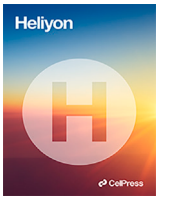
Heliyon 10 (2024) e28720

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Non-invasive glucose prediction and classification using NIR technology with machine learning

M. Naresh [a], V. Siva Nagaraju [b], Sreedhar Kollem [c], Jayendra Kumar [a], Samineni Peddakrishna [a,*]

[a] School of Electronics Engineering, VIT-AP University, Amaravti, Guntur, 522241, Andhra Pradesh, India
[b] Department of ECE, Institute of Aeronautical Engineering, Dundigal, Hyderabad, 500043, Telangana, India
[c] Department of ECE, School of Engineering, SR University, Warangal, 506371, Telangana, India

ARTICLE INFO

Keywords:
Absorbance
Spectroscopy
Glucose
Infrared
Machine learning
Noninvasive
Regression
Classification
Detectors

ABSTRACT

In this paper, a dual wavelength short near-infrared system is described for the detection of glucose levels. The system aims to improve the accuracy of blood glucose detection in a cost-effective and non-invasive way. The accuracy of the method is evaluated using real-time samples collected with the reference finger prick glucose device. A feed forward neural network (FFNN) regression method is employed to predict glucose levels based on the input data obtained from NIR technology. The system calculates glucose evaluation metrics and performs Surveillance error grid (SEG) analysis. The coefficient of determination R^2 and mean absolute error are observed 0.99 and 2.49 mg/dl, respectively. Additionally, the system determines the root mean square error (RMSE) as 3.02 mg/dl. It also shows that the mean absolute percentage error (MAPE) is 1.94% and mean squared error (MSE) is 9.16 $(mg/dl)^2$ for FFNN. The SEG analysis shows that the glucose values measured by the system fall within the clinically acceptable range when compared to the reference method. Finally, the system uses the multi-class classification method of the multilayer perceptron (MLP) and K-nearest neighbors (KNN) classifier to classify glucose levels with an accuracy of 99%.

1. Introduction

The International Diabetes Federation (IDF) has stated that diabetes mellitus is a chronic disease that significantly affects global health. Globally, diabetes caused 6.7 million deaths in 2021, with approximately 537 million adults between the ages of 20 and 79 living with the disease. The number of patients with diabetes is projected to increase by 46% by 2045 [1,2]. Regularly monitoring persons who are in a prediabetic condition, have gestational diabetes, or have reduced glucose tolerance is essential in order to avoid the development of diabetes [3]. Effective management of diabetes can be achieved through artificial insulin intake, exercise, a balanced diet, and regular monitoring of blood glucose levels by test. Hence, self-monitoring their blood glucose each day is an important part of their diabetes management. It can be monitored using various methods such as invasive (IN) [4], minimally invasive (MI), and non-invasive (NI) [5,6].

* Corresponding author.
E-mail address: krishna.samineni@gmail.com (S. Peddakrishna).

https://doi.org/10.1016/j.heliyon.2024.e28720
Received 20 October 2023; Received in revised form 12 March 2024; Accepted 22 March 2024

Available online 28 March 2024

The IN method of blood glucose detection is the most common method of blood glucose detection. A lancet pricks the fingertip to collect a drop of blood on a test strip, which is then inserted into a glucose meter to read the glucose level. However, IN methods of blood glucose detection carry risks of infection, bleeding, and nerve damage. An MI monitoring method, which causes minimal skin damage, offers an alternative way to measure glucose levels [7,8]. However, it needs periodic calibration with standard measurements and has long-term reliability and stability difficulties, making the devices costlier and unreliable for routine blood glucose monitoring. As a result, researchers have explored various NI methods for measuring blood glucose that are painless, reliable, and cost-effective [9]. This advancement has the potential to make regular blood glucose monitoring more comfortable and accessible to millions of people. Both in vitro and in vivo approaches have been proposed for NI blood glucose detection [10].

In vitro methods for NI blood glucose measurement take place outside of a live body, often in a laboratory environment [4]. On the other hand, in vivo methods are performed on a living organism and are more suitable for continuous self-monitoring of blood glucose. Optical technologies are particularly well-suited for glucose detection in the in vivo approach due to the complexity of blood and tissue properties [11]. Additionally, optical technologies are less likely to cause irritation to targeted biological tissue. Recent advancements in optical technologies, such as visible laser light, Raman spectroscopy, mid-infrared (MIR), and near-infrared (NIR) spectroscopy, have provided users with various options for glucose measurement [12]. Among these techniques, NIR spectroscopy is a precise approach for glucose monitoring, and short NIR is superior to long NIR due to its deeper penetration capability underneath the skin to detect glucose molecules.

On the other hand, in the realm of predicting glucose levels and their classification, multiple linear regression (MLR) is commonly used to study the relationship between multiple independent variables and a dependent variable [13]. However, statistical regression techniques can be inefficient when dealing with multiple features. To overcome this limitation, artificial intelligence (AI) and machine learning (ML) have made significant progress in developing models that establish relationships between attributes and instances for classification and regression tasks. The aim of ML is to create accurate predictions by extracting relevant features from existing data and building a model. Regression is a crucial component of ML that employs a trained model to make predictions [14,15].

