Heart Rate Estimation using Remote Photoplethysmography

Ali Azak - 22301042
Murat Şahin - 22301345

Presentation Structure

- Introduction
- UBFC Dataset
- Methodology
    - Region of Interest (ROI) Extraction
    - Getting Mean Signal
    - Normalizing and Detrending
    - Moving Average Filter
    - Source Separation
    - Selecting The Final Signal
    - Bandpass Filtering & HR Estimation
    - ROI Revisited
- Evaluation
- Challenges
- Conclusion

Introduction

- **Heart rate measurement is critical, important for diagnosis**

- **Timeline:**
    - Manual pulse checking
    - Stethoscopes (1800s)
    - Photoplethysmography (PPG) -> Many applications (smartwatches, fitness band etc.)
    - Remote Photoplethysmography (rPPG) -> Still developing

Just from video recordings!

Observing Color Changes

figure-ref: https://medium.com/augmented-startups/heart-rate-detection-using-eulerian-video-magnification-yolor-49818dd1b2f5

UBFC rPPG Dataset

UBFC rPPG

- Video recordings with webcam

  30 fps with a resolution of 640x480

- Ground truth heart rate recorded with CMS50E Pulse Oximeter
- All experiments are conducted indoors & they have varying illumination scenarios
- Two different sets of data
  - Stand still
  - Play a time sensitive mathematical game

S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, "Unsupervised skin tissue segmentation for remote photoplethysmography", Pattern Recognition Letters, 2017.

First Set

Second Set

Methodology

Methodology

Region of Interest(ROI) Extraction

Region of Interest(ROI) Extraction

Initial attempt -> using the whole face

- Employed Viole Jones in each frame

Region of Interest(ROI) Extraction

Several problems occurs.

Finding face in each frame yields with variational background

Region of Interest(ROI) Extraction

Several problems occurs.

Eyes and mouth significantly affects the results

Region of Interest(ROI) Extraction

First Solution:

Finding face once, extracting a fixed area from forehead & using it for the whole video

At first seems okay because subjects stands still

Forehead position and openings

are different for subjects

Region of Interest(ROI) Extraction

Region of Interest(ROI) Extraction

Second Solution:

Using Facial Landmark points

dlib library 68 facial points with additional 13 points

https://github.com/codeniko/shape_predictor_81_face_landmarks

Region of Interest(ROI) Extraction

Region of Interest(ROI) Extraction

We extracted regions from forehead, cheeks and chin.

Observed that for different subjects, different area may perform better.

Developed an innovative approach

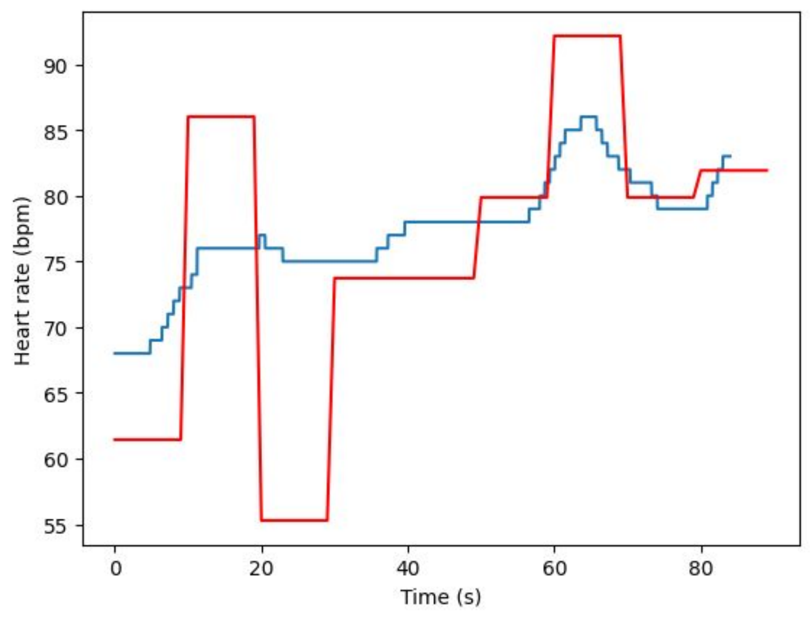
Process each ROI separately and select the signal with most power at the end.

