Signal Quality Assessment and Reconstruction of PPG-Derived Signals for Heart Rate and Variability Estimation in In-Vehicle Applications: A Comparative Review and Empirical Validation

Ruimin Gao, Carl S. Miller, Brian T. W. Lin, Chris W. Schwarz, and Monica L. H. Jones

Published in Sensors in Dec 2025

Presented by: Abdullah Mamun

Date: 4 March 2026

Email: a.mamun@asu.edu

Paper

abdullah-mamun.com

𝕏: @AB9Mamun

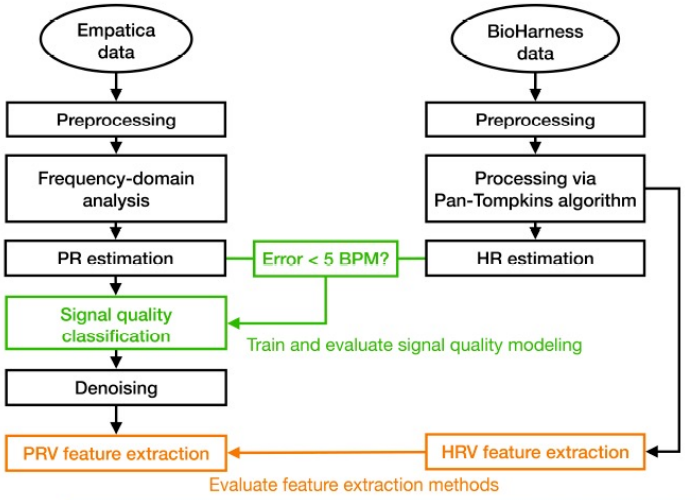


Figure 2. Data processing and evaluation flowchart. Empatica BVP-derived pulse rate (PR) estimates are compared with BioHarness ECG-derived heart rate (HR) estimates, with epochs showing <5 BPM error used to train the signal-quality classifier (green). PRV and HRV features (orange) are then extracted for method evaluation.

Motivation

Why does it matter?

- HR and HRV are important indicators of stress, fatigue, and workload
- ECG is the gold standard but requires electrodes and chest straps
- PPG is widely used in wearable devices due to:
- Comfort
- Low cost
- Easy integration
- However, PPG signals are highly sensitive to motion artifacts

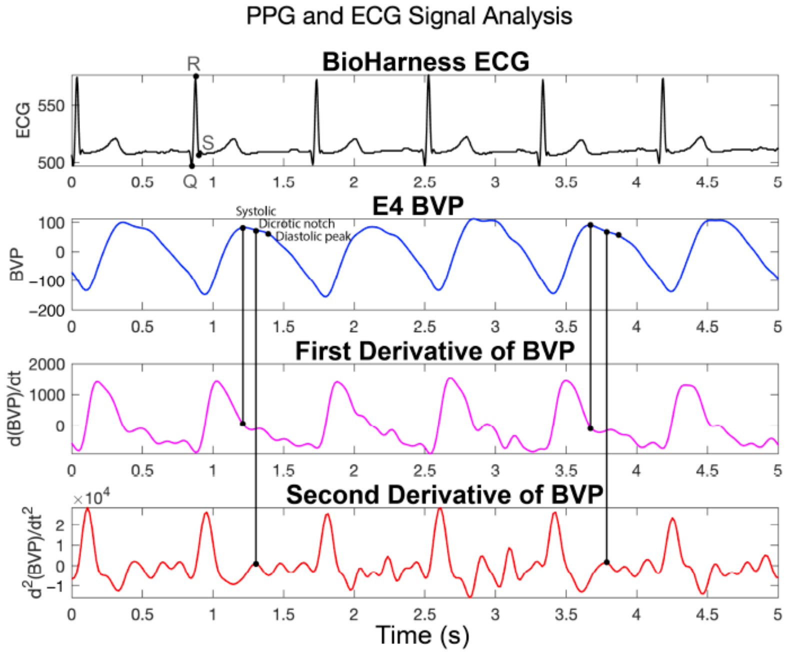


Figure 1. Representative ECG and PPG-derived signals from a single participant over a 5-s window. Panels show (top to bottom): BioHarness ECG waveform with annotated fiducial points, E4 BVP signal with systolic and diastolic peaks, the first derivative of the BVP (dBVP/dt), and the second derivative of the BVP (d^2BVP/dt^2).

