Faculty of Engineering - Cairo University
Credit Hours System - Senior Level
Department
Spring 2025

Data Analytics in medicine

Final Project Report

Prepared by:

Abdelrahman Kamal Ibrahim (ID: 4210201)

Mohamed Osama (ID: 1190234)

Ammar Yasser (ID: 4210206)

Mohamed Ehab Souka (ID: 1190234)

Submitted to:

Dr. Ibrahim Sadek

BCG Heart Rate Detection

The primary goal of this project is to measure heart and respiratory rates using signals acquired from a microbend fiber optic sensor. This sensor is placed under a subject's mattress, approximately below their chest and stomach, to capture the mechanical activity of the heart and respiratory movements.

Project Workflow :

The project follows a structured approach to process Ballistocardiography (BCG) and Reference (RR) data to ultimately compare heart rates derived from BCG signals against a reference.

1. Timestamp Standardization:

- BCG Timestamps: Raw BCG timestamps are processed and converted to UTC format. The convert_timestamp_format_BCG.py script handles this, calculating subsequent timestamps based on an initial timestamp and a sampling frequency (e.g., 140 Hz).

- RR Timestamps: Timestamps from reference heart rate data (RR files) are also converted to UTC format. This is handled by convert_timestamp_format_RR.py .

- The patient folders from 21 to 30 were not used due to absence of reference RR heart rate files.

BCG:

BCG	Timestamp	fs	Timestamp_UTC
-86	1.69902E+12	140	2023-11-03 14:35:12.866000+00:00
-90	1.69902E+12		2023-11-03 14:35:12.873143+00:00
-91	1.69902E+12		2023-11-03 14:35:12.880286+00:00
-86	1.69902E+12		2023-11-03 14:35:12.887429+00:00
-84	1.69902E+12		2023-11-03 14:35:12.894572+00:00
-88	1.69902E+12		2023-11-03 14:35:12.901714+00:00
-90	1.69902E+12		2023-11-03 14:35:12.908857+00:00
-90	1.69902E+12		2023-11-03 14:35:12.916000+00:00

RR:

Timestamp	Heart Rate	RR Interval in seconds
2023-11-03 22:39:15+00:00	85	0.708
2023-11-03 22:39:15+00:00	85	0.657
2023-11-03 22:39:16+00:00	85	0.798
2023-11-03 22:39:17+00:00	84	0.712
2023-11-03 22:39:17+00:00	84	0.73
2023-11-03 22:39:18+00:00	84	0.742
2023-11-03 22:39:19+00:00	84	0.715
2023-11-03 22:39:20+00:00	83	0.728
2023-11-03 22:39:20+00:00	83	0.736

2. Data Synchronization:

- The timestamps from the BCG and RR datasets are unified and synchronized. This crucial step ensures that data points from both sources correspond to the same moments in time. The sync_BCG_RR_timestamps.py script appears to perform this, merging datasets based on the nearest timestamps within a defined tolerance.

- The patients folders 6,13,20,32 were neglected due to the absence of any matching sensors reading at the same time stamps.

BCG	Timestamp_x	fs	Timestamp_UTC	Timestamp_y	Heart Rate	RR Interval in seconds
-410	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-510	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-118	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-743	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-98	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-1000	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-61	1.69913E+12		12:19.7	11/4/2023 19:12	99	0.613
-1174	1.69913E+12		12:19.8	11/4/2023 19:12	99	0.613
-1254	1.69913E+12		12:19.8	11/4/2023 19:12	99	0.613

