HR CALIBRATION & FORECASTING USING PPG-DALIA

OVERVIEW

This model aims to accurately calibrate and forecast an individual's heart rate (HR) using multimodal physiological data collected from wearable sensors. It addresses the inaccuracies associated with traditional estimations of HR (e.g. Blood Volume Pulse, BVP) by combining additional metrics that compensate for BVP's motion-related shortfalls. The model simultaneously performs HR calibration and HR forecasting, with forecasts extending 1 minute into the future.

METHODOLOGY

The project utilises the PPG-DaLiaA dataset, which contains 2.5 hours of continuous multimodal recordings from 15 subjects performing everyday activities under free-living conditions. Physiological signals* were collected via the Empatica E4 wristband, while the RespiBAN chest sensor provided the ECG-driven heart-rate used as the label.

*Blood Volume Pulse (BVP), Accelerometer (ACC), Electrodermal Activity (EDA), and Skin Temperature (TEMP).
Because BVP estimates are easily corrupted by motion artefacts, the inclusion of ACC, EDA, and TEMP allows the model to interpret context and correct for those artefacts when estimating true HR.

Each subject's data was segmented into 8-second windows (512 samples at 64Hz), with successive windows overlapping by 6 seconds (2-second shift). The corresponding ECG-based HR label was the mean HR within each window. To ensure temporal alignment across sensors, lower-frequency signals (ACC: 32 Hz, EDA/TEMP: 4 Hz) were interpolated to match the 512-sample length of the BVP window.

A Leave-One-Subject-Out (LOSO) approach was used to assess the model's generalizability across individuals. This means that an entire 2.5 hour reading from one subject was used as test data, with the remaining 14 subjects being used in training and validation.

To ensure exposure to a wide range of HR levels across the 2.5-hour activity sequence, the model was trained on the first 80% of each activity and validated on the remaining 20%. An 8-second (one-window) embargo was applied between training and validation segments to prevent data leakage.

MODEL DETAILS

Each input to the model is an 8-second window of wearable sensor data containing 512 samples per signal (64 Hz). Every window includes six features: Blood Volume Pulse (BVP), three-axis Accelerometer (ACCx, ACCy, ACCz), Skin Temperature (TEMP), and Electrodermal Activity (EDA). These inputs capture both physiological and motion-related patterns that influence heart rate behaviour.

The model uses a Gated Recurrent Unit (GRU) encoder to process each window sequentially. The GRU produces a hidden state at every sample, and the final hidden state acts as a condensed representation of the entire 8-second window - summarising its temporal dynamics and physiological context. This shared representation is passed to 2 independent feed-forward networks:

- Calibration Head - predicts the ECG-based heart rate for the current window.
- Forecasting Head - predicts the heart rate 60 seconds into the future, corresponding to 30 windows ahead (2 s stride × 30).

Loss calculated for calibration:

- Mean Absolute Error (average difference between predicted HR and true HR)

Loss calculated for forecasting:

- Mean Absolute Error

EVALUATION

Below are the performance metrics used to calculate the efficacy of the model during validation and testing.

Calibration	Forecasting
Mean Absolute Error	Mean Absolute Error

RESULTS

Below are the results from testing. All metrics were averaged across 15 folds to ensure robust cross-subject generalization.

Calibration	Forecasting	Both
now_MAE: 15.1138 - std 6.3127	fut_MAE: 16.1377 - std 6.2255	test_MAE: 31.2515 - std 12.5022