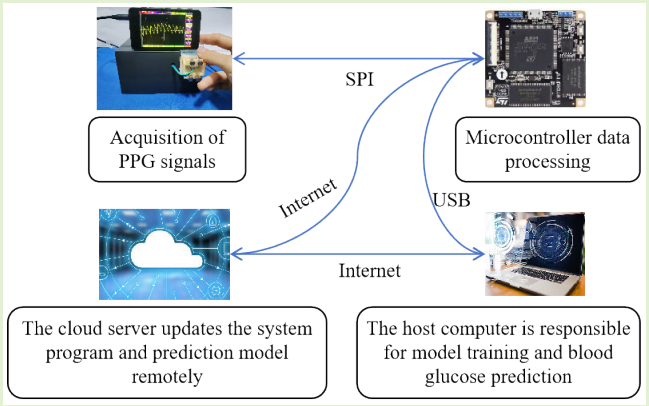
Non-Invasive Blood Glucose Based on Dual-Wavelength Transmission PPG Signal

Yulong Jia, Yu Zheng, Zhong Wang, *Member, IEEE*, and Yuee Li

***Abstract*—In this study, we design a non-invasive blood glucose (BG) monitoring device based on dual transmission wavelength photoplethysmography (PPG), aiming to provide a more convenient and comfortable BG management solution for diabetes patients. By analyzing vascular volume changes and heart rate fluctuations, combined with the near-infrared (NIR) extinction characteristics in biological tissues, we indirectly infer changes in BG levels. Here, we construct an optical instrument for collecting transmission PPG signals at 940 and 660 nm, which are particularly critical for monitoring BG levels. In addition, we design a structure to achieve precise control of the pressure at the fingertip, which is composed of a pressure-sensing membrane and an adjustable slide, adapting to the physiological conditions of different subjects. After data preprocessing, feature extraction, and feature selection, we build a BG prediction model based on machine learning algorithms. Experimental results show that our non-invasive glucose monitoring method has superior prediction performance over a wide range of BG concentrations. We verified that the clinical acceptability of the model reaches 100% through Clark's error grid analysis. The model also demonstrated high accuracy and reliability in the prediction of diabetes classification, reaching up to 97.8%, providing potent support for the practical application.**



***Index Terms*— Dual wavelength, heart rate variability (HRV), machine learning, near-infrared spectroscopy (NIRS), non-invasive blood glucose (BG), photoplethysmography (PPG) signal, pressure tunable.**

I. Introduction

DIABETES mellitus, a global health problem characterized by abnormally high blood glucose (BG) levels, poses a serious threat to human health. Currently, the mainstream BG testing methods rely on invasive BG meters [1], which not only cause pain and inconvenience to patients but also may increase the likelihood of infection and cause other complications [2], [3]. Therefore, the development of a non-invasive, accurate, and convenient BG measurement method is crucial for diabetic patients [4].

Received 22 January 2025; accepted 2 February 2025. Date of publication 17 February 2025; date of current version 2 April 2025. This work was supported by the National Natural Science Foundation of China under Grant 62227807. The associate editor coordinating the review of this article and approving it for publication was Prof. Haoran Jin. *(Corresponding author: Yuee Li.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by The First Hospital of Lanzhou University under Application No. LDYYLL 2023-233.

The authors are with the School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China (e-mail: liyuee@lzu.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/JSEN.2025.3539604, provided by the authors.

Digital Object Identifier 10.1109/JSEN.2025.3539604

Many scientists have tried to develop non-invasive approaches for BG monitoring. Carbon nanotubes [5], [6], fluorescent [7], and polymeric materials [8] were used to construct sensors to measure BG levels, but these techniques suffer from long calibration times, large size, and sensitivity to autofluorescence. Although these techniques have shown potential in some aspects, they still suffer from poor selectivity, temperature sensitivity, and skin irritation. In addition, these models do not take individual physiological characteristic differences into account, which may result in less accurate predictions due to interference from bone, tissue thickness, fat, and other unknown factors [9], [10].

Optical techniques, especially near-infrared (NIR) and mid-infrared (MIR) spectroscopies, have gained attention for their potential to track changes in BG levels [11], [12], [13], [14], and there has been an increasing number of studies on non-invasive glycemic testing using NIR spectroscopy (NIRS), facilitated by chemometric methods [15]. For example, Shih et al. [16] explained this use of Raman spectroscopy for BG measurements, and Sharma et al. [17] demonstrated the use of long-wave NIRS for BG measurements. However, these techniques usually require costly and complex laboratory equipment, limiting their usefulness in routine BG monitoring.

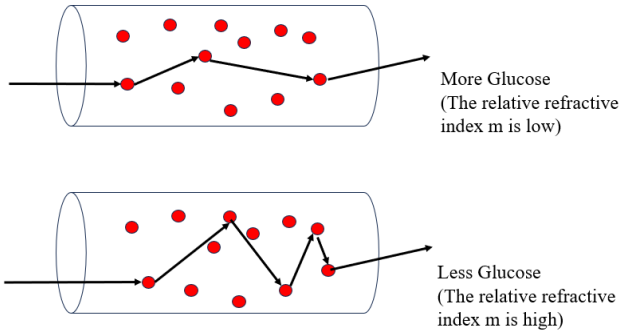


Fig. 1. Propagation of light through the tissue [4].

To address this issue, researchers have proposed simpler and more portable non-invasive glucose monitoring techniques, such as small wearable devices [18]. These technologies are typically less costly and reduce the discomfort associated with invasive measurements. Photoplethysmographic volumetric tracing, as a low-cost, non-invasive optical measurement technique, has been widely studied for estimating human physiological parameters [19]. Despite the progress of existing studies, there continue to be problems such as insufficient accuracy and susceptibility to individual differences in the field of non-invasive BG monitoring. In order to improve the accuracy of prediction, more effective signal processing methods and algorithms have yet to be explored.

