Photoplethysmogram-Based Diabetes Screening via Supervised Machine Learning: A Demographic Study on a Southeast Asian Cohort

Nazrul Anuar Nayan[1]*, Mohd Taufik Rezza Mohd Foudzi[2],
Mohd Zubir Suboh[3], Syaza Norfilsha Ishak[4], Zazilah May[5]
Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, UKM Bangi, Malaysia[1, 2, 4]
British Malaysia Institute, Universiti Kuala Lumpur, Gombak, Malaysia[3]
Department of Applied Science, Universiti Teknologi PETRONAS, Perak, Malaysia[5]

***Abstract*—Diabetes mellitus is a major chronic metabolic disorder that often leads to serious long-term vascular complications. Traditional monitoring methods focus mainly on metabolic indicators and often miss early vascular changes. This study developed and validated a non-invasive framework for classifying diabetic status based on photoplethysmogram (PPG) pulse morphology. The approach offers a scalable and affordable alternative to invasive blood tests. A dataset from 78 Malaysian participants was analyzed in five phases: signal pre-processing, feature extraction, and statistical ranking. Raw signals were filtered with a 4th-order Chebyshev Type II band-pass filter for accurate waveform analysis. From a wide set of temporal and amplitude features, key biomarkers linked to arterial stiffness and vascular compliance were identified and ranked. Six supervised machine learning models were evaluated: Logistic Regression, Decision Tree (DT), KNN, Support Vector Machine (SVM), Artificial Neural Network (ANN), and Naïve Bayes (NB). ANN and SVM models achieved the highest classification accuracy and AUC. This demonstrates effective distinction between diabetic and non-diabetic status using interpretable waveform features. Validation with a Southeast Asian cohort addresses a demographic gap in the literature. The framework shows that ranked PPG biomarkers can be used for accessible, community-level diabetes screening, especially in healthcare settings with limited resources.**

Keywords*—Photoplethysmography (PPG); diabetes prediction; supervised machine learning; signal morphology features; non-invasive screening; feature selection

I. Introduction

Diabetes mellitus is a chronic metabolic disorder characterized by elevated blood glucose levels resulting from insufficient insulin production or impaired insulin action [1]. The global burden of diabetes continues to rise, with an estimated 589 million adults affected in 2024, more than 322 million of whom reside in the Western Pacific and South-East Asia regions. This number is projected to reach 853 million by 2050, according to the International Diabetes Federation (IDF) [2]. Uncontrolled diabetes can lead to severe complications, including cardiovascular disease, kidney failure, and neuropathy, underscoring the need for early detection and continuous monitoring [3]. Current diagnostic strategies, however, have significant limitations. Standard methods are invasive and, despite the widespread use of glucose meters, provide limited insight into the underlying vascular alterations associated with diabetes. By focusing primarily on metabolic control, traditional monitoring approaches tend to overlook concomitant cardiovascular abnormalities, which restricts their ability to detect early complications. Unlike fasting glucose, oral glucose tolerance, and HbA1c tests, which require blood samples, PPG-based monitoring is more comfortable, scalable, and can be continuous for real-time diabetes screening [4]–[6].

Despite the promise of machine learning in diabetes prediction, significant research gaps remain. First, previous studies rely heavily on continuous glucose monitoring (CGM) or complex clinical datasets, which still require invasive or costly acquisition procedures. CGM refers to wearable devices that continuously measure glucose levels in the interstitial fluid beneath the skin. When PPG has been utilized, research has predominantly favored data-intensive, 'black-box' deep learning algorithms (such as convolutional neural networks, or CNNs), which are complex models whose decision processes are difficult to interpret. These models fail to clearly link predictions to underlying physiological changes. Crucially, the datasets driving these models are almost exclusively derived from Western or South Asian participants, leaving a critical demographic research gap for Southeast Asian populations. Combining interpretable PPG waveform morphology with supervised machine learning provides a meaningful improvement over existing methods; it captures the early vascular alterations and arterial stiffness associated with diabetes that standard metabolic tests often miss, offering a non-invasive, scalable, and explainable screening alternative tailored to a Southeast Asian demographic.

This study proposes a physiology-driven approach to diabetes classification using detailed morphological features extracted from PPG waveforms, directly addressing the research and demographic gaps identified above. In contrast with previous methods reliant on broader clinical datasets or limited PPG metrics, this work introduces nine statistically ranked biomarkers obtained from fiducial-point analysis. Employing a five-phase process—data collection, signal pre-processing, feature extraction, statistical ranking, and supervised machine learning classification (Logistic Regression, DT, SVM, ANN, KNN, NB)—the study demonstrates the practical utility of ranked PPG-based features in distinguishing diabetic from non-

*Corresponding author.

diabetic individuals. ANN and SVM models show superior performance. The framework thus provides an explainable, scalable, and resource-efficient pathway for community-level diabetes screening in Southeast Asia.

The main contributions of this study are threefold. First, it addresses a critical demographic gap by utilizing a uniquely collected dataset from a Malaysian population, thereby providing essential Southeast Asian representation in PPG-based diabetes research. Second, the study prioritizes physiological explainability by proposing a transparent, physiology-driven approach based on precise fiducial-point analysis, which serves as an interpretable alternative to opaque deep learning algorithms. Finally, it contributes to biomarker identification by systematically extracting, statistically ranking, and reducing a comprehensive set of PPG morphological features to pinpoint the core PPG biomarkers most indicative of diabetic vascular changes.

II. Literature Review

Recent studies have demonstrated the effectiveness of machine learning (ML) models for diabetes prediction. Hossain et al. [7] employed convolutional neural networks (CNNs) for glucose level estimation, while Gupta et al. [8] used Random Forest (RF) models with PPG signals, reporting R^2 values of up to 0.91. Zanelli et al. [9] applied CNNs combined with demographic data, achieving an area under the curve (AUC) of 75.5%. Conventional models such as KNN [10], Naïve Bayes (NB) [11], and Logistic Regression (LR) [12] have also shown strong predictive performance. Similarly, the works in [13] and [14] investigated Support Vector Machine (SVM) with radial basis function (RBF) kernels and Boosting Decision Trees (BDT), respectively, reporting accuracies above 92%. Although these approaches are promising, they primarily rely on continuous glucose monitoring (CGM) or clinical datasets, with relatively limited investigation of PPG waveform morphology for diabetes detection. Ensemble methods such as AdaBoost [15] and Random Forest [8] have also been widely explored, but often in settings that require large datasets and complex feature sets that are difficult to obtain in low-resource environments.

TABLE I. Summary of Literature Review on Diabetes Prediction Using Machine Learning.

Ref.	Year	ML Models Used	Data Type / Features	Best Result(s)
[7]	2019	CNN	Biomedical signals	95% within ±1.1 mmol/L
[8]	2020	RF	Transmissive PPG	$R^2 = 0.91$
[9]	2023	CNN	PPG + age + sex	AUC = 75.5%
[10]	2020	KNN	Biomedical features	Good accuracy for small dataset
[12]	2020	LR	Clinical variables	Effective baseline model
[11]	2023	KNN, NB	Clinical	NB > KNN in accuracy
[16]	2018	SVM, ANN, NB	UCI Pima Indians dataset	SVM & ANN high; NB > KNN
[13]	2023	SVM (RBF), BDT	CGM	SVM = 98.25%, BDT = 92.58%

Table I shows that most studies using CNN, Random Forest, and Logistic Regression for diabetes prediction rely on clinical or glucose monitoring data. These need invasive or complex testing and are mainly from non-Asian groups.

