Cross-Dataset Analysis of PPG Feature Correlations with Blood Pressure Targets

Matteo Ricci, Arturo Martínez-Rodrigo,
Raúl Alcaraz, Frida Sandberg, José J Rieta

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

Aims: Photoplethysmography (PPG) is a low-cost, non-invasive technology studied for blood pressure (BP) estimation. However, the consistency of its morphological features across diverse cohorts is unclear. This study investigates the stability of correlations between PPG features and systolic and diastolic blood pressure (SBP/DBP) across four heterogeneous clinical datasets.

Methods: Data comprise representative subsets from four public sources: 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 paired with synchronized BP targets. PPG signals were bandpass-filtered (0.5–12 Hz) and screened for physiological plausibility. We extracted 41 features per pulse, including timing, amplitude, second derivative ratios (SDPTG, a–e), and Pulse Decomposition Analysis (PDA) descriptors. Segments were discarded if the median per-feature variability ratio (intra-segment to dataset-wide standard deviation) exceeded 0.3. Spearman correlation (ρ) between features and SBP/DBP was computed and ranked by mean $|\rho|$ across the four datasets (indicating stable correlation) and min-max range.

Results: After applying the variability criterion, the algorithm retained 89.3% of the segments for PPG-BP, 55.4% for VitalDB, 62.5% for MIMIC-III, and 70.9% for MIMIC-IV. Twenty-two features showed mean $|\rho| > 0.3$ with SBP across all four datasets, and 15 with DBP. Stable SBP predictors include SDPTG ratios d/a (mean $|\rho| = 0.71$, [0.58–0.81]), b/a (mean $|\rho| = 0.59$, [0.53–0.65]), and $(b-c-d-e)/a$ (mean $|\rho| = 0.47$, [0.38–0.54]), as well as the wave reflection index AI_{gh} (mean $|\rho| = 0.55$, [0.42–0.68]) and amplitude ratio Y_{gh} (mean $|\rho| = 0.53$, [0.42–0.61]). For DBP, temporal distances from the dicrotic notch (Δn_{pf}, Δn_{gh}; mean $|\rho|$: 0.46–0.49) exhibited comparable stability. Systolic upstroke time (Δn_{up}) correlated strongly in PPG-BP and VitalDB ($|\rho| = 0.63$) but degraded on MIMIC-III ($|\rho| = 0.18$). DBP predictor pulse slope (S_{pf}) declined from $|\rho| = 0.53$ (VitalDB) to 0.12 (MIMIC-III).

Conclusions: Stable and highly correlated features provide a solid foundation for generalizable BP estimation; however, cohort-specific variability continues to challenge universal model applicability.