In recent years, many researchers have used this principle to conduct non-invasive BG measurements, but most of the current studies have certain defects. Rachim and Chung [20] collected and analyzed the photoplethysmography (PPG) signals of 12 volunteers by using vision–NIR light. According to the diagram of the instrument given in the article [20], it can be found that their LED arrangement is linear, which causes the penetration position of the LED to be inconsistent, which introduces more variables and affects the accuracy of the prediction model. Yue and Ziliang [21] used NIR light to collect PPG signals from 18 volunteers for analysis. Differences in finger width between individuals lead to variations in finger pressure; neither of the studies takes into account the effect of pressure on signals.

Research has demonstrated a significant association between heart rate variability (HRV) measures and BG levels. Low-frequency (LF) power, high-frequency (HF) power, and total power (TP) have been found to inversely correlate with BG levels, while the LF/HF ratio shows a positive correlation in diabetic populations [22]. Vishinov et al. [23] explored the relationship between short-term HRV and instantaneous BG measurements. While strong linear correlations were not observed, certain time-domain parameters such as SDNN and nonlinear metrics, such as SD1/SD2 ratio, showed promising associations.

In this article, we propose a non-invasive BG monitoring system via dual-wavelength transmission PPG features and HRV features to provide a BG management solution for diabetic patients. We built a dual-wavelength transmission instrument to indirectly infer changes in BG levels by analyzing vascular volume changes and heart rate fluctuations in the fingertip. The study underlines the importance of wavelength selection, noting that 940- and 660-nm wavelengths are particularly critical for BG monitoring, and designs a structure to precisely control pressure at the detection site. A BG prediction model based on machine learning algorithms was constructed after data preprocessing, and feature extraction and selection. The experimental results demonstrate that the model has good prediction performance over a wide range of BG concentrations, and its clinical acceptability is verified by Clark's error grid analysis, which provides support for the practical application of this non-invasive BG monitoring method.

II. Design of the Non-Invasive BG Monitoring System

A. Changes in Tissue Absorption and Scattering Caused by Glucose Concentration in NR Region

When a light beam interacts with human tissue, scattering and absorption occur, leading to a reduction in the beam's intensity. Compared to other components in the tissue, glucose has relatively low absorption in the NIR region. Therefore, the influence of glucose concentration on the absorption coefficient can be negligible, which has been demonstrated in both in vitro and in vivo experiments [24], [25], [26], [27]. The influence of glucose concentration on the output light intensity after going through the tissue is mainly due to its influence on the refractive index difference between the extracellular fluid and the cell membrane, which changes the scattering of cells and leads to the change of light path in the tissue.

The reduced scattering coefficient μ_s' of the tissue is defined as [27]

$$n\mu_s' = \frac{K_T\left[n_{\text{cell}} - n_{\text{ECF}}(0) - \delta n\right]}{n_{\text{ECF}}(0)} \tag{1}$$

$$m = \frac{n_{\text{cell}}}{n_{\text{ECF}}(0) + \delta n} \tag{2}$$

where $n \simeq n_{\text{ECF}}(0)$, K_T is a proportional factor, and n_{cell} is the refractive index of the cell membrane. Both are constants and do not affect the reduced scattering coefficient. $n_{\text{ECF}}(0)$ is the refractive index of the extracellular fluid at the baseline physiological glucose concentration, which increases to $n_{\text{ECF}}(0) + \delta_n$, when glucose concentration rises [27]. Therefore, glucose concentration can influence the refractive index of the extracellular fluid, thereby affecting the reduced scattering coefficient of the tissue. The relative refractive index of erythrocytes to plasma m is in the range of 1.04–1.05 in healthy conditions. When the concentration of glucose increases, the refractive index of plasma rises, while the refractive index of red blood cells remains constant, resulting in a decrease in relative refractive index m. This leads to a shorter transmission path for light, thereby increasing the output light intensity. As shown in Fig. 1, conversely, when the concentration of glucose decreases, the refractive index of plasma decreases, leading to a higher relative efficiency and a longer transmission path for light. As a result, the output light intensity through tissue is relatively weaker [26].

B. Introduction to Photoplethysmography (PPG) and HRV

Both optical wavelengths and optical instrument setup are important for designing a non-invasive BG monitoring system

to ensure good accuracy and reliability. In our study, we primarily rely on the analysis of PPG and HRV, collected through the transmitted light path of a finger, to indirectly infer BG levels.

PPG is a non-invasive measurement method [28], and the PPG signal can be captured by optical sensors using transmission or reflection. The alternating current (ac) component of the PPG signal primarily reflects changes in blood volume within the vessels, while the direct current (dc) component is more susceptible to factors such as skin pigmentation and tissue thickness. To improve measurement accuracy, filtering techniques are typically employed to reduce interference from nonblood-related factors, with a focus on analyzing the ac component. In addition, HRV, as an indicator of autonomic nervous system activity, has been shown to have some correlation with changes in BG levels [23]. A detailed analysis of HRV features within PPG signals can provide additional useful information for BG monitoring.

C. Wavelength Selection

Appropriate wavelength selection can enhance signal quality, thereby improving the detection accuracy. The VIS-NR light, ranging from 600 to 1300 nm, interacts with low-energy radiation, achieving maximum penetration depth through the tissue [25]. Absorption of other components in blood, such as eumelanin, pheomelanin, and red blood cells also need to be considered [29]. These components have significant absorption in the NIR region, as shown in supplementary materials Fig. S1.

Considering that water constitutes more than 90% of plasma, we conducted in vitro experiments using a simple glucose buffer solution. The absorption spectrum is shown in supplementary materials Fig. S2. It can be observed that the absorption spectrum remains consistent across different glucose concentrations, indicating that the glucose concentration in water showed a very small impact on the intensity of the absorption spectrum, which confirms our conclusion on absorption and scattering in Section II-A. To avoid large attenuation caused by water, we choose 940 nm as the detection wavelength for glucose.

Previous studies have shown that glucose concentration in the human body is relatively low, and the variation in BG levels is narrow. Moreover, due to the complexity of the human body's environment, it is extremely difficult to measure glucose concentration variation directly using absorption alone. Our approach instead focuses on the change in the refractive index between plasma and blood cells as glucose concentration varies. An increase in glucose concentration leads to a reduction in scattering light and an increase in transmitted light. Here, the selected wavelengths (660 and 940 nm) are relatively cost-effective and provide good penetration while being less affected by water absorption. The reference wavelength is used to contrast with the detection wavelength. Here, 660 nm was chosen as the reference wavelength excludes the effect of light absorption by substances other than glucose, which varies with blood volume. It is sufficient as a reference to correct for measurement errors caused by tissue differences and other physiological factors.