Most existing studies use datasets from Western or South Asian participants, creating a research gap in Southeast Asia. This study addresses that gap by using data from a Malaysian population, allowing the framework to reflect regional physiological and demographic factors. Deep learning models like CNNs are data-intensive and difficult to interpret, making them less suitable for small or resource-limited studies. The use of PPG waveform morphology for diabetes classification remains poorly studied. Most PPG-based research focuses on glucose estimation or uses black-box models [16]–[18] that do not clearly link waveform features to physiological changes. As a result, simple and interpretable features such as amplitude metrics, timing intervals, and vascular health indices have not been fully explored.

III. Methodology

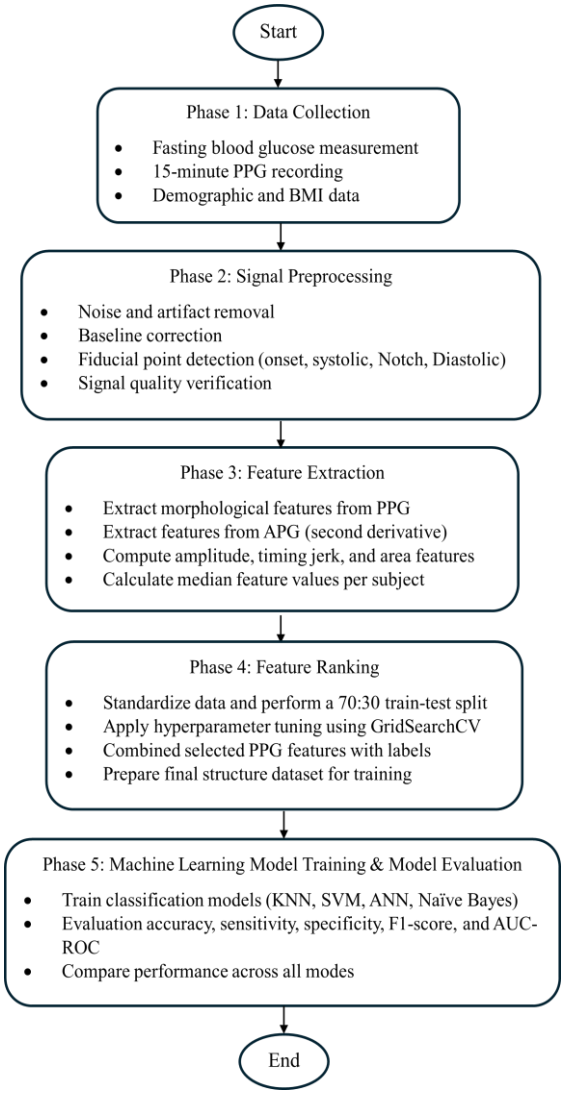

Fig. 1. The methodology of the proposed work.

Fig. 1 illustrates the methodological workflow of the proposed PPG-based diabetes prediction system. The pipeline

begins with data collection, which includes fasting blood glucose measurements, 15-minute PPG recordings, and relevant physiological parameters. The acquired signals are subsequently pre-processed using noise reduction, baseline correction, and fiducial point detection to ensure high-quality, reliable waveform morphology. Next, feature extraction is performed on both the original PPG signal and its second-order derivative (APG), yielding a comprehensive set of temporal, amplitude, Jerk, and area-related descriptors. A systematic feature ranking procedure is then applied to eliminate redundant variables and identify the most informative biomarkers. The selected features are combined to form a clean, structured dataset for training multiple machine learning classifiers, including Logistic, Regression, DT, SVM, ANN, KNN, and NB. Model performance is evaluated using standard metrics: accuracy, sensitivity, specificity, F1-score, and AUC-ROC. In the final stage, the results are compared across models, their strengths and limitations are analysed, and the best-performing classifier is selected to enable reliable diabetes prediction from non-invasive PPG signals.

A. Phase 1: Data Collection

Data for this study were collected from 78 volunteers, consisting of 45 males and 33 females, spanning a wide age range. Of these participants, 50% had a clinical diagnosis of type 2 diabetes, while the remainder had normal blood glucose levels. Volunteers attended the data collection session either after an overnight fast or after at least two hours without food to ensure stable metabolic conditions for blood testing.

A small blood sample from each participant was applied to a test strip and analyzed using a Contour TS blood glucose meter, which measured the glucose concentration within 5 seconds and displayed the result on the screen. Blood glucose measurements were categorized according to the product manual and grouped into normal, prediabetic, or diabetic ranges, as summarized in Table II. For subsequent machine learning analysis, subjects were labelled as 1 for diabetic and 0 for non-diabetic. Additional information, including demographic details, medical history, height, weight, and body mass index (BMI), was also collected during this phase. The blood glucose values were therefore classified according to the device's product manual (Table II) and coded as 1 for participants with diabetes and 0 for those without.

TABLE II. BLOOD GLUCOSE CLASSIFICATION (FASTING OR 2-HOUR POST MEAL).

Condition	Fasting Blood Glucose (mmol/L)	2 Hours After Meal (mmol/L)
Normal	3.9–5.5	< 7.8
Prediabetes	5.6–6.9	7.8–11.0
Type 2 Diabetes	≥7.0	≥11.1

To collect the PPG signals, we used a pulse oximeter (CMS50E, Camtek, China) with a sampling frequency of 60 Hz. The device was attached to each participant's index finger for a 15-minute recording. The SpO2 Assistant V3.1.0.4 software, as shown in Fig. 2, was used to monitor and record key physiological parameters in real time, including blood saturated oxygen (SpO2), pulse rate, and the PPG waveform. These signals were later analyzed to assess vascular health and extract waveform features.

Fig. 2. Real-time monitoring dashboard showing SpO2 levels (top graph), pulse rate (middle graph), and the photoplethysmogram (PPG) waveform (bottom graph).

Based on the standard reference values in Table III, the volunteer's SpO2 reading of 98% falls within the normal range (95–100%) and indicates adequate oxygenation. The pulse rate is 76 bpm, which is within the healthy resting heart rate range of 60–100 bpm. The PPG waveform at the bottom of the display is stable and continuous, reflecting consistent changes in blood volume with each heartbeat and suggesting proper circulation and vascular health. The perfusion index (PI%) is not fully evident on the display. Overall, these findings indicate normal oxygen saturation, a healthy heart rate, and good blood flow regulation, with no signs of immediate cardiovascular problems.