There are various ML regression algorithms that can be used for this purpose, including linear regression, support vector machine (SVM), K-nearest neighbor (KNN), decision tree, random forest (RF), and adaptive boosting [16]. Linear regression, SVM, KNN, and decision tree are commonly used to solve simple regression problems [17]. In contrast, ensemble methods such as RF and AdaBoost are used to combine the predictions of multiple ML models that are individually weak, resulting in a more accurate forecast [18–20]. In healthcare, ML has the potential to revolutionize disease diagnosis and personalized treatment [21,22]. Numerous ML techniques have been proposed in the literature to extract relevant features from measured or predicted data. For instance, RF has been employed to analyze continuous glucose monitoring data and predict the occurrence of hypoglycemic events in type 1 diabetes patients [23]. Photoplethysmography (PPG) based glucose detection with smartphone using principle component regression, SVM and RF [24]. Additionally, KNN algorithms have been used to classify glucose data based on their similarity to previously observed patterns [25]. Therefore, the use of ML techniques in glucose prediction and classification [26] has yielded promising results, and this approach holds significant potential for improving diabetes management.

Artificial neural networks also belong to the field of machine learning and utilize a set of algorithms to solve both linear and non-linear problems [27]. There exist different types of neural networks, each with its unique structure and application. Multiple neural network models are appropriate for solving regression problems, including the Multi-layer Perceptron (MLP), a feed forward neural network with multiple hidden layers that can perform regression [28].

This study investigates the potential of ML techniques to improve the precision of glucose level measurement using short-wave NIR absorbance and reflectance spectroscopy. We propose a dual-wavelength system that overcomes limitations of single-wavelength approaches by providing a deeper understanding of glucose levels. Specifically, we select wavelengths in both reflective and transmission modes with subtle differences for measuring glucose on human finger tissue. These wavelengths balance glucose absorption, tissue penetration, and resistance to interference from other tissue components. Specifically, light with a wavelength of 940 nm displays a higher absorption coefficient for glucose, while light at 950 nm exhibits enhanced transmission through tissue.

In addition to the dual-wavelength approach, we propose the integration of ML algorithms, particularly feed forward neural network (FFNN) to further enhance the accuracy of glucose measurements. FFNN trained on collected dataset of NIR spectra and corresponding blood glucose levels to identify complex relationships between spectral features and glucose concentration. This data-driven approach can potentially lead to more precise and personalized glucose monitoring. The performance of this ML approach is then compared to that of other fundamental models. Furthermore, two classification techniques, namely MLP and KNN, are utilized to categorize glucose levels into three categories: normal, hypoglycemia, and hyperglycemia.

2. Literature on glucose monitoring technologies

In recent years, there has been progress in glucose monitoring methods, particularly in the development of NI approaches for glucose estimation. These approaches are categorized based on their technologies, which include electromagnetic (EM) wave sensing, transdermal, and enzymatic methods. Enzymatic methods can enable NI monitoring through readily accessible body fluids such as saliva, tear drops, or exhaled breath [29]. Enzymatic methods require the use of specialized sensors that can be costly and may have limited accuracy due to variations in the composition of bodily fluids [30]. Transdermal methods, on the other hand, may be limited by the skin barrier and require special preparations to enhance skin permeability. EM wave sensing, however, does not have these limitations and is capable of providing accurate glucose measurements through various body tissues using NI techniques.

EM sensing encompasses various non-ionizing EM radiation technologies, such as ultraviolet (UV), visible light, microwave (MW)/mm-wave, optical wave spectroscopy technology (OWST). While studies on EM waves sensing with UV can lead to ioniz-

ing radiation and tissue damage [31]. On the other hand, lower frequencies, typically below 100 MHz, are utilized to study the tissue impedance and are directly related to material properties [32]. The technique used to measure this impedance is known as impedance spectroscopy, which involves measuring the skin's impedance using a known alternating current intensity. Although this technique is relatively simple and affordable in practical scenarios, it has limitations as it can be affected by factors such as sweat, water content, and temperature variation.

As an alternative to high and low frequencies, numerous studies have been conducted in MW/mm-wave spectroscopy and OWST. The MW/mm-wave frequency range corresponds to 1 mm to 1 m, and EM wave radiation carries lower photon energy, which relatively avoids tissue damage [33]. These devices are advantageous in terms of portability and cost. However, the dielectric properties of the skin are closely related to the dielectric properties of these sensors, and they vary with changes in glucose levels [34]. It is also a challenging task to obtain the relationship between glucose levels and permittivity, which is one of the cores limiting factors in their development. Therefore, researchers are currently working on improving microwave sensors' selectivity and sensitivity [35].