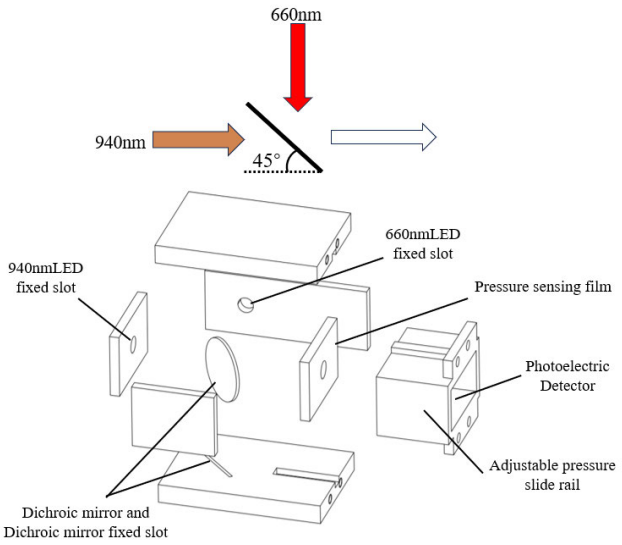


Fig. 2. Experimental optical path diagram and exploded view of the structure.

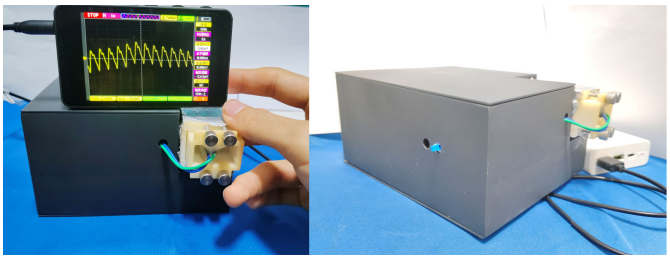


Fig. 3. Physical picture of the device we designed.

In all, the choice of 660- and 940-nm wavelengths is based on their strong penetration capabilities and significantly different absorption rates for hemoglobin, which helps mitigate hemoglobin's influence. By simultaneously analyzing PPG signals at the 940- and 660-nm wavelengths, more comprehensive information on blood flow and glucose level changes can be considered, thereby providing an effective technique for non-invasive BG monitoring.

D. Detection Site Selection

Choosing the detection site and designing suitable optical equipment are crucial for accurate BG monitoring. Physiological differences, such as skin thickness and tissue composition, affect light propagation, making site selection vital. Common sites such as the earlobe, cheek, tongue tip, finger, arm, and lips are rich in blood vessels and conducive to signal monitoring. The fingertip, with its simple structure, abundant superficial vessels, and stable PPG signals, is ideal for non-invasive BG detection, minimizing interference and enhancing signal quality.

In previous studies, non-invasive BG measurements have major problems in accuracy, and the pressure on the test portion is an easily overlooked variable. It has been shown [24], [30] that the contact pressure of glucose sensors applied to human fingers has a strong correlation with glucose measurement data. This makes the fingertip, which is convenient for applying pressure, more suitable for PPG signal extraction.

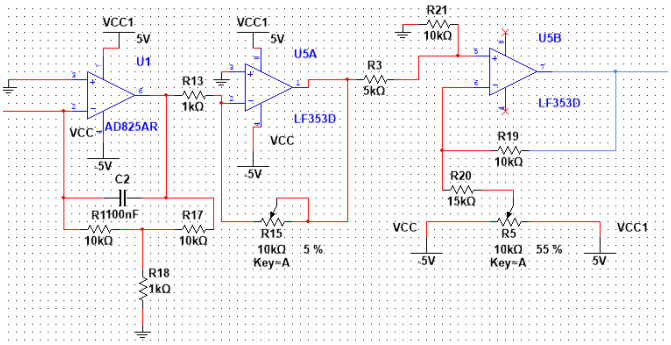


Fig. 4. Photoelectric conversion amplifier circuit.

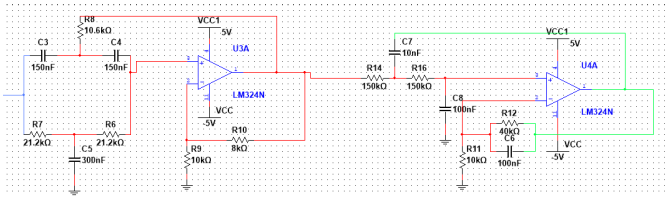


Fig. 5. Trap filter amplifier circuit.

In summary, the fingertip as the detection site, combined with the pressure-adjustable transmissive detection method, provides an effective and reliable solution for non-invasive glucose monitoring.

E. Optical Equipment Construction

In this study, we used LED940E and LED660L produced by THORLABS for the acquisition of PPG signals. In order to control the variables, we need to ensure that both LEDs incident on the finger from the same location, permitting the data collected at two wavelengths, have higher relevance. Here, we utilized dichroic mirrors, with almost 100% reflection for light below 780 nm and nearly 100% transmission for light above 850 nm, for constructing the optical instrument for glucose detection of the fingertip. As shown in the top half of Fig. 2, two beams are combined into one after passing through the dichroic mirror and incident to the fingertip with the exact same light paths.

Considering the optical path and the impact of finger pressure mentioned in Section II-D, we built a 3-D printing model for fixing all optical elements, which includes fixed holes for LEDs of two wavelengths, a slot for the dichroic mirror, a pressure-sensitive film, and an adjustable rail for pressure regulation. A p-i-n photoelectric detector (LSSPD-3.2) at 400–1100-nm range is used for the detection of transmitted light intensity. The exploded view of the overall structure is shown in the bottom half of Fig. 2, and the concrete physical device is shown in Fig. 3. The key components are wrapped in tin foil to shade light and shield electromagnetic interference.