TABLE III. PHYSIOLOGICAL PARAMETER REFERENCE RANGES

Parameter	Normal Range
SpO_2 (Oxygen Saturation)	95% – 100%
Pulse Rate (Heart Rate)	60 – 100 bpm
Respiration Rate (implied by HR context)	12 – 20 breaths/min (general resting)
PPG Waveform	Stable pattern expected
Perfusion Index (PI%)	Typically, >0.2% for detection; 3–10% good

Compared with commonly used datasets in diabetes prediction research, the dataset collected in this study presents several distinctive characteristics. Many previous studies rely on publicly available clinical datasets, such as the Pima Indians Diabetes dataset [24], which primarily include demographic and metabolic variables rather than physiological waveform signals. Other works employ datasets derived from continuous glucose monitoring (CGM) systems to develop machine learning models for glucose prediction and diabetes management [25], [26]. While these datasets provide valuable metabolic information, they do not capture vascular waveform characteristics. In contrast, the dataset used in this study consists of photoplethysmogram (PPG) waveform recordings collected directly from Malaysian participants using a non-invasive pulse oximeter. This allows the analysis to focus on vascular

morphological characteristics reflected in the PPG signal rather than solely on traditional clinical indicators. Furthermore, most existing PPG-based studies are derived from Western or South Asian populations, whereas this dataset represents a Southeast Asian cohort. The inclusion of region-specific physiological data, therefore, helps address the demographic gap in current PPG-based diabetes research.

B. Phase 2: Signal Processing

After collecting the signal, pre-processing is essential. Filtering is used at this stage to remove noise from the raw PPG data.

Low-pass filters are used to remove high-frequency noise caused by electronic interference or sudden movements. At the same time, high-pass filters help eliminate low-frequency components such as baseline drift, often caused by breathing or slight changes in sensor position. In some cases, band-pass filters are applied to retain only the frequency range of interest, typically around the heart rate, thereby preserving the most relevant information. Identifying and removing artifacts, especially those produced by motion, is essential in PPG signal processing. Adaptive filtering techniques are frequently used for this purpose, as they can adapt to signal changes and more effectively suppress noise.

Wavelet transform methods are also employed because they can decompose the signal into different scales, making it easier to isolate and remove unwanted components. In this study, a 4^{th}-order Chebyshev II band-pass filter with a passband of 0.5–8 Hz was used. Among the tested options, this 4^{th}-order Chebyshev II filter provided the most tremendous improvement in PPG signal quality. Fig. 3 shows a comparison between the raw PPG signal and the PPG signal after processing with a Chebyshev II filter. After filtering, the systolic and diastolic peaks become much clearer and easier to identify. From this filtered signal, the morphological features of the PPG waveform can then be extracted. The signal appears inverted because the reflective mode was used during data acquisition.

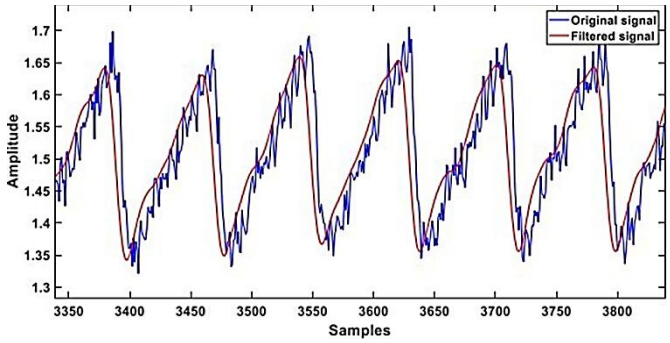

Fig. 3. Raw PPG signal vs. filtered PPG signal.

C. Phase 3: Feature Extraction

Fig. 4 shows how the PPG fiducial points are detected before proceeding with feature extraction. The pulse onset, systolic peak, diastolic peak, and dicrotic notch are all indicated. In this phase, key waveform features such as peak amplitude, inter-peak intervals, and the area under the curve are identified and measured. These features are essential for determining various physiological parameters. In Fig. 4, each fiducial point is marked with a different shape and color to make them easy to distinguish. The red circle represents the pulse onset, the green triangle indicates the systolic peak, the blue diamond marks the dicrotic notch, and the black star shows the diastolic peak. Feature extraction is an essential step before feeding data into the machine learning model and occurs after signal pre-processing.

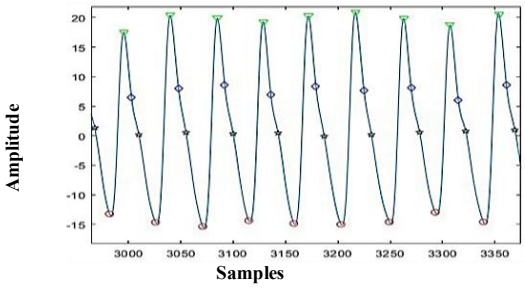

Fig. 4. PPG fiducial point.

The processed PPG signal must first have its features extracted before it can be used as input to the model. Before these features, specifically the fiducial points in the time and amplitude axes of the PPG signal, can be obtained, the morphological characteristics of the PPG waveform must be identified. During this process, several morphological features are considered, each of which plays a vital role in identifying patients with diabetes based on PPG morphology. These features differ in their properties and in how they are calculated, but together they provide valuable information about a person's cardiovascular health.

In general, the PPG signal consists of pulsations that reflect changes in local blood volume in response to fluctuations in blood pressure during the cardiac cycle. The peaks and valleys of the PPG waveform correspond to systolic and diastolic pressures, respectively. These PPG pulses are superimposed on slower variations caused by other physiological processes, such as respiration. The PPG signal also contains local minima and maxima that may have physiological significance and can be analysed by extracting specific PPG features, as described below. One crucial morphological feature of the PPG is the Relative Crest Time (RCT). According to Park et al. [19], the first derivative of the PPG signal can be used to determine the crest time, defined as either the time from the start of the original pulse to the systolic peak, or the time interval ΔT from the systolic peak to the diastolic peak. Peak time is the interval from the start of the VPG waveform (the first derivative of PPG) to its next zero crossing. De Trefford and Lafferty [20] reported that peak time is longer in patients with vascular disease or hypertension than in healthy individuals. Jerk (J) is another morphological feature of the PPG signal. Peak amplitude (PA) is also an important morphological feature of the PPG. The peak amplitude reflects changes in blood volume within the microvascular tissue during each cardiac cycle. Studies have shown that diabetic patients often exhibit distinct PPG characteristics compared with non-diabetic individuals. In diabetes detection, specific changes in the PPG signal, such as reduced peak amplitude, can indicate microvascular

abnormalities. These changes may result from increased arterial stiffness, reduced blood flow efficiency, and alterations in blood flow dynamics caused by diabetes-related complications.

The time ratio (TR) is another extractable morphological feature of the PPG signal. One indicator of aortic stiffness is the augmentation index, defined as the ratio of the diastolic to the systolic peak. The time delay between the systolic and diastolic peaks decreases with age and, when adjusted for the subject's height, provides an indicator of significant artery stiffness [21]–[23]. Both the augmentation index and the peak time delay can offer insight into an individual's cardiovascular health. Once the systolic and diastolic peak times are determined, the time ratio is calculated by dividing the diastolic peak time by the systolic peak time. This ratio is important because it reflects differences in pulse wave timing, which may be related to arterial stiffness and other vascular changes often associated with diabetes. Diabetic patients typically exhibit greater arterial stiffness and less efficient blood flow, which can alter the time profile of the PPG waveform.