OWST is another NI glucose monitoring technique that utilizes various optical techniques to measure glucose levels. It encompasses a wide range of technologies, including photoacoustic spectroscopy, occlusion spectroscopy, optical coherence tomography (OCT), optical polarimetry, Raman spectroscopy, thermal emission, MIR, and NIR. Although some of these techniques, such as photoacoustic spectroscopy, occlusion spectroscopy, and OCT, may indirectly involve the use of EM radiation, they primarily rely on other physical principles for glucose measurement. Photoacoustic spectroscopy uses acoustic waves generated by EM radiation absorption to measure glucose. This technique is simple and resistant to water absorption. However, this method is vulnerable to temperature, pressure, and environmental changes, and it has a low signal-to-noise ratio. The instrumentation required for this method is also expensive [36]. Occlusion spectroscopy relies on changes in light scattering and absorption due to changes in blood volume [37]. This technique is limited to a depth of a few millimeters in tissue, which restricts its use for the measurement of deep tissue glucose levels. The OCT uses light interference patterns to generate high resolution images of tissues [38]. The technique has a limited field of view, which means that only a small area can be imaged at once. OCT is a relatively expensive technique, which may limit its availability in some healthcare settings. Mostly, this technique has been widely utilized for various clinical applications [39].

On the other hand, optical polarimetry, thermal emission, Raman spectroscopy, MIR, and NIR directly involve the use of EM radiation. These techniques operate on the principles of light absorption, reflection, refraction, polarization, and scattering to measure glucose levels. The optical polarimetry is the measurement and interpretation of the polarization of light. It is a NI, non-destructive technique that can provide information on the molecular composition, structure, and orientation of biological tissues [40]. This technology has been used for glucose monitoring by measuring the polarimetric changes in the light as it passes through the skin. However, due to light scattering and other physiological factors, it is not suitable for skin measurements. Instead, it can be used on aqueous humor of the eye due to its high resolution and sensitivity to small changes in glucose concentration. Also, this method is sensitive to temperature changes, eye motion, and interferences from other optically active compounds in the eye. The other method called thermal emission [41] involves the measurement of infrared radiation emitted by an object, which is directly related to its temperature. The technique is highly dependent on the temperature of the skin, and any changes in skin temperature can cause variations in the measured signal, leading to inaccuracies in glucose measurement. Also, thermal emission spectroscopy requires sophisticated instrumentation and signal processing algorithms, making it difficult to implement in low-cost, portable glucose monitoring devices.

Raman spectroscopy is another technique that utilizes laser light to excite molecular vibrations in glucose molecules and measures the scattered light to determine glucose concentrations [42]. One of the main challenges is the interference from other biological molecules present in the sample, which can affect the accuracy and specificity of glucose measurements [43]. Additionally, the cost of the equipment required for Raman spectroscopy is relatively high, which may limit its accessibility for some applications [44]. It also requires careful calibration to obtain accurate measurements, which can be time-consuming and require technical expertise.

Despite facing certain challenges, infrared spectroscopy has demonstrated potential in glucose monitoring. Two emerging technologies in this field are mid-infrared (MIR) [45] and near-infrared (NIR) spectroscopy, both of which rely on light absorption and reflection to measure glucose levels. MIR spectroscopy operates in the MIR region of the EM spectrum, which spans from 2500 μm to 25 μm to detect glucose concentration in interstitial fluid through reflection [46–49]. MIR has a high sensitivity for glucose detection due to its sharp response peaks and low scattering. However, it is limited to reflection mode since it cannot penetrate human tissue well, making it incapable of determining glucose levels in blood vessels [50].

In contrast, NIR has good penetration depth in the skin and has been studied for glucose estimation for several decades. NIR spectroscopy and NIR PPG are both techniques that use NIR light to analyze biological samples or signals, but they differ in their specific applications and measurement methods. NIR PPG measures the changes in light absorption or reflection caused by blood flow in tissues. It works by shining NIR light into the tissue and measuring the amount of light that is reflected back to a detector. As blood flows through the tissue, it changes the amount of NIR light that is absorbed or reflected, allowing for the detection of glucose [51]. However, PPG signals have drawbacks such as low signal-to-noise ratio (SNR), sensitivity to motion artifacts, and interference from other physiological factors.

On the other hand, NIR spectroscopy involves illuminating a sample with NIR light and measuring the absorption or reflection of the light by the sample. This can be used to identify the molecular composition of the sample, such as the presence and concentration of different compounds. NIR spectroscopy is divided into long and short wave NIR. NIR waves fall within the EM spectrum range of 700-2500 nm and can be divided into two regions based on its bandwidths, the first overtone band (1400-2000 nm) and the second overtone band (750-1400 nm). Some researchers have presented glucose measurements using long-wave NIR spectroscopy [47]. However, the system has not been designed for continuous glucose monitoring. Another method has been utilized for continuous

glucose monitoring but the narrow penetration of long-wave NIR hinders in vivo testing. In contrast, short NIR wavelengths have excellent penetration for in-vivo testing [52]. This makes NIR a promising technology for NI glucose monitoring. However, its accuracy is hindered by scattering and the limited bandwidth, which results in a poor correlation with blood glucose levels.

