IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 6 May 2024, accepted 19 May 2024, date of publication 24 May 2024, date of current version 7 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3404971

RESEARCH ARTICLE

QU-GM: An IoT Based Glucose Monitoring System From Photoplethysmography, Blood Pressure, and Demographic Data Using Machine Learning

MD NAZMUL ISLAM SHUZAN[1], MOAJJEM HOSSAIN CHOWDHURY[1], MUHAMMAD E. H. CHOWDHURY[1], (Senior Member, IEEE), KHALID ABUALSAUD[2], (Senior Member, IEEE), ELIAS YAACOUB[2], MD AHASAN ATICK FAISAL[1], MAZUN ALSHAHWANI[2], NOORA AL BORDENI[2], FATIMA AL-KAABI[2], SARA AL-MOHANNADI[2], SAKIB MAHMUD[1], AND NIZAR ZORBA[1], (Senior Member, IEEE)

[1]Department of Electrical Engineering, Qatar University, Doha, Qatar
[2]Department of Computer Science and Engineering, Qatar University, Doha, Qatar

Corresponding authors: Muhammad E. H. Chowdhury (mchowdhury@qu.edu.qa) and Khalid Abualsaud (k.abualsaud@qu.edu.qa)

This work was made possible by National Priorities Research Program (NPRP) grant # 13S-0205-200270 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors. The Open Access publication of this article is supported by Qatar National Library.

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by Hamad Medical Corporation under Approval No. IRB-HMC-2021-011, and performed in line with the Declaration of Helsinki.

ABSTRACT Patients with hyperglycemia require routine glucose monitoring to effectively treat their condition. We have developed a lightweight wristband device to capture Photoplethysmography (PPG) signals. We collected PPG signals, demographic information, and blood pressure data from 139 diabetic (49.65%) and non-diabetic (50.35%) subjects. Blood glucose was estimated, and diabetic severity (normal, warning, and dangerous) was stratified using Mel frequency cepstral coefficients, time, frequency, and statistical features from PPG and their derivative signals along with physiological parameters. Bagged Ensemble Trees outperform other algorithms in estimating blood glucose level with a correlation coefficient of 0.90. The proposed model's prediction was all in Zone A and B in the Clarke Error Grid analysis. The predictions are thus clinically acceptable. Furthermore, K-nearest neighbor model classified the severity levels with an accuracy of 98.12%. Furthermore, the proposed models were deployed in Amazon Web Server. The wristband is connected to an Android mobile application to collect real-time data and update the estimated glucose and diabetic severity every 10-seconds, which will allow the users to gain better control of their diabetic health.

INDEX TERMS Continuous glucose monitoring (CGM), Internet of Things (IoT), machine learning, photoplethysmography (PPG), wearable device.

I. INTRODUCTION

Diabetes is one of the leading causes of death around the world and according to the International Diabetes Federation (IDF), around 537 million people around the world suffered from diabetes in 2021 [1]. This is a disease that can be predominantly seen in lower to middle-income countries and caused the death of nearly 6.7 million people in the year 2021 alone [1]. Diabetes is a pathological condition in which the body is unable to utilize glucose (body sugar) efficiently

The associate editor coordinating the review of this manuscript and approving it for publication was Jason Gu.

which leads to the accumulation of sugar in the blood. A person develops diabetes if the body cannot produce enough insulin or is resistant to it. Diabetes can mainly be divided into two broad categories: Type 1 Diabetes (T1D), which is the most lethal, and Type 2 Diabetes (T2D). There is another special form of diabetes seen only among women during pregnancy, namely Gestational diabetes. T1D results from the destruction of the cells that produce insulin, which means that the pancreas stops producing insulin permanently and the patient's life depends on insulin injections. In the case of T2D, just the pancreatic production of insulin is weak. So, T2D is less dangerous than T1D and can be prevented. Several countries have been developing strategies that aim at formulating advanced visions for diabetic patients to control diabetes and enhance global health and quality of life [2]: i) Investing significant time and attention in raising general awareness of diabetes, improving prevention, and assisting individuals in maintaining healthy lives; ii) To ensure that all segments of society have easy access to information about diabetes, and that the main messages of diabetes awareness campaigns are clear and consistent; iii) To establish and maintain a robust and empowered workforce of health professionals who possess a thorough understanding of diabetes and diabetes care options, and are capable of implementing the future model of diabetes care; iv) To offer high-quality, easily accessible services that save time and effort for every individual through the implementation of a new caring model and a strong commitment towards achieving this ambition. If all recommended steps to control diabetes are not followed properly, the number of diabetic patients in the age range of 35-60 years is projected to increase by two and a half times by 2045, as estimated by IDF [1]. Therefore, it is crucial that diabetes is being tackled on many fronts, and advanced technologies and modern engineering, which have helped improve every aspect of people's lifestyles, can play a vital role in it.

Glucose measuring devices can be broadly classified into two classes: Blood Glucose Meters (BGMs) and Continuous Glucose Meters (CGMs). BGMs are traditional blood glucose measuring techniques where the user needs to manually prick their finger to collect blood samples and for checking the glucose level. Despite the high accuracy and reliability of this system, the main cons are discontinuity and invasiveness. The invasive attribute of such devices can be especially irritating for people who are required to prick their fingers multiple times in a single day. On the contrary, CGMs can check blood glucose levels (BGLs) continuously at regular time intervals. Regardless of its ability to continuously measure glucose, it still needs finger pricking at least twice a day for calibration [3]. CGMs are not as robust as BGMs while more costly and most importantly, for CGMs, it is required to insert the sensor under the skin of the subject (e.g., abdomen, thighs, upper arms, etc.) which is even more uncomfortable and painful than BGMs. So, regardless of the cons of using BGMs, due to the lack of robust non-invasive alternatives, most clinics around the world still fully or partially rely on BGMs. The global market size for BGMs was recorded to be approximately thirteen and a half billion USD in 2020 [4]. However, amid the COVID-19 outbreak, the demand for glucose monitoring systems boosted and the global market witnessed around 14% annual growth as compared to 2020 and is expected to grow from 14 billion USD in 2021 to around 29 billion USD in 2028 [4]. Based on the growing demand for glucose monitoring devices, a reliable, non-invasive, and easy-to-use alternative to BGMs can quickly capture some of its market shares.

Several modern glucose devices are already available on the market. Widely used devices such as the Rite Aid TrueMETRIX Meter [5] or the Walgreens TrueMetrix Bluetooth Blood Glucose Meter [6] are BGM devices. Some of the CGM devices available in the market contain better technology and features than their BGM counterparts. For example, the Guardian Connect System from Medtronic [7] measures glucose upon insertion under the skin every five minutes while the sensor sends data to the central device wirelessly. The relevant app, which is available for both iOS and Android, shows detailed glucose tracking data such as current glucose level, insulin doses or meal schedules, daily glucose trendline, etc. There are devices under development which are painless invasive or non-invasive. For instance, K' watch [8], a PKVitality product, is a smart wearable that consists of micro-needles that are small enough to penetrate the skin without any pain and record levels of blood glucose while wirelessly synchronizing with smartphones. The product is doing its clinical trial now. Another research and development level product is GlucowiseTM [9], a painless, non-invasive non-wearable device. It employs non-ionizing, safe-to-use radio waves with a frequency of around 40 GHz. These waves are both sufficiently large to penetrate the tissue and sufficiently small to provide adequate resolution of the blood regions within the tissue. When a measurement is initiated, a micro-composite film temporarily renders the skin transparent to radio waves. This ensures consistent readings for all individuals, regardless of age, skin type, or skin tone. With unique sensor technology, it can monitor BGLs multiple times without having to pierce the skin. Moreover, smart Cloud technology and the App will provide personalized advice and alerts for managing the patient's conditions. Based on the current and historical data, smart analytics will analyze and forecast immediate trends in BGLs, so that it can adjust the patient's food and medication intake according to his or her activities. However, this device is still not available in the market. Aside from some benefits, there are several drawbacks among the currently available non-invasive glucose measuring devices. First of all, they are on average less accurate, and thus less reliable than their traditional counterparts. There are other challenges as well depending on the apparatus. For example, FreeStyle Libre [10], a glucose monitoring device that uses interstitial fluids instead of blood, was reported as irritating the skin of contacts [11].

So, it is required to have non-invasive glucose measuring devices which provide robust readings while free of such side effects.

There have already been researchers working on inventing non-invasive blood glucose monitoring technologies involving smart materials and nanoparticles. For example, Xu et al. [12] developed a poly(3,4-ethylenedioxythiophene) (PEDOT) by combining poly (styrene sulfonate) (PSS) conductive hydrogel with Prussian Blue nanoparticles (PBNPs), to act as an electrochemical conductive matrix for non-invasive and continuous measurement of blood glucose and showed great correlation to the readings from a commercial BGM device. Cai et al. [13] proposed an air-permeable electrode enabled by graphene fiber fabrics for monitoring glucose noninvasively. Li et al. [14] developed an electro-chemiluminescent biosensor for noninvasive glucose measurement. In an interesting study [15], Tankasala et al. developed a process to estimate glucose noninvasively by condensing exhaled breath of humans. There have also been several studies that used near-infrared (NIR) spectroscopy for the noninvasive detection of glucose. Li et al. [16] proposed a hybrid system termed wavelet prism modified uninformative variable elimination and least squares support vector machine (WPmUVE-LSSVM) to improve the glucose concentration prediction accuracy and diabetes classification performance from NIR-spectroscopy. Several studies also tried to apply classical and deep machine learning techniques to the outcomes from NIR-spectroscopy to reliably estimate blood glucose levels. To effectively analyze non-linear NIR-spectroscopy data from multiple patients and wavelengths, Han et al. [17] used deep learning techniques instead of the traditional linear regression model to improve noninvasive blood glucose measuring performance.

In recent years, some studies used noninvasively collected Photoplethysmography (PPG) raw signals or extracted features as input for machine learning (ML) algorithms to estimate BGLs. Gupta et al. [18] applied classical machine learning-based regression techniques on features extracted from transmissive and reflective PPG signals from three frequency bands (green, red and infrared (IR)) to estimate BGL. Prabha et al. [19] tried to estimate BGL from PPG signals collected from wearable wristbands using classical ML techniques. Riaz et al. [20] tried to extract some features from PPG signals and link them to corresponding glucose readings. Authors in [21], [22] and [23] attempted to estimate glucose levels from fingertip videos collected through smartphones. However, the performance of these systems is either not up to the level of acceptance yet or the dataset used for validation is very small. Moreover, most of these systems are relying on PPG signals alone and there are several other physiological and demographic factors that can lead to change in BGLs. However, from these studies, it can be ascertained that whether as signals or images, PPG signals have great promise in estimating blood glucose.

Inspired by the current literature, in this study, we propose QU-CGM, a smart, end-to-end Internet of Things (IoT) based, real-time, continuous, noninvasive, and reliable blood glucose measurement system using PPG signals, blood pressure and demographic information. The main contributions of this study are summarized as follows:

- Develop a prototype that will continuously estimate blood glucose levels noninvasively and in real-time from PPG signals collected from wristband sensors.
- An end-to-end framework for continuous data collection, sending to the backend server and reporting back to the mobile application is developed and tested.
- The developed prototype will classify a patient's blood glucose level as normal, warning, and dangerous to aid in daily care at home.
- The prototype will send data to a mobile application and show the raw data in the mobile application and alert the patient if the glucose value becomes abnormal. It will send a message to the emergency number provided by the patient for cases such as extremely high blood sugar (hyperglycemia) or dangerously low blood sugar (hypoglycemia) which might even result in a diabetic coma.