All these PPG morphological features depend on the positions of specific PPG fiducial points: onset (O), systolic peak (S), dicrotic notch (N), diastolic peak (D), and the subsequent onset (P). A MATLAB function called Pulse Wave Delineator is used to detect four fiducial points: pulse onset, systolic peak, dicrotic notch, and diastolic peak. For each PPG signal, multiple morphological feature values are obtained. Because MATLAB does not process all individual values in this context, only the median value of each feature is used in the analysis.

Table IV presents a comprehensive list of 30 features extracted from the photoplethysmogram (PPG) waveforms for each subject in the study. These features were derived using a peak detection algorithm described by Suboh et al. [21], which identifies key fiducial points in the PPG signal, including the onset, systolic peak, dicrotic notch, and diastolic peak. The extracted features cover a wide range of temporal, amplitude, Jerk, and area-based measurements that describe both the morphology and dynamics of the PPG pulse. For recordings containing multiple pulses, the median of each feature was used to represent central tendency and to reduce the influence of outliers or irregular beats.

TABLE IV. FEATURE DESCRIPTION

Feature	Description
TR_OSSP	Time interval ratio between Onset-Systolic and Systolic-Onset
TR_ODDP	Time interval ratio between Onset-Diastolic and Diastolic-Onset
TR_ONNP	Time interval ratio between Onset-Notch and Notch-Onset
TR_OSND	Time interval ratio between Onset-Systolic and Notch-Diastolic
TR_SNDP	Time interval ratio between Systolic-Notch and Diastolic-Onset
RCT_OSO	Relative crest time between Onset-Systolic
RCT_ONO	Relative crest time between Onset-Notch
RCT_ODO	Relative crest time between Onset-Diastolic
RCT_SNO	Relative crest time between Systolic-Notch
RCT_SDO	Relative crest time between Systolic-Diastolic
RCT_SPO	Relative crest time between Systolic-Onset
RCT_NDO	Relative crest time between Notch-Diastolic
RCT_NPO	Relative crest time between Notch-Peak
RCT_DPO	Relative crest time between Diastolic-Peak
PA_OS	Peak amplitude from Onset to Systolic
PA_ON	Peak amplitude from Onset to Notch
PA_OD	Peak amplitude from Onset to Diastolic
PA_NS	Peak amplitude from Notch to Systolic
PA_ND	Peak amplitude from Notch to Diastolic
PAR_ODO	Peak amplitude ratio between Onset-Diastolic and Diastolic-Onset
PAR_OSO	Peak amplitude ratio between Onset-Systolic and Systolic-Onset
PAR_NSN	Peak amplitude ratio between Notch-Systolic and Systolic-Notch
J_OS	Jerk between Onset and Systolic
J_ND	Jerk between Notch and Diastolic
J_NS	Jerk between Notch and Systolic
J_DP	Jerk between Diastolic to Peak
A_OP	Area from Onset to Peak
A_ON	Area from Onset to Notch
A_NP	Area from Notch to Peak
AR_DS	Area ratio between Diastolic and Systolic

The features are grouped into several categories: time interval ratios (TR_), relative crest times (RCT_), peak amplitudes (PA_), peak amplitude ratios (PAR_), jerk values (J_), area under the curve (A_), and area ratios (AR_).

Time interval and crest time features capture the timing relationships between fiducial points and provide insight into the temporal behavior of the cardiovascular system. Amplitude- and area-based features reflect the magnitude and distribution of the PPG waveform and may indicate changes in vascular compliance or peripheral resistance. Jerk features characterize the rate of change of the slope between specific waveform segments. Together, these features form a multidimensional representation of the PPG signal, supporting detailed analysis and potential discrimination between physiological states, such as diabetic and non-diabetic conditions.

The selected morphological features were chosen because they reflect physiological characteristics of the cardiovascular system that are known to change in individuals with diabetes. Temporal features such as the relative crest time (RCT) and time ratio (TR) describe the timing relationships between the systolic peak, dicrotic notch, and diastolic components of the pulse waveform. These timing characteristics are closely related to arterial stiffness and pulse wave reflection, which are known to be altered in patients with diabetes due to vascular remodeling and endothelial dysfunction [19], [22], [23]. Amplitude-based features, such as peak amplitude (PA), reflect variations in blood volume within peripheral microvascular tissue and can be affected by reduced vascular compliance and impaired peripheral circulation in diabetic individuals [19], [20]. Jerk-

based features (J) capture the rate of change of the waveform slope between fiducial points and therefore provide information about the dynamic behavior of blood flow acceleration during the cardiac cycle [21]. In addition, area-related features (A and AR) quantify the distribution of the PPG pulse waveform energy and may reflect changes in vascular resistance and arterial elasticity. Together, these temporal, amplitude, and area-based descriptors provide physiologically meaningful indicators of vascular health and serve as informative biomarkers for distinguishing diabetic from non-diabetic individuals using PPG morphology.

D. Phase 4: Features Ranking

To analyze differences in PPG features between diabetic and non-diabetic individuals, the dataset was first split into two groups based on the target label: 0 for non-diabetic and 1 for diabetic subjects. For each feature, descriptive statistics were calculated separately for both groups. These included the mean and standard deviation, providing measures of central tendency and variability. The results were reported as "mean ± standard deviation" to facilitate comparison between groups. To determine whether the differences in feature means were statistically significant, an independent two-sample t-test was performed for each feature. This test assesses whether the observed difference between the two groups is likely due to chance.

The resulting p-values were then evaluated, with smaller values (typically $p < 0.05$) indicating statistically significant differences between groups. This approach helps identify which features are potentially crucial for distinguishing diabetic from non-diabetic individuals. To further examine the relationship between each feature and diabetes status, a point-biserial correlation analysis was conducted. This method is appropriate when one variable is binary (e.g., diabetic vs. non-diabetic) and the other is continuous. It quantifies both the strength and direction of the linear relationship between each feature and the target label. Higher absolute correlation values indicate stronger associations, meaning those features, whether positively or negatively correlated, are more likely to help separate the two groups.

The correlation matrix was also used to detect feature redundancy. Features were considered redundant if they were highly correlated (absolute correlation coefficient ≥ 0.8), as such pairs likely carry similar information. Within each group of highly correlated features, only the feature with the lowest p-value from the t-test (i.e., the most statistically significant) was retained. This strategy reduces multicollinearity and simplifies the model without sacrificing interpretability. Comparing the results from the t-test and correlation analysis provides complementary insights. The t-test assesses whether the group means differ significantly, while correlation analysis measures the strength of the linear association between each feature and the target variable. A feature with a low p-value but a relatively weak correlation may capture a subtle but consistent difference. In contrast, a feature with a strong correlation suggests higher predictive power, even if the p-value is less extreme.

E. Phase 5: Machine Learning Model Training and Model Evaluation

This study performed classification on a dataset of 78 subjects, each represented by nine physiologically meaningful features extracted from their PPG waveforms. These features were preselected based on their statistical significance and correlation with diabetic status. To ensure consistent and unbiased evaluation, the data were first cleaned by imputing any missing values using the mean of the corresponding feature. The entire dataset was then standardized using z-score normalization. After preprocessing, the data were split into training and test sets at a 70:30 ratio. Stratified sampling based on the binary target label (diabetic vs. non-diabetic) was used to preserve the original class distribution in both subsets.

