Feature Stability of PPG Morphological Descriptors Across Diverse Clinical Acquisition Settings

Matteo Ricci, Lorenzo Fácila, Raúl Alcaraz,
José J Rieta, Frida Sandberg

BioMIT.org, Electronic Engineering Department, Universitat Politecnica de Valencia, Valencia, Spain.

Aims: Photoplethysmography (PPG) offers a promising non-invasive approach for blood pressure (BP) estimation, however machine learning models often struggle to generalize across datasets due to feature variability. Our study evaluates the consistency of PPG morphological features across four public datasets to identify robust features that may support generalizable BP models.

Methods: Analyses were conducted on representative subsets from four cohorts: PPG-BP (219 subjects, ambulatory), VitalDB (285 subjects, intraoperative), MIMIC-III (225 subjects, ICU), and MIMIC-IV (63 subjects, ICU). Extracted segments per dataset (30 s for VitalDB/MIMIC; 2.1 s for PPG-BP) were bandpass-filtered (0.5–12 Hz) and validated against three physiological criteria: heart rate plausibility (40–180 bpm), spectral power ratio (band-to-total power > 0.4), and rhythmic stability (inter-beat interval variability below a preset threshold). 41 morphological features were extracted for each pulse, including timing parameters, amplitudes, second derivative indices (SDPTG a-e), and descriptors from Pulse Decomposition Analysis (PDA). For each feature, stability was quantified as the standard deviation within a segment divided by the standard deviation of that feature over the entire dataset.

Results: Features were considered unstable if their median instability ratio across segments exceeded 0.3. Based on this threshold, the number of stable features varied significantly across datasets: 23/41 in PPG-BP, 27/41 in VitalDB, 21/41 in MIMIC-III, and 33/41 in MIMIC-IV. The most consistent features across datasets were fiducial points time interval (Δn_{up}), PDA peaks interval (ΔT_{12}), and b/a derivatives ratio, showing low instability ratios (mean [min – max]) of 0.15 $[0.13–0.18]$, 0.17 $[0.14–0.19]$, and 0.22 $[0.20–0.25]$, respectively. A_{15} (systolic-to-distal PDA waves amplitude ratio) exhibited the highest instability in PPG-BP (0.69), whereas it remained much lower in MIMIC-IV (0.24). Temporal distances from the dicrotic notch (Δn_{gf}, Δn_{pf}) showed minimum instability in PPG-BP (0.24) and peaked in MIMIC-III (0.69).

Conclusions: Amplitude ratios and dicrotic notch metrics remain dataset-dependent, whereas timing and SDPTG descriptors exhibit strong cross-dataset robustness. Prioritizing stable features can support more generalizable BP models.