The rest of this study is organized as follows: after an introduction and brief literature review in **Section I**. **Section II** describes the experimental setup which contains hardware and software parts after a brief description of the dataset collected for this study along with the methodology of prototyping of the hardware system. **Section III** discusses the experimental results and **Section IV** concludes the article.

II. MATERIALS AND METHODS

A. OVERALL SYSTEM OVERVIEW

The proposed device was created for helping diabetic patients continuously measure their blood sugar levels so that they can live normally while keeping control of their blood sugar levels. The device consists of two parts: hardware, and software. The wearable hardware system can be attached to the wrist of the patient for comfortable capturing of PPG signal and can be used all the time whereas the software part of the system consists of integrating a mobile application and backend server for continuous monitoring (every 10-second update). **Figure 1** illustrates the proposed methodology with respect to the traditional methodology.

In order to make a device that measures glucose continuously, two phases are required, the first stage is System Development (Training and Validation). Data must be collected from as many people as possible in order to complete the first phase. A dataset comprising of age, gender, height, weight, systolic and diastolic blood pressures measured using the pressure monitor, heart rate, type of diabetes, blood sugar measured using the Accu-Chek Glucometer, and a raw PPG signal using the wearable hardware device were collected from 139 subjects. The study was conducted at the Qatar Diabetic Society with the ethical approval of the Hamad Medical Corporation ethical approval committee (IRB-HMC-2021-011). Written informed consent was taken from all the subjects before collecting the data. A total of 3 minutes of PPG signal was recorded per individual.

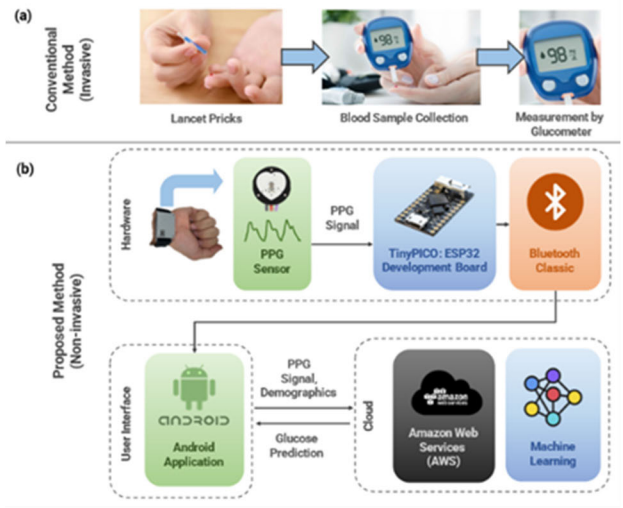

FIGURE 1. High-level Architecture of the proposed method compared to the conventional system. (a) The traditional glucose measurement method. (b) The proposed glucose measurement method.

A mobile application is developed to acquire the PPG signal, blood pressure data and demographic information to send to Amazon Web Services (AWS) platform. Using the acquired data a machine learning model is developed to estimate blood glucose value and classify the patient into different diabetic severity. These models are deployed in the AWS server for real-time and continuous glucose monitoring and severity classification and reporting back to the mobile application.

After completing the first phase and preparing the backend server, the second phase is system validation. We test the remote server by sending new data from the QU-GM device to the mobile application, which then transmits the data to the remote server, which analyzes it, determines the glucose value, performs classification, and then sends the value of glucose and classification results to the application to display to the patient.

Figure 1 (b) shows the high-level architecture diagram of the proposed system. The architecture diagram is composed of three separate blocks. First, in the ''Hardware'' block the high-level overview of the QU-GM device has been provided. The QU-GM device is composed of a PPG sensor, a microcontroller, and a battery. The PPG sensor is used to collect the PPG signal from the wrist. The signal is sampled by the microcontroller and sent to the user using Bluetooth communication.

Then ''User Interface'' block shows the Android application which is used to maintain communication between the device and the remote server. First, the application receives the PPG signal data from the hardware using Bluetooth. After that, the signal is visualized into the app's user interface. Second, there are multiple fields in the app to input the demographics information of the user. Lastly, the app sends the PPG signal and the demographics information input by the user into our remote server using WiFi or Moblie Data (3G/4G). The mobile application developed in this work also displays the estimated glucose prediction results in real time. It receives the glucose value from the remote server through WiFi or Mobile Data and shows the patient his/her glucose value is normal, warning, or dangerous.

The remote server (backend server) has been visualized in the figure inside the ''Cloud'' block. The remote server is responsible for preprocessing the data collected and making estimations and classifications. The first job of the remote server is to filter the PPG signals collected by the QU-GM device and remove any noise collected along with the signals. Then the top-ranked features that were chosen in the training phase (discussed later in detail) will be extracted. The features will be fed to the best-performing model that is developed in the first phase. To obtain the best model for both classification and regression, multiple models were trained from which the best model was chosen (discussed later in detail). Finally, it will display the glucose value, classify the patients, save the result on the remote server, and then transmit it back to the mobile application.

B. HARDWARE DESIGN

Mainly three factors were considered during designing the hardware of the data acquisition system for the QU-CGM system: Portability, Comfortability, and Safety. As mentioned earlier, this study aims to develop a reliable system that can continuously but non-invasively estimate Blood Glucose Levels (BGL) from PPG signal, blood pressure and demographic data. One of the most important requirements for the system was portability. A portable system is easy to carry i.e., small, and lightweight. While not compromising much on reliability, it is crucial to design the hardware of a system as compact and light as possible. As shown in **Figure 2**, the proposed system contains a very small, lightweight Pulse Sensor[1] (SEN-11774, Sparkfun Electronics, USA) for capturing PPG signals, a tiny microcontroller named TinyPICO[2] (Unexpected Maker, Australia) to control the circuitry, an ON/OFF switch to manually turn on and off the device, and a small 250mAh Lithium-ion (Li-ion) battery[3] (Sparkfun Electronics, USA) to power the system. The wavelength of the sensor is 609nm. The APDS-9008, a low cost true 1.8V analog-output ambient photo sensor in miniature chipLED that consists of a photo sensor was used as detector.

In this study, we chose TinyPICO as the microcontroller for controlling the circuit due to it being the world's smallest, fully featured ESP32 development board [24]. Contained in a package as small as 18mm × 32mm, TinyPICO efficiently utilizes ESP32's 240 MHz dual core while having Bluetooth and Wi-Fi connectivity options [24]. The feature of TinyPICO which was crucial while developing the prototype is its ability to efficiently manage Lithium-Polymer (LIPO) or

[1]Pulse Sensor: https://www.sparkfun.com/products/11574.
[2]TinyPico: https://www.tinypico.com/.
[3]Li-ion Battery: https://www.sparkfun.com/products/13851.

Li-ion batteries using its onboard 3.3V Low-Dropout (LDO) voltage regulator [24]. So, a Li-ion battery can be charged and discharged by the microcontroller alone without any need for external circuitry which greatly reduced the prototype size and weight. As shown in **Figure 2(a)**, during work mode (i.e., while being used), the battery supplies power to the circuit which can be turned on and off. On the other hand, as in **Figure 2(b)**, when the circuit is connected to a central power supply (e.g., personal computer, power bank, etc.), the battery gets charged instead when the switch is on. The charging process can be stopped by turning the switch off.

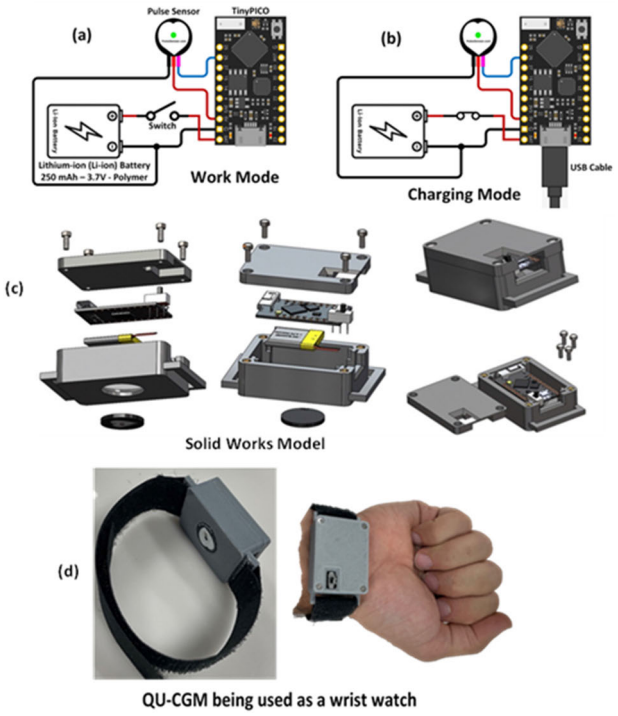

FIGURE 2. The QU-CGM prototype in (a) Work Mode (b) Charging Mode; (c) SOLIDWORKS® model for the wristband module, and (d) QU-CGM worn as a Smartwatch to collect PPG data from the Wrist.

The pulse sensor requires a supply of 5V to run while the battery can only supply 3.7V. The onboard LDO enables feeding the pulse sensor without the use of an external power boost converter, thus reducing the price while improving the portability and compactness of the device. There are various sizes and weights of Li-ion batteries available in the market, and their size and/or weight are directly correlated to their charge storage capacity (milliampere-hour or mAh) [25] while maintaining a fixed voltage output of 3.7V. So, for our prototype device, a tiny 3.7V-250mAh Li-ion battery [26] that may last up to 24 hours (it takes around 2-3 hours to be charged), weighing approximately 5.0 grams and size close to TinyPICO was chosen to power the circuitry.

A sliding switch has been incorporated into the circuit to manually turn the device on or off to prevent the battery charger from drying out. As explained earlier in this section, the PPG signals collected by the device are wirelessly transmitted to an application using the Bluetooth protocol supported by TinyPICO. The application will receive the signal and send it to a remote AWS server, where the remote server will analyze the signal and estimate the value of glucose. BLE is an energy-efficient protocol that is suitable for such biomedical devices with high portability and reliability constraints [27]. TinyPICO's Bluetooth feature perfectly suits LIPO batteries due to their ability to supply fast and high currents [28]. LIPO can manage to do it due to its superior energy efficiency [29]. LIPO batteries are also able to supply constant current during their discharging period due to their energy efficiency. They also have a longer lifetime due to less charge leaking or self-discharge rate [29]. Apart from portability benefits, safety is very crucial for biomedical devices used for humans. The LIPO battery package used in this prototype comes with safety circuitry embedded into the package [30]. The wristband module was designed using the SOLIDWORKS® software as shown in **Figure 2(c)**.