Six machine learning classifiers were evaluated: Logistic Regression, DT, KNN, SVM, ANN, and NB. For KNN, the number of neighbors (neighbors) and the weighting scheme (weights) were tuned. For SVM, the regularization parameter (C), kernel type, and kernel coefficient (gamma) were optimized. The ANN was trained with different hidden-layer configurations, activation functions, solvers, and L2 regularization strengths (alpha). Gaussian Naïve Bayes was used without hyperparameter tuning, as it has no hyperparameters.

Hyperparameter optimization for Logistic Regression, Decision Tree, KNN, SVM, and ANN was performed using GridSearchCV with 5-fold cross-validation, aiming to maximize classification accuracy. The specific tuning ranges and selected values for each model are summarized in Table V.

TABLE V. HYPERPARAMETER TUNING RANGES FOR ML MODELS

Model	Hyperparameter	Tuning Range
Logistic Regression	C	[0.01, 0.1, 1, 10]
Decision Tree	Max depth	[2, 3, 4, 5]
KNN	n_neighbors	[3, 5, 7, 9, 11]
	weights	['uniform', 'distance']
SVM	C	[0.1, 1, 10, 100]
	gamma	['scale', 'auto']
	kernel	['linear', 'rbf']
ANN	hidden_layer_sizes	[(50,), (100,), (50, 50), (100, 50)]
	activation	['tanh', 'relu']
	solver	['adam', 'sgd']
	alpha	[0.0001, 0.001, 0.01]
NB	No tuneable hyperparameters	N/A (Gaussian Naïve Bayes has no grid search parameters)

The models were evaluated on the test set using several performance metrics: mean squared error (MSE), accuracy, sensitivity (true positive rate), and specificity (true negative rate). In addition, AUC-ROC curves were generated to assess each model's ability to discriminate between classes across a range of decision thresholds.

The models demonstrated complementary strengths, with some achieving higher sensitivity and others providing better specificity. All evaluation metrics were summarized in a comparative table, offering an integrated view of each model's predictive performance on unseen data. The ROC analysis supported these findings by illustrating the trade-off between sensitivity and specificity for each model.

IV. Results and Discussion

A. Feature Analysis Results

The correlation heatmap in Fig. 5 shows the pairwise relationships among the extracted PPG features, with deeper red indicating stronger positive correlations. This matrix was used to identify groups of highly correlated features as candidates for dimensionality reduction. For example, within the first cluster of ten highly correlated features (TR_SNDP, RCT_SPO, RCT_ONO, PAR_OSO.1, TR_OSND, TR_ODDP, RCT_OSO, TR_ONNP, RCT_NPO, TR_OSSP), only RCT_ONO was retained as the representative feature. Likewise, RCT_ODO was selected to represent its correlated pair with RCT_DPO, while PA_OS and A_OP were chosen to represent their respective correlated groups. The final set of retained features and their corresponding redundant groups are summarized in Table VI.

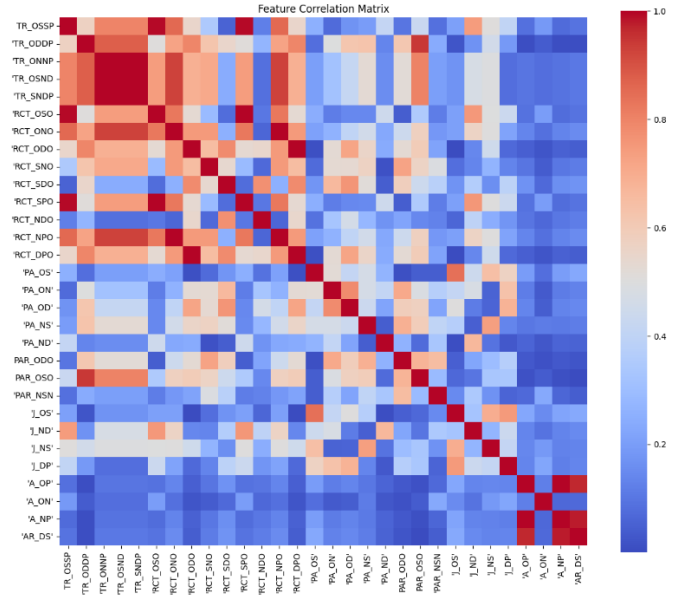

Fig. 5. Correlation heatmap for all features to target.

TABLE VI. Tuning Parameters for ML Models

Feature Groups	Selected Features
TR_SNDP, RCT_SPO, RCT_ONO, PAR_ODO, TR_OSND, TR_ODDP, RCT_OSO, TR_ONNP, RCT_NPO, TR_OSSP	RCT_ONO
RCT_ODO, RCT_DPO	RCT_ODO
PA_OS, J_OS	PA_OS
AR_DS, A_NP, A_OP	A_OP

Feature redundancy among groups likely results from physiological similarities and the strong correlation between features taken from similar fiducial points in the PPG waveform.

In the first group, features such as TR_SNDP, RCT_SPO, TR_OSND, and PAR_OSO all represent time intervals or ratios involving the systolic, notch, and diastolic phases, leading to overlapping temporal information. RCT_ONO was selected because it uniquely captures the relative timing between onset and notch, providing distinct insight while minimizing redundancy. In the second group, RCT_ODO and RCT_DPO, both quantify crest times related to the diastolic point, but RCT_ODO was retained because it is anchored at the onset, which is more consistently defined across pulses. In the third group, PA_OS and J_OS are both derived from the onset-to-systolic segment, with J_OS being derived from PA_OS. Therefore, PA_OS was retained for its greater physiological interpretability and greater robustness to noise. Finally, in the area-based group, AR_DS, A_NP, and A_OP share overlapping integration segments. A_OP was chosen because it captures the dominant systolic energy content from onset to peak, providing a reliable indicator of vascular compliance without the redundancy introduced by derived ratios or narrow bounds. In total, 13 redundant features were removed, and the 17 remaining features are listed in Table VII. Based on the comparative analysis of the non-diabetic and diabetic groups, several features showed statistically significant differences, indicating their potential usefulness for distinguishing between the two conditions. Notably, RCT_ONO, PA_OD, and PA_ND showed strong statistical significance ($p < 0.001$) with moderate absolute correlations (> 0.37), suggesting that these features capture meaningful physiological variations associated with diabetes-related vascular changes. For example, the increases in RCT_ONO and RCT_ODO among diabetic subjects may reflect delayed wave reflections or altered arterial stiffness. In contrast, the decreases in PA_OD and more negative PA_ND values may indicate reduced diastolic waveform amplitude, potentially linked to impaired vascular compliance. Features such as J_ND, PA_NS, and RCT_SNO also demonstrated significant p-values ($p < 0.005$), highlighting changes in waveform shape dynamics, particularly around the notch and diastolic phases, which are often sensitive to alterations in endothelial and arterial tone in diabetic conditions. Conversely, features such as PAR_OSO, RCT_SDO, and J_NS were statistically significant but showed relatively low correlation values. This suggests that, on their own, they have weaker discriminatory power, even though they may still play a valuable supporting role in multivariate models. In contrast, several features, including PA_OS, A_OP, and J_DP, had non-significant p-values and low correlations, indicating that they contribute little to distinguishing between diabetic and non-diabetic groups when considered individually. The feature PAR_NSN, which had the highest p-value and the lowest correlation, was identified as remarkably uninformative in this setting. Overall, these findings highlight the importance of temporal and amplitude features, especially those associated with the onset-to-notch and notch-to-diastolic segments of the waveform. These findings reinforce the value of feature selection in diabetes detection and further support the redundancy-reduction approach applied in the feature-reduction stage. Based on this analysis, the top nine features were selected because they showed statistically significant differences ($p < 0.05$) and relatively high absolute correlation values, indicating strong discriminative ability and clear relevance to diabetic status. A consistent pattern emerged: features with lower p-

