SUMMARY

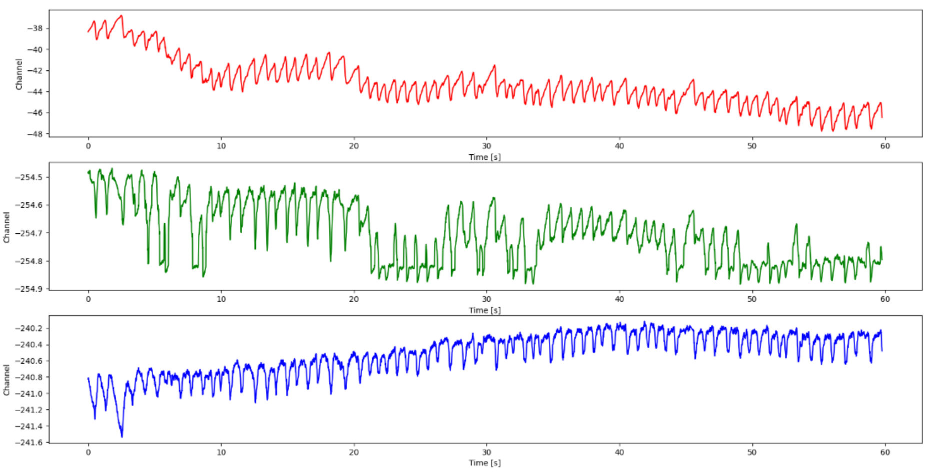
Tanya Kushwahaa

The remote photoplethysmogram (PPG) signal is widely used in the medical sector due to its non-invasiveness and cost-effectiveness in collecting signals through mobile phones and wristwatches. This signal contains valuable information on the variation of blood flow through tissues, offering insights into the cardiovascular, respiratory, and nervous systems. We can derive parameters such as heart rate, heart rate variability, and blood pressure from PPG signals which enables data collection for various cardiovascular diseases. The potential for diagnosing other respiratory or chronic diseases using PPG signals presents a promising breakthrough in medical science, especially considering the widespread availability of inexpensive devices which can be used to collect PPG signals such as mobile phones and wrist watches .

To measure cardiovascular health parameters accurately, obtaining high-quality rPPG signal is crucial. PPG signals are recorded with an oximeter while rPPG signals are typically retrieved through an RGB camera, such as the flash light of a phone camera. However, this method presents challenges such as motion artifacts, sudden movements, and background noise.

The primary challenge lies in combining RGB channels to form a high-quality rPPG signal. There are various algorithms in the literature for this purpose, including normalisation of color channels, giving weights to each channel, principal component analysis, and machine learning techniques. Nevertheless, it is a question of looking for more accurate, reliable and uniform algorithm for this conversion.

I was provided with 10 files, each containing 60 seconds of signals from red, green, and blue channels sampled at 75fps. It has been observed that the green color channel closely resembles a PPG signal compared to the red and blue channels [1]. Researchers have also hypothesized that the ratio of GR and GB or the sum of GR and GB may improve the quality of the formed rPPG signals [2]. I have employed the GRGB method to construct rPPG signals from RGB channels based on its performance [2].



Plot of unprocessed RGB channels. We can see a variety of issues from noise to baseline wander to motion artifact present in the signal.

1. What signal processing techniques would you use to clean the raw PPG signals and prepare them for feature extraction (i.e. cardiac metrics)?