Several studies have attempted to improve the accuracy of NIR for glucose monitoring by combining it with other technologies. One study that utilized short-wave NIR in combination with Raman spectroscopy to improve the specificity of glucose measurement [53]. Despite these efforts, NIR technology still faces challenges in achieving the required accuracy and sensitivity for glucose monitoring. Ongoing research and development are focused on addressing these challenges and improving the performance of NIR technology. NIR waves can be used to measure the concentration of functional groups in the blood by partially scattering or absorbing through the skin tissue. The absorption and reflection of NIR waves are related to the molecular vibrations of chemical bonds of molecules present in the medium, including glucose molecules.

Glucose molecules are made up of carbon, hydrogen, and oxygen atoms, which bond together to form C-H and C-O bonds, as well as O-H groups. These chemical bonds interact with near-infrared (NIR) waves, which are capable of penetrating the skin and reaching the blood vessels underneath. When NIR waves encounter glucose molecules, they are either absorbed or reflected by the glucose molecules in a manner that is dependent on the glucose concentration. In a research conducted by [54] examined the accuracy of NI blood glucose measurement using short wavelength NIR spectroscopy. Their results indicated that the detection of a CH vibration at 920 nm did not overlap with the OH vibration of water molecules [55]. Shorter wavelengths can result in higher selectivity to glucose molecules by minimizing the interfering effect of water [56]. In a study developed by Jain et al. used an NI glucose measurement device that uses both transmittance and reflectance modes, with three different wavelengths operating at 940 nm and 1300 nm, and achieved 94% accuracy [57].

Particularly, NI glucose meters that use NIR spectroscopy at 940 nm, show high levels of reliability in detecting glucose levels, offering an innovative way to monitor diabetes mellitus [58,59]. The study conducted by Yadav et al. utilizes a NIR LED emitting at 940 nm (TSAL5300) to minimize interference from other substances, such as water [60]. Similarly, in other works, a wavelength of 950 nm was specifically chosen for an adequate penetration depth [61–63].

Based on above work, to improve accuracy, a continuous glucose monitoring system using short-wave NIR spectroscopy was developed. It uses a dual wavelength of 950 nm and 940 nm NIR technology to achieve an accuracy level of 99% with RMSE of 3.02 mg/dl. The data collected from both the wavelengths was used in FFNN, which accurately predicted glucose levels from the input features such as voltage, absorbance, transmittance, and reflections with MSE of 9.16 $(mg/dl)^2$.

This combined approach allows for the development of a reliable and accurate glucose monitoring system, which can be employed for diabetes management.

3. Design and implementation

To address the limitations of single-wavelength technology, we propose the adoption of dual-wavelength system to achieve comprehensive glucose insights. The light at 940 nm has a higher absorption coefficient for glucose, while the light at 950 nm passes more easily through tissue, so it is possible to measure thicker tissues, such as the back of a hand, at this wavelength. As a result, they effectively reduce skin interference and provide enhanced stability, ensuring more accurate glucose readings.

In order to measure the amount of light reflected and transmitted by a finger, TCR1000 (a combination emitter-photo transistor) and TSAL6400 emitter and BPW34 photodetector are used. The system also includes a precise analog-to-digital converter (ADC) and a computation unit using a micro-controller to process the data, as shown in Fig. 1.

The TCRT1000 device is capable of detecting glucose-induced energy absorption when an infrared (IR) emitting LED comes into contact with a finger. To protect the IR LED, a current-limiting resistor $R_1 = 100\,\Omega$ is added in series. Similarly, a resistor $R_2 = 10\,k\Omega$ is connected in series with the collector through the phototransistor. In the case of the 940 nm, a current-limiting resistor $R_3 = 100\,\Omega$ is added in series with the TSAL6400 emitter to prevent damage. Additionally, a resistor $R_4 = 10\,k\Omega$ is added in series with the BPW34 photodetector. These resistance values can be calculated from the respective devices' datasheets.

In order to convert the analog signal to digital signal, the output is routed to an ADS1115 converter. Particularly the channels A_0 and A_2 are used to convert the signal from reflective and transmission modes, respectively. The micro-controller reads the digital value from ADC and then processed and analyzed to extract the features. The procedure to extract the features are given in Table 1, as Algorithm 1 and Algorithm 2 with a wavelength of 950 nm and the 940 nm, respectively.

As from the algorithms in Table 1, the detectors read a signal from finger and pass it to the micro-controller via ADC. Then, the received samples are smoothed and averaged to reduce the random variation of the extracted input features.

After the data obtained from both the wavelengths, it is possible to predict the glucose concentration using an ML regression model. The model considered four input features from each wavelength and predicts the glucose concentration by training. The proposed glucose prediction using FFNN regression model work presented in the following section.

4. Glucose prediction and evaluation metrics

To predict the blood glucose concentration, a ML model such as a neural network was used. The independent variables from each wavelength were used as input features to the model, and the blood glucose concentration was the target variable. The model would then learn the relationship between the input features and the target variable by updating the weights and biases through an iterative process such as back propagation.