values tended to have higher correlation values, suggesting stronger associations with the target condition. This relationship underscores the robustness of the selected features in capturing meaningful physiological variations in PPG waveforms related to diabetes. The final ranking of these features, as shown in Fig. 6, provides a solid basis for dimensionality reduction and model optimization. These selected features will then be used as inputs to train and evaluate machine learning classifiers to improve diagnostic performance.

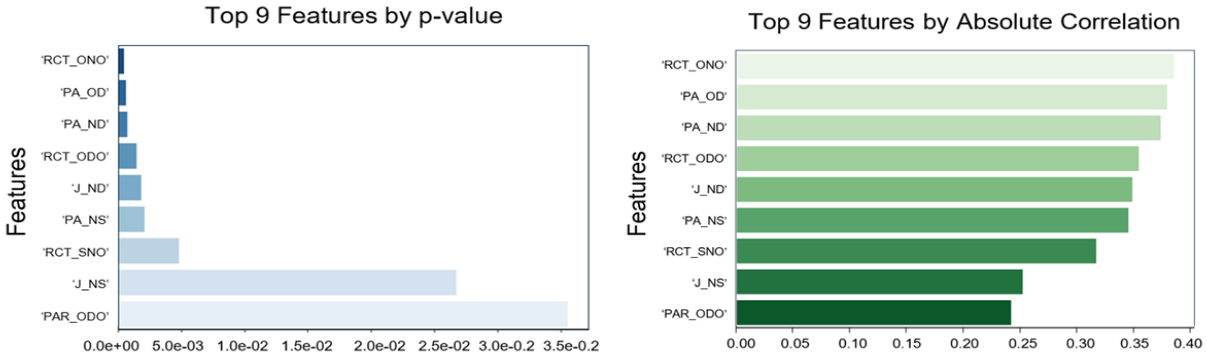

Fig. 6. Ranks for selected features based on p-value and correlation strength.

TABLE VII. FEATURE RANKING AFTER REDUNDANCY REMOVAL

Feature	Non-diabetic (mean±std)	Diabetic (mean±std)	p-value	Absolute Correlation
RCT_ONO	0.52 ± 0.04	0.56 ± 0.06	0.0006	0.384
PA_OD	16.67 ± 3.89	12.92 ± 5.31	0.0007	0.3784
PA_ND	-8.00 ± 3.14	-10.29 ± 2.61	0.0008	0.3725
RCT_ODO	0.71 ± 0.05	0.75 ± 0.06	0.0015	0.3534
J_ND	111.85 ± 37.97	144.38 ± 50.10	0.0019	0.3476
PA_NS	19.92 ± 3.46	23.31 ± 5.63	0.0021	0.3446
RCT_SNO	0.19 ± 0.03	0.21 ± 0.03	0.0048	0.3161
J_NS	217.30 ± 43.93	247.30 ± 69.96	0.0267	0.2518
PAR_OSO	1.79 ± 0.25	2.12 ± 0.92	0.0355	0.2416
RCT_SDO	0.38 ± 0.04	0.40 ± 0.06	0.0813	0.1991
A_ON	798.40 ± 150.66	860.23 ± 169.11	0.0924	0.1919
PA_OS	45.76 ± 2.62	47.47 ± 7.00	0.1596	0.1617
A_OP	1800.86 ± 662.42	2658.13 ± 4825.60	0.2784	0.1251
J_DP	127.47 ± 35.21	118.36 ± 52.05	0.3688	0.1032
PA_ON	25.08 ± 4.25	24.02 ± 6.39	0.3905	0.0987
PAR_NSN	-2.24 ± 1.33	-2.35 ± 1.50	0.7252	0.0404
RCT_NDO	0.17 ± 0.03	0.17 ± 0.04	0.8171	0.0266

Beyond their statistical significance, the selected PPG biomarkers also show meaningful physiological relevance to diabetes-related vascular alterations. Temporal features, such as relative crest time (RCT), reflect pulse wave propagation and may increase due to delayed wave reflection and arterial stiffness, which are commonly observed in individuals with diabetes [19], [22]. Amplitude-based features, such as peak amplitude (PA_OD and PA_ND), reflect variations in peripheral blood volume and may decrease when vascular compliance is reduced due to endothelial dysfunction. Meanwhile, waveform dynamics features such as jerk values (J_ND and J_NS) capture changes in blood flow acceleration and may indicate altered vascular resistance. These morphological changes are consistent with the vascular effects of diabetes, including increased arterial stiffness and impaired peripheral circulation [19], [23], supporting the interpretability of the selected PPG biomarkers for non-invasive diabetes screening.

B. Machine Learning Model Performance

To provide a clearer performance benchmark, two baseline classifiers, Logistic Regression and Decision Tree, were also evaluated using the same experimental setup. Logistic Regression achieved an accuracy of 66.67% with high specificity (91.67%) but relatively low sensitivity (41.67%), indicating limited ability to correctly identify diabetic cases. The Decision Tree model achieved 79.17% accuracy, with balanced sensitivity and specificity, demonstrating that simple tree-based models can capture some nonlinear relationships in the PPG

features. However, the classification results show that the ANN achieved the highest test accuracy (83.33%), with a good balance between specificity (75%) and sensitivity (91.67%). These results suggest that ANN generalizes well to unseen data. Both the SVM and KNN models achieved 75% accuracy. However, the SVM demonstrated higher sensitivity (91.67%) but lower specificity (58.33%), indicating a tendency to overpredict diabetic cases. In contrast, the Naïve Bayes (NB) achieved the highest specificity (83.33%) but the lowest sensitivity (58.33%), making it less reliable for correctly identifying actual diabetic cases. The AUC-ROC scores were relatively high for both SVM and NB (0.86), indicating good discriminatory ability. However, the ANN's slightly lower AUC (0.78) suggests some inconsistency in ranking positive cases, despite its strong overall accuracy and balanced sensitivity and specificity. Overall, the ANN model provides the most balanced performance among the tested algorithms for diabetes prediction using the selected PPG features. These results are summarized in Table VIII and illustrated in Fig. 7. The system's inability to surpass 90% accuracy is likely due to both physiological and data-related factors. PPG signals are indirect measures of vascular health. While features such as RCT_ONO (relative crest time) and PA_OD (pulse amplitude) show statistically significant differences between diabetic and non-diabetic groups ($p < 0.05$), they are also influenced by confounding variables, such as age, blood pressure, and arterial stiffness, which were not accounted for in this study. Although diabetes affects microvascular compliance and arterial resistance, PPG-derived features may not capture these changes with enough specificity, leading to overlapping feature distributions between groups. In addition, the relatively small dataset (n = 78) limits the models' ability to learn robust patterns, and the moderate absolute correlation values (0.24–0.38) suggest only weak linear relationships between the extracted features and diabetes status. Despite these limitations, PPG still shows strong potential for non-invasive detection of diabetes. This is supported by statistical detection of differences in selected features. For instance, a diabetic subject exhibited higher RCT_ONO and lower PA_OD, consistent with increased vascular rigidity and reduced blood flow modulation.

