



PÁZMÁNY

Pázmány Péter Catholic University

Faculty of Information Technology and Bionics

# Fusion of PPG and EMG Measurements for Feature Extraction and to Validate Muscle Contraction

Barlay Anna — JQLP8J

Celia Casanova — SBON4I

Fenyődi András — B7A02O

Pázmány Péter Catholic University  
Faculty of Information Technology and Bionics  
50/a Práter Street, 1083 Budapest, Hungary

Date: 2025.12.09.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Objectives</b>	<b>3</b>
2.1	Motivations of Parallel Measurements . . . . .	3
2.2	Measurement Protocol . . . . .	3
2.3	Expected Results . . . . .	4
2.4	Scientific Significance . . . . .	4
<b>3</b>	<b>Scientific background</b>	<b>4</b>
3.1	Electromyography (EMG) . . . . .	4
3.2	Photoplethysmography (PPG) . . . . .	5
3.3	Multimodal Fusion of EMG and PPG . . . . .	5
3.4	Effects of Load and Recovery on EMG and PPG . . . . .	5
3.5	Statistical Analysis and Mathematical Framework . . . . .	5
3.6	Signal Envelope Extraction (Muscle Activity Score) . . . . .	5
<b>4</b>	<b>Proposed Statistical Analysis</b>	<b>5</b>
4.1	Proposed Statistical Framework and Reasoning . . . . .	5
4.2	Spectral Analysis and Windowing . . . . .	6
4.3	Correlation Analysis (EMG vs. PPG) . . . . .	6
4.4	Group-Level Hypothesis Testing . . . . .	6
4.5	Signal Difference Calculation . . . . .	7
4.6	Ranking the Differences . . . . .	7
4.7	Statistical Significance . . . . .	7
4.8	Rejection Criteria . . . . .	7
4.9	Hypothesis and Application . . . . .	7
4.10	Estimation Error Metric . . . . .	7
4.11	Correlation Coefficient and Significance Testing . . . . .	8
<b>5</b>	<b>Methodology</b>	<b>8</b>
5.1	Self-collected Signal Measurements . . . . .	8
5.1.1	Finger PPG . . . . .	8
5.1.2	EMG Electrodes . . . . .	8
5.2	Recording Procedure . . . . .	9
5.3	Data Availability . . . . .	9
5.4	Complete Processing Workflow . . . . .	9
5.4.1	Data Format and Extraction . . . . .	9
5.5	Pipeline Steps . . . . .	9
5.5.1	Data Input . . . . .	10
5.5.2	Preprocessing . . . . .	10
5.5.3	Segmentation and feature extraction . . . . .	10
5.5.4	Signal Quality Index (SQI) . . . . .	11
5.5.5	Normalization . . . . .	11
5.6	Automation and Workflow . . . . .	11
5.6.1	Running the Program . . . . .	11
5.6.2	Preprocessing Automation . . . . .	11
5.6.3	Fiducial Markers Calculation . . . . .	11
5.6.4	Statistical Analysis Automation . . . . .	11
5.6.5	Results . . . . .	12
5.7	Usage and Project Maintenance . . . . .	12
5.7.1	How to Run the Program . . . . .	12
5.7.2	Code Maintenance . . . . .	12
5.7.3	Git Repository . . . . .	12

<b>6 Results</b>	<b>13</b>
6.1 Filtered PPG Signal and Derivatives . . . . .	13
6.2 PPG Signal with Fiducial Points Detected . . . . .	14
6.3 Signal Comparison between EMG and PPG during Lift and Rest Phases . . . . .	14
<b>7 Discussion</b>	<b>15</b>
7.1 Rationale for Employing PPG as a Wearable Alternative to EMG . . . . .	15
7.2 Physiological Basis of PPG Attenuation During Muscle Contraction . . . . .	16
7.3 Variability of Pulse Wave Attenuation Due to Natural Arterial Compression . . . . .	16
7.4 Motion-Induced Artifacts and Their Impact on PPG Signal Quality . . . . .	16
7.5 Comparative Considerations Between PPG-, EMG-, and Multimodal Approaches . . . . .	16
7.6 Ethical, Practical, and Data-Handling Considerations in Wearable Monitoring . . . . .	17
7.7 Implications for Future Development and Application of Wearable Muscle-Monitoring Systems . . . . .	17
<b>8 Conclusion</b>	<b>18</b>
<b>9 Deliverables and Project Outcomes</b>	<b>18</b>
<b>10 Statement of Contributions</b>	<b>19</b>
<b>11 Statement of Contributions</b>	<b>19</b>

## Abstract

This lab report contains the descriptions and results of measurements taken place at the PPKE-ITK 214 laboratory. The aim of this lab practice was to record and analyse neural signals. The protocols mostly followed the Biopac Student Lab Lessons 1-4

# 1 Introduction

This paper guides the reader through our journey with parallel EMG and PPG processing. The introduction describes the aim of our experiments, what kind of roadmap we planned for measurements and analysis. To make sure everyone is on the same terms, when we talk about biomedical signal processing, the following section introduces the theoretical basis and scientific significance of our roadmap. The experiment protocol description, detailed with expected results, is followed by the dataset description. Then, further analysis methods, mathematical apparatus, and the used statistical methods are presented under the Methods section. The Results of analysis shortly summarize the hypotheses, followed by the evaluation of both tools and data. The discussion and further perspective parts include both the retrospective wisdom, improvement possibilities, novelty, and further potentials of our experience, alongside the concluding theoretical comparison with the objective lenses of the field. The last, but not least essential part of our paper is the Acknowledgements section, where we express our gratitude towards the colleagues who supported our work.

# 2 Objectives

## 2.1 Motivations of Parallel Measurements

In this section, the expected findings are introduced alongside the use of performing simultaneous measurements on the same subject.

Nowadays, wearable wireless sensors tend to be designed to be more portable and minimized. For more convenient and everyday life use of signal acquisition and processing methods, regarding non-invasive devices, PPG has great potential.

On the market, several new wireless surface EMG sensors appear. Concurrently, they are still not the most comfortable for everyday use. Furthermore, in strict academic experimental environments, the classical convention is using wired surface EMG as the most punctual and well-built approach. When the measurements were planned, a question occurred: whether muscle activity can be estimated from a single PPG signal, or if there is a way to find a feature that can be projected to observe muscle activity levels from a single PPG.

The two main perspectives of our scope are to understand the phenomena of the drop in PPG signal by elevated muscle activation, and finding a feature engineering approach for a feature that can be projected to observe muscle activity levels from a single PPG.

## 2.2 Measurement Protocol

In this section, the reader is introduced to the measurement protocol that we planned, then how the protocol had to be fine-tuned due to device or situation limitations.

