Measurement

### 2.1.1 The Stress Response: Physiology of Stress

### 2.1.2 Photoplethysmogram-PPG

Photoplethysmogram (PPG) also known as Blood Volume Pulse (BVP) are non-invasive optical techniques used to monitor changes in blood volume within tissue. They rely on the principles of light absorption and reflection to capture valuable information about cardiovascular activity. PPG and BVP sensors are commonly found in wearable devices. PPG signals are obtained by transmitting light into the skin and measuring the amount of light either transmitted through or reflected back.
When the heart beats, it propels blood through the circulatory system, causing periodic changes in the volume of blood vessels. PPG sensors emit light into the tissue and measure the amount of light that is either absorbed or reflected back. During each heartbeat, blood absorbs more light, leading to a decrease in the amount of light detected by the sensor. Between heartbeats, when blood flow is less pulsatile, more light is detected.
The resulting waveforms from both PPG and BVP typically consist of a series of peaks and troughs, with each peak corresponding to a heartbeat (systole) and each trough representing the resting period between beats (diastole). By analyzing the time intervals between these peaks, the heart rate can be calculated. This heart rate measurement is fundamental and provides valuable information about a person's cardiovascular health and overall fitness level. It serves as a key metric in various applications, including exercise tracking, stress assessment, and medical diagnosis.
Furthermore, PPG signals enable the assessment of heart rate variability (HRV). HRV is the variation in time between successive heartbeats and is an essential indicator of the autonomic nervous system's activity. By analyzing the subtle changes in the intervals between PPG peaks, HRV can be quantified. High HRV typically indicates a healthy heart and a well-balanced autonomic nervous system, while reduced HRV can be associated with stress, illness, or various medical conditions. HRV analysis provides insights into the body's ability to adapt to different situations and is valuable for assessing stress levels, mental well-being, and overall cardiovascular health.
Other measures that can be derived from PPG data include and estimation of blood oxygen

saturation levels (SpO2), valuable for respiratory and circulatory health assessment. PPG can also be used to estimate respiration rate, reveal vasomotor activity changes associated with the autonomic nervous system, emotions, or vascular health, and provide insights into arterial stiffness and blood flow dynamics. First derivative and second derivates of PPG signals can also be analaysed. The first derivative (Velocity Plethysmogram, VPG) and the second derivative (Acceleration Plethysmogram, APG) features can be used to blood pressure estimation etc.(Suboh u. a. 2022)
PPG not only provides heart rate data but, when analyzed, can also reveal heart rate variability (HRV), which is indicative of the body's stress and relaxation levels.




(a) PPG process (b) Typical PPG Waveform

Abbildung 2.1: (Empatica o. D.) (Suboh u. a. 2022)

### 2.1.3 Electrodermal Activity-EDA

Electrodermal Activity (EDA), also known as the galvanic skin response(GSR), is a way to measure changes in how our skin conducts electricity. Even moderate amounts of sweating that are not observable at the skin surface can alter skin electrical conductivity. The more the body sweats, the more conductive the skin becomes, and this change can be measured to infer physiological or psychological states.More specifically EDA measures the skin's electrical conductance changes, which depend on the quantity of sweat secreted by eccrine sweat glands in the hypodermis of the palmar and plantar regions. Sweat secreted in the palmar and plantar regions is caused mainly by central nervous activity related to affective and cognitive states, including mental or emotional sweating. Thus EDA becomes one of the promising noninvasive methods widely used in detecting stress and emotion. EDA is a powerful method for real-time measurement and could be used as an index of emotional or cognitive stimulation related to stress.(Gellman 2020).EDA is useful in several ways: it shows how we respond emotionally, helps us see how our body reacts to stress etc. It acts as a biomarker for emotional responsiveness and serves as a key indicator for stress-related bodily responses.
Electrodermal Activity (EDA) comprises two main components: the tonic and phasic components. The tonic component, also known as skin conductance level (SCL), reflects slow and consistent changes in the signal's background. In contrast, the phasic components, referred to as skin conductance response (SCR) or spontaneous fluctuation of skin response, are the rapid and momentary fluctuations within the signal that occur within specific

time intervals (2017). SCR appears in response to stimuli activating the sympathetic nervous system. Consequently, SCR can be linked to a stimulus and can be valuable in measuring cognitive stress levels. However, directly extracting the components of EDA isn't straightforward.
When EDA sensors measure skin conductivity (SC) signals, they typically yield results in microsiemens. To extract the SCL and SCR components accurately, it is necessary to deconvolve the SC signals (Alexander u. a. 2005, postnote). Without proper separation of the original SC signals, overlapping SCRs can lead to less precise information during feature extraction . Therefore, it is crucial to perform deconvolution to distinguish the SCR and SCL signals effectively.

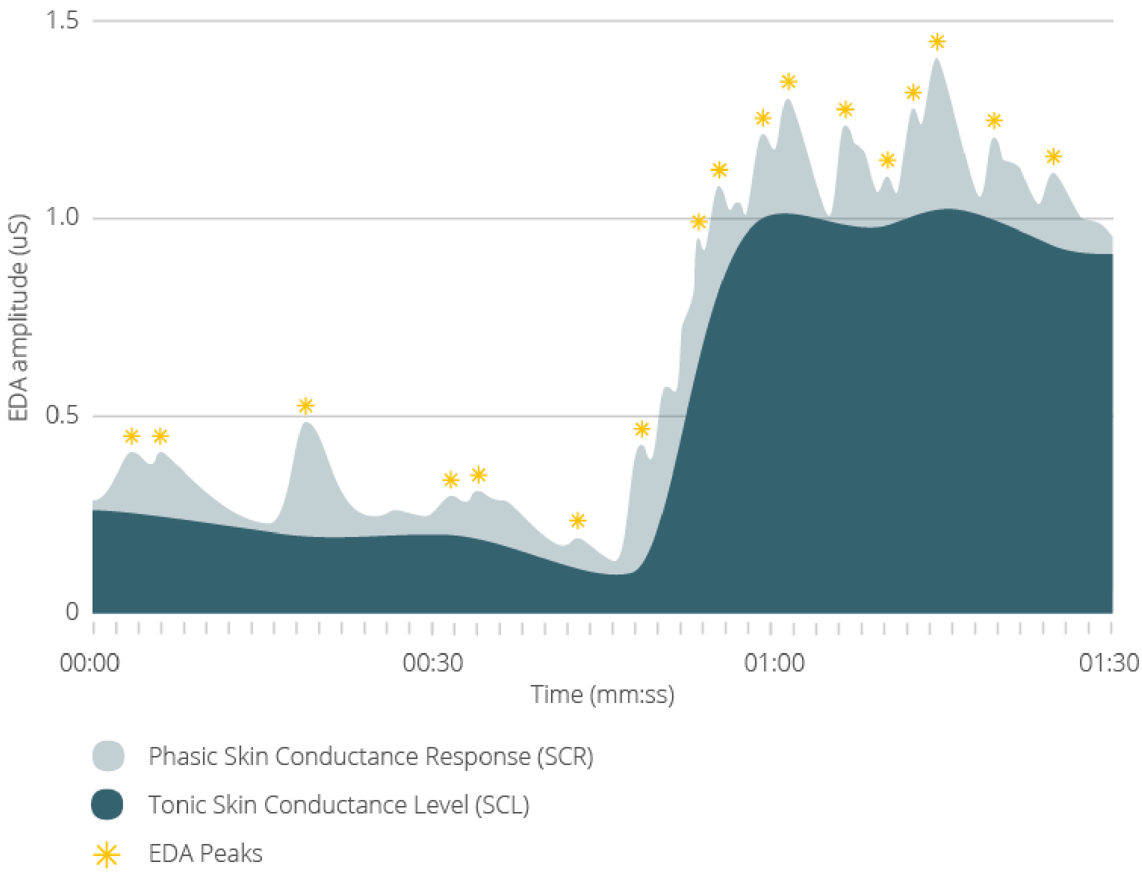



Abbildung 2.2: EDA Example signal - (imotions o. D.)