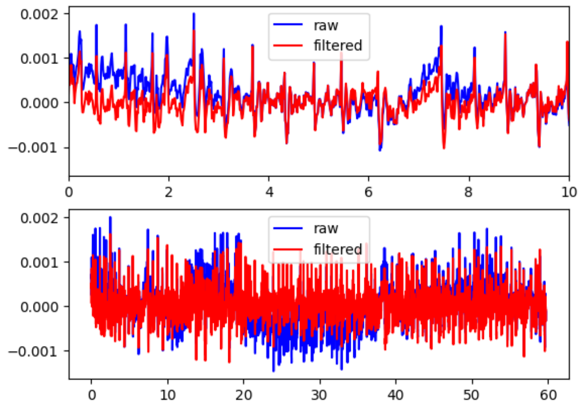
An rPPG signal can face various types of noises such as high-frequency noise, low-frequency noise, and noise due to motion artifacts, which are caused by movements. Processing a noisy RGB would result in a noisy rPPG signal. Therefore, before proceeding to calculate the heart rate parameters, we need to clean and filter the signal to obtain high-quality PPG signal.

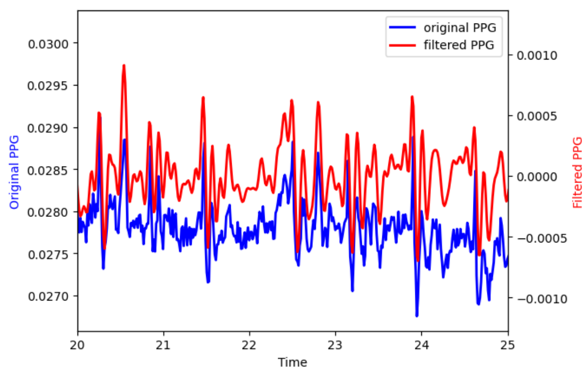
We can use the following signal processing techniques:

Filtering:

Filtering attenuates high-frequency noise and removes unwanted signal components outside the passband. Thus, effectively cleaning and preserving the interested frequency components of the signal. It is typically used to remove low- and high- frequency noise from the signal. However, filtering can also shift the signal to mean=0 by removing low- frequency components of the signal essentially noise including baseline wander.

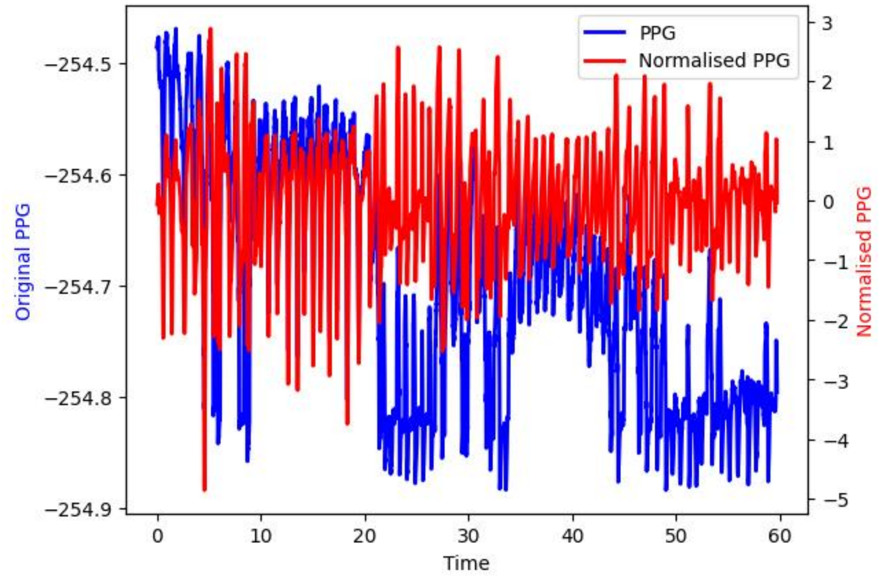
Signals below 0.5Hz and above 12 Hz are filtered out using a bandpass filter, commonly used to filter noise in PPG signals [4]. I have used a Butterworth filter with a bandpass filter to filter out the frequencies above 12 Hz and below 0.5. It also corrects the baseline wander.





Baseline correction : We can explicitly correct baseline wander in our signal by computing a baseline using polynomial fitting or cubic spline, which should be subtracted from the signal. This helps in bringing the signal to 0 and removing any unwanted variations in the signal.