When planning this very first measurement protocol, everything was orchestrated around the main hypothesis. The following conditions were set in the expected protocol: In three periods, the subject had to lift three different weights, with 10 seconds of rest after each 45-second long lifting epoch. The lower arm attached surface electrodes measured the skeletal muscle contraction, while the PPG sensor was attached to the finger on the same arm of the subject. When lifting weights, the intramuscular pressure is increased by muscle contraction, local blood perfusion is reduced, and diminished PPG amplitudes are expected. The stronger the contraction is, the greater amplitude reduction is caused.

When the device properties were introduced by the lecturer and joint decisions were made about the feasibility of the test, the fine-tuning of our submitted project proposal was necessary. The measurement could not be longer than two minutes because the wireless PPG device to application construction was built in a way that only two minutes of recording session could be stored remotely per each subject. Then, for some reason, the application did not allow us to define 10 subjects individually, so

we grouped the subjects into 3 categories according to BMI index range. The data of subjects were handled anonymously.

For a recording, the trigger, scope, and BioPac BSL Analysis software lecture measurement software were used. Three surface electrodes—ground, negative, and positive electrodes—were attached with conducting gel transfer facilitation, and the wireless PPG device was resting on the subject. The wireless PPG device was driven via application navigation under the authentication of the general lab account. Before recording a session, once the subject had the proper monitoring, the real-time detection of the PPG signal was expected on the device monitoring system. After the signal was stable and properly detected, the joint recording started, conducted by the two of us. The recording minutes for ourselves and the instructions to the subject were coordinated by one, and the recording was navigated by the other.

## 2.3 Expected Results

This chapter aims to highlight what the conditions allowed and what the expectations are holding as the invisible symbolic envelope of the final realization in the project.

- We expect that the PPG amplitude decrease will be visible.
- We expect a lot of motion artifacts.
- We expect the EMG to be in milliVolt and true to the calibrations that are automatized in the BSL Analysis software.
- We do not expect as much fatigue in muscle activity due to the short recording period time given by the PPG recording limitations.
- We expect that the EMG root mean square (RMS) value gives us stable information about muscle activity per subject.
- Our statistical test perform well on multiple subjects yet with the same sample size data.
- We expect to perform the analysis and statistics and feature engineering steps that are the scope of the course, but due to data regulations and lack of resources, machine learning investigations are out of the scope of this study.
- We expect to show the prominent correlating feature, the Muscle Activity Score, to PPG amplitude of the EMG and PPG signal.

## 2.4 Scientific Significance

As the latter section summarized the computational expectations, referring to the introduction, the pace of global technological development gives the momentum and direction of our research. The aim is to learn how physiological parameters can be algorithmically estimated, minimizing sensor size and artifacts. This research is crucial in the context of wearable health monitoring devices, where the goal is to make them more user-friendly and suitable for everyday use. Understanding and improving these technologies can provide valuable insights into monitoring muscle activity and other physiological functions in real-time.

# 3 Scientific background

## 3.1 Electromyography (EMG)

Electromyography (EMG) is a technique used to measure electrical activity in skeletal muscles to obtain information about neuromuscular function [2]. The EMG signal is composed of action potentials from motor units that originate from the contraction of individual muscle fibers, and the signal intensity is related to the level of muscle activation [9]. EMG has proven useful in both clinical diagnosis and biomechanical research, allowing the evaluation of muscle effort during various physical tasks [3].

In this study, we will use the EMG envelope as an indicator of muscle activation, since it increases with load during isometric contractions. Higher intensity muscle contractions lead to a greater number of motor units being recruited, resulting in a greater amplitude in the EMG signal [5]. Furthermore, the relationship between EMG amplitude and contraction intensity is often characterized by a nonlinear response [6].

### 3.2 Photoplethysmography (PPG)

Photoplethysmography (PPG) is a non-invasive optical technique that measures changes in blood volume, reflecting peripheral circulation. PPG is sensitive to variations in the amplitude of reflected light as blood vessels expand and contract with each pulse [1]. During muscle contraction, vasoconstriction of peripheral blood vessels occurs, resulting in a reduction in the PPG signal amplitude. This decrease is thought to be a physiological response to increase blood flow to the contracting muscles, while simultaneously limiting peripheral blood flow [10].

The PPG signal is regulated by multiple factors, including heart rate, vascular tone, and oxygen saturation [7]. During isometric contractions, the PPG signal is expected to decrease during muscle work and recover more slowly after heavy loads. This provides us with information about the physiological state of the muscle after contraction.

### 3.3 Multimodal Fusion of EMG and PPG

In recent years, the integration of EMG and PPG signals for classifying muscle effort has gained interest. This is due to the advantages offered by both modalities. EMG is excellent for directly measuring muscle activity, and PPG provides complementary information on vascular responses and peripheral circulation. The fusion of these modalities could provide a more accurate and robust measurement of muscle effort, especially in wearable devices that require reliable real-time monitoring [8].

### 3.4 Effects of Load and Recovery on EMG and PPG

As the intensity of muscle contraction increases, the EMG signal shows a monotonic increase in amplitude due to the recruitment of motor units [4]. On the other hand, PPG signals are expected to decrease in amplitude as vasoconstriction occurs, and the recovery phase is expected to take longer to reach baseline values after high-intensity contractions. This delayed recovery may reflect physiological differences in how the cardiovascular system and muscles recover from intense physical exertion [?].

### 3.5 Statistical Analysis and Mathematical Framework

This section outlines the mathematical transformations applied to the raw signals to derive the "Muscle Activity Score" and the statistical tests used to validate the physiological hypotheses.

### 3.6 Signal Envelope Extraction (Muscle Activity Score)

To transform the high-frequency EMG signal into a continuous measure of muscle exertion, known here as the Muscle Activity Score, we extracted the signal envelope. Rather than using simple rectification or low-pass filtering, we employed the Hilbert Transform.

The Hilbert Transform produces the analytic signal, from which the instantaneous amplitude (the envelope) can be computed. This method effectively captures dynamic changes in muscle activation while reducing high-frequency noise. The result is a robust "Muscle Activity Envelope" that preserves the onset and offset of contraction events.

## 4 Proposed Statistical Analysis

### 4.1 Proposed Statistical Framework and Reasoning

This section outlines the mathematical transformations and statistical tests designed to validate the physiological hypotheses quantitatively. While the current study focuses on the qualitative visual

inspection of the signals to demonstrate the inverse relationship between EMG and PPG, the following framework represents the intended validation pipeline. This theoretical approach is structured to robustly confirm the "Lift vs. Rest" hypothesis and quantify the correlation between muscle activation and vascular response, should larger datasets and automated processing become available.

## 4.2 Spectral Analysis and Windowing

After extracting the envelope, the proposed method involves segmenting the signal into 2-second windows for local frequency analysis. To minimize spectral leakage during this process, a Gaussian window function is applied to each segment before computing the spectrum.

The Gaussian window function  $g(t)$  is defined mathematically as:

$$G(t) = \pi^{-\frac{1}{4}} \exp\left(-\frac{t^2}{2}\right)$$

The Gaussian kernel is for smoothing the boundaries of the segmented signal. This ensures that the local frequency estimates remain accurate and free from artifacts that can arise due to window discontinuities.