### 2.1.4 Empatica E4

For all the physiological signals explained above, we use the the Empatica E4: The Empatica E4 wristband is a versatile and compact device designed to capture a wide range of physiological data in real-time. Equipped with sensors for measuring metrics like heart rate, skin conductance, temperature, and motion. Its sleek and unobtrusive nature makes it comfortable to wear, while its comprehensive data collection capabilities have made it an invaluable asset.

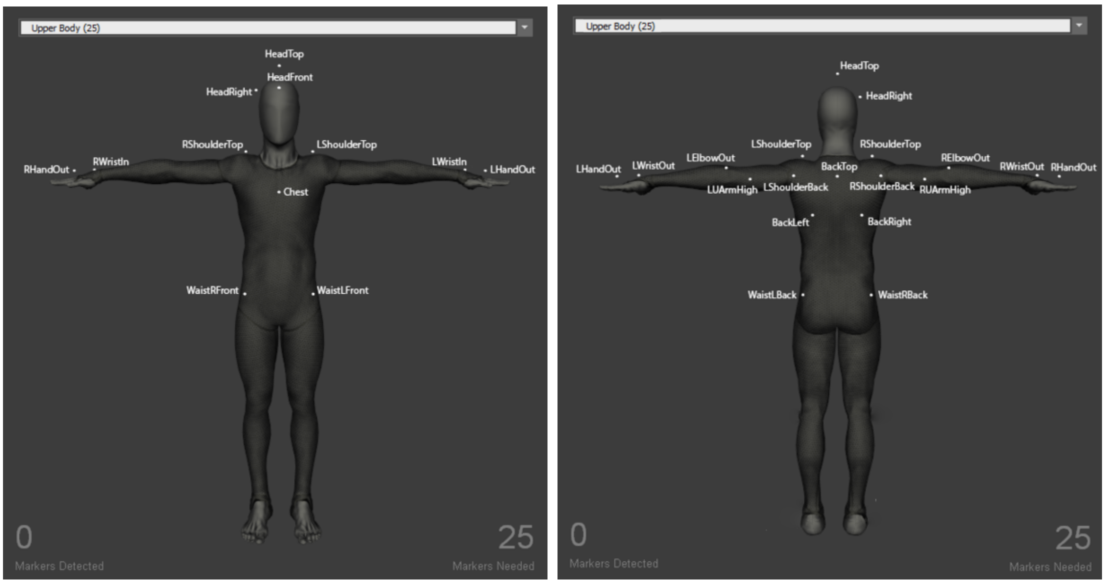
The Empatica E4 wristband collects photoplethysmography (PPG) data using a process that involves emitting green and red light from LEDs into the skin and measuring the reflected light with a sensor. The green light, which is absorbed by the blood, provides a pulsatile signal corresponding to the cardiovascular pulse wave, used to determine heartbeats. The red light acts as a reference to correct for motion artifacts. This data is then processed by algorithms within the wristband to output the blood volume pulse (BVP), from which the interbeat interval (IBI)—the time between heartbeats—is calculated, offering a non-invasive method to monitor heart rate continuously.

Electrodermal Activity (EDA) is measured by detecting the electrical conductance across the skin, which is an indirect indicator of the sweat gland activity influenced by the sympathetic nervous system. To obtain these measurements, Empatica employs a method that relies on passing a very small and electrical current between two electrodes that are in contact with the skin, typically placed on the bottom wrist.

The Empatica E4 combines these measurements with additional sensors that track body temperature and movement, providing a comprehensive overview of the wearer's physiological state. This approach allows for a more advanced understanding of the wearer's health, offering potential applications in medical research, clinical trials, and personal health monitoring.


Abbildung 2.3: Empatica E4 features (Empatica o. D.)

https://support.empatica.com/hc/en-us/articles/204954639-Utilizing-the-PPG-BVP-signal
https://support.empatica.com/hc/en-us/articles/360029719792-E4-data-BVP-expected-signal




(a) Front View (b) Back View

Abbildung 2.4: 25 Upper Body Marker Set

### 2.1.5 Motion Capture

## 2.2 Subjective Measures

Mental stress can be assessed through subjective and objective measures. Subjective ratings, such as self-report questionnaires, have been commonly used to estimate levels of mental stress in humans (Aigrain et al., 2018). Participants are asked to answer a variety of questions about their experiences in the experiment. The NASA-Task Load Index (NASA-TLX) has been utilized in numerous research studies to assess people's mental stress levels. For instance, Zheng et al. (2012) employed the NASA-TLX to investigate the mental workload experienced by surgeons during endoscopy training. In the context of smart factories, Zakeri et al. (2021) applied the NASA-TLX to examine various factors contributing to mental stress, such as task complexity, time constraints, and collaboration duration. However, it is important to acknowledge the limitations of self-reporting, as participants cannot report in real-time and may not express their true feelings (Bethel et al., 2007).

## 2.3 UR10 robot and collsion avoidance

# 3 Data Collection-Subject Study

A collaborative assembly task involving wooden pieces was designed in the robot laboratory of the Institute of Control Theory and Systems Engineering at the Technical University of Dortmund. This setup aimed to accurately replicate the types of tasks commonly seen in industrial settings where collaboration between humans and cobots is frequently observed. The Universal Robot UR10 was selected as the cobot to be used in the study. A call for participants invite was sent out, and 20 male students who all had a technical background volunteered to participate in the study. 17 of the students did not have any previous experience working with a robot in any way, whereas the remaining had some previous experience with robots. All of them were between the ages of 21-28. Before the experiment began, all participants were given a comprehensive overview of the study's objectives and procedures, along with a consent form. Only those participants who agreed to the terms and signed the consent form were permitted to proceed with the experiment. To ensure the ethical integrity of the study, a prior request for ethical approval was submitted to the appropriate ethics council (to be added) and permission was obtained to conduct the subject study.

## 3.1 Design of Tasks

Since we wanted to replicate an industrial assembly task, assembly of various mock items using wooden children's toys was selected. The effect of stress on different factors was to be investigated.
These included three distinct levels of different collaboration levels:

- **Different Workspace**: Human and cobot have no overlapping space. The cobot works in the background.The human already has all items required for the assembly task in front of him.
- **Shared Workspace**: Where the human and cobot share the same work area.The cobot brings the each item required for the assembly tasks to the human and places it on the table in front of the human.
- **Shared Workspace with Direct Collaboration**: Where the human and cobot share the same work area as well, The cobot brings the item required for the assembly tasks to the human and directly hands over the items to the human.

| **Collision Avoidance Strategy** | **Different Workspace (A)** | **Shared Workspace (B)** | **Shared Workspace with Direct Collaboration (C)** |
|---|---|---|---|
| **No Collision Avoidance (X)** | AX | BX | CX |
| **Dynamic Collision Avoidance (Y)** | Not Available | BY | CY |
| **Predictive Collision Avoidance (Z)** | Not Available | BZ | CZ |

Tabelle 3.1: Task names for different Collision Avoidance Strategies and Workspace Scenarios

As well as three robot collision avoidance strategies:

- **No Collision Avoidance**: No collision avoidance measures are in place, and the robot stops at a collision.
- **Dynamic Collision Avoidance**: The robot identifies the human as a dynamic obstacle and adjusts its trajectory to avoid collisions.
- **Predictive Collision Avoidance**: This strategy uses predictions to predict the human's future position and adjusts its trajectory to avoid collisions.

The table 3.1 shows the various combinations of factors and the naming conventions of the tasks, yielding a total of seven experimental scenarios since collision avoidance is not applicable when different workspaces are involved. For example task BZ employed a predictive collision avoidance strategy whilst the task was done in a shared workspace. So a within-subject experimental design was employed where each participant was tested on all seven different scenarios. By having the same participants perform each task, we minimized the impact of differing skill sets, experiences, and cognitive abilities, which could otherwise skew the results.