Despite the promising results, this study has limitations that present clear avenues for future work. First, the dataset is relatively small (n = 78), which limits the statistical power and generalizability of the machine learning models. Future studies must prioritize large-scale data collection to train more robust classifiers and prevent potential overfitting. Second, while focusing exclusively on a Southeast Asian (Malaysian) cohort successfully addresses a regional gap in the literature, it inherently limits the global applicability of the findings. Subsequent research should validate this framework across more diverse demographic groups, which encompass varying ethnicities, genetic backgrounds, and geographical locations, to ensure the identified PPG biomarkers are universally reliable. Beyond expanding dataset size and demographic diversity, future work will also focus on incorporating time-series analyses of PPG signals (e.g., wavelet transforms) to capture subtle waveform dynamics and on exploring hybrid models that combine PPG-derived features with other biomarkers (e.g., HbA1c) in a multimodal framework.

TABLE VIII. CLASSIFICATION PERFORMANCE OF ALL MODELS

Model	Best Parameters	Test Accuracy	Test Specificity	Test Sensitivity	Test MSE	AUC-ROC
Logistic Regression	{'C':0.01}	0.6667	0.9167	0.4167	0.3333	0.78
DT	{'max_depth':2}	0.7917	0.75	0.8333	0.2083	0.80
KNN	{'n_neighbors': 3, 'weights': 'uniform'}	0.75	0.6667	0.8333	0.25	0.83
SVM	{'C': 10, 'gamma': 'scale', 'kernel': 'rbf'}	0.75	0.5833	0.9167	0.25	0.86
ANN	{'activation': 'relu', 'alpha': 0.0001, 'hidden_layer_sizes': (50,), 'solver': 'adam'}	0.8333	0.75	0.9167	0.1667	0.78
NB	N/A (GaussianNB)	0.7083	0.8333	0.5833	0	0.86

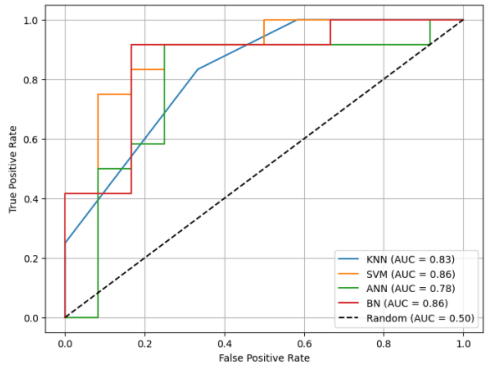

Fig. 7. AUC-ROC curves for all models.

C. Machine Learning Model Performance Comparison

To contextualize the performance of the proposed framework, Table IX compares our best results with recent machine learning approaches for diabetes prediction. Direct comparison is inherently challenging due to differences in dataset sizes, demographic profiles, and data acquisition methods (e.g., continuous glucose monitoring versus standard PPG).

TABLE IX. COMPARISON OF THE PROPOSED MODEL WITH RECENT DIABETES PREDICTION STUDIES.

Ref.	ML Models Used	Data Type / Features	Best Result(s)
[7]	CNN	Biomedical signals	95% within ±1.1 mmol/L
[9]	CNN	PPG + age + sex	AUC = 75.5%
[11]	KNN, NB	Clinical Variables	High Accuracy
[13]	SVM (RBF), BDT	CGM	SVM = 98.25%, BDT = 92.58%
Proposed Work	ANN / SVM	PPG Morphology (Malaysian Cohort)	Accuracy = 83.33% AUC = 0.86

While models using complex continuous glucose monitoring (CGM) data or deep learning algorithms (e.g., CNNs) occasionally report higher raw accuracies (e.g., >92% [13]), they often lack physiological interpretability and require invasive or costly data collection. In contrast, our proposed ANN model achieves a highly competitive accuracy of 83.33% and an AUC of 0.86 (via SVM) using only non-invasive, single-site PPG waveform features. Notably, our framework outperforms certain deep learning models utilizing PPG (e.g., Zanelli et al. [9], which reported an AUC of 75.5%). Ultimately, the proposed method provides a strong balance of predictive performance, computational efficiency, and physiological explainability, making it highly suitable for scalable community screening in Southeast Asian populations.

V. Conclusion

This study introduced a physiologically interpretable and non-invasive framework for distinguishing between diabetic and non-diabetic individuals using morphological features derived from PPG waveforms. The proposed pipeline includes signal pre-processing, detailed fiducial-point-based feature extraction, statistical feature ranking, and evaluation across multiple machine learning models. Leveraging a locally collected dataset from low-cost PPG devices, the study demonstrates practical relevance and addresses demographic gaps often ignored in the existing literature.

The results indicate that certain PPG morphological features, particularly relative crest times and amplitude-based indices, differ significantly between diabetic and non-diabetic groups. These statistically significant and physiologically meaningful biomarkers served as effective inputs for classification models. Among the evaluated algorithms, the ANN achieved the highest accuracy with balanced sensitivity and specificity, while the SVM model exhibited a high AUC, reflecting strong discriminative ability. These findings demonstrate that diabetes-related vascular alterations can be detected through carefully selected, non-invasive PPG features.

By applying this pipeline to a locally collected Malaysian dataset, the study provides a structured and explainable approach to diabetes prediction that addresses demographic gaps often overlooked in Western-centric research. The results highlight the potential for developing affordable, community-oriented tools tailored to the Malaysian healthcare context.

Overall, this work presents an interpretable framework for diabetes detection based on detailed PPG morphology, moving beyond conventional clinical variables and black-box models. Practically, the reliance on standard PPG signals means this framework is highly translatable to existing wearable health monitoring systems, such as smartwatches and fitness trackers. By integrating these trained ML models into wearable devices, at-risk individuals could benefit from continuous, passive, and non-invasive diabetes screening without requiring specialized clinical hardware. Furthermore, in clinical workflows, particularly in resource-limited or rural healthcare settings, this method could serve as a rapid, low-cost triage tool. Using basic, commercially available pulse oximeters, clinicians can flag patients exhibiting diabetes-related vascular alterations for confirmatory invasive testing (such as HbA1c or OGTT), thereby optimizing resource allocation. Future research should focus on validating these findings in larger, more diverse cohorts, integrating additional physiological signals, and exploring real-time implementation on edge devices to further advance non-invasive diabetes monitoring.