III. Auxiliary Circuit Design

Since the PPG signal is susceptible to interference from motion artifacts, baseline drift, and other noises, we need to complete the signal amplification and filtering of the PPG signal before feature extraction and analysis [31]. The required circuit mainly includes a photoelectric conversion amplifier circuit and a trap filter amplifier circuit.

The photoelectric conversion amplification circuit is shown in Fig. 4, and the principle of the trap filter amplifier circuit is shown in Fig. 5. It is important to note that the impact of different skin pigmentation levels primarily affects the dc signal, which does not influence the ac signal. The useful information we extract is from the ac signal, which represents the periodic variations in blood flow within the arteries during the cardiac cycle (systole and diastole), after removing the dc component. Therefore, we have designed a dc bias circuit, which allows for dc compensation and ensures that the system operates properly despite variations in skin pigmentation. The PPG signals themselves are weak physiological signals with low frequencies distributed in the range of 0.1–10 Hz; therefore, the high-frequency signals are further filtered out here by second-order active filtering. Filtering also improves the signal-to-noise ratio of the PPG signal and highlights useful physiological information. It makes the subsequent feature extraction, such as the peaks and troughs of the pulse wave, more precise, which is crucial for calculating the characteristic parameters and then achieving an accurate prediction of the BG level.

TABLE I
Statistical Characteristics of the Datasets

	Age	Weight(kg)	High(cm)
Mean	40.9	68.6	167.4
Variance	252.2	80.67	46.89
Sex Rate(1 male 0 female)		Incidence Rate(1 sick and 0 not sick)	
1:0=12:8		1:0=8:12	

IV. Feature Extraction and Training

In this study, feature extraction and training are key steps in constructing a BG prediction model. Our goal is to predict human BG levels by analyzing features extracted from PPG signals. In this section, a description of the feature extraction and training process is presented in detail.

A. Source of Datasets

Our data consist of 540 datasets collected from 20 individuals, with statistics on weight, height, age, and the mean and variance of the BG labels (LABEL) for the entire 540 samples. Ethical Statement: the patient datasets are acquired in the first hospital of Lanzhou University. The ethics approval ID is LDYYLL 2023-233. To mitigate the effects of spatial and temporal variations, most of our data were continuously tracked and collected over two years. The statistical characteristics of the dataset we collected are given as follows.

Table I shows the statistical characteristics of the dataset. Among them, the average age is 40.9 years with a variance of 252.2; the average weight is 68.6 kg with a variance of 80.67; the average height is 167.4 cm with a variance of 46.89. In terms of gender ratio, the ratio of male to female is 12:8; in terms of incidence rate, the ratio of sick to nonsick is 8:12. These data reflect the basic information distribution of the population involved in the study and provide basic data support for subsequent analysis and research. During

PPG signal measurement, we first measure at 660 nm and then switch to 940 nm by pressing a button to toggle the wavelength.

B. Data Preprocessing

To remove any residual HF jitter in the signals after circuit processing, we applied a low-pass filter to the PPG signal for filtering out signals above 10 Hz since the heart rate features of the PPG signal are mainly concentrated below 10 Hz. Next, we used digital wavelet transform to filter out LF baseline drift in PPG signals.

To eliminate the influence of different signal magnitudes, we normalized the PPG signal, bringing its amplitude range between 0 and 1. The preprocessing process diagram of the PPG signal is shown in supplementary materials Fig. S9.

After the previous three steps, we used a peak detection algorithm to find the peaks and troughs in the PPG signal, as shown in supplementary materials Fig. S10. Then, the first derivative is computed to identify inflection points in the signal. In addition, to filter out insignificant fluctuations, we set a width threshold, where only points with a distance greater than the threshold are considered valid peaks or troughs.

Through the above preprocessing steps, high-quality PPG signal data are ready for the feature extraction phase, laying a solid foundation for building an accurate BG prediction model.

C. Feature Extraction and Selection

In this study, we extracted over a hundred features, some from the time domain, some from the frequency domain, and some from the energy domain. Some important features are listed in the following.

We calculated HRV parameters and extracted a total of 32 features from the PPG signals at two wavelengths from the time-domain PPG signal. These features include the standard deviation of RR intervals, the root mean square of successive differences (RMSSD) between RR intervals, the standard deviation of successive differences between RR intervals, and others. Then, we obtained a total of 60 time-domain features at two wavelengths, including peak-to-peak value, standard deviation, area, impulse factor of the waveform, and so on. We obtained a total of 20 frequency-domain features at two wavelengths, including frequency centroid, root mean square frequency, frequency variance, and TP. These features can reveal the stability and noise level of the signal and the energy distribution of different frequency components within the signal. Finally, we extracted energy-domain features such as spectral entropy, Teager energy, and Kaiser–Teager energy.

D. Feature Selection

Here, we used the minimum redundancy maximum relevance (MRMR) method [33] for feature selection. To ensure both the accuracy and simplicity of the model, we selected the top 40 features in the ranking for model training. Some representative features, including the difference between the maximum power frequencies of the two wavelengths (PS_MPFdifference), the median of the mean consecutive differences in RR intervals of 940-nm data (Data2_HRV_MCVNN), the margin factor of 940-nm data (Lteager2), the percentage of consecutive RR interval differences exceeding 20 ms in 940-nm data (Data2_HRV_pNN20), and the interquartile range of RR intervals in 660-nm data (Data_HRV_IQRNN), are shown in Table II.

TABLE II
REPRESENTATIVE FEATURES AND MRMR

Feature	MRMR
PS_MPFdifference	0.3954
Data2_HRV_MCVNN	0.3810
Lteager2	0.3045
Data2_HRV_pNN20	0.2161
Data_HRV_IQRNN	0.1631

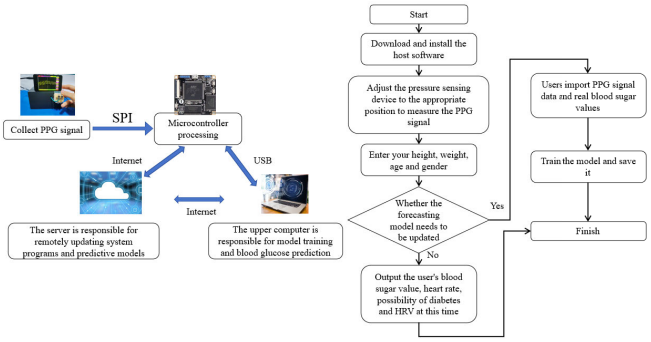


Fig. 6. Photoelectric conversion amplifier circuit.