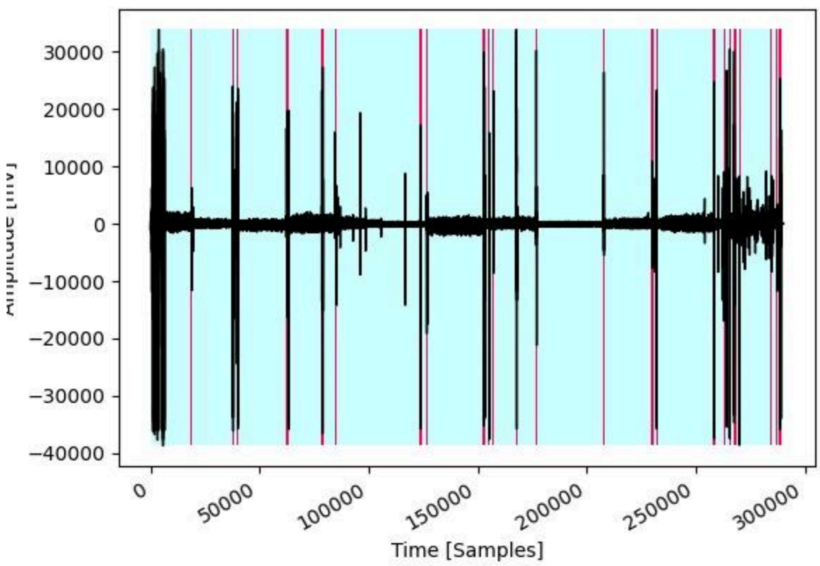
3. BCG Signal Resampling:

- The BCG signal, originally sampled at a higher frequency (e.g., 140 Hz), is resampled to a lower frequency (e.g., 50 Hz). This is a common preprocessing step to standardize the data or reduce computational load. Scripts like resample_BCG.py or resresspppple.py (likely a typo for resample) are used for this, employing methods like Fourier-based interpolation.

BCG_downsampled	Timestamp_x_downsampled	Heart Rate_downsampled
-328.5836716	1.69913E+12	99
-360.1971764	1.69913E+12	99
-706.0926754	1.69913E+12	99
-632.6275959	1.69913E+12	99
56.84153293	1.69913E+12	99
-21.05895873	1.69913E+12	99
-323.9227668	1.69913E+12	99
-1152.241436	1.69913E+12	99
-475.2189988	1.69913E+12	99

4. Signal Processing and Feature Extraction (primarily in main.py):

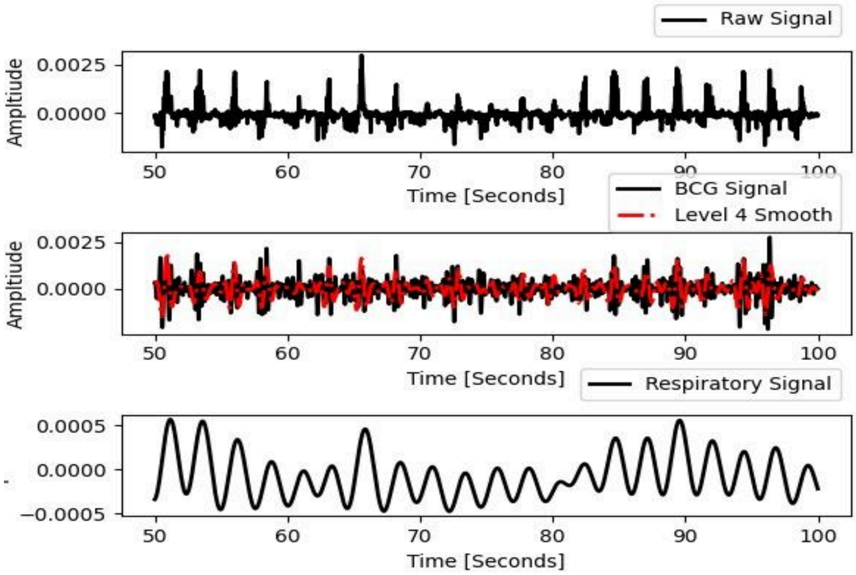
- Data Loading: The main.py script iterates through patient data folders, loading CSV files containing BCG and RR data. and on processing a single, pre-processed (synchronized and resampled) CSV file.
- Body Movement Detection: The detect_patterns function (from detect_body_movements.py) is used to identify and potentially segment or handle periods of significant body movement in the BCG signal, as these can interfere with



heart rate extraction.

- Band-Pass Filtering: The raw BCG data is filtered to isolate specific frequency bands corresponding to BCG (heart activity) and respiratory signals. This is done by the band_pass_filtering function.

- Wavelet Transform: Maximal Overlap Discrete Wavelet Transform (MODWT) and its Multi-Resolution Analysis (MRA) are applied to the filtered BCG signal (movement component). This is handled by modwt (from modwt_matlab_fft.py) and modwtmra (from modwt_mra_matlab_fft.py). The 'bior3.9' wavelet is used, and a specific level (e.g., level 4) of the MRA decomposition (dc[4]) is selected as the wavelet_cycle for heart rate estimation.