Research Gap

- Motion artifacts distort PPG waveforms
- Lack of standardized signal processing protocols
- Most datasets collected under **resting conditions**
- HRV estimation from PPG remains unreliable in real-world environments
- Need for **robust signal quality assessment and reconstruction pipeline**

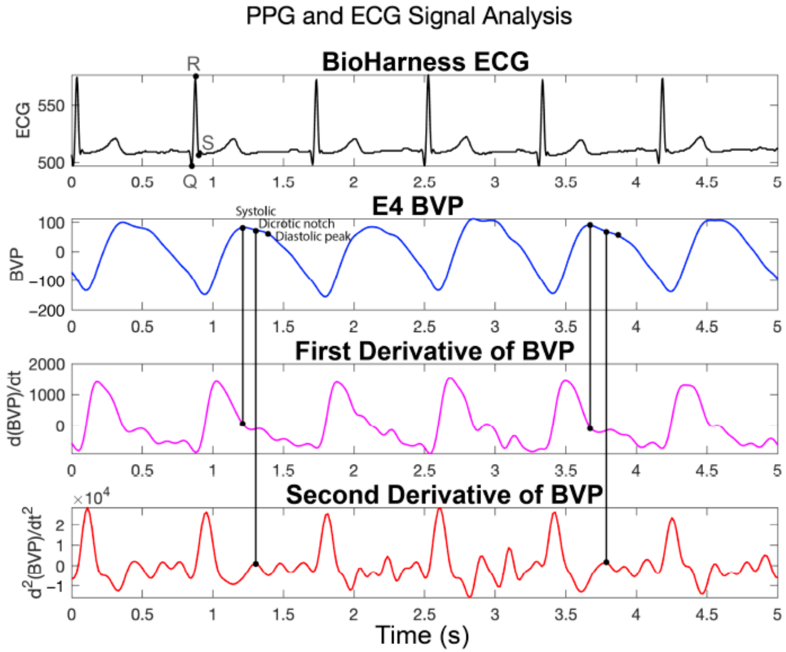


Figure 1. Representative ECG and PPG-derived signals from a single participant over a 5-s window. Panels show (top to bottom): BioHarness ECG waveform with annotated fiducial points, E4 BVP signal with systolic and diastolic peaks, the first derivative of the BVP (dBVP/dt), and the second derivative of the BVP (d^2BVP/dt^2).

Objectives of the paper

- Evaluate accuracy of **PPG-derived HR and HRV** in dynamic in-vehicle settings
- Develop a **signal quality assessment (SQA) model**
- Apply **spectral signal reconstruction**
- Compare different processing pipelines
- Identify which HRV metrics are reliable from PPG