## 4.3 Correlation Analysis (EMG vs. PPG)

To test the hypothesis that increased muscle activity leads to decreased peripheral blood volume (vasoconstriction), the framework proposes analyzing the relationship between the Muscle Activity Score (EMG envelope) and PPG peak amplitudes.

Pearson's correlation is deemed not suitable in this case, as it assumes a strict linear relationship ( $y = mx + c$ ), which does not account for the often curvilinear or saturation trajectories observed in physiological responses to load. Instead, Spearman's Rank Correlation ( $\rho$ ) is selected as the non-parametric alternative suitable for monotonic relationships (i.e., confirming that as one variable increases, the other consistently decreases).

The Spearman coefficient  $\rho$  ranges from -1 to +1, where:

- **Positive correlation ( $\rho > 0$ ):** Both variables increase together.
- **Negative correlation ( $\rho < 0$ ):** As one variable increases, the other decreases.

We hypothesize a strong negative correlation ( $\rho \approx -1$ ) between the EMG envelope and PPG amplitude.

To determine if the observed correlations are statistically significant (i.e., different from zero), the framework utilizes the t-statistic for the correlation coefficient using the standard formula:

$$t = r \cdot \sqrt{\frac{n - 2}{1 - r^2}}$$

Where:

- $r$  is the correlation coefficient.
- $n$  is the sample size (number of paired data points).

This test evaluates the null hypothesis that there is no linear or monotonic relationship between the variables.

## 4.4 Group-Level Hypothesis Testing

Due to the small sample size ( $N = 10-14$ ) and the likelihood that biological signal magnitudes are not normally distributed, parametric tests (such as the standard Student's t-test) are deemed inappropriate for group comparisons in this framework.

To validate the "Lift vs. Rest" hypothesis, the framework employs the Wilcoxon Signed-Rank test, a non-parametric alternative to the paired t-test that is more robust for data that may not follow a normal distribution.

## 4.5 Signal Difference Calculation

Let the difference between the paired measurements for the  $i$ -th subject be represented as:

$$d_i = \text{Lift Amplitude}_i - \text{Rest Amplitude}_i$$

Here, Lift Amplitude and Rest Amplitude represent the measurements of the PPG signal (or other relevant signal) during the lifting and resting phases, respectively. The protocol excludes pairs where  $d_i = 0$  (i.e., no difference between the two conditions).

## 4.6 Ranking the Differences

The absolute differences  $|d_i|$  are ranked from smallest to largest, with  $R_i$  representing the rank of  $|d_i|$ . The signed rank for each observation is then given by:

$$\text{sgn}(d_i) \cdot R_i$$

The test statistic  $W$  is calculated as the sum of the positive ranks:

$$W = \sum_{i:d_i>0} R_i$$

## 4.7 Statistical Significance

The z-score to test for statistical significance is calculated as:

$$z = \frac{W - \mu_W}{\sigma_W}$$

Where the mean  $\mu_W$  and standard deviation  $\sigma_W$  are given by the formulas:

$$\mu_W = \frac{N(N + 1)}{4}$$

$$\sigma_W = \sqrt{\frac{N(N + 1)(2N + 1)}{24}}$$

## 4.8 Rejection Criteria

The framework rejects the null hypothesis (which assumes there is no difference between the "Rest" and "Lift" states) if  $|z| > z_{\text{critical}}$  for a chosen significance level (e.g.,  $z_{\text{critical}} \approx 1.96$  for  $\alpha = 0.05$ ).

## 4.9 Hypothesis and Application

The test evaluates whether the distribution of differences between paired observations (Rest Amplitude vs. Lift Amplitude) is symmetric around zero.

In this report, we present the comparison of EMG and PPG signal side-by-side to show the intuitive validation for the hypothesis, while the statistical tests above serve as the theoretical validation model.

## 4.10 Estimation Error Metric

To quantify the precision of the Muscle Activity Score in tracking the contraction profile, the framework utilizes the Root Mean Squared Error (RMSE). RMSE is a widely used metric to measure the goodness of fit between two signals, aggregating the magnitudes of prediction errors into a single value.

The RMSE is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{\text{emg}}[i] - y_{\text{ppg,inverted}}[i])^2}$$

Where: -  $N$  is the total number of samples in the synchronized time series. -  $y_{\text{emg}}[i]$  is the normalized Muscle Activity Score at time  $i$ , which is derived from the EMG signal. -  $y_{\text{ppg,inverted}}[i]$  is the normalized, inverted PPG amplitude at time  $i$ . The inversion is applied because increased muscle activity causes a decrease in PPG amplitude, and we aim to align it with the increasing EMG amplitude.

A lower RMSE value indicates a stronger relationship between the vascular response (represented by PPG) and skeletal muscle activation (represented by EMG), supporting the hypothesis that PPG can serve as a proxy for muscular effort.

## 4.11 Correlation Coefficient and Significance Testing

For the correlation of grouped data, the calculation is defined by the following formula:

$$r = \frac{\Sigma(fxy) - \frac{(\Sigma fx)(\Sigma fy)}{N}}{\sqrt{\left[\Sigma fx^2 - \frac{(\Sigma fx)^2}{N}\right] \left[\Sigma fy^2 - \frac{(\Sigma fy)^2}{N}\right]}}$$

We expect EMG and PPG data to be correlated, as the EMG amplitude increases while PPG amplitude decreases during muscle contraction. A t-test is proposed to further assess the significance of this correlation.

The null hypothesis of a two-tailed t-test states that there is no correlation (i.e., no linear relation) between the variables  $x$  and  $y$ . The alternative hypothesis states that there is a significant correlation (i.e., a linear relationship) between  $x$  and  $y$ .

The t-test is a statistical test for the correlation coefficient. It is used when:

- $x$  and  $y$  are linearly related.
- The variables are random variables.
- The population of the variable  $y$  is normally distributed.

## 5 Methodology

### 5.1 Self-collected Signal Measurements

#### 5.1.1 Finger PPG

A finger photoplethysmography (PPG) sensor from HeartReader (<https://readingtheheart.com/heartrader/>) was placed on the subject's index finger. The measurements were recorded at an effective sampling rate of 200 Hz, calculated from a 140 s segment containing 28,057 samples.

The recording software used by the device is proprietary and was not disclosed to us; therefore, acquisition settings beyond the sampling rate could not be verified. For each subject, the PPG sensor was placed on the dominant hand.

#### 5.1.2 EMG Electrodes

Surface EMG measurements were prepared and recorded using a standardized laboratory setup. The equipment used included a BIOPAC MP36 data acquisition amplifier, which was connected to the computer via USB, along with SS2LB electrode cables, three disposable adhesive electrodes for each subject, electrode paste, and tissues with alcohol for skin cleaning.