E. Workflow

The workflow of our system is depicted in Fig. 6. Users initiate the process by downloading and installing the application software on their devices. Subsequently, they capture photoplethysmogram (PPG) waveform signals, which are uploaded to the software. These signals undergo data processing and analysis, culminating in glucose-level predictions. In addition, users have the option to upload their known actual glucose values during the application usage, thereby retraining the predictive model. This personalized approach not only enhances the model's adaptability but also significantly reduces prediction errors for individual users. Furthermore, our system architecture facilitates remote updates of the user's program code via a server, ensuring continuous optimization of system performance.

V. RESULTS AND DISCUSSION

A. Regression Model

In this part, we analyze fingertip transmission PPG signals at 660 and 940 nm to complete the prediction of BG levels in the human body. Using glucose meter measurements (Accu-Chek Guide) as labels, we trained different regression models to evaluate the accuracy of BG predictions.

After feature selection, we used four different machine learning models, including support vector regression (SVR), ensemble bagged trees algorithm (EBTA), elaboration tree (ET), and Gaussian process regression (GPR), to train these features for predicting BG levels. Each model was trained using tenfold cross-validation.

TABLE III
PREDICTION PERFORMANCE OF FOUR REGRESSION MODELS

	$RMSE$(mmol/L)	R^2	MSE(mmol/L)	MAE(mmol/L)
SVR	2.1813	0.37	4.7580	1.3061
EBTA	1.3553	0.79	1.8368	1.0199
ET	1.6229	0.70	2.6338	1.1281
GPR	2.0188	0.53	4.0755	1.3501

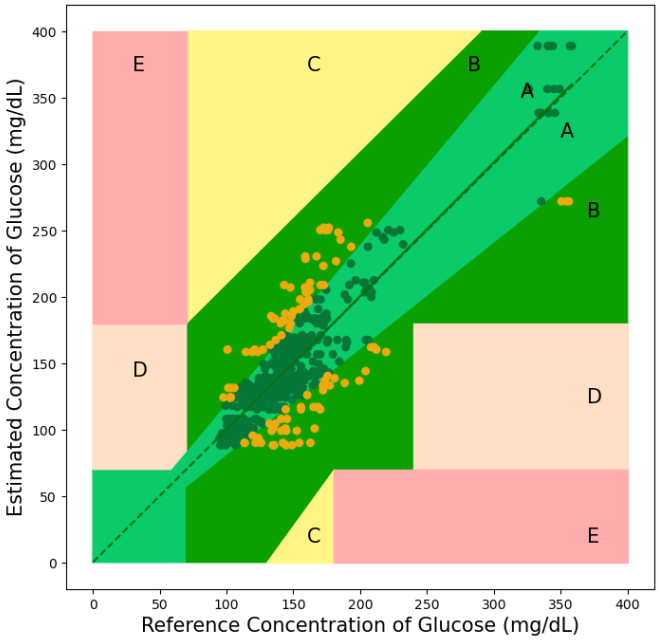


Fig. 7. Clark's network for the EBTA model.

To evaluate the prediction performance of the models, we calculated four parameters, including root mean square error (RMSE), mean square error (mse), R-squared (R^2), and mean absolute error (MAE). The results were shown in Table III.

It can be observed that the EBTA model performed the best in predicting BG levels, with the lowest mse and MAE and the highest R^2 value. This indicates that the EBTA model not only predicts BG levels more accurately but also better captures the trends in BG level changes. The SVR and GPR models have lower R^2 values, which may suggest that they did not fully utilize the information in the data during modeling or that the models were not complex enough to capture the intricate changes in BG levels. In addition, the ET model has a reasonable R^2 value, mse, and MAE, indicating that it may provide stable predictions in certain situations. In all, the EBTA model would be the preferred model to achieve BG prediction with high accuracy and reliability. We also only used the results of the EBTA model for comparison with previous studies.

The Clarke error grid (CEG) is a tool used to assess the accuracy of BG predictions by categorizing prediction errors into different zones to determine the clinical acceptability of the results. Predictions that fall within the A and B zones are considered clinically acceptable [34], [35]. We applied the CEG for evaluating the performance of our EBTA model, as shown in Fig. 7. The EBTA model Clark network fallout results are shown in Table IV, where 80.71% of the EBTA model's predictions fall within region A and 19.29% within region B. The results falling within the AB region reached 100%. This excellent prediction performance within the clinically acceptable range is of extraordinary practical value. It is notable that the upper limit of our prediction model reaches 380 mg/dL, which is about 21 mmol/L. This permits an accurate prediction of BG levels in patients with hyperglycemia.

TABLE IV
EBTA MODEL CLARK NETWORK FALLOUT RESULT

Area	A	B	C	D
Percentage(%)	80.71	19.29	0.00	0.00

Further analysis of the CEG results reveals that the EBTA model maintains low errors when predicting both hypoglycemia and hyperglycemia, which is particularly important in clinical practice. Accurate hypoglycemia predictions can prevent potential dangers, while accurate hyperglycemia predictions help in timely adjusting treatment plans.

We compared our results with the previous study; the results are shown in Table V. It indicates that our EBTA model has significant advantages in non-invasive BG prediction compared to other methods in the existing literature [2], [32], [36], [37]. OrSense's requires high-cost equipment to achieve 95% [36], while we achieved a 100% AB zone coverage rate using low-cost equipment. In addition, although the method proposed by Guo et al. [2] is theoretically advanced and shows high MAE, it has high computational complexity and its feasibility in practical devices has not been fully verified. Chu et al. [37] calibrated BG testing by regularly drawing blood for HbA1c testing, which improved accuracy at the cost of introducing intrusive detection. The researchers used microwave technology for the measurement of BG in [38] and [39] and got good results. However, in [38], only in vitro simulation experiments were carried out using serum solutions with different glucose concentrations, and no real human experiments were carried out. Thus, the practicability of Abdolrazzaghi's method has yet to be verified. In the article [39], they did not train the general model but only trained the model on four volunteers, respectively, which made the model too specific and hardly universal.