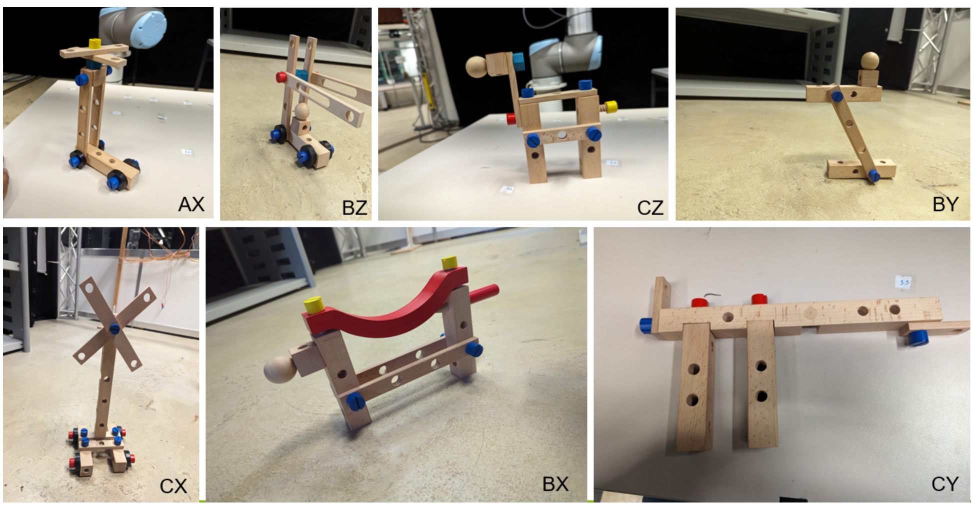



Abbildung 3.1: 7 different assembly tasks

To avoid any potential learning effects and ensure that each task measures the intended variables accurately, we had to design 7 different assembly tasks. Each task involved assembling a unique item, carefully chosen of similar complexity and required effort. This similarity in task difficulty was crucial to avoid the introduction of other variables such as familiarity with the task which could affect the results. Figure 3.1 shows the final assemblies completed in each of the seven tasks. Each task had approximately 4-6 different parts which was supposed to be screwed together to complete the tasks. By standardizing the complexity across tasks, we aimed to isolate the impact of the collision avoidance strategies and collaboration levels. We also had to consider the order in which the task was administered for each participant, ensuring that learning or fatigue did not affect the outcomes in any way. Each participant experienced the tasks in a unique order, balancing out any potential biases introduced by the order of task presentation.

## 3.2 Apparatus and Experimental Setup

The experiment was set up in a specially designated area of our laboratory. At the center of this arrangement was the collaborative workspace, featuring a table and chairs for the human participant (Area B in Fig 3.2), positioned directly opposite the robot's dedicated workspace. Adjacent to this, on the right side of the collaborative workspace, a table was placed to hold the various items needed for the assembly tasks(Area A Fig3.2). The robot would pick the necessary items from this table for each specific task and deliver them to the human participant. In front of the human participant, a mobile device was also placed. This device was key to the experiment, as it presented the participant with concise, step-by-step instructions for each assembly task. These instructions were visually displayed, offering clear and easy-to-follow guidance. The participant would start the assembly task as soon as the first item is delivered to the human participant. The robot would then proceed to the next item, and so on until all items are delivered to the human participant. Figure 3.2 shows the experimental setup.

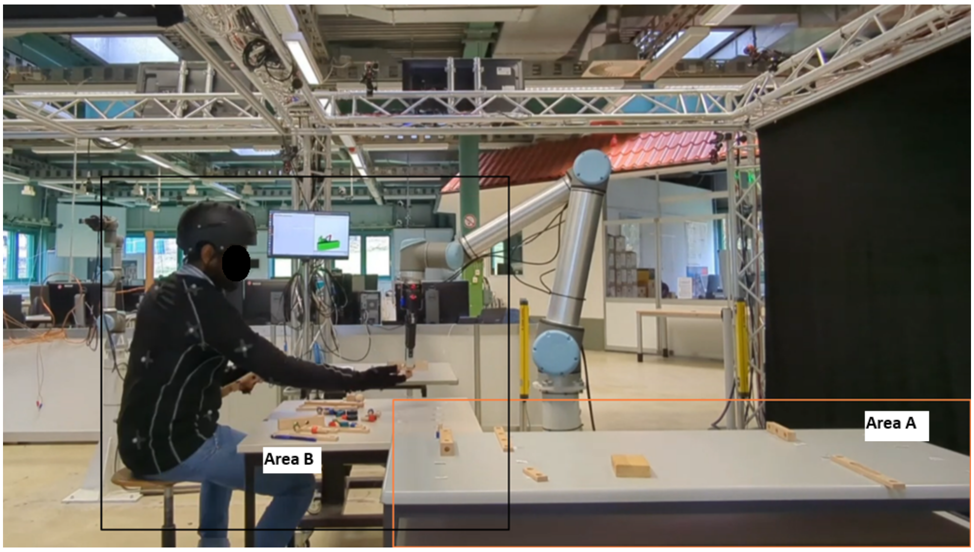



Abbildung 3.2: The experimental setup

The entire experimental area had an advanced OptiTrack motion capture system, outfitted with 12 high-precision cameras fitted across all 4 sides as seen in Fig 3.3. These OptiTrack cameras, known for their exceptional accuracy and low latency are used to capture every detail of the human participant's movements. This setup was necessary in providing a detailed and continuous record of the participant's interactions with the robot, also aiding the collision avoidance trajectory planning for the robot. The use of the OptiTrack system enabled us to gather precise data on human motion. The participants were equipped with the motion capture suit which had 25 distinct marker points which were used to capture the human's head and upper body. To prioritize participant safety, especially given the proximity to a large robotic arm, a helmet was provided to each participant.

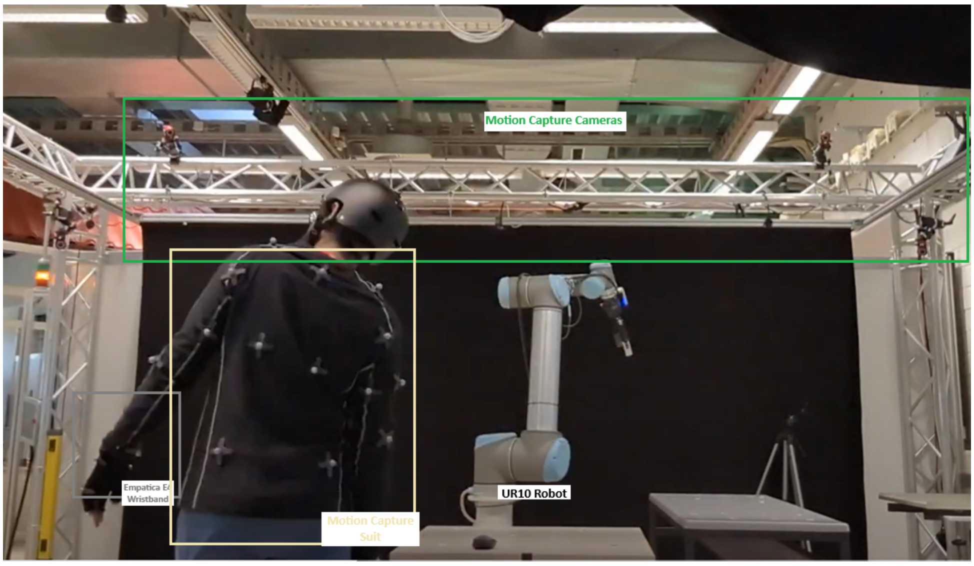



Abbildung 3.3: Apparatus used

For capturing the physiological signal of the participants, the Empatica E4 wristband was equipped in the participant's non-dominant hand. The participant's physiological signals such as GSR, EDA, HR and temperature, among others are captured by the Empatica E4 wristband, which transmits data wirelessly via Bluetooth to a Windows PC. This PC runs the E4 streaming server, facilitating the real-time transfer of this data. The various motion capture cameras recording the participant's movements are synced together and are connected to the Windows PC as well running the motion capture software, Motive.

The physiological data from the Empatica E4 and the motion data from the motion capture system are then streamed to a Linux PC running the Robot Operating System (ROS) specifically ROS1 Melodic. The motion capture data is published to /tf topic. Whereas the Empatica E4 node is available as a ROS2 node running inside a docker container interfacing with the ROS1 using a ROS bridge. Then a data synchronization script is used to create a ROS node that subscribes to the multiple topics from various sources, synchronizes the incoming data, and publishes a compiled message to the /aggregated_-data topic. A more detailed description of the data synchronization process is provided in Section **??**. This synchronized data topic is then recorded to a rosbag. A general schematic of this is shown in Fig 4.1.