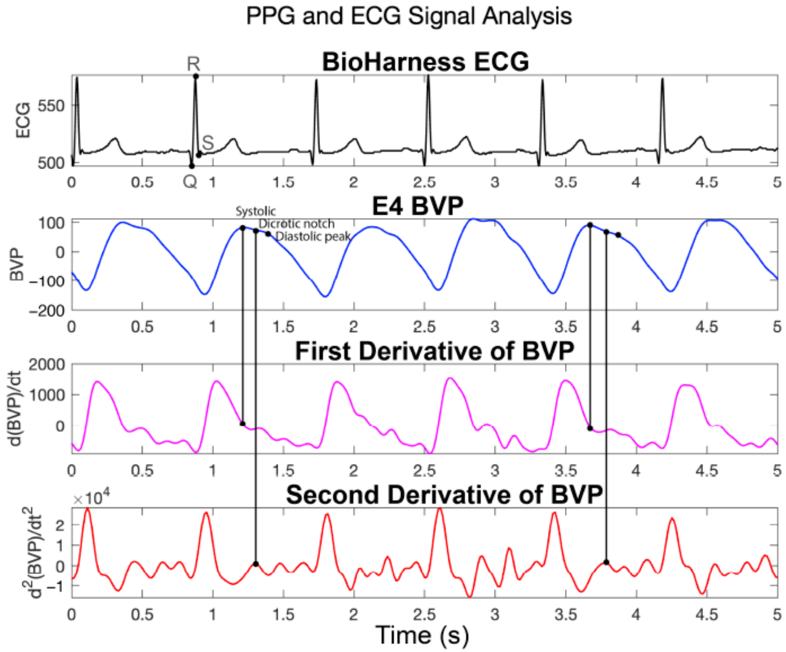


Figure 1. Representative ECG and PPG-derived signals from a single participant over a 5-s window. Panels show (top to bottom): BioHarness ECG waveform with annotated fiducial points, E4 BVP signal with systolic and diastolic peaks, the first derivative of the BVP (dBVP/dt), and the second derivative of the BVP (d^2BVP/dt^2).

Experiment: Dataset

- **114 trials (110 valid trials)**
- Participants seated in **automated vehicle**
- Exposure to **acceleration–braking events**
- ~20 minutes per participant

Sensors

- ECG: Zephyr BioHarness (250 Hz)
- PPG: Empatica E4 wristband (64 Hz)
- Accelerometer for motion detection

Each participant wore two physiological monitoring devices to enable simultaneous ECG and PPG data collection. ECG signals were acquired using the Zephyr™ Bioharness™ 3 (Medtronic, Minneapolis, MN, USA) device, a wearable chest strap positioned immediately below the pectoral muscle, with the Biomodule aligned under the left arm. The BioHarness device employs a single-lead ECG configuration using conductive chest-strap electrodes and records R-R intervals at 250 Hz. Strap tension was standardized across participants using the integrated tension indicator loop to ensure consistent electrode contact. PPG signals, recorded as BVP, were collected using the Empatica E4 wristband (Empatica, Inc., Cambridge, MA, USA) worn on the non-dominant wrist just proximal to the wrist joint. The PPG sensor, embedded in the underside of the wristband, captured vascular signals at 64 Hz by measuring variations in reflected light, as blood volume changed with each heartbeat. In addition to PPG, the Empatica E4 recorded three-axis accelerometer data at 32 Hz, providing contextual information about movement during the in-vehicle protocol (see: https://www.empatica.com/en-int/store/e4-wristband/, accessed on 4 September 2025). It should be noted that, as of February 2025, the Empatica E4 and its associated software suite were formally discontinued (see: https://www.empatica.com/research/e4-sunset/, accessed on 4 September 2025). This discontinuation does not affect the integrity of data collected prior to the sunset period.

Experiment: Proposed Processing Pipeline

<u>Steps:</u>

1. Signal preprocessing
2. Frequency-domain analysis
3. Pulse rate estimation
4. Signal quality classification
5. Signal reconstruction
6. HRV feature extraction