The device itself has been designed like a smartwatch, as shown in **Figure 2 (d)**. So, the outer cover of the device has been designed aesthetically pleasing in appearance with no indication that it is designed for diabetic patients. The QU-CGM device is around 30mm×42mm in size. Wearing the device on the wrist while it collects wrist PPG using the pulse sensor boosts its comfortability manifold. The proposed QU-CGM prototype is more comfortable than traditional BGM and CGM devices due to their non-invasive nature. The purpose of the CGM device is to display the blood sugar level without causing pain. It does not require the use of a needle or a blood sample to measure the blood sugar level. On the other hand, it is cost-effective in the long run since, in the existing devices on the market, the patient must change the needles and/or strips regularly. Since the QU-CGM device will not require such periodic costs, it can be purchased and used for a long time. Due to using very cheap components, the prototype will be very affordable and competitive in the market due to its robustness.

C. MACHINE LEARNING SYSTEM DEVELOPMENT

This part of the methodology explains the process of developing the machine learning model for this system. Exploratory data analysis is first performed on the dataset used in this work. Then the preprocessing steps taken are described. Meaningful feature engineering of PPG signals is then extracted, which is followed by two feature reduction techniques (removing highly correlated features and selecting features using feature selection algorithms). ML models used in this work are then explained along with the metrics used to judge the performance of the said models.

In this work, two types of ML tasks are performed: estimating blood glucose levels (regression) and predicting diabetic severity (classification). The rationale is to use these two separate outputs to ensure that the user knows not only the glucose level but also the severity of diabetes. **Figure 3** depicts the overall ML system development and deployment.

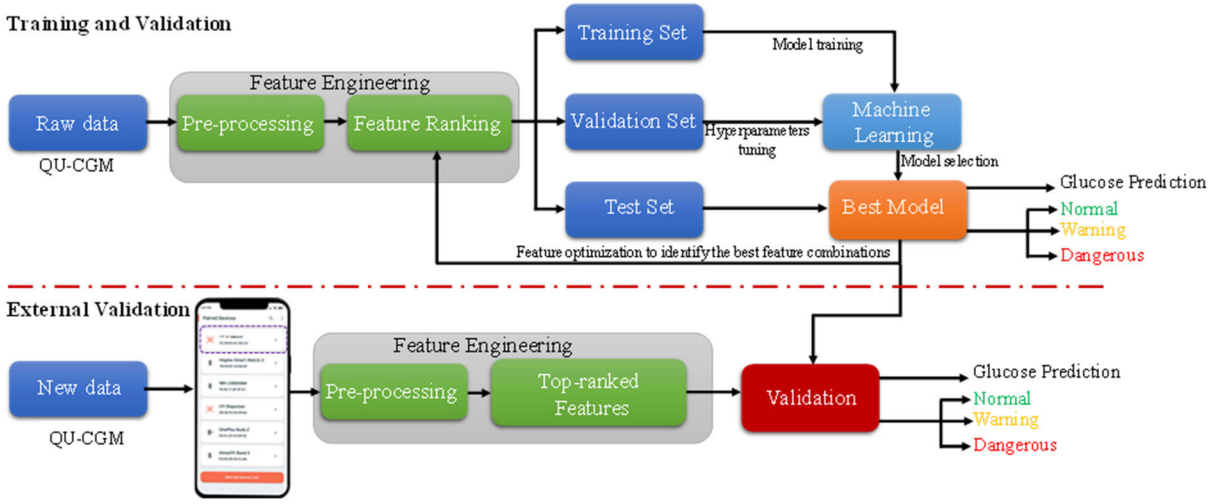

FIGURE 3. A Comprehensive Look at the Development and Deployment of the Proposed Machine Learning Systems.

TABLE 1. Summary of statistics for the collected data.

Parameter	Statistics	Overall
Subjects	Count	139
Age (years)	Min	13
	Median	43
	Max	87
BMI (kg/m2)	Min	16.61
	Median	26.67
	Max	45.99
Systolic BP (mmHg)	Min	90
	Median	135
	Max	234
Diastolic BP (mmHg)	Min	60
	Median	93
	Max	169
Pulse Rate (beats per second)	Min	53
	Median	79
	Max	128
Glucose Level (mg/dL)	Min	66
	Median	118
	Max	600
Gender	% Female	41.84
Diabetes	% Diabetes	49.65

During training and validation, features are extracted and then reduced using a feature selection algorithm. The best model and feature set combination is then found which would be used for prediction during model deployment.

1) DATASET DESCRIPTION

For 139 subjects, the developed wristband was used to record PPG signal data. The systolic and diastolic blood pressure (SBD & DBP) data and heart rate (HR) were recorded from the Omron[4] automatic blood pressure monitor device. The glucose readings were collected from an ACCU-Check Guide meter. The PPG signals were collected during the resting period and sampled at a rate of 125Hz. **Table 1** provides a summary of the dataset and **Supplementary Figure S1** shows the distribution of the blood glucose level. The PPG signal was collected using a 4.096-second window of 512 samples in one segment, which enables enough data to be obtained to train the ML models.

The dataset collected is well-balanced in terms of the number of diabetic patients and non-diabetic patients. Furthermore, the gender distribution is not very skewed as there are more than 40% females. In terms of Age and Body mass index (BMI), it can be observed that there is a good distribution amongst the subjects as age ranges from 13 to 87 and BMI ranges from 16.61 to 45.99. Hence it can be stated that the dataset is quite diverse.

Before embarking on any ML project, it is important to analyze the data beforehand. The features that will be fed to the model must have a strong connection to the output. Since a large number of features will be extracted from the PPG signal itself, it is necessary to see whether there is any change in PPG morphology across various glucose levels. To do so four subjects with various glucose levels were chosen. Each beat was separated and then an average morphology of the signal is plotted in a time-normalized and amplitude-normalized fashion. The mean morphology along with the standard deviation is depicted in **Figure 4**. In lower levels of glucose, the average morphology shows more prominence in notch areas. This disappears at higher glucose levels. Furthermore, it seems that systolic peak time is larger with higher blood glucose levels. In **Figure 4**, the power spectrum density of the signals is also shown at the bottom. It is interesting to note that there are only 2 spikes with a glucose level of 450 mg/dL compared to the other three signals.

2) SIGNAL PREPROCESSING

Motion artifacts and high-frequency noise are seen in the PPG waveform in the dataset. The possible sources for these

[4]https://omronhealthcare.com/products/7-series-wireless-upper-arm-blood-pressure-monitor-bp7350/

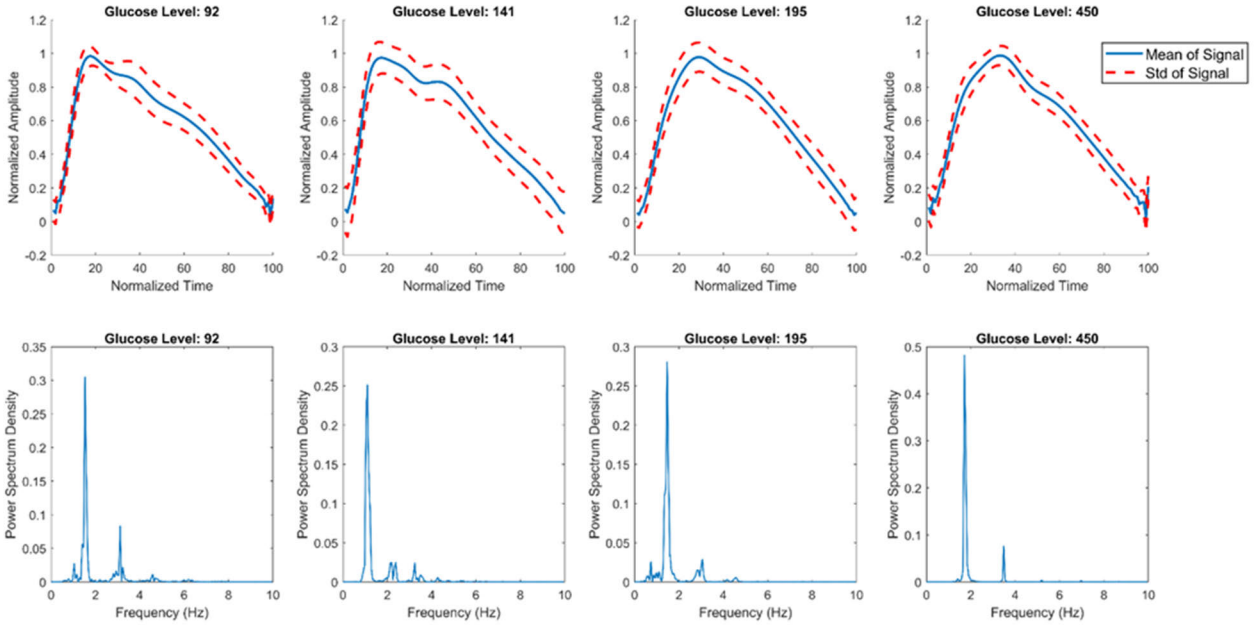

FIGURE 4. Difference in morphology PPG signal for different glucose levels.

noise and artifacts are powerline noise, breathing, etc. The extraction of features might be hampered by these disturbances. As a result, a low-pass Butterworth Infinite Impulse Response (IIR) Zero-Phase Filter was used for eliminating the high-frequency disturbances from the PPG waveforms. **Figure 5** depicts the raw PPG signal with high-frequency noise along with the filtered signal. In Python, a sixth order IIR filter was used with a cut-off frequency of 25 Hz to remove the high-frequency noise.

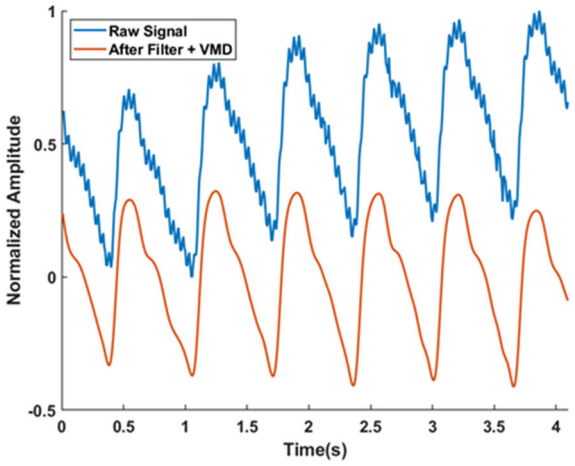

FIGURE 5. Preprocessing steps involving the removal of high-frequency noise and motion artifacts.

PPG signals are frequently affected by motion artifacts (MA) in the real world. The signal is distorted by MA, which creates spikes and other deformation. This makes extracting useful time-domain characteristics extremely challenging. For reducing motion distortions from a one-dimensional signal, several signal processing algorithms have been used. The motion artifacts in PPG signals were recently removed using Variational Mode Decomposition (VMD) [31]. After filtering, the quality of the segmented signals was assessed, however, none of the segments was determined to be unfit. VMD is a robust algorithm that can filter out noise and other distortions. VMD can decompose a signal into multiple modes. Empirically, it has been found that decomposing the signal into five modes would be optimal for this study. The last mode was found to include most of the motion artifacts that corrupt the signal. The first four modes were used to rebuild the PPG signal as proven in [31]. The reconstructed signal is depicted in **Figure 5**.

3) FEATURE ENGINEERING

Different types of features retrieved in this investigation are summarized in **Supplementary Table S1-S5**. PPG waveforms are quite detailed and have a lot of interesting properties. It has characteristics such as the systolic peak, waveform foot, pulse width, peak-to-peak interval, the height of the foot, systolic peak time, etc. We employed the feature extraction approaches described in [31] and [32] to extract the significant characteristics. **Figure 6** describes the feature extraction process in this work.