Acknowledgment

This work was supported by the Higher Education (MOHE) Malaysia under the Fundamental Research Grant Scheme (FRGS) with the reference code FRGS/1/2024/TK07/UKM/02/4.

References

[1] L. Dilworth, A. Facey, and F. Omoruyi, "Diabetes Mellitus and Its Metabolic Complications: The Role of Adipose Tissues," Int. J. Mol. Sci., vol. 22, no. 14, 2021.

[2] I. D. Federation, IDF Diabetes Atlas, 11th ed. Brussels: International Diabetes Federation, 2024. [Online]. Available: https://diabetesatlas.org/

[3] S. A. Antar, N. A. Ashour, M. Sharaky, M. Khattab, N. A. Ashour, R. T. Zaid, E. J. Roh, A. Elkamhawy, and A. A. Al-Karmalawy, "Diabetes mellitus: Classification, mediators, and complications; A gate to identify potential targets for the development of new effective treatments," Biomed. Pharmacother., vol. 168, p. 115734, 2023.

[4] I. M. Magodoro, K. A. Wilkinson, B. L. Claggett, N. A. B. Ntusi, M. J. Siedner, M., and R. J. Wilkinson, "Discordance between measures of Mycobacterium tuberculosis sensitization and type 2 diabetes mellitus in the United States (NHANES): A population-based cohort study," J. Infect., vol. 90, no. 6, 2025.

[5] Y. Thewjitcharoen, A. Jones Elizabeth, S. Butadej, S. Nakasatien, P. Chotwanvirat, E. Wanothayaroj, S. Krittiyawong, T. Himathongkam, and T. Himathongkam, "Performance of HbA1c versus oral glucose tolerance test (OGTT) as a screening tool to diagnose dysglycemic status in high-risk Thai patients," BMC Endocr. Disord., vol. 19, no. 1, p. 23, 2019.

[6] D.-L. Kim, S.-D. Kim, S. K., Kim, S., Park, and K.-H. Song, "Is an Oral Glucose Tolerance Test Still Valid for Diagnosing Diabetes Mellitus?" Diabetes Metab J, vol. 40, no. 2, pp. 118–128, Apr. 2016.

[7] S. Hossain, B. Debnath, S. Biswas, M. J. Al-Hossain, A. Anika, and S. K. Z. Navid, "Estimation of Blood Glucose from PPG Signal Using Convolutional Neural Network," in 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), 2019, pp. 53–58.

[8] S. S. Gupta, S. Hossain, C. A. Haque, and K.-D. Kim, "In-Vivo Estimation of Glucose Level Using PPG Signal," in 2020 International Conference on Information and Communication Technology Convergence (ICTC), 2020, pp. 733–736.

[9] S. Zanelli, M. A. E. Yacoubi, M. Hallab, and M. Ammi, "Type 2 Diabetes Detection With Light CNN From Single Raw PPG Wave," IEEE Access, vol. 11, pp. 57652–57665, 2023.

[10] H. Tjahjadi and K. Ramli, "Non-invasive Blood Pressure Classification Based on Photoplethysmography Using K-Nearest Neighbors Algorithm: A Feasibility Study," Information, vol. 11, p. 93, Feb. 2020.

[11] M. E. Febrian, F. X. Ferdinan, G. P. Sendani, K. M. Suryanigrum, and R. Yunanda, "Diabetes prediction using supervised machine learning," Procedia Comput. Sci., vol. 216, pp. 21–30, 2023.

[12] Y. Qawqzeh, A. Bajahzar, M. Jemmali, M. Otoom, and A. Thaljaoui, "Classification of Diabetes Using Photoplethysmogram (PPG) Wavefom Analysis: Logistic Regression Modeling," Biomed Res. Int., vol. 2020, pp. 1–6, Aug. 2020.

[13] C. Dewi, J. Zendrato, and H. J. Christanto, "Improvement of support vector machine for predicting diabetes mellitus with machine learning approach," J. Auton. Intell. Vol 7, No 2 (2024)DO - 10.32629/jai.v7i2.888, Dec. 2023. [Online]. Available: https://jai.front-sci.com/index.php/jai/rticle/view/888

[14] B. Alkalifah, M. T. Shaheen, J. Alotibi, T. Alsubait, and H. Alhakami, "Evaluation of machine learning-based regression techniques for prediction of diabetes levels fluctuations," Heliyon, vol. 11, no. 1, p. e41199, 2025.

[15] J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, "Boosting methods for multiclass imbalanced data classification: an experimental review," J. Big Data, vol. 7, no. 1, p. 70, 2020.

[16] H. Kaur and V. Kumari, "Predictive modelling and analytics for diabetes using a machine learning approach," Appl. Comput. Informatics, vol. 18, no. 1/2, pp. 90–100, Jan. 2022.

[17] S. Sen Gupta, T.-H. Kwon, S. Hossain, and K.-D. Kim, "Towards non-invasive blood glucose measurement using machine learning: An all-purpose PPG system design," Biomed. Signal Process. Control, vol. 68, p. 102706, 2021.

[18] [18] M. Zeynali, K. Alipour, B. Tarvirdizadeh, and M. Ghamari, "Non-invasive blood glucose monitoring using PPG signals with various deep learning models and implementation using TinyML," Sci. Rep., vol. 15, no. 1, p. 581, 2025.

[19] J. Park, H. S. Seok, S.-S. Kim, and H. Shin, "Photoplethysmogram Analysis and Applications: An Integrative Review," Front. Physiol., vol. 12, 2022.

[20] J. de Trefford and K. Lafferty, "What does photoplethysmography measure? " Med. Biol. Eng. Comput., vol. 22, no. 5, pp. 479–480, 1984.

[21] M. Z. Suboh, R. Jaafar, N. A. Nayan, N. H. Harun, and M. S. F. Mohamad, "Analysis on Four Derivative Waveforms of Photoplethysmogram (PPG) for Fiducial Point Detection," Front. Public Health, vol. 10, 2022. [Online]. Available: https://www.frontiersin.oHealthurnals/health/articles/10.3389/fpubh.2022.920946

[22] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, and C. Ibukiyama, "Assessment of Vasoactive Agents and Vascular Aging by the Second Derivative of Photoplethysmogram Waveform," Hypertension, vol. 32, no. 2, pp. 365–370, Aug. 1998.

[23] S. Millasseau, R. Kelly, J. Ritter, and P. Chowienczyk, "The vascular impact of aging and vasoactive drugs: Comparison of two digital volume pulse measurements," Am. J. Hypertens., vol. 16, pp. 467–472, Jun. 2003.

[24] J. W. Smith, J. E. Everhart, W. C. Dickson, W. C. Knowler, and R. S. Johannes, "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus," Proceedings of the Annual Symposium on Computer Application in Medical Care, pp. 261–265, 1988.

[25] I. Contreras and J. Vehi, "Artificial intelligence for diabetes management and decision support: Literature review," Journal of Medical Internet Research, vol. 20, no. 5, 2018.

[26] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "A deep learning algorithm for personalized blood glucose prediction using continuous glucose monitoring data," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1486–1494, 2019.