- More Robust
- Better Performance

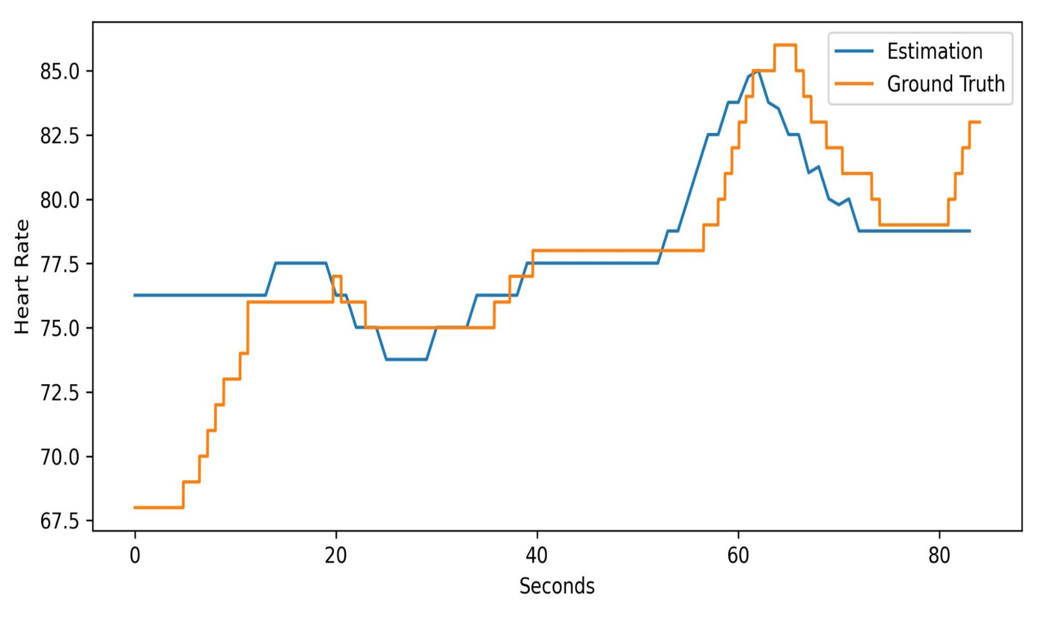
Region of Interest(ROI) Extraction

Improvement

Estimation in Progress Presentation



Estimation in This Presentation

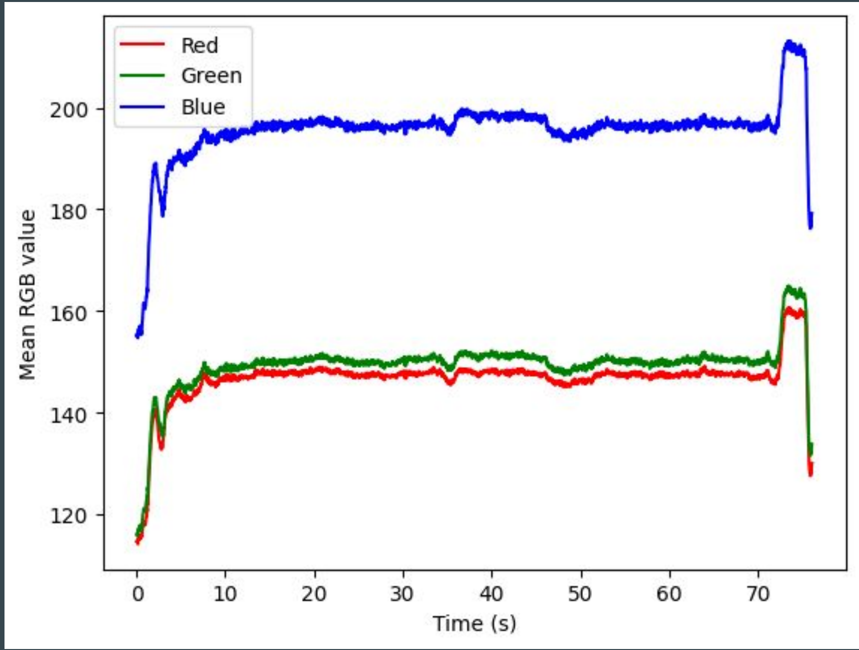


Methodology

Getting Mean Signals

Getting Mean Signals

Intuition is spatially averaging all color channels.



Getting Mean Signals

In literature, some different methods employed

Verkruysse et al. found that the green channel in RGB produces the strongest HR signal

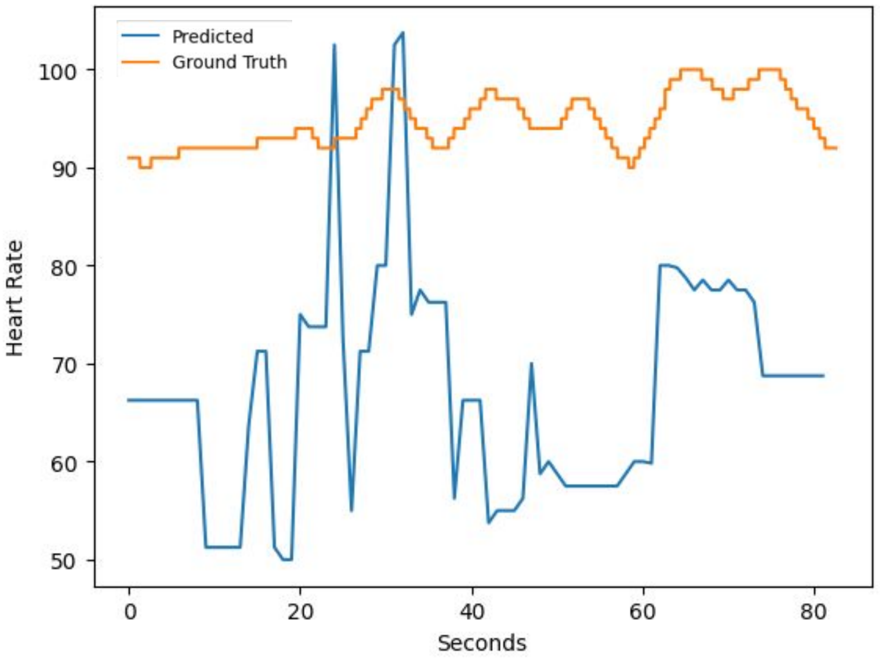
Pal et al. states that red channel in RGB gives better results.