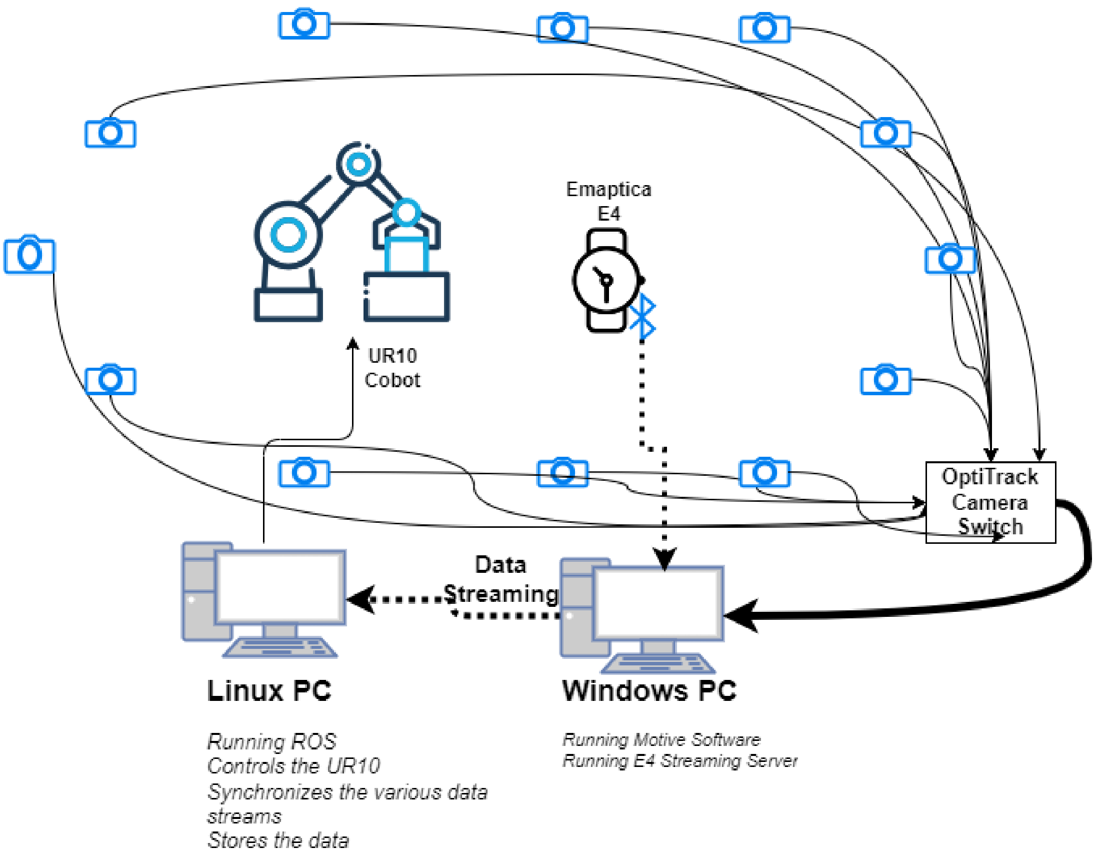



Abbildung 3.4: Schematic of the experiment setup

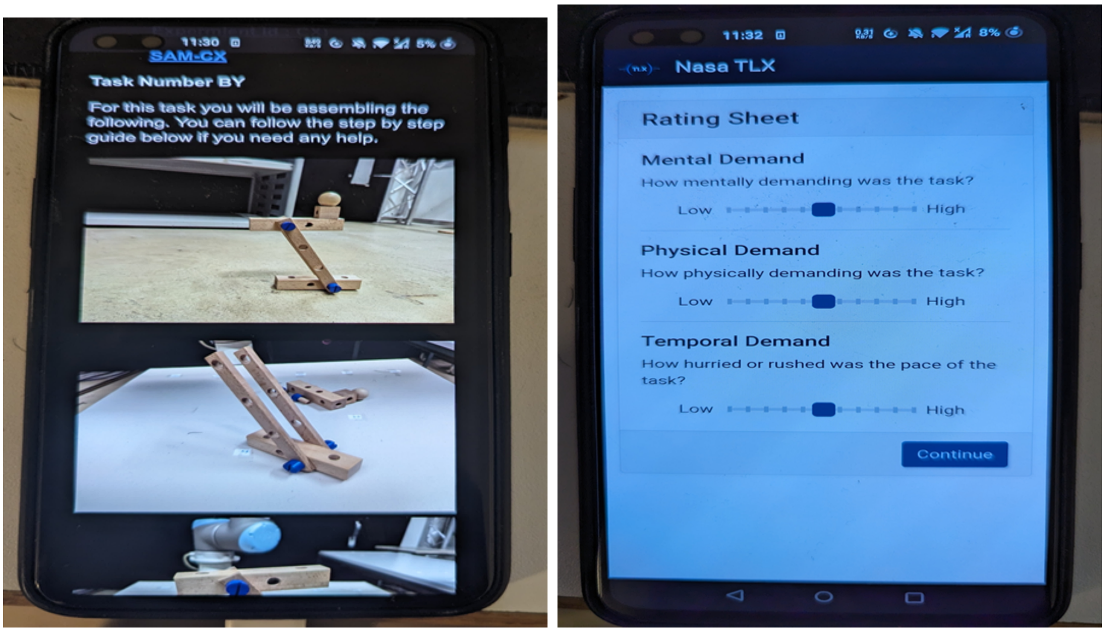
## 3.3 Experimental Procedure

The participants visited the experiment room in the time slot they had selected. Upon arrival, they were greeted and guided through a comprehensive orientation that explained the experimental procedures. This session included detailed descriptions of the various equipment involved, such as the Empatica E4 wristband for monitoring physiological responses and the motion capture system for observing and recording precise movement.Each participant was fitted with the motion capture suit and the Empatica E4 wristband, which was placed on their non-dominant hand, and both were carefully calibrated for accurate data collection.

Participants then were given an initial questionnaire that included a consent form, general information, and questions about their prior experience with collaborative robots as well as the General Attitudes Towards Robots Scale (GAToRS) questionnaire.

Once the preliminary documentation was complete, we established a baseline of physiological signals for each participant, which involved recording data for 2 minutes without any interaction with the cobot. This step ensured that we had a standard reference point for each participant's physiological state prior to beginning the tasks.

The main experimental procedure involved a sequence of seven distinct tasks, with the sequence randomized for each participant to control for learning effects. Before the start of each task, a two-minute briefing was provided. This briefing not only outlined the objectives




(a) Step by Step Task instructions (b) NASA-TLX Questionnaires

Abbildung 3.5: Screenshots from mobile device

and requirements of the task but also walked the participant through the instructions displayed on a mobile device, ensuring clarity and preparedness.Figure 3.5 illustrates the interface that participants encounter on the mobile device during the experiment. After the introduction, participants performed the task (Task i), during which both physiological and motion data were recorded. Each task lasted for about 5 minutes in average.

Upon completion of each task, participants were asked to fill out a post-task questionnaire. This included theNational Aeronautics Space Administration–Task Load Index (NASA-TLX)to assess cognitive workload and the Self-Assessment Manikin (SAM) to measure emotional response. These instruments were crucial for evaluating the impact of the task on the participant and infer subjective stress levels and emotional well being from each task. Whilst the participants were filling the questionnaires, the experimental setting was reset to their original position in preperation for the next task.

After a participant had completed all tasks, we conducted a debriefing session. During this session, the participant could provide feedback and discuss their experiences as well explain the whole aim of the study and research. The whole session lasted for 45-60 mins on average

The structured design of this protocol ensured the collection of consistent and reliable data on human-robot interaction, with careful consideration of participant engagement and task impact.