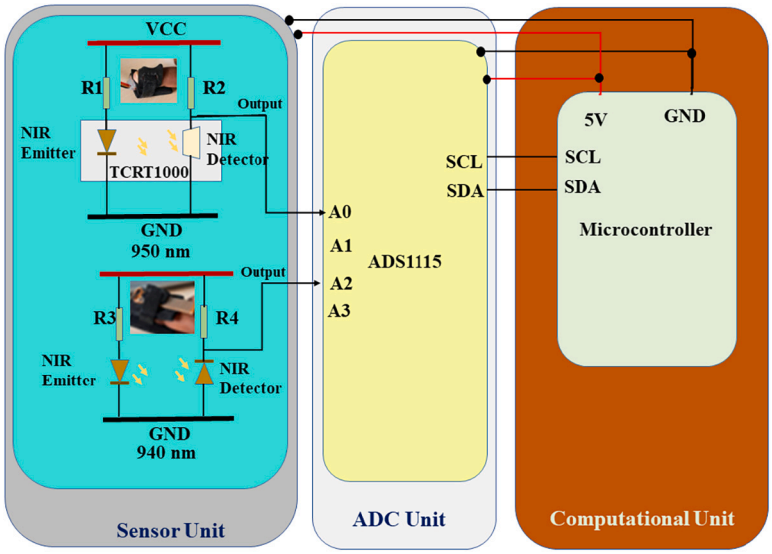


Fig. 1. Block diagram of sensor and computational unit.

Table 1
Algorithm 1 and Algorithm 2 for 950 nm and 940 nm detector respectively.

Algorithm-1	Algorithm-2
Input: adc0 Output: X_1, X_2, X_3, X_4 Input features from 950 nm	Input: adc2 Output: X_5, X_6, X_7, X_8 Input features from 940 nm
Step 1: Read the detector value from finger to adc0 ADS1115 Step 2: initialize mv ← 0, R ← 4.08 (offset value) Step 3: convert detector value into millivolts. mv ← (adc0*0.1875)/100 Step 4: Get 10 sample value from the detector for smoothing. a[i] ← mv Step 5: sort the data from small to large. B ← a[i] a[i] ← c[j] c[j] ← B Step 6: take the average value of 6 center samples. d+ ← a[i] e ← d/6 Step 7: to get average voltage value in V. X1 ← e Step 8: Calculate Transmittance. $X_2 \leftarrow X_1/R$ Step 9: Calculate Absorbance of signal. $X_3 \leftarrow -\log(X_2)$ Step10: Calculate Reflectance of signal. $X_4 \leftarrow 1 - X_2 - X_3$	Step 1: Read the detector value from finger to adc2 ADS1115 Step 2: initialize mv ← 0, T ← 3.78 (offset value) Step 3: convert detector value into millivolts. mv ← (adc0*0.1875)/100 Step 4: Get 10 sample value from the detector for smoothing. a[i] ← mv Step 5: sort the data from small to large. B ← a[i] a[i] ← c[j] c[j] ← B Step 6: take the average value of 6 center samples. d+ ← a[i] e ← d/6 Step 7: to get average voltage value in V. X5 ← e Step 8: Calculate Transmittance. $X_6 \leftarrow X_5/T$ Step 9: Calculate Absorbance of signal. X7 ← -log(X_6) Step10: Calculate Reflectance of signal. $X_8 \leftarrow 1 - X_6 - X_7$

4.1. Predictive model for blood glucose concentration

The FFNN has been designed to predict the glucose concentration and trained on a dataset collected from the detectors. The data is divided using a non-stratified approach, with 75% used for training and 25% for testing.

The FFNN consists of an input layer, two hidden layers, and an output layer. The input layer has 8 neurons corresponding to each of the input features. The hidden layers consist of 64 neurons in the first hidden layer, and 16 neurons each in the second hidden layers. The number of neurons in each layer was determined using a grid search approach, which involves testing multiple different network architectures with different numbers of layers and nodes to determine which performs best on a validation set. This approach is time-consuming but can be effective for finding the optimal network architecture.

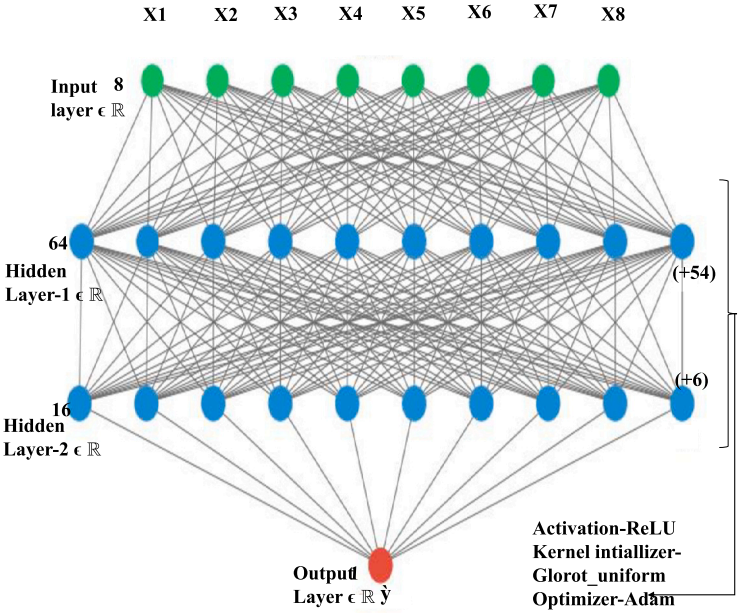


Fig. 2. Proposed FFNN glucose prediction.