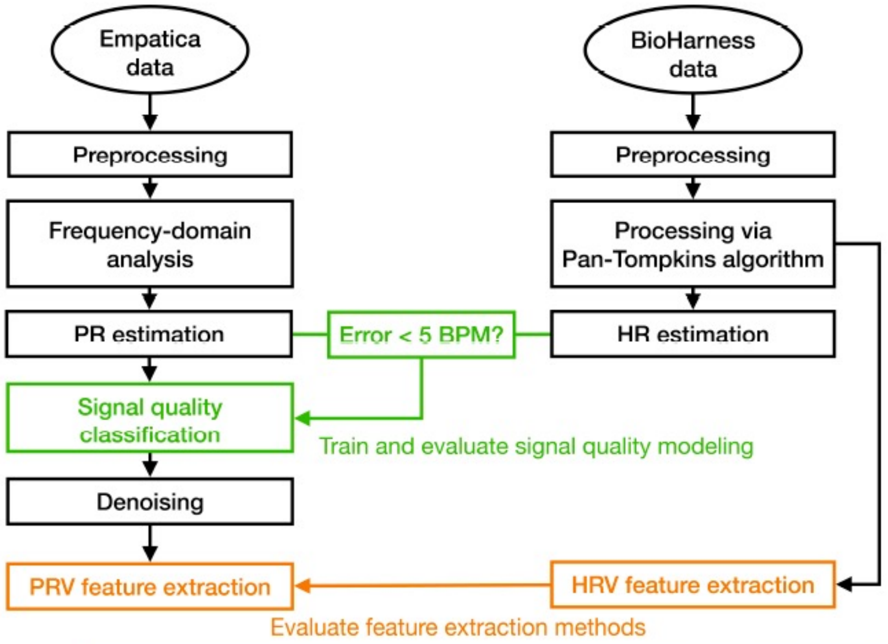


Figure 2. Data processing and evaluation flowchart. Empatica BVP-derived pulse rate (PR) estimates are compared with BioHarness ECG-derived heart rate (HR) estimates, with epochs showing <5 BPM error used to train the signal-quality classifier (green). PRV and HRV features (orange) are then extracted for method evaluation.

Signal Quality Assessment Model

SQA classifier built using **12 signal features**

Feature categories:

Time-domain features

- Kurtosis
- Skewness
- Shannon entropy

Frequency-domain features

- Spectral peak power
- Harmonic power

Motion features

- Accelerometer amplitude
- Cross-bicoherence

Table 3. Feature extraction of PPG-derived and accelerometer sensor measures for the SQA classification model.

Source	Feature	Description	References
BVP time domain	Kurtosis	Scaled version of the fourth moment of the PPG distribution, representing the tailedness of the PPG signal distribution. Measures peakedness of the waveform; high kurtosis may indicate clean, sharp pulse waves	[30]
	Skewness	Measure of the asymmetry of the PPG signal around zero. Describes waveform asymmetry; deviations from symmetry may signal noise or distortion	[30]
	Shannon entropy	Measure of the disorder in the PPG signal probability distribution. Quantifies signal complexity; higher entropy may suggest irregularity due to noise	[30]
BVP frequency domain	Spectral kurtosis	Scaled version of the fourth moment of the PPG spectral distribution, representing the tailedness of the PPG frequency-domain signal. Detects spectral sparsity; flatter spectra may indicate noise or artifact	[30]
	Relative power of dominant peak	Power ratio of the dominant peak in the PPG spectrum compared to the total power. Power of the peak frequency in the heart rate band; used to confirm signal periodicity	[67]
	Relative power of harmonics	Power ratio of the 2nd and 3rd harmonics of the PPG spectral dominant peak compared to the total power. Power in harmonic components; supports waveform integrity checks	[67]
	HRF deviation from moving median	Absolute difference between the spectral peak of the current PPG epoch and the median spectral peak of the nearest 1 min segment. Measures abrupt change in pulse frequency; used to detect transient noise	
	Bispectral self-coupling	Number of self-coupling events among the three most prominent peaks (f0, f1, f2) in the diagonal slice of the bispectrum. Assesses cross-frequency coupling around HRF; reduced coupling may signal distortion	[68]
Accelerometer time domain	Amplitude mean	Average magnitude of the accelerometer data. Indicates overall movement intensity; elevated values may suggest a potential motion artifact	[38]
	Amplitude SD	Standard deviation of the accelerometer magnitude. Captures motion variability; high standard deviation often correlates with motion-induced noise	[38]
Accelerometer frequency domain	Maximal cross-bicoherence to PPG	Maximum bicoherence between the PPG signal and the accelerometer data from the x-, y-, or z-axis. Measures motion energy overlapping the HR band; used to detect confounding artifact sources	[66]
	Relative power of heart rate frequency band	Relative power of the [2/3, 10/3] Hz band, corresponding to the heart-rate frequency band ranging from 40 BPM to 200 BPM. Estimates nonlinear coupling between motion and the pulse signal; high values imply motion contamination	Inspired by [69]