**Setup for Each Subject:** The SS2LB cable was connected to Channel 1 of the MP36 amplifier (which was powered off during setup). The electrodes were wetted with saline and placed on the subject's dominant arm at the locations shown in the figure (electrode placement). The SS2LB cable was then connected to the correctly positioned electrodes on the dominant arm. The order of the connections is illustrated in the figure (cable connections).

## 5.2 Recording Procedure

All measurements were performed with subjects standing in an upright, stationary position. The hand equipped with the PPG sensor was positioned on the handle of a weight; however, the sensor-bearing finger did not actively grip the handle, aiming to reduce the effects of forceful contact or mechanical compression on the PPG signal.

Immediately after starting the recordings on the devices, the subjects executed a forearm contraction task designed to activate the muscle group instrumented with EMG electrodes. For this task, the subject's hand was oriented palm upward with a loose grip on the handle, and the elbow remained extended (not flexed). Subjects rested for approximately 5 seconds, then performed an isotonic lower-arm lift by raising the weight through forearm flexion and subsequently holding the lifted position for an estimated 15 seconds to maintain sustained muscle activation. This action required visible contraction of the forearm flexor muscles targeted by the EMG configuration. The 5-second rest and 15-second lift were repeated 2 times after the initial lift.

## 5.3 Data Availability

During recording, the PPG device transmitted recordings to a smartphone application provided by the manufacturer, which then uploaded the measurement data to a remote cloud storage system. These data were not accessible to users in their raw form. As a result, the raw measurement files could not be retrieved locally after data collection.

To obtain the recordings, a data request was submitted to the manufacturer, who provided the unmodified raw data in CSV format. No preprocessing, filtering, or transformation was performed by the manufacturer prior to delivery; the supplied files represent the original output generated by the device.

## 5.4 Complete Processing Workflow

### 5.4.1 Data Format and Extraction

The raw PPG recordings provided by the manufacturer were delivered in CSV format, where each row contained three fields: a subject identifier, a timestamp, and a hexadecimal-encoded string sequence representing the PPG measurement data.

Each hexadecimal sequence represented a contiguous block of samples but varied in length across subjects. Before analysis, the data required conversion into numerical sample arrays and restructuring into a standardized tabular form.

To perform this transformation, a MATLAB script was used to:

- Read each line of the CSV file.
- Identify entries containing valid hexadecimal data.
- Extract the timestamp and isolate the hexadecimal portion of the string.
- Convert the hexadecimal sequence i

## 5.5 Pipeline Steps

The data analysis followed a structured pipeline with each step aimed at ensuring clean, standardized, and comparable data for further analysis. First the EMG and PPG had their separate preprocessing steps, then at a point based on event timings the two types of time series were merged. The main steps were as follows:

- creating structural class of Subject's data, and further Time Series Tools for design pattern conventions
- The process has a comprehensive Facade Design Pattern with Modular Pipeline attachments in the back of it.
- PPG Raw segment plot, built-in preprocessing

- PPG filtered plotting
- PPG fiducial points, SQI and biomarkers
- then save filtered PPG to MatLab struct
- create EMG time axis
- calculate EMG bandpass filtering and rectify, considering Nyquist-frequency
- create EMG Hilbert-envelope, computationally but also using NeuroKit2 signal cleaning
- EMG resampling (500 Hz to 200 Hz) and common time axis with PPG
- create EMG active and rest intervals chunking
- calculate EMG windowing of chunks with 2 s windows, 50% overlap
- calculate EMG gaussian weighting of each windowed signal part
- calculate RMS and MAV for each windowed segment
- calculate PPG max amplitude per windowed segment
- saving all features - mainly chunks and values - for the correlation analysis and plotting
- plotting extracted features and the signals

### 5.5.1 Data Input

The process began by collecting the raw EMG and PPG data from the subjects. These signals were then loaded into the system, ready for preprocessing and analysis. Both data were collected by different applications. About the EMG we know BSL analysis did notch filtering and integrated EMG envelope on the raw signal. Both PPG and EMG were imported from csv files.

### 5.5.2 Preprocessing

In the preprocessing step, the following procedures were applied to the raw signals to clean and standardize them. The time axis and sampling frequency were standardized across all subjects to ensure consistency in the data.

Filters were applied to remove noise from both the EMG and PPG signals. The EMG signal underwent band-pass filtering to remove unwanted low and high-frequency noise, while the PPG signal was filtered to focus on the relevant frequency range. The EMG signal was rectified, and its Hilbert envelope was calculated to extract the amplitude variations that reflect muscle activation over time. For merging the two modalities to the same time axis, re-sampling was executed on the signals to 200 Hz. when aligning the signals to eachother, the first 5 seconds of PPG were not taken into consideration, as they were some initial noise before the actual measurement have been begin.

### 5.5.3 Segmentation and feature extraction

The next step involved segmenting the signals into meaningful periods. A standard protocol was followed to define the rest and lift periods based on the markers recorded during the experiment. Within these segments, sliding windows of 2 seconds with 50% overlap were applied to both the EMG and PPG signals. This allowed for more granular analysis and ensured that overlapping data points were captured for accurate feature extraction. All in all, once the signals were aligned, the 3 sets of lift and rest periods were cut into separate intervals based on the integrated EMG signal activity to rest periods. In this way the PPG amplitude drop and the EMG MAV were significantly separable from rest period values. The aim of 2 seconds windowing was to prepare the data for correlation analysis supporting the core idea of the coupled measurements, that the PPG amplitude has a negative linear correlation to the EMG muscle activity value.

**EMG Features:** Several time-domain features were constructed for extraction, including RMS (Root Mean Squared), MAV (Mean Absolute Value), variance, Welch spectra analysis, and zero-crossings, as well as frequency-domain features such as median frequency and mean frequency. These

features provide insights into muscle activation intensity and frequency content. For power spectral density we have a too short period of measurement windows, thus we did not take it into consideration.

**PPG Features:** For the PPG signal, features like the amplitude (max-min), mean amplitude, the slope of the systolic upstroke, and variance were calculated to quantify blood volume changes and vascular response during muscle contractions.

#### 5.5.4 Signal Quality Index (SQI)

To assess the quality of the signals, several Signal Quality Indices (SQIs) were computed. The fraction of saturated samples was calculated to check for signal clipping. The ratio of band-pass power to wideband power was also assessed to ensure the signal contained the relevant frequencies.

#### 5.5.5 Normalization

After extracting the features, normalization was applied to ensure comparability across all subjects or per subject. Each feature was z-score normalized, ensuring that the extracted values from all subjects were on a comparable scale. This normalization process allowed for accurate statistical analysis and was particularly important for the subsequent machine learning models.

### 5.6 Automation and Workflow

The program's workflow is designed to process EMG and PPG signals from the raw data, with automation integrated into the preprocessing, fiducial calculation, statistical analysis, and result generation stages.

#### 5.6.1 Running the Program

The program is executed by running the main.py script, which sets up key parameters like sampling frequencies, signal lengths, and output directories. Once the script is run, it processes each subject's data sequentially.