We experimented three methods

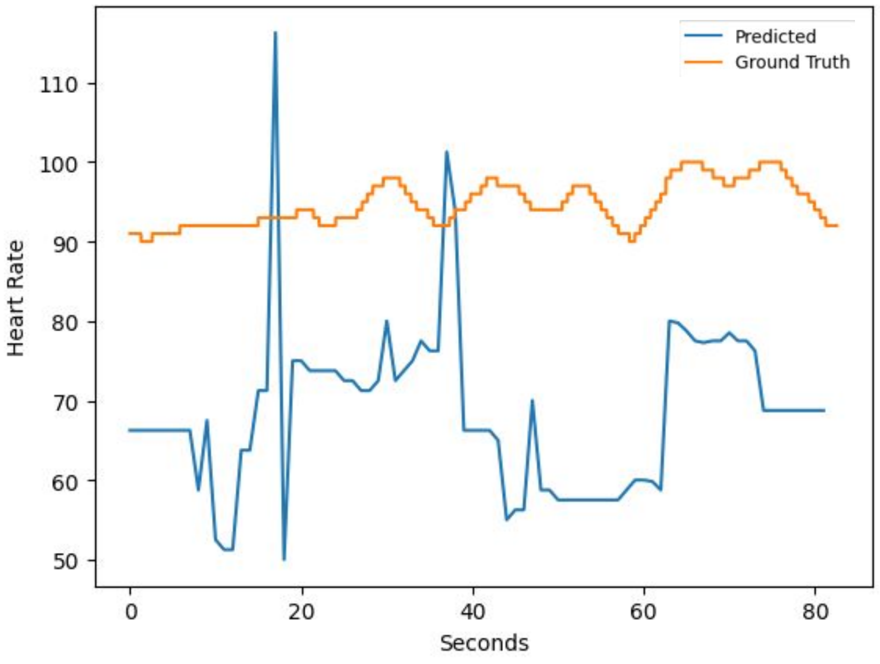
Verkruysse W, Svaasand LO, Nelson JS. Remote plethysmographic imaging using ambient light. Opt Express. 2008 Dec 22;16(26):21434-45. doi: 10.1364/oe.16.021434. PMID: 19104573; PMCID: PMC2717852.
Arpan Pal, Aniruddha Sinha, Anirban Dutta Choudhury, Tanushyam Chattopadyay, and Aishwarya Visvanathan. 2013. A robust heart rate detection using smart-phone video. In Proceedings of the 3rd ACM MobiHoc workshop on Pervasive wireless healthcare (MobileHealth '13). Association for Computing Machinery, New York, NY, USA, 43–48. https://doi.org/10.1145/2491148.2491156

Getting Mean Signals

Only Red Channel



Only Green Channel



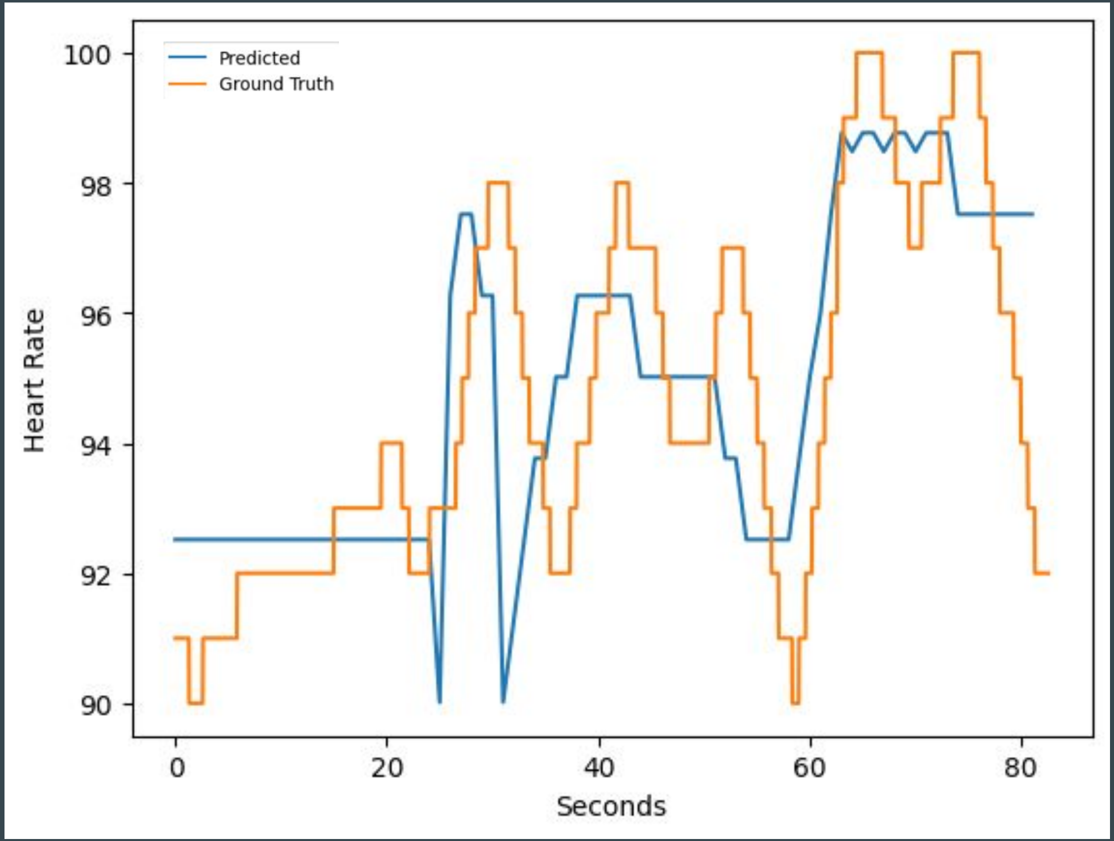
Getting Mean Signals

We conclude that

RGB channels combined

(With ICA)

produces the best results.