Signal Quality Assessment Method

Key idea:

- Remove motion artifacts around HR frequency

Method:

- Identify dominant heart rate frequency
- Apply **notch filters around HR and harmonics**
- Subtract noise components from raw signal

Advantages:

- Works with **single-channel PPG**
- Computationally lightweight

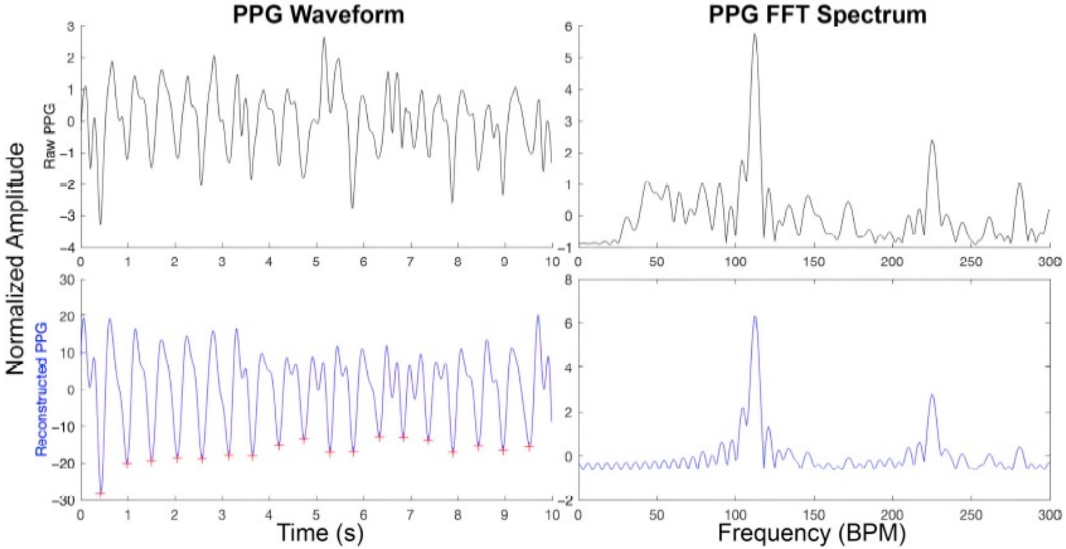


Figure 4. An illustration of PPG signal reconstruction using spectral peak information. The first row shows the original PPG waveform in the time-domain (left) and frequency-domain (right). The second row presents the reconstructed signal, with only information corresponding to the heart rate frequency and its harmonics retained. Red plus signs indicate the detected systolic peaks (beat locations) in the reconstructed waveform.

Results: SQA Model Performance

Best configuration:

- Epoch length: **10 seconds**
- Classifier: **SVM**
- Selected features: **8**

Performance on test data:

- Accuracy: **0.888**
- Specificity: **0.892**
- Sensitivity: **0.887**
- AUC: **0.96**

Table 4. Cross-validation performance of signal quality assessment (SQ) classification models. Bolded entries indicate the best-performing classifier, as determined by the highest accuracy and balanced sensitivity/specificity.