#### 5.6.2 Preprocessing Automation

The pyPPG library plays a central role in processing PPG signals. The function `process_ppg_subject()` is used to handle the filtering, cleaning, and segmentation of the data. A fourth-order band-pass Butterworth filter is applied with a lower cutoff of 0.5 Hz and an upper cutoff of 12 Hz to remove baseline drift and high-frequency noise.

After filtering, the first three derivatives of the signal (VPG, APG, JPG) are calculated to capture the dynamic changes in the signal related to muscle activation.

Additionally, the Signal Quality Index (SQI) is calculated automatically for each PPG signal. This index assesses the quality of the signal and ensures that only reliable data is used in further analysis.

The Signal Quality Index (SQI) is calculated automatically for each PPG signal to assess its quality and ensure reliable data is used in further analysis.

#### 5.6.3 Fiducial Markers Calculation

The fiducial markers are automatically calculated using the pyPPG function '`compute_fiducials()`'. These markers include key points such as systolic peak, pulse onset, dicrotic notch, and diastolic peak for the original signal, as well as derivatives for the first, second, and third derivatives of the signal. These markers help in segmenting the signal and aligning the data for further analysis.

#### 5.6.4 Statistical Analysis Automation

After feature extraction, statistical tests such as Pearson's correlation and Wilcoxon tests are automated to evaluate relationships between the EMG and PPG signals. These tests are performed to check how muscle activity and vascular responses are correlated across subjects.

### **5.6.5 Results**

The results, including PPG SQI, fiducial markers, and statistical analysis, are automatically saved in a structured format. The results are saved as graphs and CSV files, which can be used for further analysis or future comparisons.

## **5.7 Usage and Project Maintenance**

### **5.7.1 How to Run the Program**

To run the program, execute the main.py script. It automatically loads the raw data, preprocesses the signals, extracts features, performs statistical analysis, and saves the results in the specified output directory.

### **5.7.2 Code Maintenance**

The code is documented with comments that explain each function and processing step. This makes it easy for collaborators to understand and extend the project. The project uses standard Python libraries like numpy, pandas, and pyPPG to ensure compatibility with other tools.

### **5.7.3 Git Repository**

The project code is stored on GitHub, where the team and collaborators can access and contribute. The repository includes:

- **Code:** Python scripts for data processing and analysis.
- **Data:** Folder for storing raw and processed data.
- **Documentation:** Instructions and explanations of methods used.

The project is version-controlled using Git, making it easy to track changes and maintain.

<https://github.com/barlaya/PPG-EMG-fusion-BSPT1ACA.git>

## 6 Results

### 6.1 Filtered PPG Signal and Derivatives

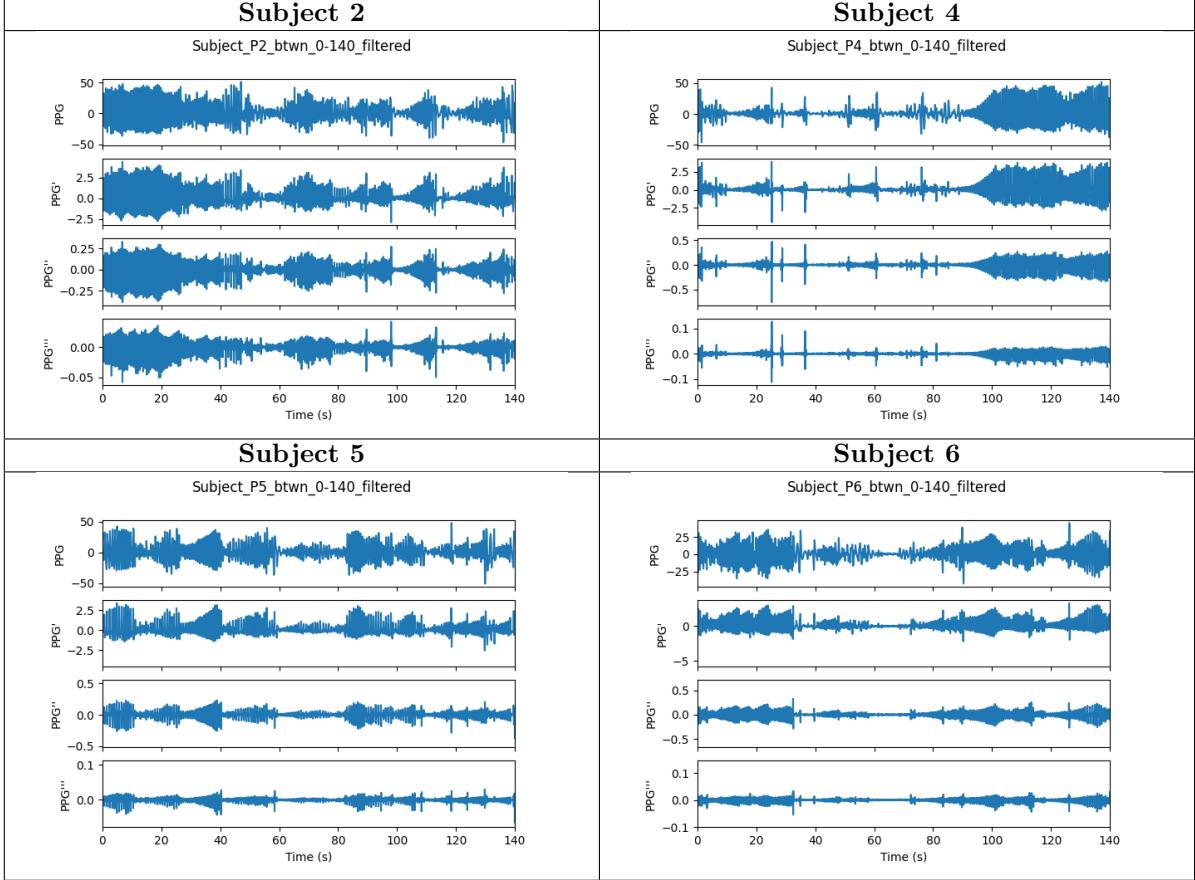


Table 1: PPG signal from subjects 2, 4, 5, and 6. The plots show the raw PPG signal acquired from each subject.

As seen in [Table 1](#), the beginning of the PPG signal has a significant amount of noise, which could interfere with accurate EMG measurements. To avoid this, we decided to start recording the EMG signal 5 seconds after the PPG signal. This delay allowed us to bypass the noisy initial part of the PPG signal, ensuring cleaner and more accurate data for the EMG analysis..