Methodology

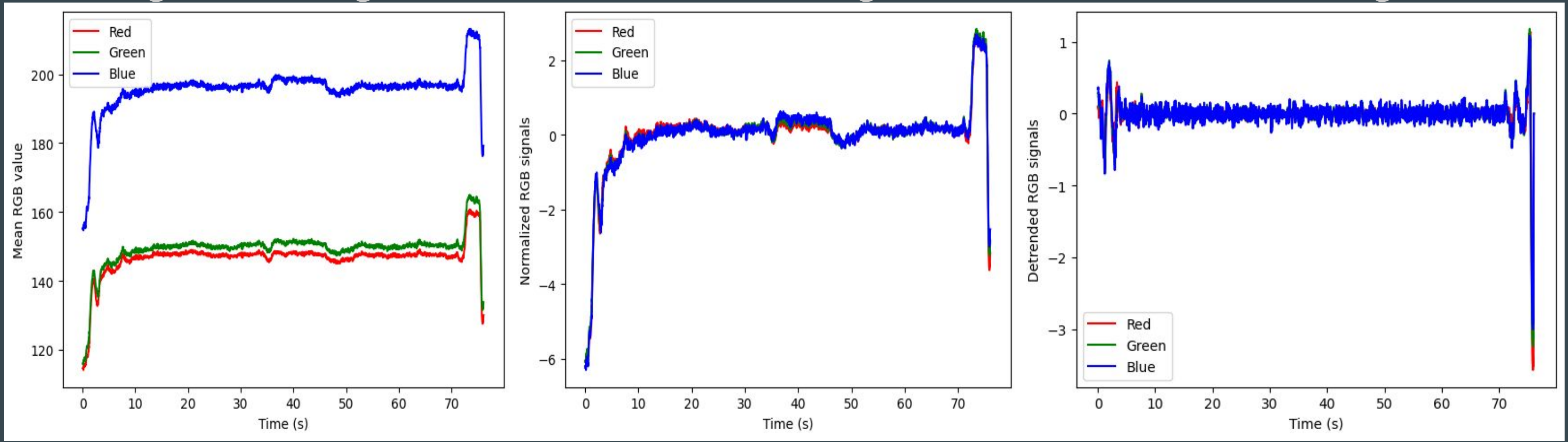
Normalizing and Detrending

Normalizing and Detrending

Environmental Artifacts (Illumination, Noise)