The statistical characteristics are calculated using the preprocessed signal, whereas the time-domain features are retrieved from the PPG signal and its first and second derivatives. The first peak and first trough of the signal were the most prominent elements of the derivatives. Following that, the time and amplitude characteristics were determined. Most time-domain characteristics have their mean, standard

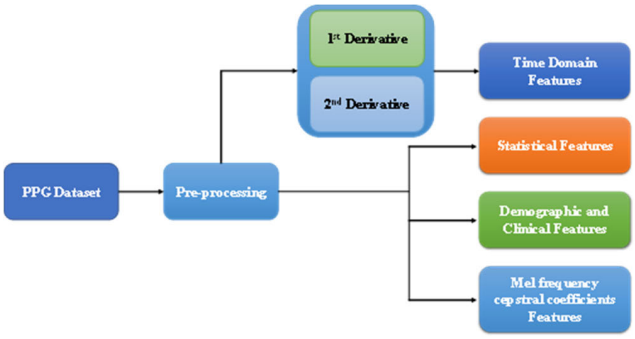

FIGURE 6. Description of the procedure for extracting features.

deviation, and variance computed. This is because these characteristics are critical for capturing the information induced in PPG. These time-domain characteristics were useful in prior studies. The statistical characteristics utilized in this study were taken from [31] and [32].

Furthermore, Mel-frequency cepstral coefficients (MFCC) features were extracted from the signals to capture the signal amplitude spectrum [33]. It has been found that the diabetic condition affects the blood flow in the body, along with heart rate variability (HRV). Specifically, hemodynamics has been shown to be reduced in diabetic patients, meaning that blood pressure is decreased, and heart rate variability is increased [34]. This information can be extracted as a feature if the PPG signal is analyzed in the frequency domain. In this case, MFCC features can be used to represent these features and help in the prediction of BGL. MFCC features are already used in [35], [36], [37], and [38] to extract features from physiological signals for use in multiple applications. We extracted 24 MFCC features using a window length of 512, an overlap length of 200 and a filter bank with a frequency range from 1.5 Hz to 2 Hz. A total of 139 features including signal and demographic features were extracted.

4) REMOVAL OF HIGHLY CORRELATED FEATURES

Feature engineering sometimes creates features which are highly correlated to each other. This creates multicollinearity in the model. When features that are highly correlated to each other are added to the model, the model is weakened. To remove these features, Pearson's correlation coefficients were calculated between each feature. If two features had more than a 0.85 correlation coefficient, one of them would be removed, which resulted in 78 features. These features were then further reduced by feature selection algorithms.

5) FEATURE SELECTION

Feature selection is a critical step in machine learning that involves reducing the dimensionality of the training data by selecting the most relevant features. This not only simplifies the data but also helps in avoiding overfitting, which occurs when the model is too complex and starts to fit the training data too well, resulting in poor performance on new and unseen data. Therefore, feature selection is an essential technique that improves the efficiency and accuracy of the model by selecting only the most relevant features for the given problem. Feature Ranking Library (FSLib) is a widely used library for feature ranking [32]. In this study, three feature selection algorithms were tested from that library to find the best feature selection algorithm (FSA), and the optimal feature ranking technique is reported.

a: RELIEFF FEATURE SELECTION (RFS)

ReliefF is an instance-based technique that gives each feature a relevance weight depending on its ability to distinguish across class values. The lowest-ranked features were progressively deleted until the best result was attained to choose the best subset with ReliefF from the ranking features. For distance-based supervised models that employ pairwise distances between data to forecast, RFS performs considerably better to estimate the importance of the function.

b: CORRELATION-BASED FEATURE SELECTION (CFS)

Correlation is a test that determines whether or not a characteristic is highly correlated with the class or with any of the other features. CFS is an algorithm that couples this evaluation formula with an appropriate correlation measure and a heuristic search strategy.

c: MINIMUM REDUNDANCY MAXIMUM RELEVANCE (MRMR)

The MRMR method is a feature selection algorithm used to identify the most relevant and distinct set of features to describe a response variable accurately. The method aims to reduce the duplication of the feature set while simultaneously increasing its relevance to the response variable, thereby improving the overall performance of a predictive model. In this method, the mutual information of variables is used to quantify the redundancy and relevance of the features. The pairwise mutual information of each feature with the response variable is calculated to determine the relevance of the feature. In contrast, the mutual information between the features is calculated to determine their redundancy. The algorithm then selects the features with the highest relevance to the response variable and the lowest redundancy with other selected features. This results in a set of features that are efficient in describing the response variable while minimizing the risk of overfitting and improving the overall predictive power of the model.

6) ESTIMATING BLOOD GLUCOSE

The data set was partitioned into three sets to ensure accurate and reliable results. The first set, which comprised 60% of the data, was used for training. The second set comprised 20% of the data and was allocated for validation, while the last 20% was reserved for testing. This approach was implemented to guarantee the reliability of the evaluation process by ensuring that each set had a distinct group of subjects. The distinct groups helped reduce potential bias during the

model's development and testing, as the model was trained on one set, validated on another, and tested on a different set. This tactic allowed the model to perform well on new data, improving its accuracy and reliability in estimating blood glucose levels. For the estimation of blood glucose, the results of three top-performing ML models are reported in this work. The models are described below.

a: GAUSSIAN PROCESS REGRESSION (GPR)

GPR is a modeling technique that relies heavily on the Bayesian approach. Many popular supervised machine learning algorithms learn exact values for every parameter in a function, but the Bayesian approach infers a probability distribution over all possible values. For convenience, the prior distribution and the likelihood are assumed to be Gaussian. Using that assumption and solving for the predictive distribution, we get a Gaussian distribution, from which we can predict points using its mean and quantify uncertainty using its variance. GPR calculates the probability distribution over all admissible functions that fit the data because it is nonparametric.

b: BAGGED ENSEMBLE TREES (BET)

The outputs of multiple regression trees are combined using a weighted sum in BET. This paradigm's fundamental concept is to combine several weak learners' strengths to produce a strong learner. Bagging is used in this model to further improve the results.

c: DECISION TREE (DT)

A Decision Tree is a model that tries to predict the output by learning simple rules about the features. They can be seen as piece-wise function approximators. So, to create a Decision Tree that can model the relationship between features and output, more depth for the trees is needed. Increasing the depth means more complex rules will be learned. This, however, sometimes causes the model to overfit as it tries to learn too much from the training data. But the biggest pros of Decision Trees are that they are very easy to interpret. Furthermore, they do not require any feature normalization or categorical encoding.

7) ML MODEL DEVELOPMENT FOR CLASSIFICATION

For classifying the diabetic severity of the subjects, a three-class approach is taken. The subjects are classified using the ground truth blood glucose level. In **Table 2**, the glucose level is mentioned for each class. The ML models were trained in a five-fold stratified cross-validation fashion. Stratification ensures that the distribution of the classes in all the folds is the same. Ensuring this distribution means that the model should perform almost equally in all the folds. So, the feature matrix is divided into five different folds and the models were trained five times; each time four folds are used for training while the out-of-fold data is used for testing. Then the overall evaluation metrics were reported for each model.

TABLE 2. Diabetic severity classes based on glucose level.

Normal	Warning	Dangerous
66 - 109.8 mg/dL	109.8-140.4 mg/dL	140.4 mg/dL and above

For classification, the results of three models are reported in this work: Decision Tree, K-Nearest Neighbor and Support Vector Machine. The structure of the Decision Tree for classification is the same as that of regression. The difference is that the model predicts a label instead of a number. The other two models are described below.

a: K-NEAREST NEIGHBOUR (KNN)

KNN is a type of ML model that uses non-generalizing learning. This means that it does not create a general internal model rather it uses the neighbour points to classify the data. For new data, the distance is calculated for each neighbour point. The one with the minimum distance is then examined and the new data is assigned the label which is most common in that neighbouring point. The biggest advantage of this model is that it is easy to implement and deploy. However, the biggest downfall is that with the higher number of features, the model will overfit.

b: SUPPORT VECTOR MACHINE (SVM)

It is a supervised learning algorithm where the model is trained using a symmetrical loss function. SVM works by creating hyperplanes in a higher dimensional space. This hyperplane allows for more accurate separation of the data. Its utilization of various kernels allows for very adaptive decision boundaries. Furthermore, SVM is very effective with large-dimensional data, i.e., feature matrix with a high number of features.

8) EVALUATION CRITERIA

Three performance matrices were employed in this investigation for regression. Here, Xp represents anticipated data, X represents ground truth data, and n represents the number of samples or recordings.

a: MEAN ABSOLUTE ERROR (MAE)

The Mean Absolute Error is the mean of the absolute of the predicted errors.

$$MAE = \frac{1}{n}\sum_{n} |X_p - X| \tag{1}$$

b: ROOT MEAN SQUARED ERROR (RMSE)

The standard deviation of the prediction error, or residuals, is measured by the RMSE, with residuals indicating the distance between data points and the regression line. As a result, RMSE is a metric for determining the spread of residuals, the smaller the spread, the better the model.

$$RMSE = \sqrt{\frac{\sum |X_p - X|^2}{n}} \tag{2}$$

c: CORRELATION COEFFICIENT (R)

It is a statistical technique that also tells us how well the prediction matches the ground truth. The R statistic is used to measure how closely two variables are related.

$$R = \sqrt{1 - \frac{MSE(Model)}{MSE(Baseline)}} \tag{3}$$

where MSE (baseline) $= \frac{\sum |X - mean(X)|}{n}$

Five metrics are used to assess the classifiers' performance as a model. The terms true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP) all refer to the percentage of correct diagnoses. Following is a breakdown of the five metrics [31], [32]:

d: ACCURACY

When evaluating the efficacy of a classification model, one of the fundamental criteria that is examined is its level of accuracy. A straightforward definition of accuracy would be the total number of right guesses divided by the total number of samples. The metric is calculated using equation (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

e: PRECISION

Precision is a statistic that determines how accurately numerous positive classes can be identified based on the total number of instances that were anticipated to be positive. This statistic is utilized in an effort to cut down on the number of instances of false positives. Using the following equation (5), we can determine this metric.

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

f: RECALL

A statistic known as recall determines how many positive classes out of all the positive instances were properly identified. This number is calculated as a percentage. This is particularly significant in the medical field since it is critical to avoid missing any good examples in the majority of situations. Using the following equation (6), we can determine this metric.

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

g: F1-SCORE

A statistic known as the F1-score takes into consideration both false positives and false negatives in its calculations. It is possible to express it as a weighted average of both recall and accuracy. When evaluating models using unbalanced datasets, this statistic is much more relevant than usual to consider. This metric is calculated as shown in equation (7).

$$\begin{aligned} F1 - score &= 2 \times \frac{Precision \times Recall}{Precision + Recall} \\ &= \frac{2 \times TP}{2 \times TP + FN + FP} \end{aligned} \tag{7}$$

h: RECEIVER OPERATING CHARACTERISTICS–AREA UNDER THE CURVE (ROC-AUC)

The true positive rate is displayed against the false positive rate on a graph known as a ROC curve. When evaluating the effectiveness of the model, it is necessary to compute the area under the curve (AUC). The ROC-AUC score can have a value anywhere from 0.0 to 1.0, with 0.0 indicating the poorest possible performance and 1.0 indicating the greatest possible outcome.