## 6.2 PPG Signal with Fiducial Points Detected

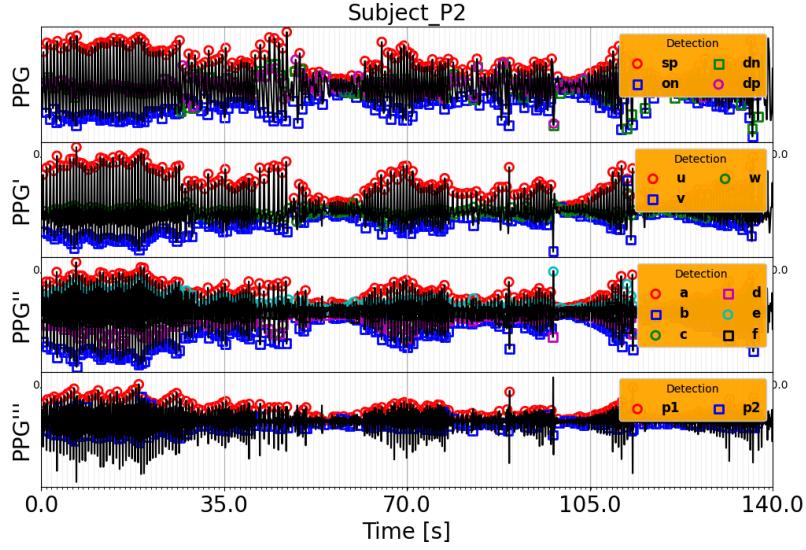


Figure 1: PPG signal with fiducial points for Subject 2. Despite the noise in the raw signal, key fiducial points are detected, demonstrating the effectiveness of the preprocessing techniques.

As shown in Figure 1, although the raw PPG signal exhibits significant noise, key fiducial points, such as systolic peaks and dicrotic notches, can still be accurately detected. This is possible due to the use of signal derivatives (PPG', PPG'', PPG''') that highlight important features in the signal despite the noise. The fiducial points are marked with different symbols, demonstrating the effectiveness of the preprocessing techniques in extracting meaningful physiological information.

## 6.3 Signal Comparison between EMG and PPG during Lift and Rest Phases

As shown in **Table 2**, the raw PPG signal (Signal 1) exhibits significant fluctuations. These fluctuations are linked to the physiological changes that occur during muscle activity. The relationship between the PPG and EMG signals is also evident in these plots. While the PPG signal decreases during the lift phase, the EMG signal increases, indicating muscle activation. This inverse relationship demonstrates the potential of using PPG as a proxy for muscle effort.

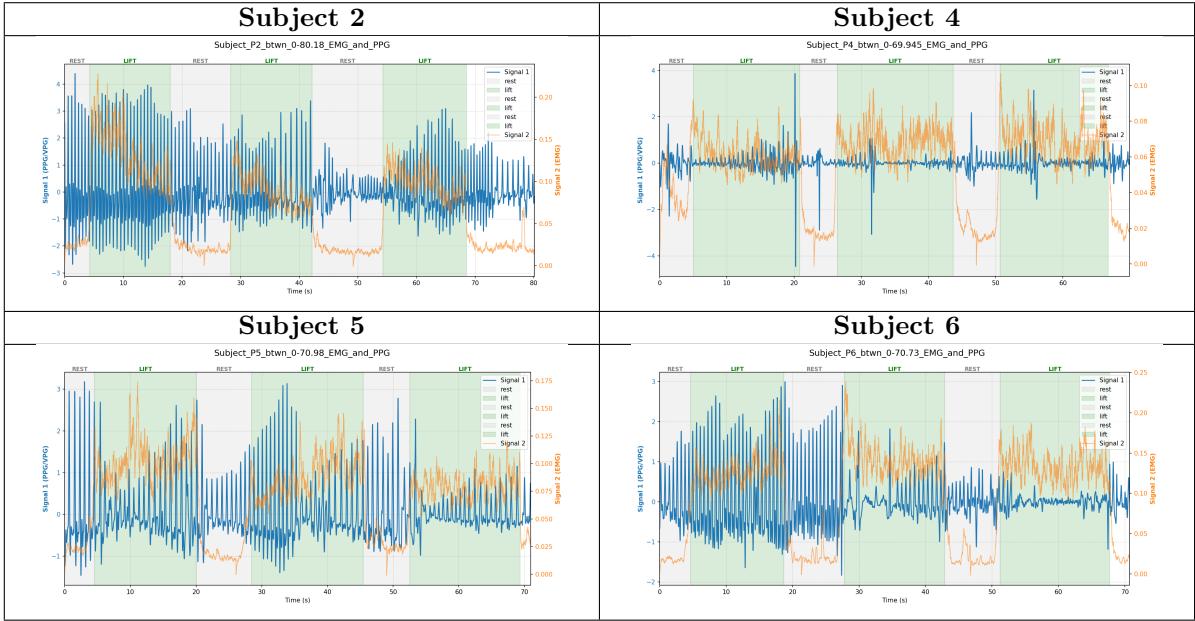


Table 2: PPG signal from subjects 2, 4, 5, and 6. The plots show the raw PPG signal acquired from each subject.

The figure above shows the comparison between the PPG (Signal 1) and EMG (Signal 2) signals for Subject P9 during a weight-lifting task. The data is displayed over a period of 67.735 seconds, with the phases of rest and lift clearly marked.

The PPG signal reflects changes in peripheral blood volume, which typically decreases as muscle activity increases. During the lift phase, fluctuations in the amplitude of the PPG signal can be observed, which correlate with the level of muscle exertion.

The EMG signal measures the electrical activity in muscles during contraction. As the subject lifts the weight, there is a marked increase in EMG amplitude, reflecting higher muscle activation. Conversely, the EMG signal remains relatively low during the rest phase, indicating minimal muscle activity.

The observed inverse relationship between the PPG and EMG signals—where the PPG amplitude decreases and the EMG amplitude increases during the lift phase—supports the hypothesis that PPG can be used as a proxy for muscle effort. This interaction between the two signals highlights the effectiveness of PPG in monitoring vascular responses during physical exertion.

## 7 Discussion

In this study, we examined the potential of photoplethysmography (PPG) as a simpler, more accessible complement, or even alternative, to traditional electromyography (EMG) for monitoring muscle activity during grip exercises. Our interest in PPG arises from the increasing availability of wearable devices equipped with optical sensors, which offer an attractive opportunity to assess physiological signals in everyday contexts.

By contrast, EMG, although highly reliable and well established, often requires more complex preparation, additional equipment, and controlled conditions. This contrast makes the combination of EMG with PPG a promising avenue for more practical and scalable muscle monitoring systems.

### 7.1 Rationale for Employing PPG as a Wearable Alternative to EMG

EMG has long been the reference method for assessing muscle activation because it directly measures electrical activity in the muscles. However, surface EMG requires precise electrode placement, conductive gel, and a relatively controlled setup, which can be uncomfortable or impractical for long-term or daily monitoring.

PPG, on the other hand, is easier to use. Since many commercial wearables already include PPG sensors, it provides a convenient way to capture physiological changes indirectly associated with muscle effort. Muscle contractions compress peripheral blood vessels, reducing blood flow and causing a detectable drop in PPG amplitude. This means PPG can give insight into muscle effort without the need for electrodes or specialized setup.