Irrelevant Fluctuations

Original Mean Signal | Normalized Signal | Detrended Signal



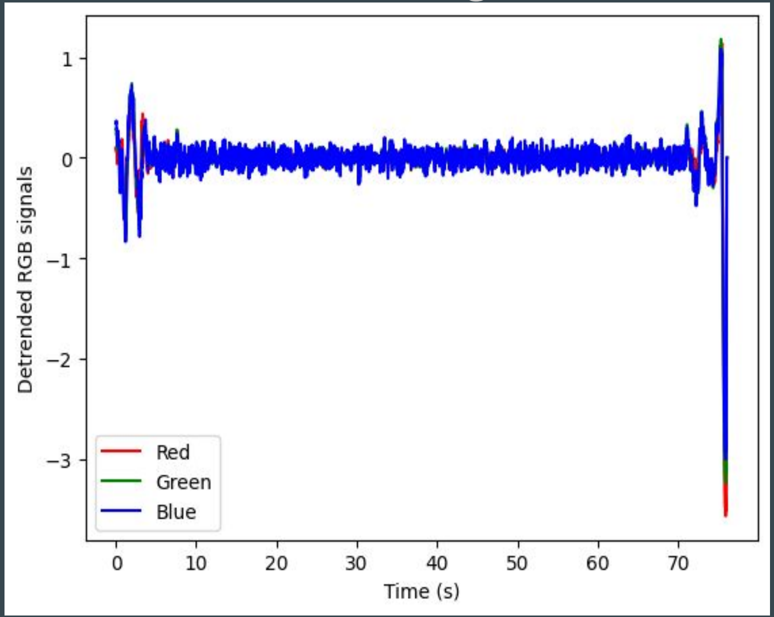
Methodology

Moving Average Filter

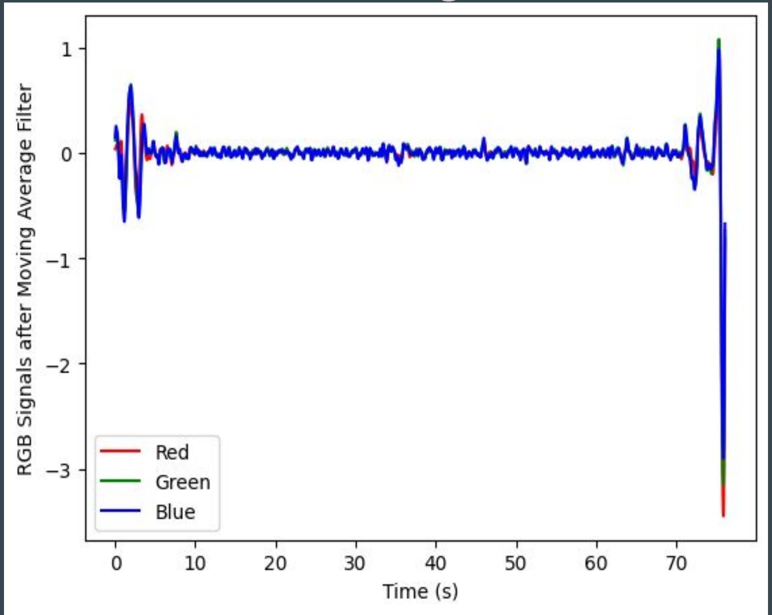
Moving Average Filter

Used for obtaining smoother signal, different window sizes tested

Detrended Signal



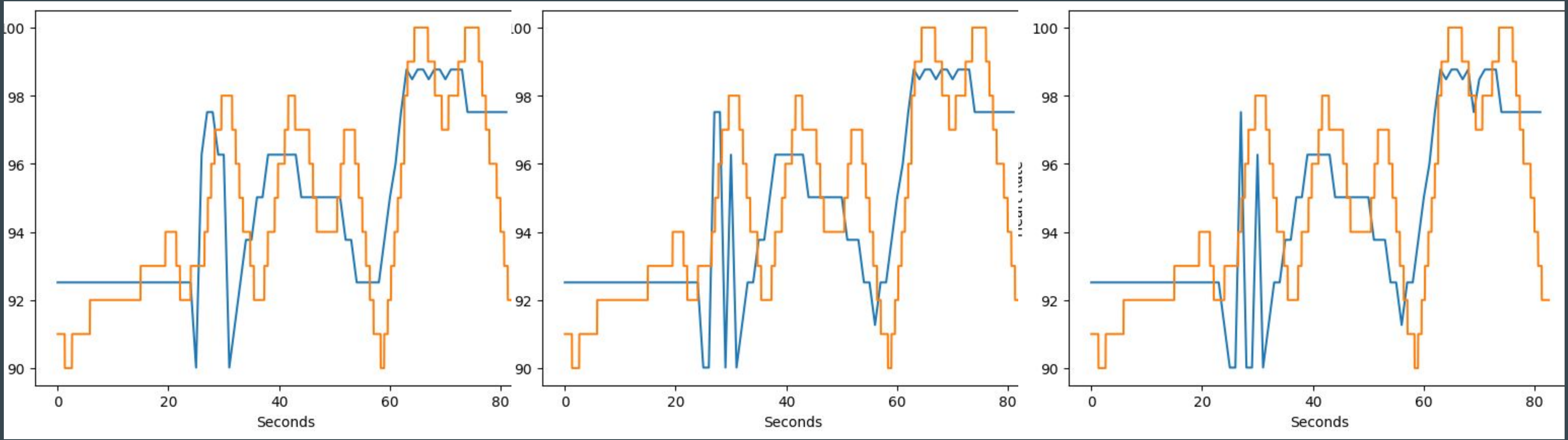
Filtered Signal



Moving Average Filter

Used for obtaining smoother signal, different window sizes tested

Window Size = 5 (0.18s) | Window Size = 7 | Window Size = 9



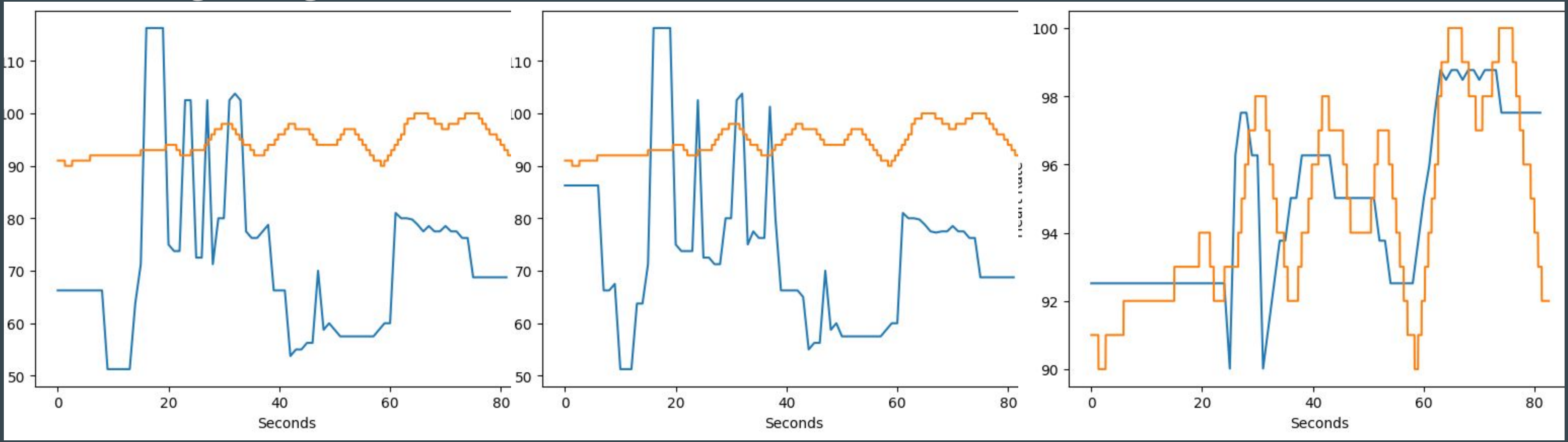
Methodology

Source Separation

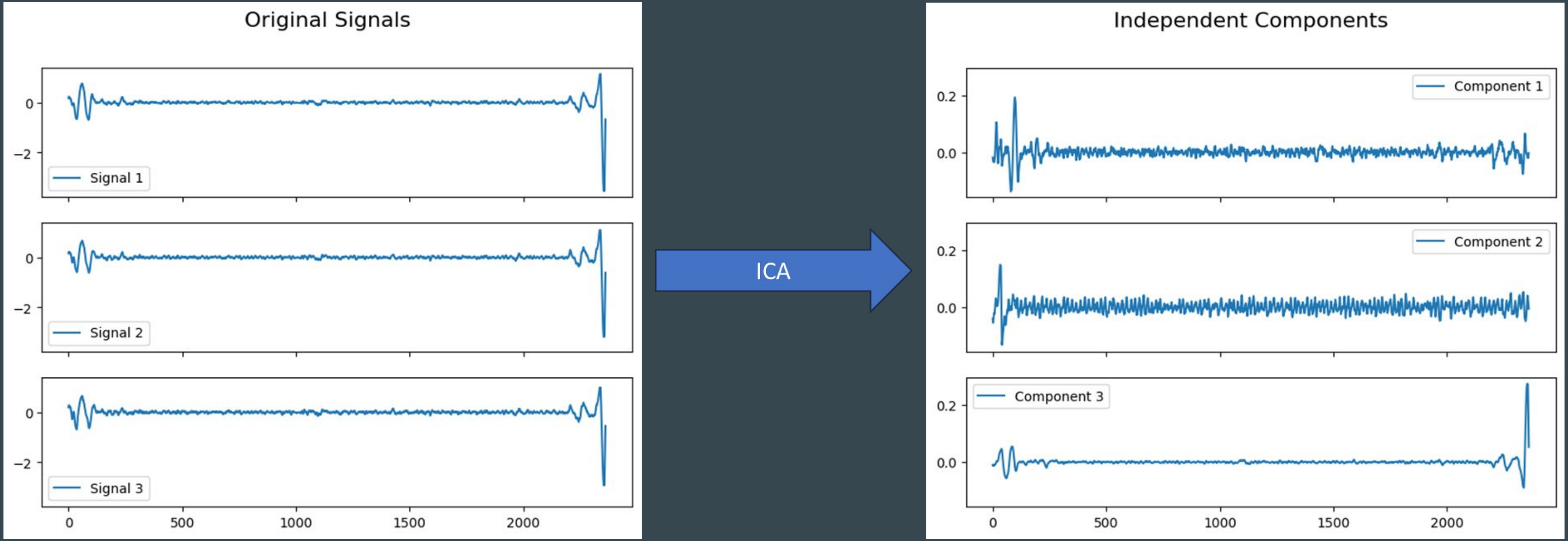
Source Separation

- Different source separation techniques are widely used
- PCA and ICA are the most popular

Original Signals PCA ICA



Source Separation

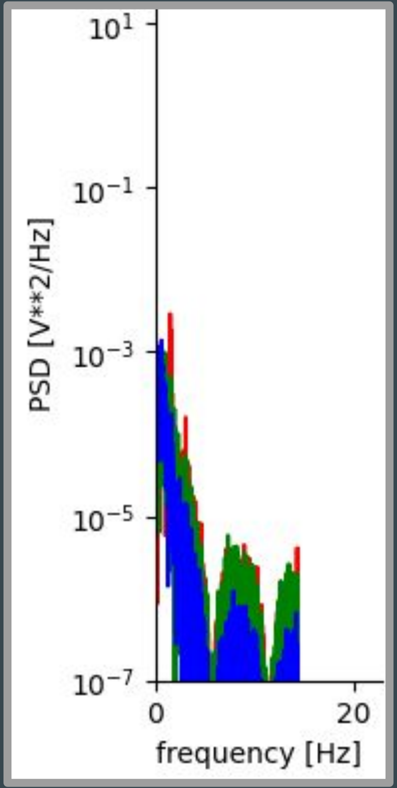


Methodology