D. SOFTWARE DESIGN

To make the algorithm accessible to the user, an Android application was developed. Android platform was chosen because it can run on mobile devices, and it is also the most popular smartphone operating system in the world. However, the application can be developed for iOS users as well. We have developed a mobile application for the Android operating system, and it has support for devices running Android version 5.0 (Lollipop) or newer. The app has been written in the Kotlin programming language instead of the traditional Java programming language to support more advanced features and architectural patterns. The purpose of the mobile application is to collect the data coming from the PPG sensor wirelessly and then send it to the server where it will be processed using the machine learning model to provide the glucose prediction and severity classification. The basic functionality of the app can be divided into two parts:

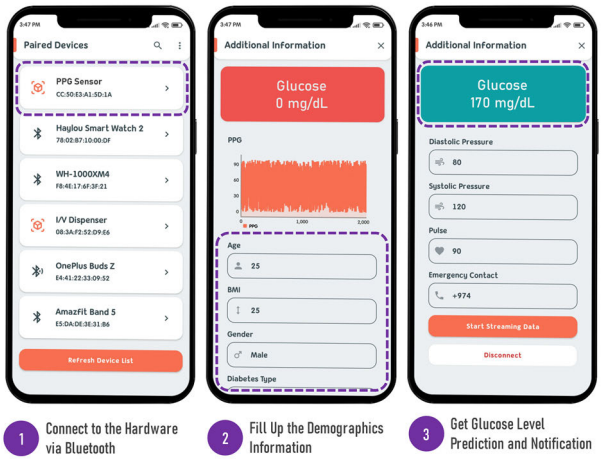

FIGURE 7. The user interface of the different screens of the developed Android application. The figure also includes basic instructions for using the app.

Bluetooth Communication: To wirelessly transfer the data from the PPG sensor, we are using the Bluetooth Classic protocol. To support continuous data transmission at a high sampling rate, we have chosen Classical Bluetooth over the Bluetooth Low Energy (BLE) protocol. When connected to the sensor, the app receives the PPG data at a 125 samples/sec rate. Upon receiving the data, it is stored in a temporary buffer. Once the buffer is full, the data is then handed over

to the network module of the application which handles communication with the server.

Network Communication: This part of the application handles the traffic between the backend server and the application. Firstly, the PPG data along with the blood pressure and demographic information of the user are sent to the server via an HTTP POST request. The body of the request contains all the outgoing data in JavaScript Object Notation (JSON) format. Once the request is received by the server, it transfers the data to the machine learning models to calculate glucose prediction and classify diabetic severity. Once the predicted value and classification result is available, the server responds with a JSON object containing the glucose data. The app then receives the response from the server and displays it to the user. If anything goes wrong in the in-between process, e.g., network error, invalid data, etc., the server throws an error which is then raised to the user. The purpose of the application is to provide an intuitive way for the user so they can make full use of the proposed glucose level prediction and classification algorithm.

The following instructions should be followed while using the app for the very first time:

1. First, the user needs to pair the PPG device to their smartphone using the Bluetooth connectivity section of the device.
2. Once the pairing process is done the user can open the app. On first-time use, the application will ask for the required permissions. The user needs to approve all the permissions in this part otherwise the application will not function.
3. After granting the permission, the user will be prompted to turn on Bluetooth if not enabled already.
4. The user will be presented with a list of paired devices. If the desired device is not on the list, the user can click the ''Refresh Device Button'' which will refresh the list. Upon clicking on the button for the desired device, a Bluetooth connection process will take place. Upon successful connection, the list will become invisible, and the user will be presented with the form to fill up the demographic information and blood pressure values.
5. Once the demographic information is correctly filled up, the user can click on the ''Start'' button which will start sending the data to the server. Initially, it will wait for the PPG buffer to fill up and after that, the data will be transmitted every 10 seconds.
6. Once the response from the server is available, it will be displayed in the UI in blue colour at the top of the screen. The glucose prediction will be updated automatically as new data are transferred to the server.
7. After using the app, the user can close the app by clicking back on the navigation menu which will close the Bluetooth socket and server-side communication.

The UI of the Android application and the user instructions are shown in **Figure 7**.

TABLE 3. Performance of various feature selection algorithms with different ML models.

FSA	ML	Num. of Features	R	MAE	RMSE
MRMR	BET	44	0.90	22.90	30.74
	GPR	51	0.80	30.12	42.14
	DT	56	0.76	35.73	49.44
Relieff	BET	52	0.88	24.34	32.15
	GPR	48	0.81	30.11	40.77
	DT	6	0.38	59.88	79.94
CFS	BET	53	0.87	23.70	32.56
	GPR	51	0.78	32.01	47.35
	DT	53	0.64	35.96	55.13

III. RESULTS AND DISCUSSION

This section shows the experimental results of this work along with a detailed discussion of the performance. This section is divided into primarily three parts: estimation of blood glucose, classification of diabetic severity and comparison of current work with the literature.

A. ESTIMATION OF BLOOD GLUCOSE

1) FINDING THE BEST FEATURE SELECTION ALGORITHM (FSA) AND ML COMBINATION

Table 3 shows the result from the combinations of three different ML models and three different FSA. In total nine experiments are reported. It can be seen that the Bag Ensembles Tree (BET) model consistently outperforms the Decision Tree (DT) and Gaussian Process Regression (GPR) in all cases. But the combination of BET with the MRMR feature selection algorithm gives the best model. The combination of 44 top features gives an R of 0.90, MAE of 22.90 and RMSE of 30.74.

Figure 8 shows the feature importance plot of the top 44 features as decided by the MRMR technique. It can be observed that demographic and clinical features are among the top features. Various coefficients of MFCC features are also present in the top 44 features. On top of that, statistical features such as the Kurtosis of the signal were also deemed important. This is the main novelty of this work, as it has for the first time reported clinically relevant and important features in subject-independent blood glucose estimation. All other recent works are relying on PPG features alone which are not sufficient to estimate blood glucose reliably.

2) PERFORMANCE ANALYSIS OF THE BEST MODEL

To analyze the performance of the best model for the estimation of blood glucose, three different methods are used: regression plot, Bland-Altman plot and Clarke's Error Grid Analysis. With the help of a trendline, the regression plot, as shown in **Figure 9**, allows us to assess how near the predictions are to the ground truth. The better the model, the

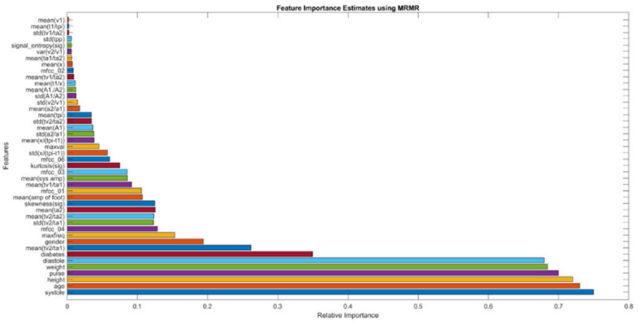

FIGURE 8. Feature importance plot for MRMR algorithm.

closer the trendline is to the y = x line. The regression plot, in this case, shows that the predictions are close to the ground truth which is validated as it has an R of 0.90.

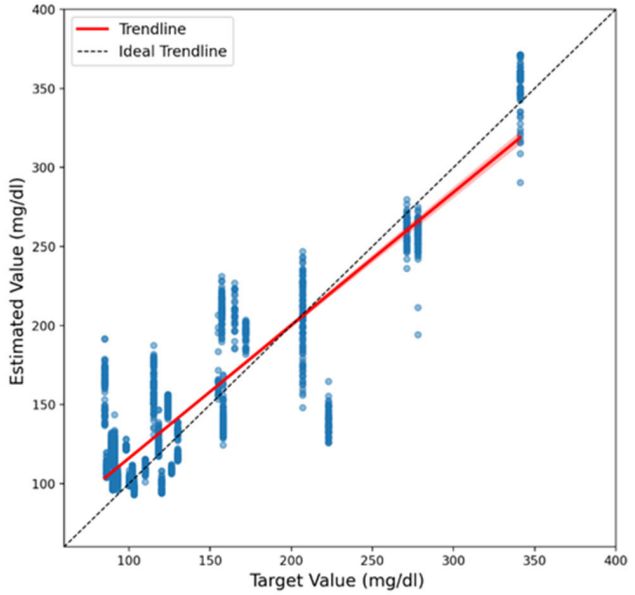

FIGURE 9. Regression plot for the predicted blood glucose level.

As illustrated in **Figure 10**, the Bland-Altman plot shows the data spread as well as the 95 percent limit of agreement (LOA), with a smaller LOA indicating a better model. In the Bland-Altman plot, the LOA is from 47.41 to −66.89 mg/dl. The spread is smaller considering the model was tested on completely independent test subjects.

To assess the clinical precision of BGL monitoring devices, the Clarke error grid analysis was proposed [38]. It is considered a gold standard for assessing the performances of BGL measuring devices. There are five zones (A to E) on a Clarke Error Grid. Each zone is interpreted differently in relation to the prediction error of the model. Zone A, for instance, reflects prediction errors of fewer than 20% or clinically accurate choices. The Clarke error grid analysis of the proposed system is shown in **Figure 11**. The total number and percentage of points in each region are given in **Figure 11**. Using the proposed method, 61.88 % of points are located in zone A. The remaining 38.12% of the data falls into zone B, which refers to an estimated BGL that differs from the actual BGL by more than 20%. This zone leads to incorrect clinical decisions, but they are uncritical. Zone C indicates hyperglycemia or hypoglycemia false positive findings. This zone's points result in unneeded treatment. The areas D and E are crucial.

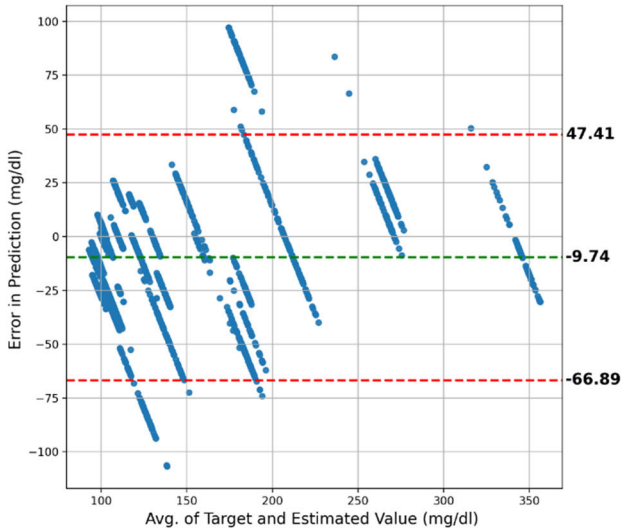

FIGURE 10. Bland-Altman analysis for the predicted blood glucose.

