Noninvasive Glucose Prediction and Diabetes Classification With ECG and PPG Derived Domain-Specific Features*†

Su-Jin Park[1], Se-Yeon Lee[2], Gyu-Hyeok Lee[2], and Gun-Woo Kim[2‡]

[1] Department of Management Information Systems
[2] Department of Computer Science and Engineering
Gyeongsang National University, Jinju, Republic of Korea
{park_11412, 2023014356, gyuhyeok, gunwoo.kim}@gnu.ac.kr

Abstract

This study systematically extracts domain-specific features physiology-based, morphological, temporal, frequency or wavelet domain, and contextual from PPG and ECG signals. We quantitatively evaluate performance on noninvasive continuous glucose (CGM) regression and WHO-based three-class classification using subject-wise GroupKFold cross-validation. The proposed pipeline controls data leakage, inter-subject variability, and confounding factors, and applies a per-fold weighted ensemble of the top two models. Consequently, PPG achieved lower MAE/RMSE in regression, whereas ECG (HR/HRV) outperformed PPG in classification, yielding higher Macro-F1 and Balanced Accuracy.

Keywords: Noninvasive glucose monitoring · PPG · ECG · Domain-specific features

1 Introduction

Noninvasive blood glucose prediction is vulnerable to model instability due to numerous confounders, including autonomic responses, physical activity and meals, and ambient temperature. Photoplethysmography (PPG) [**?**] is sensitive to peripheral hemodynamics and arterial stiffness, whereas the electrocardiogram (ECG) [**?**] provides electrophysiology-derived heart rate and heart rate variability (HR/HRV) indices that reflect autonomic function. In this study, we systematically extract domain-specific features from both signals and compare their performance using subject-wise cross-validation in continuous glucose regression and WHO-criteria three-class glycemic-status classification. We further discuss which modality is better suited to each task.

2 Proposal Method

2.1 Data Processing and Feature Extraction

Raw ECG (250 Hz) and PPG (64 Hz) signals are summarized into 5-minute windows and aligned to the 5-minute grid of the continuous glucose monitoring (CGM) data. Data quality is controlled via R–R interval validity checks, mean heart-rate range filtering, and CGM validity-range filtering. To prevent data leakage, we remove columns that are trivially correlated with

*This research was supported by the Regional Innovation System Education(RISE) program through the RISE Center, Gyeongsangnam-do, funded by the Ministry of Education(MOE) and the Gyeongsangnam-do Provincial Government, Republic of Korea.(2025-RISE-16-001)

†Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-38, December 16-18, 2025, Sapporo, Japan.

‡Corresponding author

time or with the target. From the preprocessed signals, we extract features across four domains as follows. All segment-level features are then aggregated into 5-minute averages, standard deviations, and validity ratios for use as model inputs.

1. **Time domain**: mean, standard deviation, RMS, skewness, kurtosis, etc.
2. **Morphology**: QRS duration, R-wave amplitude, ST-segment slope, AI (augmentation index), IPA (inflection point area), amplitude/slope ratios, etc.
3. **HR/PI/HRV**: autonomic indices such as SDNN, RMSSD, pNN50, LF/HF ratio, etc.
4. **Wavelet/Frequency**: DWT/WPD sub-band energy and entropy, FFT bandpower.

2.2 Learning and Evaluation Framework

This study addresses two tasks: $\pm5/\pm10/\pm15$-minute horizon prediction for continuous glucose monitoring (CGM) and WHO-based three-class classification (hypoglycemia/normal/hyperglycemia). Data were split with subject-wise GroupKFold (k=5), and preprocessing, training, and tuning were performed strictly within each fold. For regression, we considered ElasticNet, HistGradientBoostingRegressor, and RandomForestRegressor; for classification, LogisticRegression, HistGradientBoostingClassifier, and a multilayer perceptron (MLP). Per fold, the top two models by validation performance were combined via a weighted ensemble to derive the final out-of-fold (OOF) estimates. Evaluation metrics were MAE/RMSE for regression and Macro-F1/Balanced Accuracy for classification.

3 Experiment and Results

Experimental results showed that PPG consistently achieved lower errors in continuous blood glucose regression, with an out-of-fold (OOF) MAE of 16.44 ± 2.80 mg/dL and RMSE of 22.55 ± 3.63 mg/dL. The MAE remained nearly unchanged at 16.4 mg/dL across $\pm5/\pm10/\pm15$-minute horizons, indicating short-horizon stability. In the WHO three-class classification, ECG outperformed PPG, achieving a Macro-F1 of 0.91 ± 0.06 and a Balanced Accuracy of 0.96 ± 0.03, compared with PPG's 0.33 ± 0.00 and 0.37 ± 0.07, respectively. Based on these findings, future work will investigate synchronized multi-modal late fusion and refined context modeling (meals, exercise, insulin) to further improve performance.

Table 1: Subject-wise CV Regression (Ensemble OOF).

Modality	Horizon (min)	MAE (mg/dL)	RMSE (mg/dL)
ECG	5	59.22 ± 15.12	75.99 ± 21.58
ECG	10	59.29 ± 15.14	76.04 ± 21.68
ECG	15	59.66 ± 15.84	76.20 ± 22.23
PPG	5	**16.44** ± 2.66	**22.56** ± 3.50
PPG	10	**16.46** ± 2.56	**22.56** ± 3.46
PPG	15	**16.44** ± 2.80	**22.55** ± 3.63

Table 2: WHO 3-class Classification (Subject-wise CV).

Modality	Macro-F1	Balanced Acc.
ECG	**0.91** ± 0.06	**0.96** ± 0.03
PPG	0.33 ± 0.00	0.37 ± 0.07