Selecting Best Component

Selecting Best Component

- After the source separation, we are selecting the most powerful signal
- Power Spectral Density (PSD) is used
- Describes how the power of a signal is distributed across different frequencies



Methodology

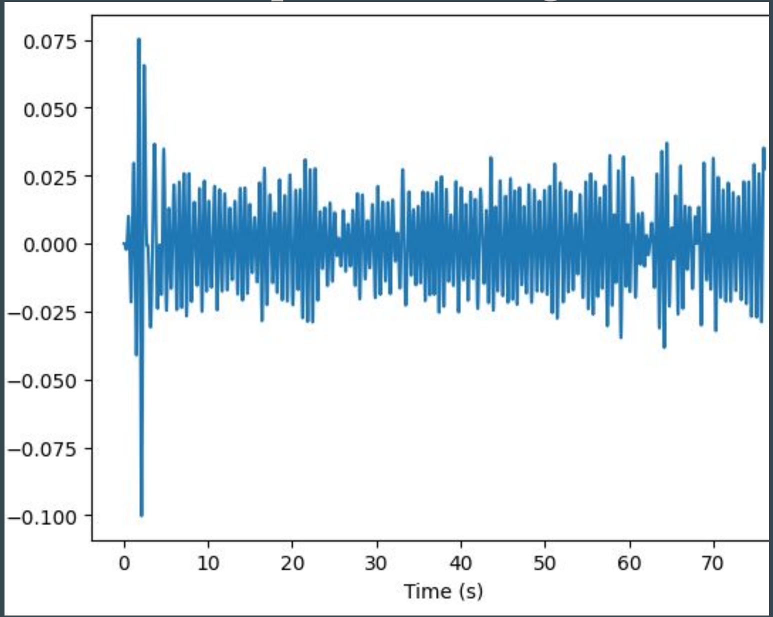
Bandpass Filtering & HR Estimation

Bandpass Filtering & HR Estimation

- Human heart rate is in a specific range
- Bandpass filter is applied between 0.5 Hz to 2 Hz

- Afterwards, most dominant frequency in the final signal (Using PSD) is the heart rate signal

Bandpass Filtered Signal



Methodology

ROI Revisited

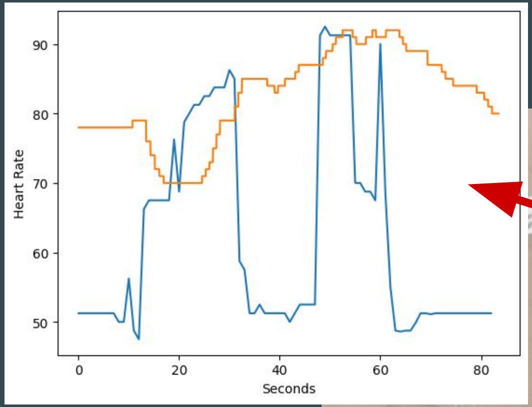
ROI Revisited

- Remember how different ROIs were selected

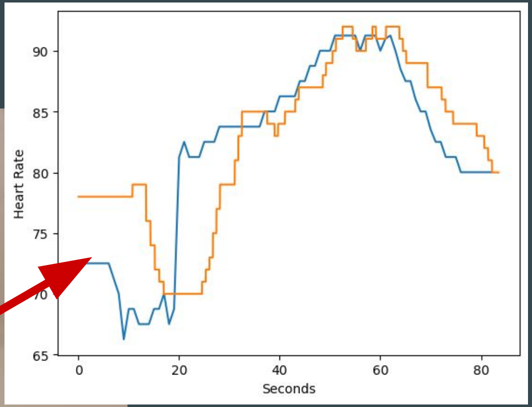
- For each ROI, a final signal is extracted and the signal that has the most power on a frequency is selected