Based on this approach a total of 80 neurons in the two hidden layers were used. The output layer has one neuron, which represents the predicted glucose concentration denoted by the vector y' as shown in Fig. 2. Each layer applies a linear transformation to the output of the previous layer, followed by a non-linear activation function. The rectified linear unit (ReLU) activation function was used in each layer to achieve a non-linear transformation of the data and linear activation is used for the output layer. Mathematically, the output of each layer can be calculated using equation (1) to (3).

$$Z_F = ReLU(W_F * X + b_F) \tag{1}$$

$$Z_s = ReLU(W_S * Z_F + b_S) \tag{2}$$

$$\hat{y} = W_L * Z_S + b_L \tag{3}$$

Where Z_F and Z_S is the output of the first and second layers, W_F and W_S are the weight matrix connecting the input layer to the first layer and from first layer to the second layer respectively. b_F and b_S are the bias vector of the first and second layers. $\hat{y}$ is the predicted glucose concentration, W_L is the weight matrix connecting the second layer to the output layer, and b_L is the bias vector of the output layer. X is the input vector containing the received data from detectors. The input vector (X), weight matrices (W_F, W_S and W_L) and bias vectors (b_F, b_S and b_L) are represented in following manner.

$$X_{8\times1} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ \vdots \\ X_8 \end{bmatrix}$$

$$W_{F_{8\times64}} = \begin{bmatrix} W_{1,1} & W_{1,2} & .... & W_{1,64} \\ W_{2,1} & W_{2,2} & .... & W_{2,64} \\ W_{3,1} & W_{3,2} & .... & W_{3,64} \\ \vdots & \vdots & .... & \vdots \\ \vdots & \vdots & .... & \vdots \\ \vdots & \vdots & .... & \vdots \\ X_{8,1} & W_{8,2} & .... & W_{8,64} \end{bmatrix}$$

$$b_{F1\times8} = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & ..... & b_{1,64} \end{bmatrix}$$

$$W_{S_{64\times16}} = \begin{bmatrix} W_{1,1} & W_{1,2} & .... & W_{1,16} \\ W_{2,1} & W_{2,2} & .... & W_{2,16} \\ W_{3,1} & W_{3,2} & .... & W_{3,16} \\ : & : & .... & : \\ : & : & .... & : \\ : & : & .... & : \\ X_{64,1} & W_{64,2} & .... & W_{64,16} \end{bmatrix}$$

$$b_{S_{1\times16}} = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & ..... & b_{1,16} \end{bmatrix}$$

$$W_{L16\times1} = \begin{bmatrix} W_{3,1} \\ W_{3,2} \\ W_{3,3} \\ : \\ : \\ W_{3,16} \end{bmatrix}$$

$$b_L = \begin{bmatrix} b_3 \end{bmatrix}$$

The network uses glorot initialization method with adam optimizer to update the weights with l_2 regularization method. The purpose of using this method is to prevent the vanishing or exploding gradients problem, which can occur in deep neural networks. This problem can cause the gradients to become too small or too large, which can negatively impact the training of the neural network. The glorot uniform initialization method was shown to be effective in mitigating this problem. This method generates initial values for the weights from a uniform distribution. The range of this distribution depends on the number of input and output units in each layer as calculated in equation (4).

$$range = \sqrt{\frac{6}{n_{inputnodes} + n_{outputnodes}}} \tag{4}$$

where $n_{inputnode}$ is the number of input units and $n_{outputnodes}$ is the number of output units in the layer. For the first layer in our FFNN, which has 8 inputs and 64 neurons, the range for initialization is calculated. This means that the weights in the first layer was initialized within the range (-0.2886, 0.2886) using a uniform distribution.

Similarly, for the second layer, the input dimensions are determined by the number of neurons in the previous layer. In this case, the second layers have 16 neurons, and the previous layer has 64 neurons, so the range for initialization is calculated as (-0.2738, 0.2738) using a uniform distribution. Finally, the output layer has only one neuron and is linearly activated. Hence, there is no need to specify any range for initialization. Using this initialization method the weights in the FFNN start at reasonable values and can be effectively trained.

After initializing the weights using the glorot uniform initialization method, the FFNN is ready to be trained on the dataset obtained from both the detectors. During training, the FFNN is fed with the input data through the network and comparing the predicted output to the true output using a loss function. The loss function measures how well the network is performing and is used to update the weights in the network through an optimization algorithm, such as Adam optimizer. Adam is used due to its efficiency in terms of memory and computational requirements. The optimizer aims to minimize the loss function, such as mean squared error (MSE) by calculating the gradient of the loss function with respect to the weights in the network through backpropagation. Backpropagation is an algorithm that calculates the partial derivative of the loss function with respect to each weight in the network, and it does so by propagating the error from the output layer back to the input layer. The gradient of the loss function was computed using equation (5).

$$\frac{\partial L}{\partial y} = 2(\hat{y}_{batch} - y_{batch}) \tag{5}$$

Where $\frac{\partial L}{\partial y}$ is the gradient of the loss function, $\hat{y}_{batch}$ is the predicted output for given input and y_{batch} is the actual output for the same input.