Normalization: As the RGB channels vary quite significantly in their amplitudes, normalising them would make them easier for visualisation as all the channels will be scaled around 0. This will not only help in visualisation but also prevent any single channel from dominating due to large amplitudes.



2. How would you deal with the missing data issue?

There are a few ways in which we can deal with missing data issue:

1. One approach to deal with missing data in PPG datasets is by interpolating or extrapolating between the neighbouring data points. There are many interpolation algorithms such as linear interpolation, cubic spline interpolation.

2. Another approach could be approximating the missing date using a dummy model signal that mimics the characteristics of the PPG signal. However, this method can become complex when dealing with variations in the PPG signals.

3. In cases, where numerous data points are missing and the overall signal is not significantly affected, we can completely remove the corresponding missing time period of the signal.

3. How would you go about implementing changes to your code or pipeline if your estimated cardiac metrics do not match the provided expected results?

Quality assessment of the signal [5], checking the signal to noise ratio and enhancing any attenuated signal to prevent loss of peaks caused by sensor issues.

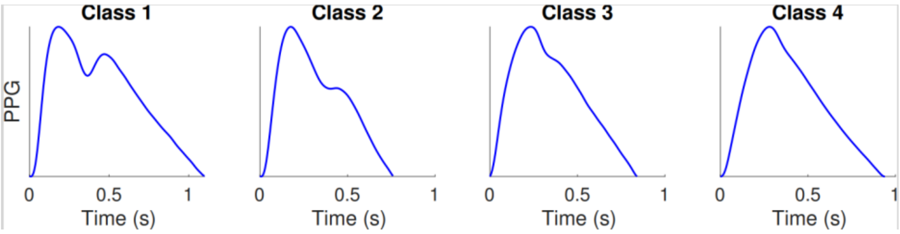
Optimising the filtering process to avoid losing important components of the PPG signal.

Variation in the signal could also cause problems with peak detection. Therefore, it is crucial to optimise the peak detection algorithm for a variety of signals and consider any variations in the signal being used.

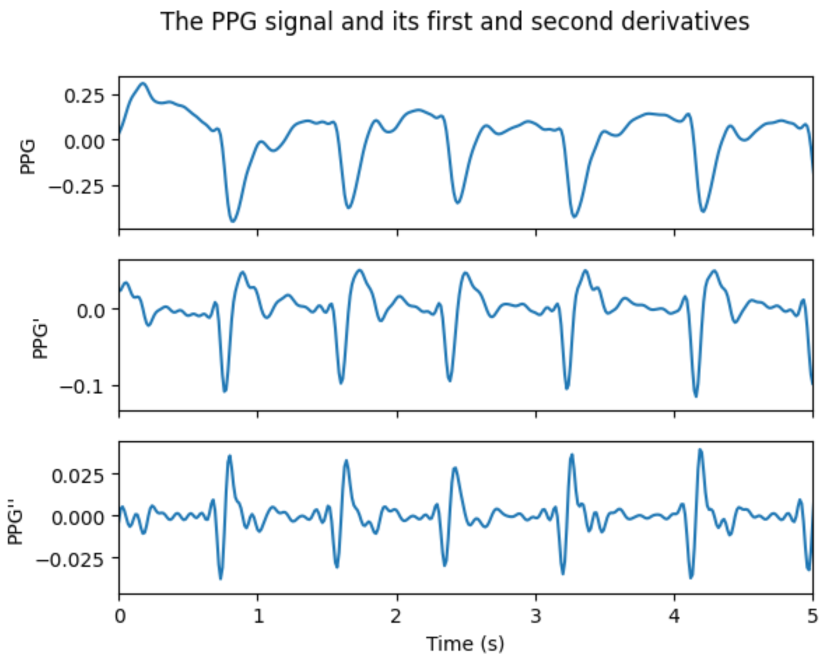
Compare the obtained results with known results from literature or other datasets.