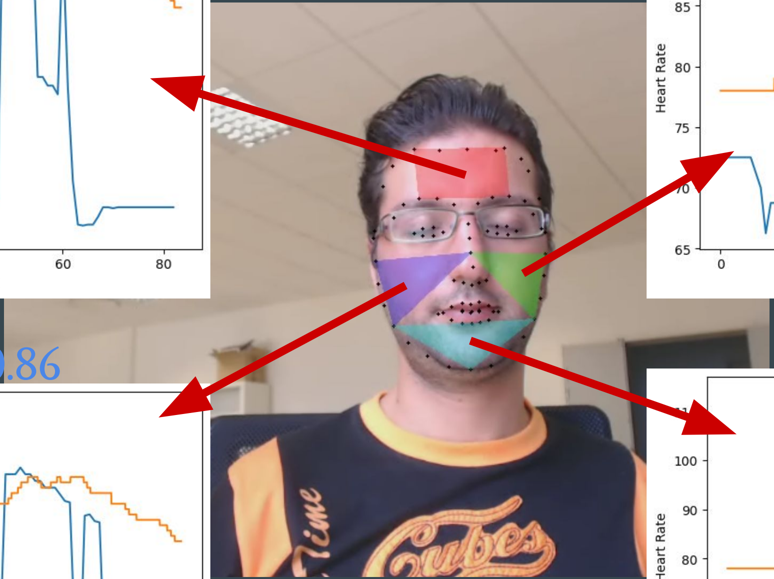
ROI Revisited

Power: 1.89

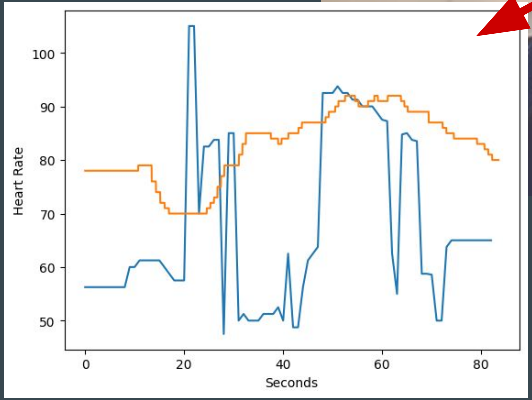


Power: 6.71

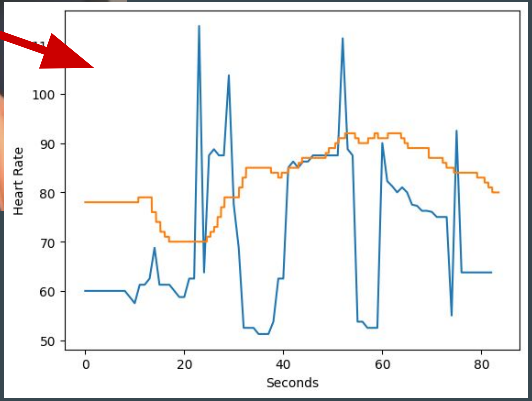




Power: 0.86



Power: 0.50

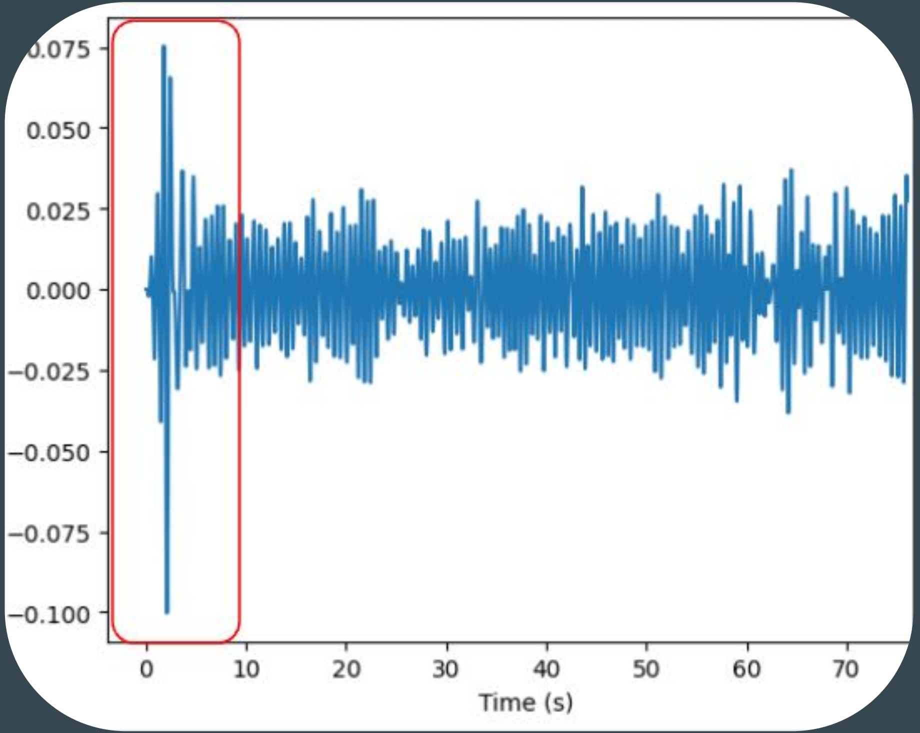


Evaluation

Evaluation

- For each video in our dataset, various metrics are calculated

- For each second, a prediction is made by using a 12 seconds window



Evaluation

- Mean Difference is for the whole video, others are calculated per second
- Results for the first dataset:

Subject	Mean Difference	MAE	RMSE	MAPE	Pearson Corr.
10-gt	0.521349	2.661177	3.427617	3.654258	-0.21588
11-gt	0.764481	1.151195	1.493356	1.474624	0.768908
12-gt	0.486518	1.743347	2.267038	1.838583	0.574649
15-gt	0.617016	1.754289	2.771182	2.365135	0.764313
6-gt	0.824021	3.763775	5.126114	4.833165	0.752028
7-gt	0.121575	1.449995	1.902859	1.530778	0.739751
8-gt	0.030669	6.224504	7.117295	8.955437	0.505299
Mean of subjects	1.337947	2.678326	3.443638	3.521711	0.555581

- Results for the second dataset:

Mean of subjects	5.702311	7.717295	11.248894	133.724075	0.367681

Evaluation (Comparisons)

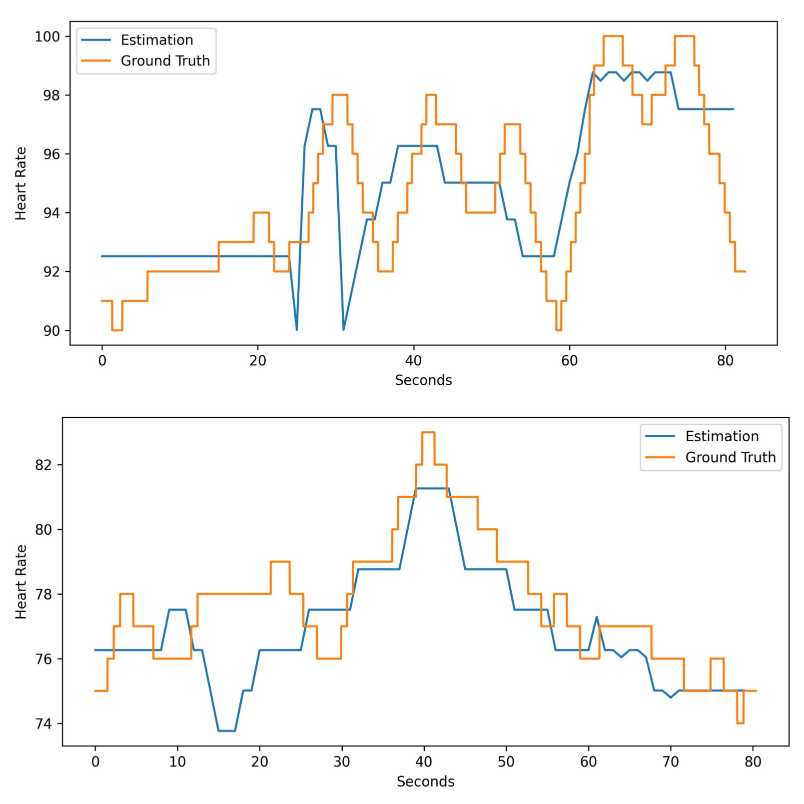
- Average results for both datasets:

Method	MAE	RMSE	Pearson Corr.
POS (Classical)	8.35	10.00	0.24
CHROM (Classical)	8.20	9.92	0.27
Ours	6.98	10.11	0.40
Green (Classical)	6.01	7.87	0.29
META-rPPG (CNN-LSTM)	5.97	7.42	0.53
SynRhythm (CNN)	5.59	6.82	0.72
PulseGAN (GAN)	1.19	2.10	0.98
Dual-GAN (GAN)	0.44	0.67	0.99

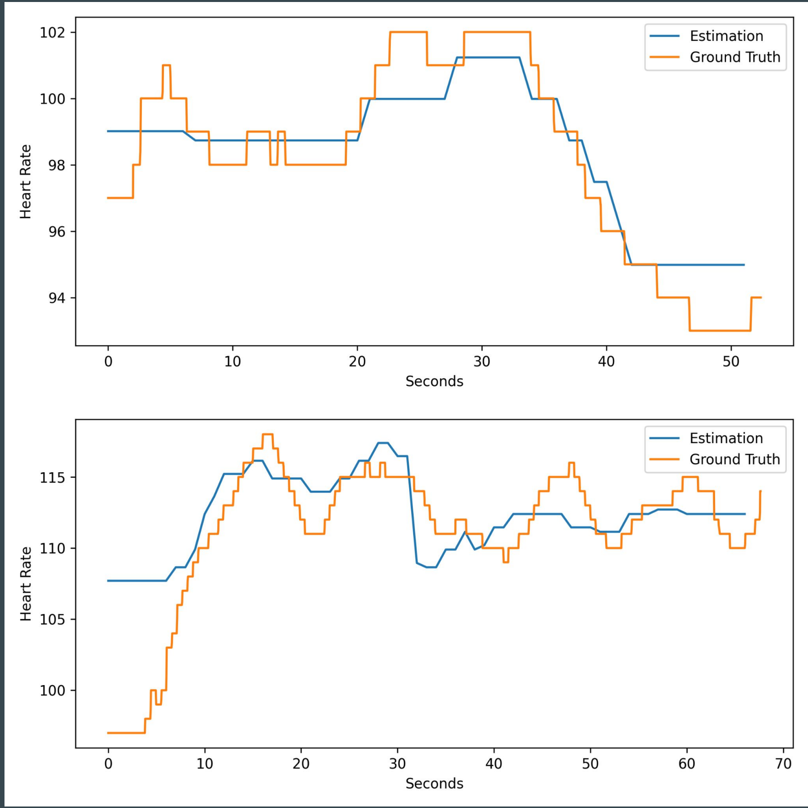
Table 1: Performance in Both Datasets

Evaluation

First Dataset



Second Dataset



Challenges

- Subject movement is an issue
- Illumination changes are causing problems
- Results may be perfect or they may not reflect the reality

Conclusion

- rPPG has potential for non-invasive heart rate measurement
- Our results are promising, deep learning can improve further
- Robust methods are required for widespread usage

THANKS FOR LISTENING