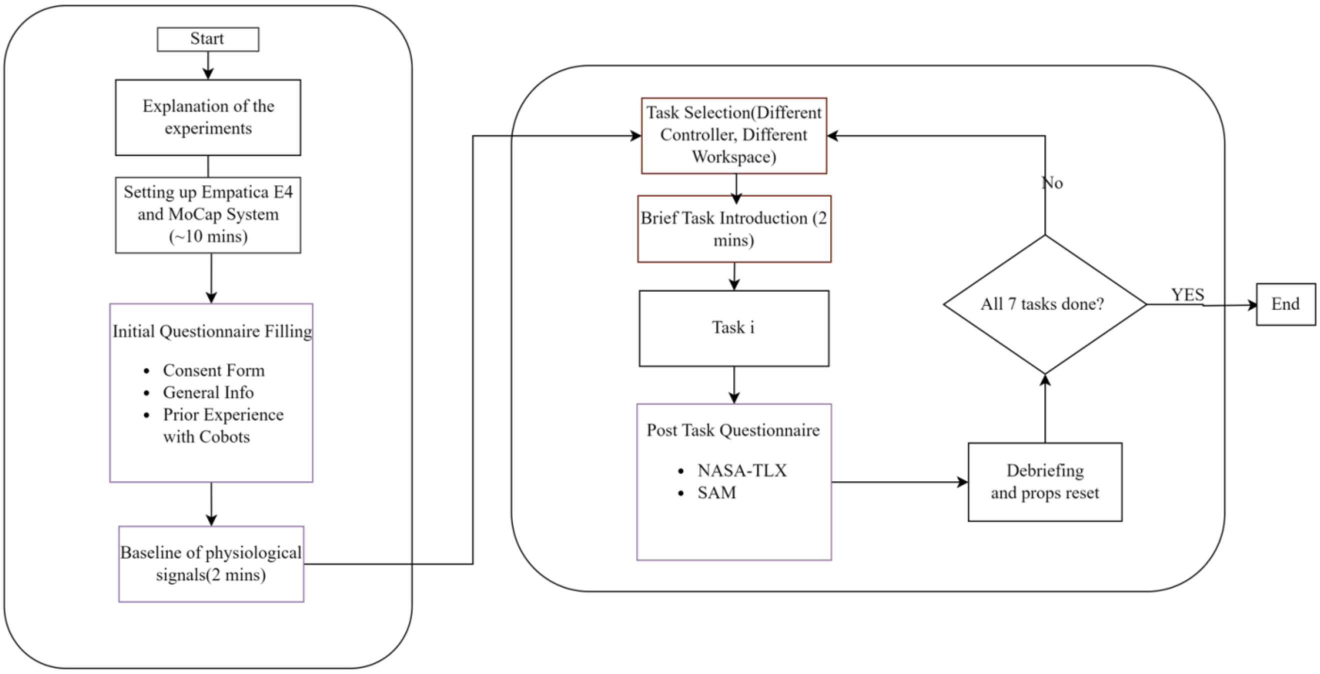



Abbildung 3.6: Schematic of the experiment protocol

# 4

# Stress Detection Methodology

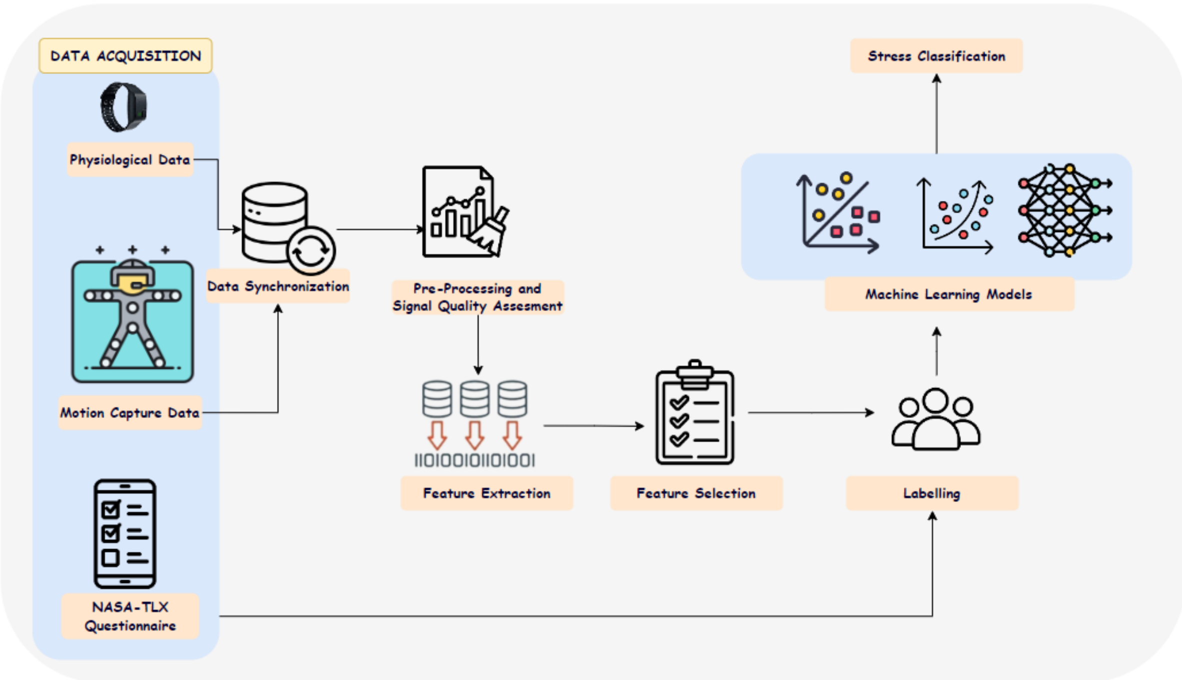



Abbildung 4.1: Schematic of the experiment setup

## 4.1 Data Synchronization

‡MO As previously eluded in section 3.2, data synchronization is crucial to our experimental framework, ensuring the consistency of data gathered from all the different sensors. In our setup, since the Empatica E4 wristband collects various physiological signals, including Blood Volume Pulse(BVP), Electrodermal Activity (EDA), Heart Rate (HR), and skin temperature at different frequencies as well as the motion capture system that tracks the participant's movements at a higher frequency. Given the varying sampling rates of these data sources, it's crucial to synchronize them to ensure they are comparable and accurate. The Empatica E4 wristband samples at different rates for different signals: the Blood Volume Pulse (BVP) data is transmitted at 64Hz, while other metrics like EDA and

temperature are transmitted at a lower rate of 4Hz. In contrast, the motion capture data, processed by the Motive software, is typically captured at a higher rate, around 120Hz. These discrepancies in sampling rates necessitate a careful synchronization approach.

Our synchronization strategy focuses on aligning diverse data streams to a unified frequency. Considering the need for detailed data and practical data processing constraints, we standardize all data streams to a frequency of 64Hz, matching the Blood Volume Pulse (BVP) rate from the Empatica E4. This standardization process involves downsampling the motion capture data, which originally records at a higher frequency of 120Hz. Simultaneously, for other signals like the Electrodermal Activity (EDA) and temperature, updating at a lower frequency of 4Hz, and acceleration data at 32Hz, we use a forward-filling approach, carrying their most recent values until an update occurs.

A specialized ROS node is responsible for this synchronization. By using the BVP rate as the master reference, the node ensures that both physiological and motion data are synchronized in time. This alignment is key for an integrated and comprehensive analysis of the participant's responses, providing a dataset that accurately reflects both the physiological states and physical movements.

Once synchronized, the data is then published to a ROS topic, typically /aggregated_data. This topic carries a comprehensive stream of information that combines detailed physiological measures with precise motion data. This rich dataset forms the foundation for subsequent phases of analysis in our study, such as feature extraction and stress detection.

## 4.2 Pre-Processing

[‡MO] Pre-processing is a crucial step in the analysis of data from physiological sensors and motion capture systems, as it prepares the raw data for subsequent processing and analysis phases. This section outlines the key aspects of pre-processing in our study.

### Data Filtering

In the pre-processing phase of our study, data filtering plays a crucial role in refining the quality of the signals gathered from the physiological sensors and motion capture systems. This step involves applying specific filtering techniques to the raw data to remove any unwanted data, thereby enhancing the signal's clarity and usability for further analysis.

The primary objective of data filtering is to isolate the significant aspects of the signal while eliminating any unwanted noise or interference. Depending on the nature of the signal and the type of noise present, different filtering methods are employed. For instance, we use N-th order Butterworth filters, which are effective in retaining the desired frequency range while attenuating frequencies outside this range. The Butterworth filter is known for its smooth frequency response and is particularly useful in physiological signal processing where preserving the integrity of the signal is crucial.