In summary, we have achieved wide-ranging BG assessment using low-cost equipment applicable to both diabetic and nondiabetic individuals. It is particularly noteworthy that our proposed model achieved a 100% AB zone coverage rate in the CEG evaluation, a result rarely seen in previous related studies. In addition, our model is trained using all the data of 20 people, which has a stronger generalization compared with the previous study. At the same time, to further improve the accuracy of the prediction model due to the small volunteer sample size, we also allow users to train the model by themselves by adding more individual data through the corresponding PC APP to make the model more reliable.

B. Classification Model

Similarly, we also can build a classification model to judge whether a subject has diabetes based on extracted features

TABLE V
COMPARISON WITH PREVIOUS STUDIES

	Method	$RMSE$(mmol/L)	R^2	MAE(mmol/L)	CEG Zone AB	Deficiency
OrSense et al. [36]	Occlusion NIR Spectroscopy	2.4244	-	-	95%	Equipment expensive
Guo et al. [2]	PPG, Slant Transform,HRV	1.3802	0.69	0.9694	-	Computationally complex
Chu J et al. [37]	PPG, Feature extraction	-	-	-	Zone A:90%	No preprocessing, human activity impact on data
Zihan Nie et al. [32]	IPPG,NIR, Feature extraction, RFR	1.2528	0.60	-	100%	Most of the data from healthy people
Abdolrazzaghi et al. [38]	Microwave	9.35(mg/dL)	0.98	-	99%	Only simulations, no volunteers
Kazemi et al. [39]	Microwave	-	0.9937	-	100%	There were only four volunteers, and the training model was personal
This article	PPG,NIR,HRV, Feature extraction, machine learning	1.3553	0.79	1.0199	100%	-

TABLE VI
COMPARISON WITH PREVIOUS STUDIES FOR CLASSIFICATION

	Method	$AUC-ROC$	$TPR(\%)$	$Accuracy$	Deficiency
E. M. Moreno et al. [40]	PPG, Feature extraction, machine learning	0.694	-	-	Low AUC-ROC; poor distinguishing
Srichan et al. [41]	NIR,PMF,SDNN	-	98.5	-	Stability and reliability not mentioned
Zhang er al. [42]	Smart phone,PPG,KNN	-	-	86.2%	Narrow age range of volunteers (20-33 years)
Susana et al. [43]	PPG, Feature extraction, machine learning	-	-		98% High costs; Less data
This article	PPG,NIR,HRV, Feature extraction, machine learning	0.9944	98.3	97.8%	-

when BG levels are unknown. Consistent with the regression models, each model here also used the MRMR feature extraction to select 40 features, and tenfold cross-validation was used.

We selected three models, including SVR, EBTA, and ET, for classification training. Their confusion matrices are shown in supplementary materials Fig. S11. EBTA shows the best performance, with high accuracy and good discrimination.

Ideally, all predictions should be located on the diagonal of the matrix. The ROC curves are depicted in supplementary materials Fig. S12, and the EBTA model still performed the best, demonstrating a very high AUC value of 0.9944 and a TPR of 98.3%. This indicates that this model has very high accuracy and reliability in distinguishing between different BG levels.

Table VI shows the comparison with previous research, our research demonstrates exceptional classification performance [40], [41], [42], [43]. It provides 97.8% accuracy at such a low cost.

Not only does our model perform excellently in terms of receiver operating characteristic (ROC) curves and true positive rates (TPRs), compared to the method proposed by Moreno et al. [40], our model also shows a significant improvement in AUC-ROC, increasing from their 0.694 to our 0.9944, indicating that our model has superior discrimination ability. In addition, although the model by Srichan et al. [41] achieved a TPR of 98.5%, they did not mention the stability and reliability of the sensor and model over long-term use. In contrast, our study, through two years of long-term data tracking, ensures the stability and reliability of the model, which is a crucial advantage of our research.

While Zhang et al. [42] used smartphone PPG signals and a KNN classifier to achieve 86.2% accuracy, their study was limited by the narrow age range of the volunteers, which may affect the model's applicability to a broader population. Our study, however, includes an elderly population and has validated the model's generalizability through long-term tracking.

In summary, our research surpasses other methods in the existing literature in multiple aspects. Our model is not only more technologically advanced but also more reliable and accurate in practical applications. Through long-term data tracking and analysis, we have demonstrated that our model has significant clinical value and practical application potential in the field of non-invasive BG monitoring.

VI. CONCLUSION

In this article, we put forward a non-invasive BG monitoring technique that combines optical instruments at selected two wavelengths and artificial intelligence of PPG signals to accurately predict BG levels. For developing this approach, the absorption of various components in tissue in the visible and NIR range and the red cell scattering influenced by environmental glucose levels were discussed. A wide range of the prediction of BG levels is achieved based on innovative hardware design and machine learning algorithms. The pressure-adjustable detection structure design used in the

study ensures the consistency and reliability of the monitoring results among different subjects. In addition, through the training and validation of machine learning algorithms, this study demonstrated high accuracy in diabetes prediction and classification tasks, with the BG prediction model achieving an RMSE of 1.3553 mmol/L and an R^2 of 0.79. The CEG of the prediction model achieved a 100% drop point within the clinically acceptable area. The TPR and AUC-ROC of the diabetes classification model also reached 98.3% and 0.9944%, respectively.

However, there are still some shortcomings in our work. First, the measurement device is currently not small enough to be wearable, and the pressure adjustment device requires manual calibration. Second, the user app is only available in a Windows version and has not yet been developed for other platforms. For future work, we plan to further optimize the device structure to make it more compact, incorporate more automation to eliminate the need for manual adjustments, and develop the user app for multiple platforms to enhance accessibility.

Overall, this study contributes a new methodology to the field of non-invasive glucose monitoring, which is expected to improve the life quality of diabetic patients and provide innovative solutions for diabetes management.

REFERENCES

[1] A. Hina and W. Saadeh, "Noninvasive blood glucose monitoring systems using near-infrared technology—A review," *Sensors*, vol. 22, no. 13, p. 4855, Jun. 2022.