(Many deep learning models have been used in the literature for optimisation [3])

4. What approaches could you use to discriminate between subjects in sinus rhythm and in atrial fibrillation?



(From [6])



A healthy individual typically exhibits two peaks in their PPG signal: the systolic peak and diastolic peak. However, factors like age and illness can cause variation in these peaks and diastolic peak can even disappear. The variations in the intervals between systolic peaks may indicate abnormalities such as atrial fibrillation.

The first and second derivatives of the PPG signal can provide valuable information about sinus rhythms and detect any variations from the normal. The first maxima corresponds to the systolic peak while the next maxima is the diastolic peak. Measuring the strengths of the peaks from the baseline or calculating the slope from the baseline can give information about pulse dynamics.

To differentiate between SR and AF subjects in the PPG signal. It is essential to separate the two components from an rPPG signal. Analysing them in the frequency space could provide details about different frequencies associated with varying pulses. AF may show varying pulse amplitudes and irregularities whereas SR tends to be more regular and evenly spaced. We can study features such as beats, R-R interval variations, slope and spectral characteristics from it.

Machine learning models including supervised or unsupervised learning algorithms can be used with input features such as beats, R-R interval, and slope. Deep learning models like CNNs can also be used for classification process [5].

Analysing the signal in frequency domain can also help in understanding pulse variability.

5. How would you visualise the results of this project?

Line & scatter plots of the signal : temporal and frequency space

Histogram of amplitudes of signal to show any variation in the beats

Plot of R-R intervals to show the heart rate variability

More interactive plots like 3D plots or dynamical simulations of PPG signals to see any irregularities in the signals

Phase space representation of PPG signals to see dynamics and any non-dynamics present in the signal

********************** Citations *************************

[1] Verkruysse, W.; Svaasand, L.O.; Nelson, J.S. Remote plethysmographic imaging using ambient light. Opt. Express 2008, 16, 21434–21445.

[2] Haugg F, Elgendi M, Menon C. GRGB rPPG: An Efficient Low-Complexity Remote Photoplethysmography-Based Algorithm for Heart Rate Estimation. Bioengineering (Basel). 2023 Feb 12;10(2):243. doi: 10.3390/bioengineering10020243. PMID: 36829737; PMCID: PMC9952130.

[3] Cheng CH, Wong KL, Chin JW, Chan TT, So RHY. Deep Learning Methods for Remote Heart Rate Measurement: A Review and Future Research Agenda. Sensors (Basel). 2021 Sep 20;21(18):6296. doi: 10.3390/s21186296. PMID: 34577503; PMCID: PMC8473186.

[4] Márton Á Goda *et al* 2024 *Physiol. Meas.* **45** 045001

[5] Pereira, T., Tran, N., Gadhoumi, K. *et al.* Photoplethysmography based atrial fibrillation detection: a review. *npj Digit. Med.* **3**, 3 (2020). https://doi.org/10.1038/s41746-019-0207-9

[6] Desquins, T.; Bousefsaf, F.; Pruski, A.; Maaoui, C. A Survey of Photoplethysmography and Imaging Photoplethysmography Quality Assessment Methods. *Appl. Sci.* **2022**, *12*, 9582. https://doi.org/10.3390/app12199582

[6] Charlton PH, Paliakaitė B, Pilt K, Bachler M, Zanelli S, Kulin D, Allen J, Hallab M, Bianchini E, Mayer CC, Terentes-Printzios D, Dittrich V, Hametner B, Veerasingam D, Žikić D, Marozas V. Assessing hemodynamics from the photoplethysmogram to gain insights into vascular age: a review from VascAgeNet. Am J Physiol Heart Circ Physiol. 2022 Apr 1;322(4):H493-H522. doi: 10.1152/ajpheart.00392.2021. Epub 2021 Dec 24. PMID: 34951543; PMCID: PMC8917928.