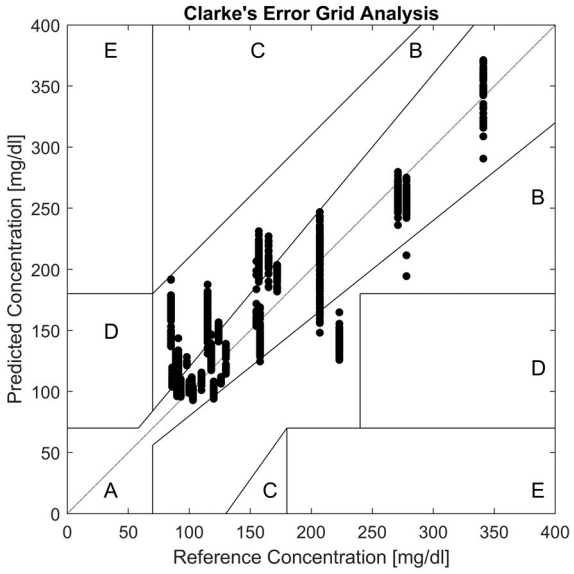

FIGURE 11. Clarke's Error Grid Analysis for estimating blood glucose.

When BGL is in the hyperglycemic or hypoglycemic range, the points in area D correspond to the prediction of normal BGL. This means that the patient is flagged as healthy despite having problems with their blood glucose. Zone E denotes predictions that result in the opposite BGL treatment need. No points fall in zones C, D, and E. Given that every point falls inside regions A and B, the suggested system is clinically acceptable.

TABLE 4. Clarke's error grid analysis - zone breakdown.

Zone	A	B	C	D	E
Number of Samples	1552	956	0	0	0
Percentage of Samples (%)	61.88	38.12	0	0	0

TABLE 5. Performance of various feature selection algorithms with different ML models for classification.

FSA	Num Features	KNN	SVM	DT
Relieff	36	97.28	96.13	95.59
MRMR	17	98.28	97.31	97.77
CFS	26	96.28	95.34	95.98

B. CLASSIFICATION OF DIABETIC SEVERITY

1) FINDING THE BEST ML AND FSA COMBINATION

Table 5 shows the result from the combinations of three different ML models and three different FSA. In total nine experiments are reported. The testing accuracy of the three classifiers (KNN, SVM, and DT) was 98.28%, 97.31%, and 97.77%, respectively. A total of 12,399 samples (from 139 individuals) were used to train KNN, SVM, and DT classifiers. The labels of the training samples were set to be 0, 1, and 2, representing the normal (G1), warning (G2) and dangerous (G3) blood glucose groups, respectively. Stratified five-fold cross-validation was used in the experiment to reduce the impact of sample partitioning. The implementation result is listed in the Table. The number of neighbours parameter "k" is the most essential adjustable parameter for the KNN algorithm. This parameter's tested values were 1, 3, 5, 7, 15 and 30 where the best-performing parameter is k=7. Observing the feature importance plots (as shown in **Figure 12**) for multi-class classification tasks, time domain, statistical, MFCC and demographic features of the PPG signal are the dominant features in determining the severity of diabetes.

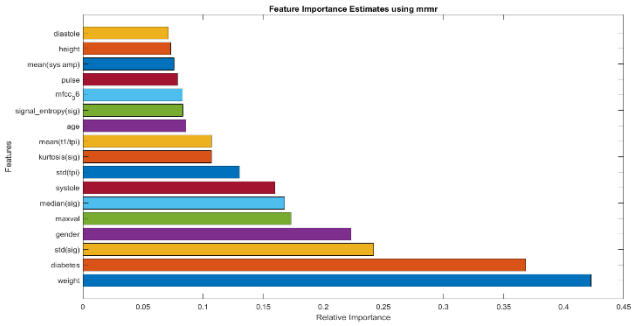

FIGURE 12. Top 17 features as chosen by MRMR along with their feature importance.

2) PERFORMANCE ANALYSIS OF THE BEST MODEL

One of the best ways of analyzing the performance of classification models is to plot the ROC curve. A ROC curve with a higher area under the curve means better performance. In this paper, the ROC curves are plotted for the top 20 important features in **Figure 13**. This plot is to show the impact of the newly added features on the model.

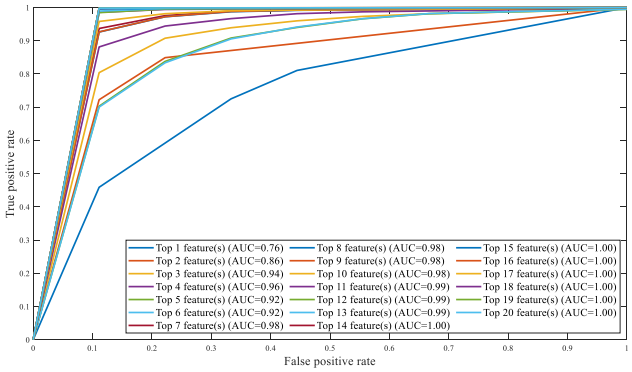

FIGURE 13. ROC curve showing the change in performance as new features are added.

3) PERFORMANCE COMPARISON

The proposed system is compared with the current work in the literature for estimating blood glucose and classifying diabetic severity. When comparing the performances of state-of-the-art technologies for estimating glucose, most of the works are using random stratification, which leaks the training and testing dataset. None of these studies reported subject-wise stratification, which can make their investigation inaccurate and the model performance on unseen subjects unreliable. However, in this work, it has been strictly ensured that the training and testing sets have separate subjects, which ensures the robustness of the model. We have tried random stratification which results in $R = 0.99$ whereas, on an independent test set, the performance cannot be improved more than 0.90 with additional clinical information. Without the extra clinical data, the performance of the model drops significantly. Since the other authors neither shared their data nor shared the frameworks they used for the investigation, it is not possible to independently evaluate their proposed algorithm. Since we have developed the end-to-end framework of data collection and real-time evaluation, it is easy for us to assess our model performance on real-world data.

In the case of Diabetic severity classification, only one work has shown promising results. Zhang [23] used 21-time domain features along with Gaussian SVM to achieve an accuracy of 81.49% in the three-class problem. Our approach outperforms it significantly as we achieved an accuracy of 98.12%. Estimating glucose values accurately is a quite challenging problem compared to classifying the subject into a group. Therefore, our model performed significantly well in this problem. This model should, therefore, be quite accurate in informing users about the severity of their diabetes.

There are some limitations of the current work. In the proposed work, we have used PPG signal-based features which can be extracted in real-time from the patients using the prototype we have made, however, the SBP and DBP values need to be entered into the mobile application using arm or wrist-worn blood pressure device by the user, which limits its application as a wearable device in its current form. However, the authors have recently developed [42], [43], [44] several deep-learning frameworks to estimate arterial blood

TABLE 6. Comparison with current work on estimating blood glucose.

Authors	Year	Data Description	Feature Used	Model	Performance
E. Monte More [39]	2011	"1 minute PPG using the iPod Digital Oximeter fingertip pulse oximeter. 5-second segments are used to extract features. 410 participants (213 M,197 F) Age: 9-80."	"Kaiser–Teager energy, HR statistics, and log energy profile of the PPG signal, Oxygen saturation range, autoregressive (AR) coefficient of the PPG, age, weight, and BMI (21 features)"	RFR	R = 0.90
S. Habbu et al., [40]	2019	"1 minute of PPG signal is used for feature extraction. 611 individuals (344 men and 267 women) aged 4 to 70."	"Time and frequency domain features and single pulse analysis features of the PPG signal constitute two feature sets (28 features)"	ANN	R = 0.95
Sumaiya et al., [45]	2020	"Used a Google Pixel 2 smartphone to record 10-plus seconds of footage at 60 frames per second; only 600 frames were saved from each clip (10 subject)."	"Time domain features with 1st and derivative of PPG waveform."	DNN	R= 0.92
Gupta et al., [18]	2021	"Four minutes' worth of both reflective and transmitting PPG signals are taken. At the same time, glucose levels are checked with a device called Caresens II Plus, which is invasive (26 subject)."	"Kaiser–Teager energy, HR statistics, and log energy profile of the PPG signal, Oxygen saturation range, autoregressive (AR) coefficient of the PPG, HR, breathing rate and BMI."	XGB regressor	R= 0.94
G.A. Alonso Silverio et al., [41]	2021	"Light-dependent resistor detects 650-nm laser radiation via fingertip. Used 4s signal for feature extraction. 122 people (58 M, 64 F)."	"Three feature sets: MFCC, PCA and ICA. MFCC- 13 coefficients PCA-First 3 components ICA- First 3 components"	Adaboost	MAE= 11.62
Proposed Method	**2022**	**Finger PPG captured by Pulse Sensors. 4 second segments are used for feature extraction. 139 subject (58 female, 81 male)**	**25 Time domain features PPG, 1st, and 2nd derivative features. 5 MFCC features and 7 demographic features (age, height, weight, gender, systole, diastole, and diabetes) (Total 44 features)**	**BET**	**R = 0.90 MAE = 22.90 RMSE=30.74**

TABLE 7. Comparison of diabetic severity classification with recent works.

Authors	Feature Used	Model	Precision	Sensitivity	F1	AUC	Accuracy
Gaobo Zhang et al., [23]	21 Time domain features	GSVM	80.84%	79.58%	83.19%	80.2%	81.49%
Proposed Method	**10 PPG signal and 1st and 2nd derivative features. 7 demographic features**	**KNN**	**98.12%**	**98.12%**	**98.12%**	**0.998**	**98.12%**

pressure (ABP) or SBP and DBP values from the PPG signals. This feature can be easily embedded in the current end-to-end framework to estimate SBP and DBP from the PPG signals. We could not embed this feature in the current work as we need a benchmark wrist PPG and corresponding ABP dataset to train the deep-learning model, which is the future direction of our work.

IV. LIMITATION

The main limitation of this study is the small sample size, which consisted of only 139 subjects, on which the proposed best machine learning model was evaluated. As a result, the generalizability of the findings may be limited. To overcome this limitation, we plan to collect more data in future studies to improve the model's performance and generalization to a broader population. Scattering of light in biological tissues is a complex phenomenon that can significantly affect PPG signals. At 609 nm, both absorption and scattering contribute to the signal. Techniques like adaptive filtering, wavelet transforms, or Fourier analysis can be employed to filter out the noise and enhance the signal quality for avoiding the scattering effect which is different from MA of PPG Signals. Using multiple wavelengths or multi-site PPG measurements can help differentiate between scattering effects and actual physiological changes. Analyzing the system response across genders might reveal physiological differences. For instance, factors like skin thickness, vascular structure, and hemoglobin concentration can vary between males and females. This will be considered in the future studies.

V. CONCLUSION

Patients diagnosed with hyperglycemia need to have their glucose levels monitored on a regular basis for their disease to be adequately treated. The methods that are currently used to test for glucose continue to rely on invasive procedures that

are not only uncomfortable but also raise the risk of infection. We have developed a wristband device to collect PPG signals from the wrist of the user. We were able to collect PPG signals, blood pressure and demographic data from 139 control and diabetic subjects. After that, PPG signals along with other parameters were used to estimate blood glucose and the severity of diabetes complications. This paper proposes the design of a blood glucose estimation system that does not involve any invasive procedures. The system made use of new Mel frequency cepstral coefficients, time, frequency, and time-frequency domain features from PPG and their derivative signal, as well as physiological and demographic parameters. We employed approaches for feature selection to reduce the amount of processing that was required and to protect against overfitting the machine learning algorithms. Machine learning algorithms were developed and tested with regression and classification features. These features were used in the training process. In estimating blood glucose level (BGL), the Bagged Ensemble Trees (BET) method, in conjunction with the MRMR feature selection technique, surpasses previous algorithms by achieving a correlation coefficient of 0.90. The predictions from the models were 100% within the acceptable range as defined by Clarke Error Grid. In addition to that, the proposed model (KNN) can classify the severity level (normal, warning, and dangerous) with an accuracy of 98.12%. As the best-performing models were ultimately implemented in an AWS server along with an android mobile application, the users can use it as a wearable system with an IoT application. This way the users will be able to acquire a better picture of their health thanks to the combination of estimating their blood glucose level and the severity of their diabetes condition.