[2] W. Guo, B. W. Ling, and Y. Wei, "Denoising of photoplethysmograms for non-invasive blood glucose estimation via slant transform based bit plane method," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Apr. 2023, pp. 1–4.

[3] K. Mosaddequr and T. Rahman, "A novel multipurpose device for dataset creation and on-device immediate estimation of blood glucose level from reflection ppg," *Heliyon*, vol. 9, no. 9, Sep. 2023, Art. no. e19553.

[4] J. Yadav, A. Rani, V. Singh, and B. M. Murari, "Near-infrared LED based non-invasive blood glucose sensor," in *Proc. Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2014, pp. 591–594.

[5] M. B. Lerner, N. Kybert, R. Mendoza, R. Villechenon, M. A. Bonilla Lopez, and A. T. Charlie Johnson, "Scalable, non-invasive glucose sensor based on boronic acid functionalized carbon nanotube transistors," *Appl. Phys. Lett.*, vol. 102, no. 18, May 2013, Art. no. 183113.

[6] N. A. Eracle, R. Lavinia, and V. Monica, "A non-invasive glucose analysis model with a carbon nanotube sensor," *ARS Medica Tomitana*, vol. 25, no. 4, pp. 189–192, Nov. 2019.

[7] K. Girigoswami and N. Akhtar, "Nanobiosensors and fluorescence based biosensors: An overview," *Int. J. Nano Dimension*, vol. 10, no. 1, pp. 1–17, 2019.

[8] M. R. Rakhshani, A. Tavousi, and M. A. Mansouri-Birjandi, "Design of a plasmonic sensor based on a square array of nanorods and two slot cavities with a high figure of merit for glucose concentration monitoring," *Appl. Opt.*, vol. 57, no. 27, p. 7798, 2018.

[9] J. Jiang and K. Xu, "Monte Carlo simulation on the effect of dermal thickness variances on noninvasive blood glucose sensing," *Proc. SPIE*, vol. 8580, pp. 142–149, Mar. 2013.

[10] W. Liu, A. Huang, and P. Wan, "Overcoming individual discrepancies, a learning model for non-invasive blood glucose measurement," *Appl. Sci.*, vol. 9, no. 1, p. 192, Jan. 2019.

[11] B. Paul, M. P. Manuel, and Z. C. Alex, "Design and development of non invasive glucose measurement system," in *Proc. 1st Int. Symp. Phys. Technol. Sensors (ISPTS)*, Mar. 2012, pp. 43–46.

[12] Y. T. Lan et al., "Noninvasive monitoring of blood glucose concentration in diabetic patients with optical coherence tomography," *Laser Phys. Lett.*, vol. 14, no. 3, Mar. 2017, Art. no. 035603.

[13] P. Jain, R. Maddila, and A. M. Joshi, "A precise non-invasive blood glucose measurement system using NIR spectroscopy and Huber's regression model," *Opt. Quantum Electron.*, vol. 51, no. 2, p. 51, Feb. 2019.

[14] I. L. Jernelv, K. Strøm, D. R. Hjelme, and A. Aksnes, "Mid-infrared spectroscopy with a fiber-coupled tunable quantum cascade laser for glucose sensing," *Proc. SPIE*, vol. 11233, pp. 105–113, Feb. 2020.

[15] Y. Yu, J. Huang, J. Zhu, and S. Liang, "An accurate noninvasive blood glucose measurement system using portable near-infrared spectrometer and transfer learning framework," *IEEE Sensors J.*, vol. 21, no. 3, pp. 3506–3519, Feb. 2021.

[16] W.-C. Shih, K. L. Bechtel, and M. V. Rebec, "Noninvasive glucose sensing by transcutaneous Raman spectroscopy," *J. Biomed. Opt.*, vol. 20, no. 5, Feb. 2015, Art. no. 051036.

[17] S. Sharma, M. Goodarzi, L. Wynants, H. Ramon, and W. Saeys, "Efficient use of pure component and interferent spectra in multivariate calibration," *Analytica Chim. Acta*, vol. 778, pp. 15–23, May 2013.

[18] A. M. Joshi, P. Jain, S. P. Mohanty, and N. Agrawal, "IGLU 2.0: A new wearable for accurate non-invasive continuous serum glucose measurement in IoMT framework," *IEEE Trans. Consum. Electron.*, vol. 66, no. 4, pp. 327–335, Nov. 2020.

[19] S. Chatterjee, Z. Patel, M. A. Thaha, and P. A. Kyriacou, "In silico and in vivo investigations using an endocavitary photoplethysmography sensor for tissue viability monitoring," *J. Biomed. Opt.*, vol. 25, no. 2, p. 1, Feb. 2020.

[20] V. P. Rachim and W.-Y. Chung, "Wearable-band type visible-near infrared optical biosensor for non-invasive blood glucose monitoring," *Sens. Actuators B, Chem.*, vol. 286, pp. 173–180, May 2019.

[21] Z. Yue and W. Ziliang, "Non-invasive blood glucose estimation using near-infrared spectroscopy based on SVR," *Sci. Rep.*, vol. 2017, pp. 594–598, Jan. 2017.

[22] L. J. Rothberg, T. Lees, R. Clifton-Bligh, and S. Lal, "Association between heart rate variability measures and blood glucose levels: Implications for noninvasive glucose monitoring for diabetes," *Diabetes Technol. Therapeutics*, vol. 18, no. 6, pp. 366–376, Jun. 2016.

[23] I. Vishinov, M. Gusev, and M. Vavlukis, "Correlating short-term heart rate variability and instantaneous blood glucose measurements," in *Proc. 28th Telecommun. Forum (TELFOR)*, Nov. 2020, pp. 1–4.

[24] J. Y. Qu and B. C. Wilson, "Monte Carlo modeling studies of the effect of physiological factors and other analytes on the determination of glucose concentration in vivo by near infrared optical absorption and scattering measurements," *Proc. SPIE*, vol. 2, no. 3, pp. 319–325, 1997.

[25] P. Sridevi, A. S. Arefin, and A. S. Md Ibrahim, "A feasibility study of non-invasive blood glucose level detection using near-infrared optical spectroscopy," *Bangladesh J. Med. Phys.*, vol. 14, no. 1, pp. 1–13, Dec. 2021.