One of the main motivations behind this study was therefore to assess whether PPG could help “democratize” muscle monitoring by taking advantage of technology that many users already wear.

## 7.2 Physiological Basis of PPG Attenuation During Muscle Contraction

The attenuation of the PPG signal during contraction is rooted in basic physiology. Muscle fibers, when activated, increase internal tension and exert pressure on surrounding arteries. Depending on the intensity of the contraction, this compression can range from a mild reduction in blood flow to a substantial limitation of pulse amplitude.

As contraction intensity increases, more motor units are recruited, generating stronger compression and a clearer drop in the PPG waveform. This relationship allowed us to observe not only how peak amplitudes changed during contraction but also how the signal recovered afterward. The recovery profile (how quickly the PPG signal returns to baseline) can reflect vascular and muscular responses to effort.

In our study, the recovery phase played a relevant role in comparing low versus high effort conditions, since high-intensity contractions tended to produce slower returns to baseline.

## 7.3 Variability of Pulse Wave Attenuation Due to Natural Arterial Compression

In our study, the attenuation of the PPG signal was entirely caused by the natural compression of the arteries during muscle contraction, without the use of any external band. As participants increased their grip intensity, the forearm muscles generated internal pressure on the surrounding vessels, which weakened the pulse wave detected at the finger. However, because this compression depends solely on the participant’s own anatomy and contraction patterns, the degree of attenuation varied between individuals and even between repetitions. Factors such as muscle mass, tissue composition, arm position, and subtle differences in how the grip was executed all influenced how strongly the vessels were compressed. This natural variability makes the PPG signal less uniform but does not diminish the overall trend we observed: higher muscular effort consistently resulted in a reduced PPG amplitude, confirming the expected physiological mechanism.

## 7.4 Motion-Induced Artifacts and Their Impact on PPG Signal Quality

Movement artifacts represented another major challenge in our PPG recordings. Even when participants maintained a relatively stable posture, small vibrations, sensor shifts, and variations in skin, sensor contact introduced considerable noise into the signal. These artifacts were especially noticeable when subjects performed tasks involving external loads, such as holding a 5 kg dumbbell, which not only produced arm vibrations but also caused natural micro-adjustments in arm position.

The presence of such noise highlights the need for more robust signal processing approaches.

In future studies, integrating accelerometer or gyroscope data could help identify and compensate for movement, while adaptive filtering or machine learning-based denoising techniques could further enhance the reliability of PPG-based muscle monitoring.

## 7.5 Comparative Considerations Between PPG-, EMG-, and Multimodal Approaches

The central question in this project was whether combining EMG and PPG would improve the classification of muscular effort.

EMG provides precise information about the electrical activity produced by muscle fibers, while PPG offers a vascular perspective on the same event. Individually, each modality captures only one aspect of muscle function; together, they provide a more complete picture.

Our findings support the idea that multimodal integration is beneficial. The fused EMG–PPG model demonstrated better ability to classify low versus high effort than either modality alone. This improvement is likely due to the complementary nature of the signals. EMG reflects neural activation and motor unit recruitment, whereas PPG reflects vascular compression and hemodynamic changes.

Despite this improvement, several factors such as sensor placement, variations in compression, and movement artifacts were shown to influence the reliability of the combined data. Strengthening these aspects in future implementations could further enhance classification accuracy.

## 7.6 Ethical, Practical, and Data-Handling Considerations in Wearable Monitoring

Throughout the study, participant privacy and data security were treated with care. All datasets were anonymized prior to analysis, and data storage followed appropriate security and confidentiality protocols. These measures are especially important as wearable-based physiological monitoring becomes more widespread, raising ethical questions about long-term data handling and user autonomy.

## 7.7 Implications for Future Development and Application of Wearable Muscle-Monitoring Systems

The insights gained from this study provide a valuable foundation for future research on wearable muscle monitoring. One promising direction is refining the mechanical setup, specifically improving band tightness control and ensuring better sensor alignment.

Another avenue is developing more advanced preprocessing pipelines to isolate physiological changes from motion-induced artifacts. Expanding the dataset to include different muscle groups, more varied movement patterns, and a larger participant pool could also help validate the method's robustness.

From an application standpoint, enhanced PPG-based monitoring could be useful in numerous fields: fitness tracking, rehabilitation, injury prevention, ergonomics, and remote health assessment. If the accuracy and stability of PPG measurements continue to improve, the method could support real-time feedback during training or rehabilitation, monitor muscle fatigue, or help clinicians track progress in recovery programs.

## 8 Conclusion

The results of this project align with the physiological expectations that guided our study. EMG and PPG both exhibited clear, load-dependent behaviours that match what is known about neuromuscular activation and vascular response during isometric effort. Across all participants, EMG amplitudes increased in proportion to grip intensity, reflecting the progressive recruitment of motor units as muscular demand rose.

Similarly, the PPG signal showed pronounced attenuation during contraction, especially at higher loads. This attenuation corresponds to the compression of arteries by contracting muscle fibers. The recovery phase further reinforced these patterns: after high-intensity contractions, the PPG signal returned to baseline much more slowly, consistent with a greater cardiovascular adjustment.

When the two modalities were combined, classification performance improved noticeably. The integrated approach captured both the electrical and vascular dimensions of muscle activity, resulting in higher macro-F1 scores and clearer separation between effort levels. These findings not only validate our experimental design but also emphasize the potential of PPG as a practical and meaningful contributor to muscle effort assessment.

Overall, the study demonstrates that with adequate processing and improved experimental conditions, PPG could evolve into a valuable tool for wearable monitoring systems. Its simplicity, low cost, and compatibility with everyday devices make it an appealing candidate for future applications in sports performance, rehabilitation, and continuous health tracking.

## 9 Deliverables and Project Outcomes

The following table summarizes the alignment between the initial project proposal and the final deliverables presented in this report. Deviations were necessary due to technical constraints and the prioritization of a robust signal processing pipeline over machine learning classification.

Table 3: Comparison of Proposed Objectives vs. Final Deliverables

Component	Proposed Objective	Actual Deliverable
Data Collection	N=10–14 subjects with 20%, 40%, 60% MVC loads.	N=9 subjects (P2–P10) recorded using a binary "Rest vs. Lift" protocol.
Signal Processing	Standard processing steps for feature extraction.	Pipeline implemented: Hilbert Envelopes for EMG, Gaussian windowing (not applied yet), and PPG derivative (VPG/APG/JPG) fiducial detection
Statistical Analysis	Hypothesis testing using Page's L trend test and Wilcoxon signed-rank test.	Signal matching for visual validation of signal correlation; Definition of a "Proposed Statistical Framework" for future quantitative validation.
Classification of signal types	Binary classification (Low vs. High Effort) using Logistic Regression (F1-score).	Descoped due to time constraints.
Software Development	Python scripts for processing and reporting.	Facade Design Pattern with Modular Pipeline, Utilizing pyPPG to automate multi-patient analysis ( <code>PPG_handle</code> , <code>backbone</code> , <code>timesrsutils</code> ); Preliminary work for statistical thesis pipeline.