ACKNOWLEDGMENT

This work was made possible by NPRP grant # 13S-0205-200270 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

ETHICAL CONSIDERATION

The study was conducted with the ethical approval of Hamad Medical Corporation (IRB-HMC-2021-011). Written informed consent was taken from all the subjects before collecting the data.

REFERENCES

[1] (2022). *IDF Diabetes Atlas | Tenth Edition*. Accessed: Aug. 23, 2022. [Online]. Available: https://diabetesatlas.org/

[2] (2022). *Ministry of Public Health*. Accessed: Aug. 23, 2022. [Online]. Available: https://www.moph.gov.qa/english/strategies/Supporting-Strategies-and-Frameworks/nationaldiabetesstrategy/Pages/default.aspx

[3] (2022). *5 of the Best Glucose Monitors: Products, Choosing, and More*. Accessed: Aug. 24, 2022. [Online]. Available: https://www.medicalnewstoday.com/articles/best-glucose-monitors#types

[4] *Blood Glucose Monitoring System Market Size & Growth [2028]*. Accessed: Aug. 24, 2022. [Online]. Available: https://www.fortunebusinessinsights.com/industry-reports/blood-glucose-monitoring-market-100648

[5] (2022). *Rite Aid TrueMETRIX Meter*. Accessed: Aug. 24, 2022. [Online]. Available: https://www.riteaid.com/shop/rite-aid-truemetrix-meter-0367379

[6] Walgreens. (2022). *Walgreens TrueMetrix Bluetooth Blood Glucose Meter*. Accessed: Aug. 24, 2022. [Online]. Available: https://www.walgreens.com/store/c/walgreens-truemetrix-bluetooth-blood-glucose-meter/ID=prod6270727-product

[7] Medtronic Diabetes. (2022). *Guardian Connect Continuous Glucose Monitoring System*. Accessed: Aug. 24, 2022. [Online]. Available: https://www.medtronicdiabetes.com/products/guardian-connect-continuous-glucose-monitoring-system

[8] *K'Watch Glucose—PKVitality*, PKvitality, Le Kremlin-Bicêtre, France, Nov. 2021.

[9] *GlucoWise—Non-invasive Blood Glucose Meter By Mediwise: Medicalexpo*. Accessed: Nov. 4, 2021. [Online]. Available: https://www.medicalexpo.com/prod/mediwise/product-105186-686902

[10] Healthline. (2022). *Abbott's FreeStyle Libre Diabetes Tech Hits Milestone Moment*. Accessed: Aug. 24, 2021. [Online]. Available: https://www.healthline.com/diabetesmine/abbott-freestyle-libre3-technology

[11] Healthline. (2022). *7 Best Glucose Monitors and Meters of 2022*. Accessed: Aug. 24, 2021. [Online]. Available: https://www.healthline.com/health/diabetes/best-glucose-monitors#comparison

[12] C. Xu, D. Jiang, Y. Ge, L. Huang, Y. Xiao, X. Ren, X. Liu, Q. Zhang, and Y. Wang, "A PEDOT:PSS conductive hydrogel incorporated with Prussian blue nanoparticles for wearable and noninvasive monitoring of glucose," *Chem. Eng. J.*, vol. 431, Mar. 2022, Art. no. 134109.

[13] S. Cai, C. Xu, D. Jiang, M. Yuan, Q. Zhang, Z. Li, and Y. Wang, "Air-permeable electrode for highly sensitive and noninvasive glucose monitoring enabled by graphene fiber fabrics," *Nano Energy*, vol. 93, Mar. 2022, Art. no. 106904.

[14] D. Li, R. Tan, X. Mi, C. Fang, and Y. Tu, "An electrochemiluminescent biosensor for noninvasive glucose detection based on cluster-like AuAg hollowed-nanoparticles," *Microchemical J.*, vol. 167, Aug. 2021, Art. no. 106271.

[15] D. Tankasala and J. C. Linnes, "Noninvasive glucose detection in exhaled breath condensate," *Transl. Res.*, vol. 213, pp. 1–22, Nov. 2019, doi: 10.1016/j.trsl.2019.05.006.

[16] Z. Li, G. Li, W.-J. Yan, and L. Lin, "Classification of diabetes and measurement of blood glucose concentration noninvasively using near infrared spectroscopy," *Infr. Phys. Technol.*, vol. 67, pp. 574–582, 2014.

[17] G. Han, S. Chen, X. Wang, J. Wang, H. Wang, and Z. Zhao, "Noninvasive blood glucose sensing by near-infrared spectroscopy based on PLSR combines SAE deep neural network approach," *Infr. Phys. Technol.*, vol. 113, Mar. 2021, Art. no. 103620, doi: 10.1016/j.infrared.2020.103620.

[18] S. S. Gupta, T. Kwon, S. Hossain, and K. Kim, "Towards non-invasive blood glucose measurement using machine learning: An all-purpose PPG system design," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102706.

[19] A. Prabha, J. Yadav, A. Rani, and V. Singh, "Intelligent estimation of blood glucose level using wristband PPG signal and physiological parameters," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103876.

[20] F. Riaz, M. A. Azad, J. Arshad, M. Imran, A. Hassan, and S. Rehman, "Pervasive blood pressure monitoring using photoplethysmogram (PPG) sensor," *Future Gener. Comput. Syst.*, vol. 98, pp. 120–130, Sep. 2019.

[21] M. A.-U. Golap, S. M. T. U. Raju, M. R. Haque, and M. M. A. Hashem, "Hemoglobin and glucose level estimation from PPG characteristics features of fingertip video using MGGP-based model," *Biomed. Signal Process. Control*, vol. 67, May 2021, Art. no. 102478.

[22] T. T. Islam, M. S. Ahmed, M. Hassanuzzaman, S. A. B. Amir, and T. Rahman, "Blood glucose level regression for smartphone PPG signals using machine learning," *Appl. Sci.*, vol. 11, no. 2, p. 618, Jan. 2021.

[23] G. Zhang, Z. Mei, Y. Zhang, X. Ma, B. Lo, D. Chen, and Y. Zhang, "A noninvasive blood glucose monitoring system based on smartphone PPG signal processing and machine learning," *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 7209–7218, Nov. 2020.

[24] (2022). *TinyPICO*. Accessed: Aug. 16, 2022. [Online]. Available: https://www.tinypico.com/

[25] (2022). *3.7V Rechargeable Li-Polymer Battery LP954575 4000mAh Without PCM and Wires*. Accessed: Aug. 26, 2022. [Online]. Available: https://www.li-polymer-battery.com/3-7v-rechargeable-li-polymer-battery-lp954575-4000mah-without-pcm-and-wires/

[26] (Jan. 16, 2017). *Lithium-ion Polymer Battery Specification*. [Online]. Available: https://www.olimex.com/Products/Power/BATTERY-LIPO250mAh/resources/JA602025P-Spec-Data-Sheet-3.7V-250mAh-170116.pdf
[27] Android Developers. (2022). *Bluetooth Low Energy| Android Developers*. Accessed: Aug. 26, 2022. [Online]. Available: https://developer.android.com/guide/topics/connectivity/bluetooth/ble-overview
[28] Z. Abbas, K. Kanav, R. Mortensen, and A. Libanabil. (2022). *BU-409: Charging Lithium-ion*. Accessed: Aug. 26, 2022. [Online]. Available: https://batteryuniversity.com/article/bu-409-charging-lithium-ion
[29] G. Free. (2022). *BU-808c: Coulombic and Energy Efficiency With the Battery*. Battery University. Accessed: Aug. 26, 2022. [Online]. Available: https://batteryuniversity.com/article/bu-808c-coulombic-and-energy-efficiency-with-the-battery
[30] (2022). *Li-Ion & LiPoly Batteries*. Accessed: Aug. 26, 2022. [Online]. Available: https://learn.adafruit.com/li-ion-and-lipoly-batteries/protection-circuitry
[31] M. N. I. Shuzan, M. H. Chowdhury, M. S. Hossain, M. E. H. Chowdhury, M. B. I. Reaz, M. M. Uddin, A. Khandakar, Z. B. Mahbub, and S. H. M. Ali, "A novel non-invasive estimation of respiration rate from motion corrupted photoplethysmograph signal using machine learning model," *IEEE Access*, vol. 9, pp. 96775–96790, 2021.
[32] M. H. Chowdhury, M. N. I. Shuzan, M. E. H. Chowdhury, Z. B. Mahbub, M. M. Uddin, A. Khandakar, and M. B. I. Reaz, "Estimating blood pressure from the photoplethysmogram signal and demographic features using machine learning techniques," *Sensors*, vol. 20, no. 11, p. 3127, Jun. 2020.
[33] B. Logan, "Mel frequency cepstral coefficients for music modeling," in *Proc. Int. Symp. Music Inf. Retrieval*, 2000, p. 11.
[34] R. Zatz and B. M. Brenner, "Pathogenesis of diabetic microangiopathy. The hemodynamic view," *Amer. J. Med.*, vol. 80, no. 3, pp. 443–453, Mar. 1986.
[35] Y. Arpitha, G. L. Madhumathi, and N. Balaji, "Spectrogram analysis of ECG signal and classification efficiency using MFCC feature extraction technique," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 2, pp. 757–767, Feb. 2022.
[36] G. Rajesh, "Analysis of mfcc features for EEG signal classification," *Int. J. Adv. Signal Image Sci.*, vol. 2, no. 2, p. 14, Dec. 2016.
[37] D. M. Nogueira, C. A. Ferreira, E. F. Gomes, and A. M. Jorge, "Classifying heart sounds using images of motifs, MFCC and temporal features," *J. Med. Syst.*, vol. 43, no. 6, p. 168, Jun. 2019.
[38] W. L. Clarke, D. Cox, L. A. Gonder-Frederick, W. Carter, and S. L. Pohl, "Evaluating clinical accuracy of systems for self-monitoring of blood glucose," *Diabetes Care*, vol. 10, no. 5, pp. 622–628, Sep. 1987.
[39] E. Monte-Moreno, "Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques," *Artif. Intell. Med.*, vol. 53, no. 2, pp. 127–138, Oct. 2011.
[40] S. Habbu, M. Dale, and R. Ghongade, "Estimation of blood glucose by non-invasive method using photoplethysmography," *Sadhana*, vol. 44, no. 6, p. 135, 2019.
[41] G. A. Alonso-Silverio, V. Francisco-García, I. P. Guzmán-Guzmán, E. Ventura-Molina, and A. Alarcón-Paredes, "Toward non-invasive estimation of blood glucose concentration: A comparative performance," *Mathematics*, vol. 9, no. 20, p. 2529, Oct. 2021.
[42] S. Mahmud, N. Ibtehaz, A. Khandakar, M. S. Rahman, A. J. Gonzales, T. Rahman, M. Shafayet Hossain, M. S. A. Hossain, M. A. A. Faisal, F. F. Abir, F. Musharavati, and M. E. H. Chowdhury, "NABNet: A nested attention-guided BiConvLSTM network for a robust prediction of blood pressure components from reconstructed arterial blood pressure waveforms using PPG and ECG signals," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104247.
[43] N. Ibtehaz, S. Mahmud, M. E. H. Chowdhury, A. Khandakar, M. Arselene Ayari, A. Tahir, and M. Sohel Rahman, "PPG2ABP: Translating photoplethysmogram (PPG) signals to arterial blood pressure (ABP) waveforms using fully convolutional neural networks," 2020, *arXiv:2005.01669*.
[44] S. Mahmud, N. Ibtehaz, A. Khandakar, A. M. Tahir, T. Rahman, K. R. Islam, M. S. Hossain, M. S. Rahman, F. Musharavati, M. A. Ayari, M. T. Islam, and M. E. H. Chowdhury, "A shallow U-Net architecture for reliably predicting blood pressure (BP) from photoplethysmogram (PPG) and electrocardiogram (ECG) signals," *Sensors*, vol. 22, no. 3, p. 919, Jan. 2022.
[45] J. Sumaiya, M. R. Hasan, and E. Hossain, "Noninvasive blood glucose measurement using live video by smartphone," in *Proc. IEEE 8th R10 Humanitarian Technol. Conf.*, Dec. 2020, pp. 1–6.