After the gradient of the loss function is calculated using backpropagation, the optimizer updates the weights in the network to move in the direction that minimizes the loss function. The optimizer uses two estimates, namely momentum and second moment estimates, to update the weights. The momentum estimate represents the exponentially weighted average of the gradients, and the second moment estimate represents the exponentially weighted average of the squared gradients. These estimates are used to update the weights and biases, which helps the optimizer avoid getting trapped in local optima and allows for faster and more reliable convergence than other optimization algorithms. The equations for updating the momentum and second moment estimates are as given in equation (6) and (7), respectively.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\frac{\partial L}{\partial y_t} \tag{6}$$

Where m_t, is the momentum estimate at time step t, β_1 is a hyperparameter that controls the exponential decay of the moving average, m_{t-1} is the momentum estimate at the previous time step, and $\frac{\partial L}{\partial y_t}$ is the gradient at time step t.

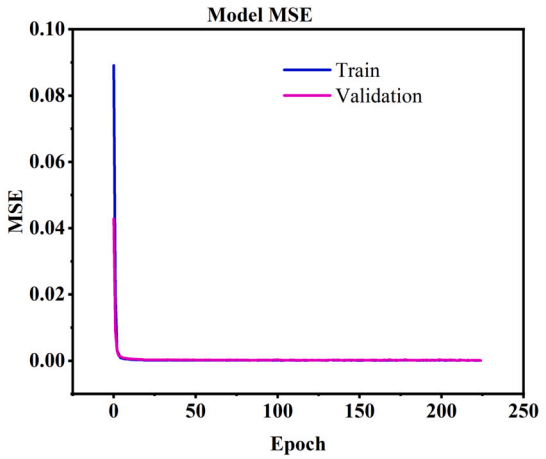


Fig. 3. MSE with respect to epochs.

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)(\frac{\partial L}{\partial y_t})^2 \tag{7}$$

Where v_t is the second moment estimate at time step t, β_2 is a hyperparameter that controls the exponential decay of the moving average, v_{t-1} is the second moment estimate at the previous time, and $\frac{\partial L}{\partial y_t}^2$ is the element-wise squared gradient at time step t. The new weights and biases are computed using the update rule given in equation (8).

$$\theta_t = \theta_{t-1} - \left(\frac{\alpha}{\sqrt{\hat{v}_t} + \epsilon}\right)\hat{m}_t \tag{8}$$

Where θ_t is the weight or bias at time step t, α is the learning rate, and ϵ is a small constant to prevent division by zero. $\hat{m}_t$ and $\hat{v}_t$ are bias-corrected estimates of the momentum and second moment, respectively. These terms are used to correct the bias towards zero of the momentum and second moment estimates at the start of the training process. They are calculated using equations (9) and (10).

$$\hat{m}_t = \frac{m_t}{1-\beta_1} \tag{9}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2} \tag{10}$$

After updating the weights, the next batch of data is fed through the network and the process is repeated for a specified number of epochs. An epoch is a full iteration through the entire dataset. After training, the performance of the network was evaluated on a separate test dataset to assess the network generalizes to new or unseen data. This step is crucial to ensure that the model can generalize well on unseen data. The evaluation results help to understand the performance of the model and determine if it has overfit or underfit the training data. This evaluation involved the feeding of test dataset through the network and comparing the predicted output to the true output using a performance metrics. This has been discussed in the following sub section.

4.2. Evaluation metrics for predicted glucose

The performance metrics were evaluated from the predicted blood glucose concentration with respect to reference glucose. Dr. Trust's conventional intrusive glucometer was used to measure reference blood glucose. The metrics calculated for 575 samples with 289 subjects from the age group of 19–69. The data was collected from individuals in the random blood glucose test mode. To evaluate the performance of a model, the entire dataset acquired through experiments was separated from the one used for training dataset as 460 samples, and the test dataset of 115 samples are used to evaluate its performance. Once the model is trained, the evaluation results help to understand the performance of the model. To understand the performance, various metrics such as mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R^2) was evaluated on both the training and test sets. One way to do this is by plotting a learning curve. A learning curve is a plot that shows the performance of the model on the training and testing data as a function of the number of epochs.

The results displayed in Fig. 3 indicate that the MSE decreases for both the training and testing datasets as the number of epochs increases. It indicates that the model is learning from the data and generalizing well to new, unseen data. Similar kind of response is obtained for MAE curve shown in Fig. 4. On the other hand the learning curve for the R^2 increases for both datasets as the number of epochs increases as shown in Fig. 5.