[26] K. Sarkar, D. Ahmad, S. K. Singha, and M. Ahmad, "Design and implementation of a noninvasive blood glucose monitoring device," in *Proc. 21st Int. Conf. Comput. Inf. Technol. (ICCIT)*, 2018, pp. 1–5.

[27] J. S. Maier, S. A. Walker, S. Fantini, M. A. Franceschini, and E. Gratton, "Possible correlation between blood glucose concentration and the reduced scattering coefficient of tissues in the near infrared," *Opt. Lett.*, vol. 19, no. 24, p. 2062, 1994.

[28] R. G. Priyadarshini, M. Kalimuthu, S. Nikesh, and M. Bhuvaneshwari, "Review of PPG signal using machine learning algorithms for blood pressure and glucose estimation," *IOP Conf., Mater. Sci. Eng.*, vol. 1084, no. 1, Mar. 2021, Art. no. 012031.

[29] S. Delbeck, T. Vahlsing, S. Leonhardt, G. Steiner, and H. M. Heise, "Non-invasive monitoring of blood glucose using optical methods for skin spectroscopy—Opportunities and recent advances," *Anal. Bioanal. Chem.*, vol. 411, no. 1, pp. 63–77, Jan. 2019.

[30] J. Y. Chen et al., "Non-invasive blood glucose measurement of 95% certainty by pressure regulated mid-IR," *Talanta*, vol. 197, pp. 211–217, May 2019.

[31] A. Hina and W. Saadeh, "A noninvasive glucose monitoring SoC based on single wavelength photoplethysmography," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 3, pp. 504–515, Jun. 2020.

[32] Z. Nie, M. Rong, and K. Li, "A blood glucose prediction method based on imaging photoplethysmography in combination with machine learning," *Biomed Signal Process Control*, vol. 79, Jan. 2023, Art. no. 104179.

[33] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.

[34] M. F. Brereton et al., "Reversible changes in pancreatic islet structure and function produced by elevated blood glucose," *Nature Commun.*, vol. 5, no. 1, p. 4639, Aug. 2014.
[35] M. Siervo, J. Lara, S. Chowdhury, A. Ashor, C. Oggioni, and J. C. Mathers, "Effects of the dietary approach to stop hypertension (DASH) diet on cardiovascular risk factors: A systematic review and meta-analysis," *Brit. J. Nutrition*, vol. 113, no. 1, pp. 1–15, Jan. 2015.
[36] T. Lin, A. Gal, Y. Mayzel, K. Horman, and K. Bahartan, "Non-invasive glucose monitoring: A review of challenges and recent advances," *Curr. Trends Biomed. Eng. Biosci.*, vol. 6, no. 5, pp. 1–8, 2017.
[37] J. Chu, W.-T. Yang, W.-R. Lu, Y.-T. Chang, T.-H. Hsieh, and F.-L. Yang, "90% accuracy for photoplethysmography-based non-invasive blood glucose prediction by deep learning with cohort arrangement and quarterly measured HbA1c," *Sensors*, vol. 21, no. 23, p. 7815, Nov. 2021.
[38] M. Abdolrazzaghi, N. Katchinskiy, A. Y. Elezzabi, P. E. Light, and M. Daneshmand, "Noninvasive glucose sensing in aqueous solutions using an active split-ring resonator," *IEEE Sensors J.*, vol. 21, no. 17, pp. 18742–18755, Sep. 2021.
[39] N. Kazemi, M. Abdolrazzaghi, P. E. Light, and P. Musilek, "In–human testing of a non-invasive continuous low–energy microwave glucose sensor with advanced machine learning capabilities," *Biosensors Bioelectron.*, vol. 241, Dec. 2023, Art. no. 115668.
[40] E. M. Moreno et al., "Type 2 diabetes screening test by means of a pulse oximeter," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 341–351, Feb. 2017.
[41] C. Srichan, W. Srichan, P. Danvirutai, C. Ritsongmuang, A. Sharma, and S. Anutrakulchai, "Non-invasively accuracy enhanced blood glucose sensor using shallow dense neural networks with NIR monitoring and medical features," *Sci. Rep.*, vol. 12, no. 1, p. 1769, Feb. 2022.
[42] Y. Zhang, Y. Zhang, S. A. Siddiqui, and A. Kos, "Non-invasive blood-glucose estimation using smartphone PPG signals and subspace kNN classifier," *Electrotechnical Review*, vol. 86, nos. 1–2, pp. 68–74, 2019.
[43] E. Susana, K. Ramli, H. Murfi, and N. H. Apriantoro, "Non-invasive classification of blood glucose level for early detection diabetes based on photoplethysmography signal," *Information*, vol. 13, no. 2, p. 59, Jan. 2022.

Yulong Jia received the B.S. degree in communication engineering from Lanzhou University, Lanzhou, China, where he is currently pursuing the M.S. degree with the Department of Information and Communication Engineering.

His current research interests include optical sensing, particularly in biomedical devices for healthcare applications.

Yu Zheng received the B.S. degree in communication engineering from Lanzhou University, Lanzhou, China, where he is currently pursuing the M.S. degree with the Department of Information and Communication Engineering.

His current research interests include optical communication and optical sensing, especially visible light communication systems and positioning systems.

Zhong Wang (Member, IEEE) received the Ph.D. degree in computer application technology from Lanzhou University, Lanzhou, China, in 2021.

He is currently an Associate Professor with the School of Information Science and Engineering, Lanzhou University. He was also a Visiting Scholar with University at Buffalo, Buffalo, NY, USA, and the State University of New York, Albany, NY, USA. His research interests include machine learning, facial feature extraction, behavior recognition, and medical image processing.

Yuee Li received the B.Sc. and M.Sc. degrees in communication engineering, and the Ph.D. degree in physics from Lanzhou University, Lanzhou, China, in 2001, 2004, and 2010, respectively.

She is a Professor with the School of Information Science and Engineering, Lanzhou University. Her group is working on developing a photoelectric sensing system combined with optical micro–nano structures and artificial intelligence models. Her research interests include optical sensing and communication for applications, especially in biomedical devices.