Each signal type dictates specific filter parameters like filter order, cutoff frequencies, and filter type (lowpass, highpass, or bandpass). This careful selection ensures the final signal is representative of the true physiological data crucial for accurate analysis.

### Data Normalization

Normalization is a critical step in data pre-processing aswell, particularly when dealing with signals of varying magnitudes or scales. Our approach involves applying a normalization process to each input signal, which standardizes the range of data values. This step is essential for comparing and combining data from different sensors effectively.
The method we use for normalization is primarily the 'z-score' method. This technique transforms the data to have a mean of zero and a standard deviation of one. By doing so, it ensures that each signal contributes equally to the analysis, irrespective of their original scale or distribution. This standardization is crucial for machine learning models, as it enhances algorithm performance and prevents any single feature from dominating due to its scale.
Normalization also aids in mitigating the impact of outliers, as it brings all data points onto a common scale, making them more suitable for analysis.

### Signal Quality Assessment

Signal quality assessment is an integral part of the preprocessing phase to ensure the reliability and accuracy of data collected from sensors, which is fundamental for accurate analysis. Various methods are employed to assess the quality of signals, each targeting specific types of anomalies or artifacts.

**Detection of Clipped Segments:** This method involves identifying segments in the signal that are clipped or truncated. Clipping often occurs when the signal amplitude reaches the sensor's recording capacity limits. By setting thresholds for positive and negative clipping, the method detects and marks these segments, facilitating their exclusion or correction .

**Detection of Flatline Segments:** This method identifies flatline segments where the signal shows minimal variation over a period. Such segments can indicate sensor displacement or malfunction. The method identifies these periods by assessing the duration of flatness and the threshold for change in signal amplitude, helping exclude non-physiological data from the analysis.
Each method plays a crucial role in verifying the integrity of the signal data. Identifying and addressing issues like clipping, flatlining, and inconsistent patterns, signal quality assessment ensures that subsequent data analysis stages are based on accurate and reliable data.

### Baseline Correction

Baseline correction forms a pivotal part of our data normalization strategy, particularly tailored for participant-specific physiological data. This approach is centered around the concept of adjusting the data relative to each participant's baseline physiological state, typically captured during a rest period prior to the experimental tasks. This preparatory measure establishes a reference point against which subsequent physiological responses are compared. In the baseline correction process, we begin by computing the average values of physiological signals recorded during the baseline phase before the start of the experiment. This baseline phase is critical as it represents a period of rest where the participant's physiological state is unaffected by experimental stressors. By establishing this baseline, we are able to set a reference point that reflects the participant's normal physiological state. Subsequently, we adjust the data points collected during the active phases of the experiment relative to these baseline averages. This adjustment is a normalization process that centers the data around a personalized zero point, effectively accounting for individual physiological variations. The core advantage of baseline correction lies in its ability to mitigate the influence of inter-individual variability on the physiological measurements. Since each participant's baseline state can vary significantly due to factors like inherent physiological differences, stress levels, and environmental conditions, normalizing data against this baseline ensures a more accurate and personalized assessment of stress responses. This method ensures that the changes observed in the physiological data during the experiment are indicative of the participant's response to the experimental conditions, rather than being a reflection of their baseline physiological state.

### Signal Segmentation

In our research, the segmentation of physiological and motion capture data into windows was a crucial part of the preprocessing. This process involved breaking down the continuous data streams into smaller, manageable windows for detailed analysis. The selection of window size and step size was critical, and was tailored based on the characteristics of the signal and the objectives of our analysis.
The window size was carefully chosen to capture relevant physiological and behavioral patterns within each segment, balancing the need to encapsulate meaningful data against the computational demands of processing. The step size determined the overlap between these windows, ensuring continuity and that no significant transient events were missed.
Accounting for the sampling rate of each signal was vital in customizing the segmentation process appropriately. This flexible approach was key to accommodating different types of signals, ensuring that the window size was appropriate for the length of the signal and that the segmentation parameters were compatible with each signal's nature.
Segmenting data into windows enabled us to convert the ongoing data streams into a format suitable for comprehensive analysis. This structured approach facilitated subsequent computational processes, including feature extraction and pattern recognition, essential for robust stress detection and analysis. This method of using windows in data segmentation is fundamental in ensuring that each part of the continuous data stream is analyzed effectively, allowing for a thorough understanding of stress indicators within the dataset.

## 4.3 Feature Extraction

‡MO Feature Extraction and Selection play a pivotal role in the effectiveness of machine learning models, especially in the context of human stress recognition. Feature extraction involves deriving meaningful attributes from the raw data collected. The type of features extracted can vary widely, including statistical features, time-domain, frequency-domain, as well as linear versus non-linear features.

The complexity of these features can range from basic statistical measures like mean, median, minimum, and maximum, to more intricate features based on specific data modalities. Each sensor used in stress detection may yield a unique set of features, contributing to the overall data analysis and model accuracy. The selection and application of these features are crucial, as they directly impact the classification stage, ultimately influencing the model's performance in stress recognition.

### 4.3.1 EDA-Electrodermal Activity

Electrodermal Activity (EDA), also known as galvanic skin response (GSR), is a sensitive measure of emotional and physiological arousal. It primarily consists of two components: tonic (Skin Conductance Level, SCL) and phasic (Skin Conductance Response, SCR).As explained in Section 2.1.3 the tonic component represents baseline levels of skin conductance, reflecting slow changes in arousal state. The phasic component, on the other hand, captures rapid fluctuations in response to specific stimuli or events.
After the usual process of filtering and normalizing the signal aswell checking the quality of the signal we first decompose the EDA signal into its tonic and phasic components using continuous decomposition analysis. This process allows us to separately analyze the steady-state (SCL) and transient (SCR) aspects of skin conductance. The decomposition is typically carried out using highpass filtering techniques, ensuring that each component accurately represents the underlying physiological processes.

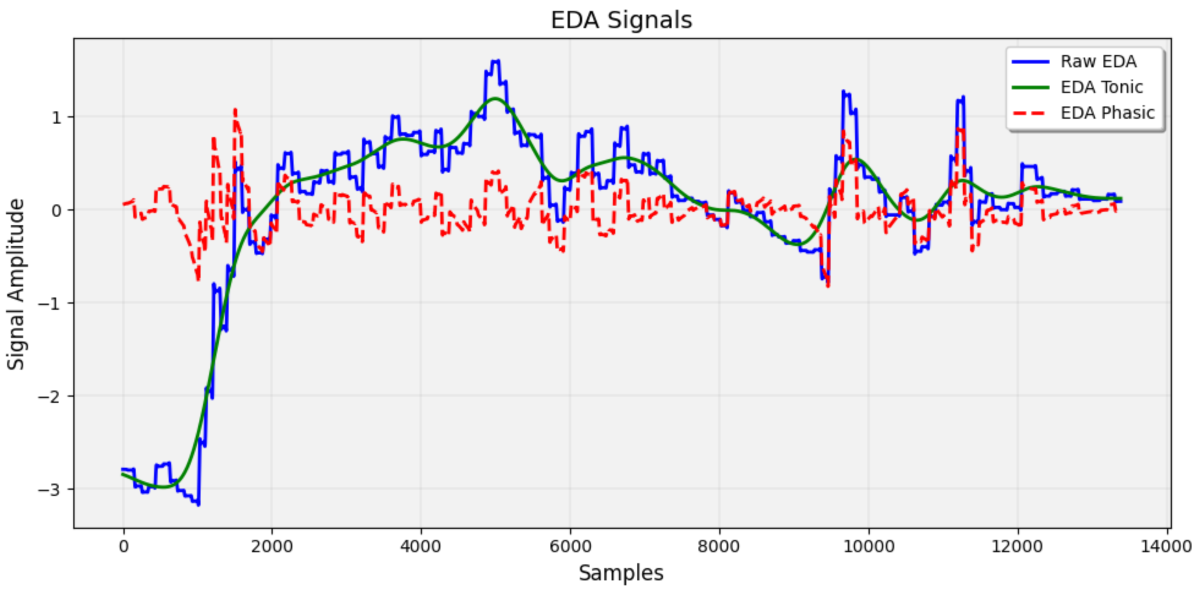



Abbildung 4.2: EDA

We conducted a comprehensive review and comparison of various EDA features, ultimately selecting those most relevant to our research objectives. This meticulous selection process, influenced by a thorough examination of nearly 40 distinct EDA features as outlined by previous studies Shukla u. a. (2021), enabled us to tailor our feature set.
From the decomposed EDA data, a variety of time domain statistical features can be extracted. These include mean ($\mu$), standard deviation ($\sigma$), coefficient of variance (CV),