Epoch (s)	Classifier	Optimal Feature N	Optimal Cost	Accuracy	Specificity	Sensitivity	Min of Specificity and Sensitivity
5	Logistic	13	5	0.874	0.885	0.871	0.871
	SVM	10	5	0.868	0.888	0.862	0.862
	NB	9	9	0.872	0.874	0.872	0.872
	LDA	9	10	0.877	0.876	0.878	0.876
10	Logistic	4	3	0.890	0.878	0.892	0.878
	SVM	**8**	**3**	**0.885**	**0.884**	**0.885**	**0.884**
	NB	6	2	0.870	0.873	0.869	0.869
	LDA	4	7	0.890	0.881	0.892	0.881
20	Logistic	12	2	0.810	0.811	0.810	0.810
	SVM	7	2	0.799	0.824	0.789	0.789
	NB	4	2	0.805	0.804	0.805	0.804
	LDA	9	3	0.798	0.844	0.780	0.780
30	Logistic	1	1	0.750	0.699	0.780	0.699
	SVM	1	1	0.745	0.739	0.749	0.739
	NB	1	1	0.721	0.715	0.725	0.715
	LDA	1	1	0.747	0.715	0.765	0.715

14 of 23

0.892, a sensitivity of 0.887, and an AUC of 0.960. These results demonstrate the model's robustness and generalizability in distinguishing usable versus unusable PPG epochs for subsequent analysis.

Results: SQA Model Performance

Best configuration:

- Epoch length: **10 seconds**
- Classifier: **SVM**
- Selected features: **8**

Performance on test data:

- Accuracy: **0.888**
- Specificity: **0.892**
- Sensitivity: **0.887**
- AUC: **0.96**

3.1. Signal Quality Assessment Classification Model: Performance

A detailed summary of the cross-validation results for the SQA classification models is presented in Table 4. Among the evaluated models, the classifier trained on 10 s epochs outperformed those trained on shorter (5 s) or longer (20 s and 30 s) time windows. This model achieved the highest balance between accuracy, specificity, and sensitivity. The best performance was observed using an SVM classifier with eight features selected via mRMR and a misclassification cost ratio of 3, resulting in a minimum specificity and sensitivity of 0.884 during five-fold cross-validation. This configuration was selected as the final model and subsequently evaluated on the 20% participant-level holdout test set. On this unseen data, the model achieved an overall accuracy of 0.888, a specificity of

0.892, a sensitivity of 0.887, and an AUC of 0.960. These results demonstrate the model's robustness and generalizability in distinguishing usable versus unusable PPG epochs for subsequent analysis.

Results: SQA Model Performance

Best configuration:

- Epoch length: **10 seconds**
- Classifier: **SVM**
- Selected features: **8**

Performance on test data:

- Accuracy: **0.888**
- Specificity: **0.892**
- Sensitivity: **0.887**
- AUC: **0.96**

3.1. Signal Quality Assessment Classification Model: Performance

A detailed summary of the cross-validation results for the SQA classification models is presented in Table 4. Among the evaluated models, the classifier trained on 10 s epochs outperformed those trained on shorter (5 s) or longer (20 s and 30 s) time windows. This model achieved the highest balance between accuracy, specificity, and sensitivity. The best performance was observed using an SVM classifier with eight features selected via mRMR and a misclassification cost ratio of 3, resulting in a minimum specificity and sensitivity of 0.884 during five-fold cross-validation. This configuration was selected as the final model and subsequently evaluated on the 20% participant-level holdout test set. On this unseen data, the model achieved an overall accuracy of 0.888, a specificity of

0.892, a sensitivity of 0.887, and an AUC of 0.960. These results demonstrate the model's robustness and generalizability in distinguishing usable versus unusable PPG epochs for subsequent analysis.