MD NAZMUL ISLAM SHUZAN received the B.Sc. degree in electrical and electronics engineering from North South University, Bangladesh. He is currently pursuing the M.Sc. degree with Universiti Kebangsaan Malaysia (UKM). He is also a Research Assistant with Qatar University (QU). His current research interests include biomedical signal processing, machine learning, and data science.

MOAJJEM HOSSAIN CHOWDHURY received the B.Sc. degree in electrical and electronics engineering from North South University, Bangladesh. He is currently pursuing the M.Sc. degree with Universiti Kebangsaan Malaysia (UKM). He is also a Research Assistant with Qatar University (QU). His current research interests include biomedical signal processing, machine learning, and data science. He placed in the top 30 in Bangladesh Physics Olympiad 2014.

MUHAMMAD E. H. CHOWDHURY (Senior Member, IEEE) received the Ph.D. degree from the University of Nottingham, U.K., in 2014. He was a Postdoctoral Research Fellow with the Sir Peter Mansfield Imaging Centre, University of Nottingham. He is currently an Assistant Professor and a Program Coordinator of the Department of Electrical Engineering, Qatar University. He has filed several patents and published more than 180 peer-reviewed journal articles, more than 30 conference papers, and several book chapters. His current research interests include biomedical instrumentation, signal processing, wearable sensors, medical image analysis, machine learning and computer vision, embedded system design, and simultaneous EEG/fMRI. He is also running NPRP, UREP, and HSREP grants from Qatar National Research Fund (QNRF) and internal grants (IRCC and HIG) from Qatar University along with academic projects from HBKU and HMC. He is a member of British Radiology, ISMRM, and HBM. He is serving as a Guest Editor for *Polymers*, an Associate Editor for IEEE Access, and a Topic Editor and a Review Editor for *Frontiers in Neuroscience*. He has recently won the COVID-19 Dataset Award, AHS Award from HMC, and National AI Competition Awards for his contribution to the fight against COVID-19. His team is the gold-medalist in the 13th International Invention Fair in the Middle East (IIFME). He has been listed among the Top 2% of Scientists in the World List, published by Stanford University.

KHALID ABUALSAUD (Senior Member, IEEE) received the B.S. degree in systems and computer engineering, in 1988, and the M.S. and Ph.D. degrees in computer science from University Utara Malaysia, in 2009 and 2015, respectively. Currently, he is a Lecturer in CE with the College of Engineering, Qatar University. He has more than 25 years of professional experience in information technology. His research work has been presented in international conferences and journals. He has authored more than 95 publications in refereed conferences and journals. His research interests include health systems, wireless sensors for IoT applications, cybersecurity, cloud computing, and computer networks. He received several awards from different local and international organizations, in addition to four IEEE Conference Best Paper Awards (IWCMC'19, ICIoT20, IWCMC'23, and ISNCC'23). He served as a technical reviewer for different international IEEE, ACM, and Elsevier journals. He is active in getting research funding from various sources, including the QNRF, SCDL (FIFA'2022), and some other organizations inside Qatar. He was the LPI of NPRP projects and also a PI in numerous internal research grants, which achieved significant outcomes. Recently, he served as a Guest Editor in Connected Healthcare Special Issue of *IEEE Network Magazine*. He is currently an Associate Editor of the *IET Quantum Communication* journal.

ELIAS YAACOUB received the B.E. degree in electrical engineering from Lebanese University, in 2002, and the M.E. degree in computer and communications engineering and the Ph.D. degree in electrical and computer engineering from American University of Beirut (AUB), in 2005 and 2010, respectively. He was a Research Assistant with AUB, from 2004 to 2005, and Munich University of Technology, in Spring 2005. From 2005 to 2007, he was a Telecommunications Engineer with Dar Al-Handasah, Shair and Partners. From November 2010 to December 2014, he was a Research Scientist/Research and Development Expert with the Qatar Mobility Innovations Center (QMIC), where he led the Broadband Wireless Access Technology Team. Afterwards, he joined Strategic Decisions Group (SDG), where he was a Consultant, until February 2016. Then, he joined Arab Open University (AOU) as an Associate Professor and a Coordinator of the M.Sc. Program in Information Security and Forensics. Between February 2018 and August 2019, he was an Independent Researcher/Consultant. He was also affiliated with AUB as a part-time Faculty Member. He joined Qatar University as an Associate Professor with the Computer Science and Engineering Department, in August 2019, and has been a Full Professor, since July 2023. His research interests include wireless communications, 5G/6G and beyond wireless networks, the Internet of Things (IoT), mobile Health (mHealth), antenna theory, and physical layer security.

MD AHASAN ATICK FAISAL received the Bachelor of Science and Master of Science degrees in electrical and electronic engineering from the University of Dhaka, Bangladesh. He is currently pursuing the Ph.D. degree in computer vision with Qatar University. His research interests include gait analysis, sign language recognition (SLR), and the practical applications of machine learning. Additionally, he possesses expertise in Internet of Things (IoT)-based solutions, mobile application development, and MLOps.

MAZUN ALSHAHWANI received the B.Sc. degree in computer engineering from the Department of Computer Science and Engineering, Qatar University, in December 2022. Presently, she serves as a Projects Manager at the Ministry of Social Development and Family (MSDF), Qatar. Her role has been started in June 2023 at the MSDF. She was on the College of Engineering Dean's list recognition two times during her studies at Qatar University and holds one invention. The invention was a Wearable Non-invasive Blood Glucose Monitoring System that received approval from the United States Patent and Trademark Office. This invention won a Gold Medal at the International Invention Fair in the Middle East, Kuwait, in February 2023. She exceptional skills and commitment have been recognized as the author of the invention. She uses her experience to modernize operations and enhance the productivity of her organization's performance. Eventually, her technology work field is evident in her published research work, which demonstrates her technical expertise and explores the limitless possibilities of technology.

NOORA AL BORDENI received the B.Sc. degree in computer engineering from the Department of Computer Science and Engineering, Qatar University, in June 2022. Currently, she is a Section Head of IT Operations at the Ministry of Education and Higher Education since January 2024. She worked as a System Engineer from August 2023 to December 2023. She was working as an IT Teacher at Amna Bint Wahab Middle School from September 2022 until June 2023. She leverages her experience to modernize processes, boost efficiency, and enhance organizational performance. She was on the College of Engineering Dean's list recognition three times during her studies at Qatar University and holds two inventions. The first invention was a Wearable Non-invasive Blood Glucose Monitoring System that received approval from the United States Patent and Trademark Office. This invention won a Gold Medal at the International Invention Fair in the Middle East, Kuwait 2023. The second invention was the Safety Wristband: Detecting Falls and Ensuring Safety in Emergencies. Intellectual property rights were secured from the Intellectual property Rights Protection Department, Ministry of Commerce and Industry, Qatar. This invention earned first place among all schools, in the State of Qatar, received a special award from Shell, and participated in the ISEF 2024 exhibition in the United States of America in May 2024.

FATIMA AL-KAABI holds the Bachelor of Science degree in computer engineering from Qatar University in 2022. She has worked as an Author, Project Manager and Software Developer at the General Authority of Customs, and is recognized for his renowned contribution. She was selected to participate in judging the Schools Robotics Championship 2023, as well as the 2024 Qatar University students' graduation projects. She won the gold medal in the International Competition for Inventions in the Middle East. She has experience in the field of cybersecurity after completing the course "Cybersecurity: Information Management in the Information Age" from Harvard University.

SARA AL-MOHANNADI received the B.Sc. degree in computer engineering from the Department of Computer Science and Engineering, Qatar University, in December 2022. She was on the College of Engineering Dean's list recognition two times during her studies at Qatar University. Currently, she has been an IT System Engineer at Public Works Authority (Ashghal) since July, 2023. She works at Networks and Telecommunications Department. Her graduation invention won a Golden Medal at the International Invention Fair in the Middle East, Kuwait, in February, 2023. She was a robotics arbitrator for Championships for schools in March 2023. Moreover, she has Cisco Certified Network Associate (CCNA), in 2023. Eventually, she eager to apply expertise in software development, hardware design, and problem-solving to drive technological advancements.

SAKIB MAHMUD received the Bachelor of Science (B.Sc.) (Hons.) and Master of Science (M.Sc.) degrees in electrical engineering from Qatar University (QU), in June 2020 and May 2023, respectively. As an undergraduate student, he received the Dean's Award for three consecutive years (2016–2019). While with QU as an Undergraduate, he participated in two Undergraduate Research Experience Program (UREP) projects funded by Qatar National Research Fund (QNRF). Subsequently, he secured employment for more than five high-impact grants from Qatar University (QU), Hamad Medical Corporation (HMC), and Qatar National Research Fund (QNRF). He successfully published around 38 peer-reviewed journal articles in high-impact journals (mostly Q1 and some Q2) and three conference papers. His field of expertise mainly include, but is not limited to, machine learning, biomedical instrumentation, signal processing, image processing, bioinformatics, and 3D modeling.

NIZAR ZORBA (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from JUST University, Jordan, in 2002, and the Ph.D. degree in signal processing for communications from UPC Barcelona, Spain, in 2007. He is currently a Professor with the Electrical Engineering Department, Qatar University, Doha, Qatar. He has authored six international patents, two books, contributed to seven book chapters, and coauthored over 160 papers in peer-reviewed journals and international conferences. He is also an IEEE ComSoc Distinguished Lecturer (2024–2025). He is the Former Chair of the IEEE Communication Society Technical Committee on Communications Systems Integration and Modelling. He is also the Associate Editor-in-Chief of *IEEE Communications Magazine*, an Area Editor of the IEEE COMMUNICATIONS LETTERS, and an Associate Editor of IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING.

Open Access funding provided by 'Qatar National Library' within the CRUI CARE Agreement