variance ($\sigma^2$), and kurtosis ($\beta$) from the phasic component.
Mean ($\mu$) provides a measure of the central tendency of the SCR amplitudes. Standard deviation ($\sigma$) and variance ($\sigma^2$) capture the variability or dispersion around the mean. The coefficient of variance (CV) offers a normalized measure of dispersion relative to the mean. Kurtosis ($\beta$) evaluates the peakedness or flatness of the distribution of SCR amplitudes.
In addition to time-domain features, we analyze the EDA signal in the frequency domain. Furthermore, frequency-domain features like spectral power in specific bands (f1sc, f2sc, f3sc) and the overall energy and entropy of the signal gave us a spectrum-based view of the EDA responses. By analyzing these features, we could discern patterns and rhythms in the EDA that are not immediately apparent in the time-domain.
Detecting and analyzing peaks in the phasic component (SCRs) is crucial. Peak amplitude, frequency, and their inter-relationships can be strong indicators of emotional and cognitive stress responses.
By examining both tonic and phasic components, we can understand the sustained arousal level (SCL) and the specific responses to stimuli (SCR). The correlation between these components can provide valuable insights into how sustained stress levels influence responses to immediate stimuli. From the tonic component, which encapsulates the underlying level of arousal, we calculated the mean, capturing the central tendency over time, and the standard deviation, offering insights into the variability around this mean. The maximum and minimum values, along with the range, provided us with the extremes of the EDA signal, painting a picture of the breadth of responses.
From the phasic component, we focused on the Skin Conductance Responses (SCRs) to discern more rapid changes associated with specific stimuli. We extracted features like SCR amplitude, which reflects the intensity of the response, and the frequency of these SCRs, indicating how often these responses occur. The kurtosis of SCR amplitudes, a measure of the 'tailedness' of the distribution, gave us an understanding of how peaked or flat the distribution of responses was, while the skewness indicated any asymmetry, offering clues about the predominant direction of the response distribution.
We also looked at the root mean square (SCR RMS), which is a measure of the signal's magnitude, and the arc length (SCR Arc Length), providing a summative measure of the signal's complexity over a given period. The integral of the SCR signal (SCR Integral) was calculated to understand the total magnitude of these phasic responses over time. These features, alongside others like SCR momentum, which is akin to the second moment of the distribution, provided a comprehensive statistical breakdown of the phasic EDA signals.
Shukla u. a. (2021) we reviewed and compared amlost 40 different EDA features, the features extracted selected with the help of this.

### 4.3.2 Body Features

Since we captured human motion using the motion capture system, we selected 13 key points from the 25 marker points used by the system, focusing on the upper body. The chosen points were Hip, Ab, Chest, Neck, Head, Left Shoulder (LShoulder), Left Upper Arm (LUArm), Left Forearm (LFArm), Left Hand (LHand), Right Shoulder (RShoulder), Right Upper Arm (RUArm), Right Forearm (RFArm), and Right Hand (RHand). These points were strategically selected to comprehensively capture upper body movements. The arrangement of these points is depicted in Fig 4.3

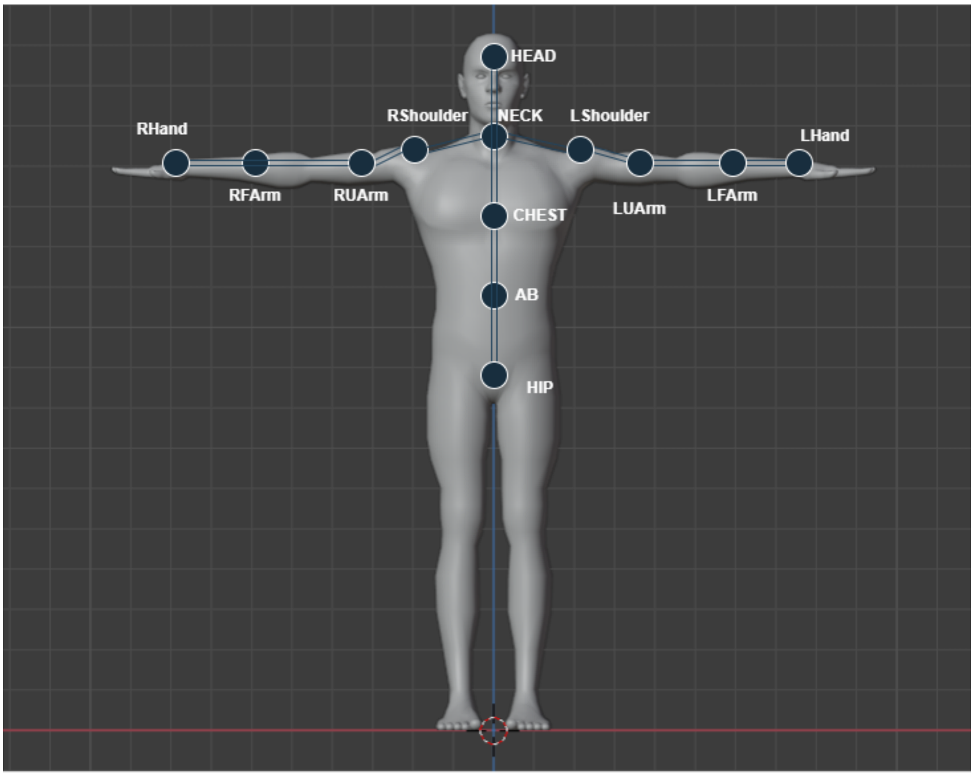



Abbildung 4.3: 13 points from the motion capture

## Self Touching

Previous studies, such as the one referenced in Giakoumis u. a. (2012), indicate that body posture and body language can be valuable indicators of stress. In line with this, we also explored body language cues such as self-touching, which Harrigan (1985)suggests can be indicative of negative affect, such as anxiety or discomfort. Specifically, we focused on face and head touching as potential stress indicators.
To determine if a person is touching her face, we compute the hand-head and the hand-neck distances. If one of these distances is below a given threshold, we consider that the person is face touching. The number of occurrences (FTC) and the average duration (FTMD) of these face touching events are used as features.
To determine gestures such as face touching, `/tf` data, which typically includes the position and orientation of each joint in space, can be utilized to calculate the distance between any two points. For example, if you have the coordinates of the right hand (`RHand`) and the head (`HEAD`), you can compute the Euclidean distance between these two points at each time frame to detect when the hand is close enough to the head to indicate potential face touching.
Let's define the 3D coordinates for the right hand and head at time $t$ as:

- $P_{\text{RHand}}(t)$ for the position of the right hand at time $t$, with coordinates $(x_{\text{RHand}}(t), y_{\text{RHand}}(t), z_{\text{RHand}}(t))$,
- $P_{\text{HEAD}}(t)$ for the position of the head at time $t$, with coordinates $(x_{\text{HEAD}}(t), y_{\text{HEAD}}(t), z_{\text{HEAD}}(t))$.

The distance between the right hand and the head at time $t$ is then calculated with the formula:

$$D_{\text{RHand-HEAD}}(t) = \sqrt{(x_{\text{RHand}}(t) - x_{\text{HEAD}}(t))^2 + (y_{\text{RHand}}(t) - y_{\text{HEAD}}(t))^2 + (z_{\text{RHand}}(t) - z_{\text{HEAD}}(t))^2} \tag{4.3.1}$$

If this distance, $D_{\text{RHand-HEAD}}(t)$, is less than a certain threshold, denoted as $\theta$, it suggests that the right hand is in proximity to the head, indicating potential face touching.