- Heart Rate Estimation:
- In main.py , the heart_rate function is called to estimate beats per minute (BPM) from the wavelet_cycle .

- Reference Heart Rate Processing: The reference heart rates (rr_heart_rates) are read from the input CSV.

- Resampling/Alignment for Comparison (in main.py): To compare the BCG-derived heart rates with the reference RR heart rates, the RR heart rates are resampled (or averaged within windows) to match the length/timing of the BCG-derived heart rate array. This ensures a fair comparison.

5. Performance Evaluation and Analysis :

- Metrics Calculation: Statistical metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) are calculated to quantify the accuracy of the BCG-derived heart rate against the reference RR data for each patient seperately.

- Statistical Summary: Minimum, maximum, and average heart rates are printed for both BCG and reference signals, along with the absolute difference between their means.

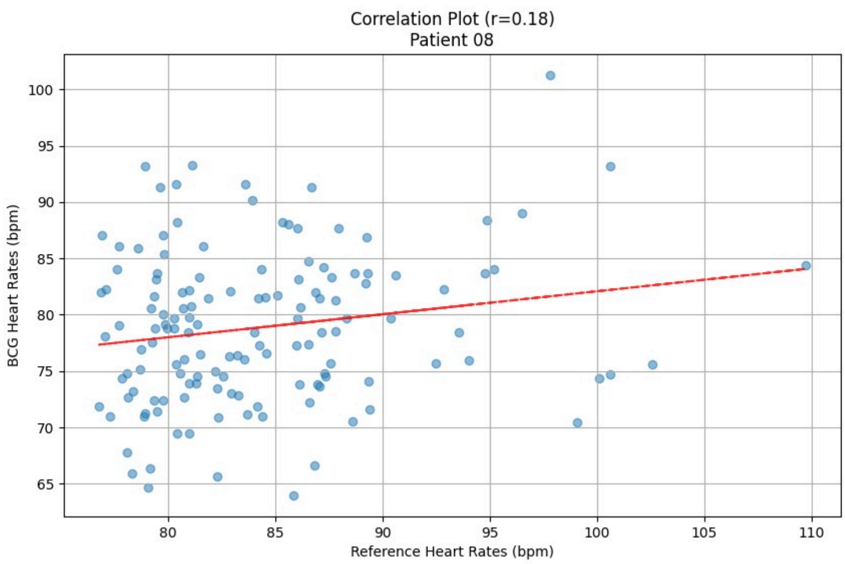
```
BCG Heart Rate Information
Minimum pulse :  62.0
Maximum pulse :  99.0
Average pulse :  81.0

Reference Heart Rate Information
Minimum pulse :  72.0
Maximum pulse :  113.0
Average pulse :  80.0

Heart Rate Difference (BCG vs Reference): 1.0

Heart Rate Information:
========================================================================================================

Mean absolute error:  8.883680385996668
========================================================================================================
Root mean square error:  10.933793702617082
========================================================================================================
Mean absolute percentage error:  0.11342903397329349
========================================================================================================
========================================================================================================
```

```
BCG Heart Rate Information
Minimum pulse :  58.0
Maximum pulse :  103.0
Average pulse :  78.0

Reference Heart Rate Information
Minimum pulse :  56.0
Maximum pulse :  126.0
Average pulse :  72.0

Heart Rate Difference (BCG vs Reference): 6.0

Heart Rate Information:
========================================================================================================

Mean absolute error:  6.62517676595728O4
========================================================================================================
Root mean square error:  8.472748112520238
========================================================================================================
Mean absolute percentage error:  0.08745461208559091
========================================================================================================
========================================================================================================
```

- Data Logging: Key information, including patient ID, average RR heart rate, average BCG heart rate, and the calculated error metrics, are appended to a dataInfo list, which is later converted to a Pandas DataFrame and can be saved to a CSV file (e.g., patientinfo.csv).

- Plotting: The create_analysis_plots function generates and saves:

- Correlation Plots: Scatter plots of reference heart rates vs. BCG heart rates, with a line of best fit and Pearson correlation coefficient.



- Bland-Altman Plots: Plots showing the agreement between the two measurement methods (BCG vs. Reference) by plotting the difference against the mean of the two heart rates.