## 10 Statement of Contributions

The project was executed collaboratively with specific roles assigned to maximize efficiency in software development, data acquisition, and reporting. The table below details the specific responsibilities of each team member.

## 11 Statement of Contributions

The following table details the specific responsibilities of each team member throughout the project lifecycle.

Table 4: Project Contributions

Task	Contributor(s)	Task	Contributor(s)
Project ideation	Andris, Celia, Anna	Per subject data structure and EMG processing	Anna
Project first summary	Anna	PPG pipeline structure	Andris
Project presentation for submission	Celia	Documentation structurer, first two chapters	Anna
Measurements of PPG with EMG	Celia, Anna	Aligning EMG and PPG axis, extract common metrics (MAV, RMS, amp)	Anna
Data Acquisition from sensor apps	Anna	Correlation and plotting	Andris
PPG signal structuring	Andris	Documentation methodology	Andris, Celia, Anna
EMG signal structuring	Anna	Documentation discussion and conclusion, tables and comprehension	Celia
PPG signal processing	Andris	Profound literature research	Celia

## References

- [1] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3):R1–R39, 2007.
- [2] J. V. Basmajian and C. J. De Luca. *Muscles Alive: Their Functions Revealed by Electromyography*. Williams & Wilkins, 5th edition, 1985.
- [3] E. A. Clancy, E. L. Morin, and R. Merletti. Electromyography signal processing and analysis. *Journal of Electromyography and Kinesiology*, 2013.
- [4] E. Criswell. *Criswell's Clinical Anatomy and Physiology for Allied Health*. Pearson Education, 2011.
- [5] C. J. De Luca. The use of surface electromyography in biomechanics. *Journal of Applied Biomechanics*, 18(2):135–163, 2002.
- [6] C. J. De Luca and B. Mambrizo. Voluntary control of human arm muscles: Coordination of the antagonistic muscle pairs during isometric contractions. *Journal of Neurophysiology*, 1987.
- [7] M. Grau, R. Engel, and R. Koch. Real-time monitoring of cardiovascular health using photoplethysmography. *Journal of Clinical Monitoring and Computing*, 29(1):57–64, 2015.
- [8] L. Liu, J. Zhang, and C. Wu. Multimodal machine learning for wearable health monitoring. *IEEE Transactions on Biomedical Engineering*, 66(3):745–754, 2019.
- [9] R. Merletti and D. Farina. Advances in electromyography and muscle physiology: The impact of new technologies on the study of muscle function. *Journal of Electromyography and Kinesiology*, 2016.
- [10] M. Schroeder, B. Frick, and K. Seebauer. Ppg signal amplitude and muscle contraction: Implications for peripheral circulation. *Journal of Vascular Research*, 48(2):133–143, 2011.

# All Generated Plots:EMG and PPG Comparison

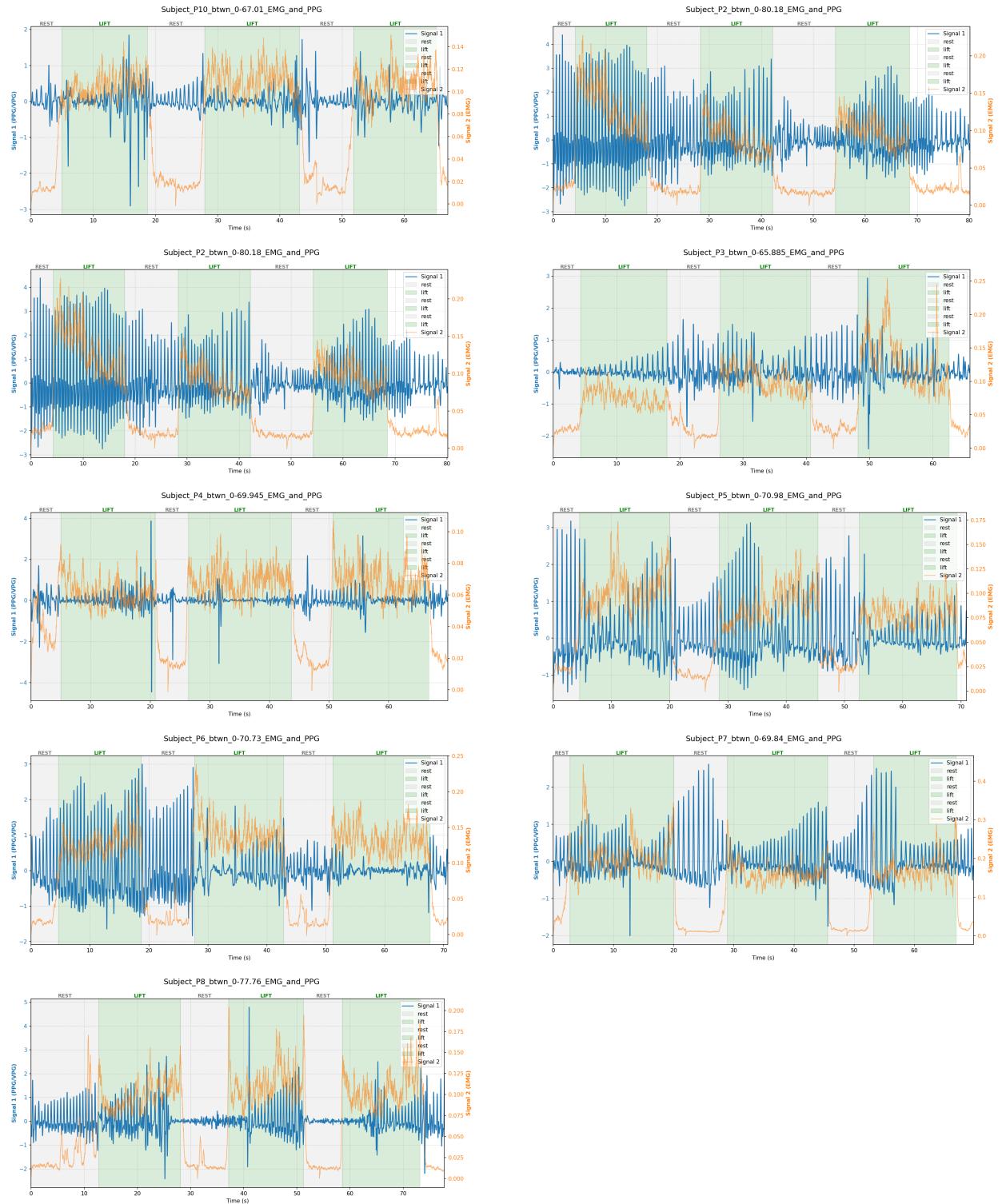


Figure 2: Comparison of EMG and PPG signals for different subjects.