To determine the occurrence of face touching, you would track when this distance becomes less than the threshold $\theta$ and also ensure that the hand remains within this threshold for a certain duration to count as an occurrence. The distance for the left hand can be calculated in a similar manner.

To compute the number of occurrences (`FTC-Face Touching Count`) and the average duration (`FTMD-Face Touching Mean Duration`) of face touching is shown in Algorithm 4.3.1.

Algorithm 4.3.1.: Face Touching Detection

**Require:** Set of time-stamped positions from the `/tf` topic, Threshold distance $\theta$

**function** DetectFaceTouching
    Initialize $FTC \leftarrow 0$ ▷ Occurrences of face touching
    Initialize $TotalDuration \leftarrow 0$ ▷ Total duration of face touching events
    Initialize $FTMD \leftarrow 0$ ▷ Mean duration of face touching events
    **for** each time stamp $t_i$ in `/tf` topic **do**
        Calculate $D_{RHand-Head}(t_i)$ and $D_{LHand-Head}(t_i)$
        **if** $D_{RHand-Head}(t_i) < \theta$ **or** $D_{LHand-Head}(t_i) < \theta$ **then**
            Start duration counter
            $FTC \leftarrow FTC + 1$
        **if** either distance exceeds $\theta$ **then**
            Stop duration counter
            Add duration to $TotalDuration$
    **if** $FTC > 0$ **then**
        $FTMD \leftarrow \frac{TotalDuration}{FTC}$ ▷ Calculate mean duration
    **return** $FTC$, $FTMD$

### Sudden Movement

Lagomarsino u. a. (2022) suggested another way in which motion data can be used to assess stress. Sudden movements, or abrupt changes in body motion, can be indicative of stress responses. These movements are characterized by significant deviations from a person's regular movement patterns and can be quantitatively assessed using motion capture data. The following equations describe the computational process used to analyze sudden movements and infer stress.

$$m_j^k = \sum_{i=0}^{\tau-1} d_j^{k-i,k-i-1} \tag{4.3.2}$$

Equation 4.3.2 defines the movement of the $j^{th}$ joint within a time window $\tau$ as the sum of the displacements between consecutive frames.

$$\Delta_j^k = m_j^k - \mu_j \tag{4.3.3}$$

In Equation 4.3.3, $\Delta_j^k$ represents the deviation of the $j^{th}$ joint's movement from its baseline mean motion $\mu_j$, calculated during a calibration phase.

$$a_j^k = \begin{cases} \frac{\Delta_j^k}{\sigma_j} - 1 & \text{if } \Delta_j^k > \sigma_j \\ 0 & \text{otherwise} \end{cases} \tag{4.3.4}$$

Equation 4.3.4 assesses the activity level $a_j^k$ for the $j^{th}$ joint, taking into account the standard deviation $\sigma_j$ as a threshold for sudden movement.

$$a_k = \min\left(1, \frac{1}{N}\sum_{j=1}^{N} a_j^k\right) \tag{4.3.5}$$

Finally, Equation 4.3.5 calculates the overall level of sudden movement at time instance $k$ by averaging the activities across all joints, thus providing a descriptor of hyperactivity or sudden movement.
This method allows for a comprehensive analysis of the motion data to identify periods of high activity that may correlate with stress responses.
While our study focused on the analysis of upper body movements for stress detection, it is important to note that other bodily cues can also be significant indicators of stress or anxiety. One such example is the rapid tapping or bouncing of one's feet, which is often a subconscious response to nervous energy or unease. Such movements are typically a form of self-soothing behavior that occurs when an individual is experiencing discomfort or stress.
Unfortunately, due to the scope of our motion capture system setup, we restricted our tracking to the upper body and therefore could not capture lower body movements, such as foot tapping or leg bouncing. These actions could potentially provide additional insights into a participant's stress levels and offer a more comprehensive understanding of physical stress responses.
Including lower body data in future studies could enhance the detection and analysis of stress indicators, allowing for a fuller picture of the physiological and behavioral state of an individual under stress. This would enable us to capture a wider range of stress-related behaviors and potentially increase the accuracy and reliability of stress detection in real-time scenarios.

## 4.4 Classification /Stress Detection/

# 5 Result

# 6
# Discussion and Conclusion

# Literatur


**Alexander, D. M., C. Trengove, P. Johnston, T. Cooper, J. August und E. Gordon (2005):** „Separating individual skin conductance responses in a short interstimulus-interval paradigm“. In: *Journal of neuroscience methods* 146.1, S. 116–123.

**Empatica (o. D.):** *Empatica E4 specs*. Available under: `https://box.empatica.com/documentation/20141119_E4_TechSpecs.pdf`.

**Gellman, M. D. (2020):** „Behavioral medicine“. In: *Encyclopedia of behavioral medicine*, S. 223–226.

**Giakoumis, D., A. Drosou, P. Cipresso, D. Tzovaras, G. Hassapis, A. Gaggioli und G. Riva (Sep. 2012):** „Using Activity-Related Behavioural Features towards More Effective Automatic Stress Detection“. In: *PLOS ONE* 7.9, S. 1–16. URL: `https://doi.org/10.1371/journal.pone.0043571`.

**Harrigan, J. A. (1985):** „Self-touching as an indicator of underlying affect and language processes“. In: *Social Science and Medicine* 20.11, S. 1161–1168. URL: `https://www.sciencedirect.com/science/article/pii/0277953685901935`.

**Hernando-Gallego, F., D. Luengo und A. Artés-Rodríguez (2017):** „Feature extraction of galvanic skin responses by nonnegative sparse deconvolution“. In: *IEEE journal of biomedical and health informatics* 22.5, S. 1385–1394.

**imotions (o. D.):** *EDA Example signals*. Available under: `https://imotions.com/blog/learning/research-fundamentals/eda/`.

**Lagomarsino, M., M. Lorenzini, E. De Momi und A. Ajoudani (2022):** „An Online Framework for Cognitive Load Assessment in Industrial Tasks“. In: *Robotics and Computer-Integrated Manufacturing* 78, S. 102380. URL: `https://www.sciencedirect.com/science/article/pii/S0736584522000679`.

**Lasota, P. A. und J. A. Shah (2015):** „Analyzing the Effects of Human-Aware Motion Planning on Close-Proximity Human–Robot Collaboration“. In: *Human Factors* 57.1. PMID: 25790568, S. 21–33.

**Nahavandi, S. (2019):** „Industry 5.0—A Human-Centric Solution“. In: *Sustainability* 11.16. URL: `https://www.mdpi.com/2071-1050/11/16/4371`.

**Sauppé, A. und B. Mutlu (2015):** „The Social Impact of a Robot Co-Worker in Industrial Settings“. In: URL: `https://doi.org/10.1145/2702123.2702181`.

**Shukla, J., M. Barreda-Ángeles, J. Oliver, G. C. Nandi und D. Puig (2021):** „Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity“. In: *IEEE Transactions on Affective Computing* 12.4, S. 857–869.

**Suboh, M. Z., R. Jaafar, N. A. Nayan, N. H. Harun und M. S. F. Mohamad (2022):** „Analysis on Four Derivative Waveforms of Photoplethysmogram (PPG) for Fiducial Point Detection“. In: *Frontiers in Public Health* 10. URL: https://www.frontiersin.org/articles/10.3389/fpubh.2022.920946.

# 7 Appendix

Das ist der Anhang (siehe Abschnitt **??**) / This is the appendix (see section **??**)

## 7.1 Usage of generative AI - Affidavit

- ☐ not at all
- ☒ for correcting, optimizing, or restructuring the entire work (This eliminates the need for explicit marking of individual passages or sections, as this type of usage refers to the entire written work. Explicit marking in the text is not necessary, as this serves as the global indication.)
- ☐ Code optimization: Optimization or restructuring of software function
- ☐ Code generation: Creating entire software functions from a detailed functional description.
- ☐ Substance generation in code: Generating entire software source code
- ☐ Media optimization: Correction, optimization, or restructuring of entire passages
- ☐ Media generation: Creating entire passages from given content.
- ☐ Substance generation in media: Generating entire sections
- ☐ More, namely:

I assure that I have provided all usages completely. Missing or incorrect information may be considered an attempt to deceive.

place, date

Jane Doe