Results: Correlation

Heart Rate

- Strong agreement between PPG and ECG
- Accurate even in dynamic environments

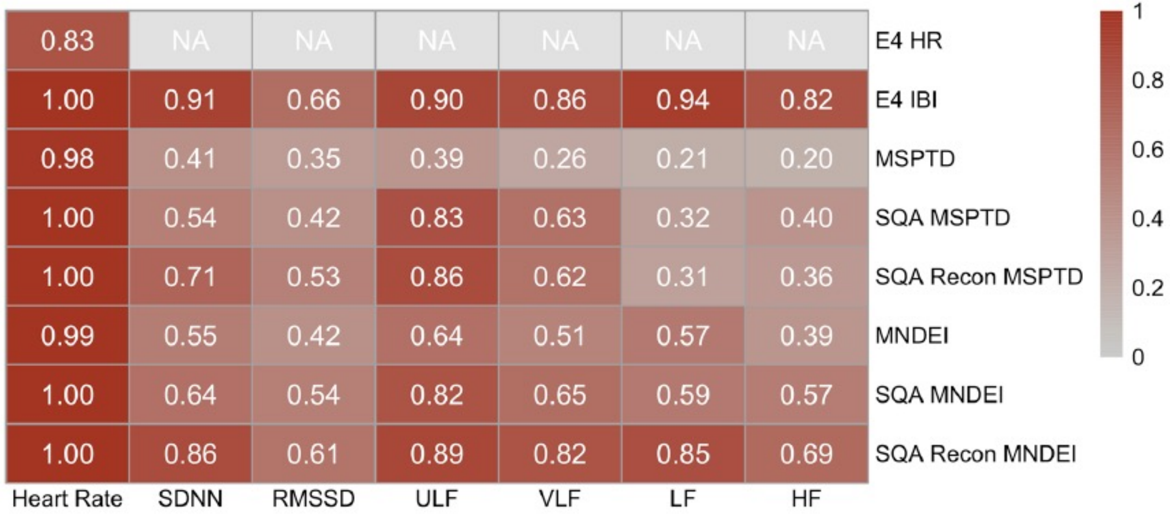
HRV Metrics

Moderate agreement for:

- SDNN
- LF
- VLF

Poor agreement for:

- RMSSD
- HF



	Heart Rate	SDNN	RMSSD	ULF	VLF	LF	HF
E4 HR	0.83	NA	NA	NA	NA	NA	NA
E4 IBI	1.00	0.91	0.66	0.90	0.86	0.94	0.82
MSPTD	0.98	0.41	0.35	0.39	0.26	0.21	0.20
SQA MSPTD	1.00	0.54	0.42	0.83	0.63	0.32	0.40
SQA Recon MSPTD	1.00	0.71	0.53	0.86	0.62	0.31	0.36
MNDEI	0.99	0.55	0.42	0.64	0.51	0.57	0.39
SQA MNDEI	1.00	0.64	0.54	0.82	0.65	0.59	0.57
SQA Recon MNDEI	1.00	0.86	0.61	0.89	0.82	0.85	0.69

Figure 6. Pearson correlation between ECG-derived HRV and PPG-derived PRV metrics. The heatmap matrix of Pearson correlation coefficients across multiple methods. Cells labeled "NA" indicate HRV metrics that are not available as outputs from Empatica E4. All Pearson correlation tests have a *p*-value of < 0.001.

Conclusion

Key Contributions

1. Proposed **interpretable single-channel PPG processing pipeline**
2. Developed **signal quality assessment classifier**
3. Introduced **spectral reconstruction for motion artifact removal**
4. Demonstrated **improved HRV estimation in dynamic environments**
5. Identified **limitations of PPG for high-frequency HRV metrics**

Conclusion

- PPG is reliable for **heart rate monitoring**
- HRV estimation from wrist PPG remains **limited under motion**

Thank You!

Email: a.mamun@asu.edu

Paper

abdullah-mamun.com

𝕏: @AB9Mamun

Backup – Reconstructing PPG.

Estimate heart rate frequency (HRF):
Apply FFT to the PPG signal and identify the dominant spectral peak corresponding to heart rate.

Identify harmonics:
Detect the **second harmonic** of the heart rate frequency to capture additional pulse information.

Apply notch filters:
Use **4th-order Butterworth band-stop filters** centered at the **HR frequency and its harmonic (±20%)** to isolate motion-related noise.

Extract noise component:
The filtered signal represents **noise near the heart-rate frequency band**.

Reconstruct clean signal:
Subtract the noise component from the raw PPG signal to obtain a cleaner waveform.

Result:
The reconstructed signal improves **pulse peak detection and HRV feature extraction** under motion conditions.