In contrast, overall evaluation metrics provide a summary of the model's performance across the entire dataset. These metrics are typically calculated after the model has been trained and tested on the entire dataset. For this purpose, the evaluation metrics

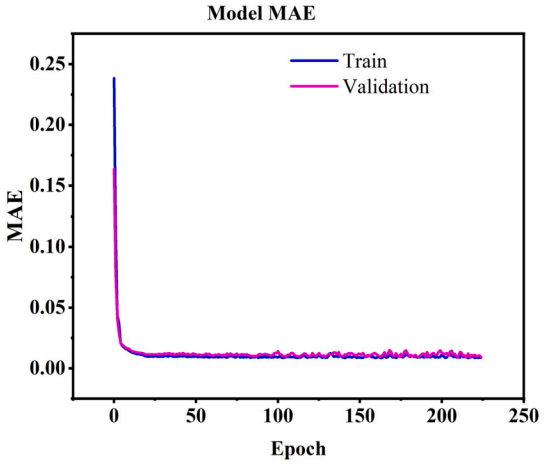


Fig. 4. MAE with respect to epochs.

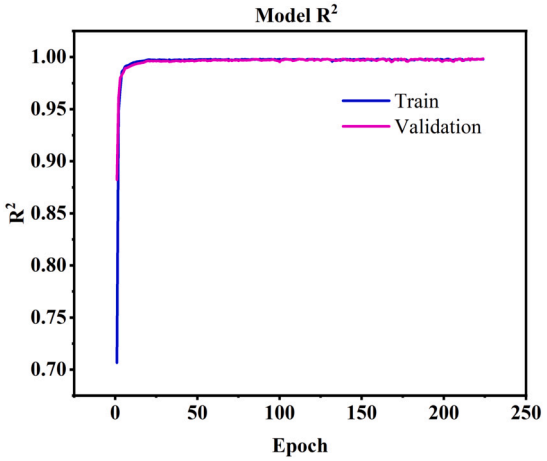


Fig. 5. Coefficient of determination (R^2) with respect to epochs.

include the MSE, MAE, and mean absolute percentage error (MAPE) were evaluated. These metrics provide a single number that summarizes the overall performance of the model. This overall metrics are determined by using equations (11) to (13). The MSE loss function is used to calculate the difference between the predicted output and the actual output. It is given by equation (11).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}|y - \hat{y}|^2 \tag{11}$$

The MAE measures the average absolute difference between the predicted output and the actual output. It is given by equation (12).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y - \hat{y}| \tag{12}$$

Where y is the actual output, $\hat{y}$ is the predicted output, and n is the total number of samples. The MAPE is given in equation (13).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{y - \hat{y}}{y} \times 100 \tag{13}$$

Finally, the R^2 is evaluated in addition to the above metrics. It measures the proportion of the variance in the dependent variable that can be explained by the independent variable(s). Mathematically, it is calculated as in equation (14).

$$R^2 = 1 - \left(\frac{SS_{res}}{SS_{tot}}\right) \tag{14}$$

Table 2
Comprehensive performance evaluation metrics comparison of various ML models.

Metrics	FFNN	LR	RF	DT	KNN	XGB	SVM
Layers	64X16	-	-	-	-	-	-
Optimizer	Adam	-	-	-	-	-	-
Epoch	200	-	-	-	-	-	-
Batchsize	32	-	-	-	-	-	-
HL activation	Relu	-	-	-	-	-	-
O/p activation	Linear	-	-	-	-	-	-
R^2	0.99	0.99	0.99	0.99	0.99	0.99	0.93
MSE $(mg/dl)^2$	9.16	9.74	10.52	15.08	11.50	12.4	509.3
MAPE (%)	1.94	2.04	1.88	2.15	2.00	1.93	18.06
MAE (mg/dl)	2.49	2.60	2.48	2.82	2.72	2.5	20.95
RMSE (mg/dl)	3.02	3.12	3.24	3.88	3.39	3.5	22.56

Table 3
Performance comparison of the current work with existing studies.

	Proposed	[57]	[61]	[64]	[65]	[66]	[67]
R^2 value	0.99	0.908	0.96	0.95	-	-	0.93
mAPE (%)	1.94	3.25	-	-	7.01	-	5.43
MSE $(mg/dl)^2$	9.16	-	-	-	-	-	-
MAE (mg/dl)	2.49	3.87	-	-	5.23	-	-
RMSE	3.02	5.61	11	-	7.64	-	-
Technology	NIR	NIR	NIR	OCT	Photoacoustic	Visible	NIR

Where SS_{res} is the sum of squares of residuals (i.e., the sum of squared differences between the predicted and actual values) and SS_{tot} is the total sum of squares (i.e., the sum of squared differences between the actual values and the mean of the dependent variable). A higher R^2 value indicates a better fit between the predicted values and the actual values, and hence a better performance of the model.

To demonstrate the effectiveness of the proposed FFNN system, we conducted a comprehensive evaluation comparing it with various ML algorithms, including LR, RF, DT, KNN, extreme gradient boost (XGB), and SVM. Across the evaluation metrics, the R^2 value remained consistently high for all algorithms, ranging from 0.93 to 0.99. This robust R^2 performance signifies the ability of these models to effectively explain the variance in glucose levels. In terms of the MAPE, results ranged from 1.88% to 18.06%. Especially, the FFNN, LR, and XGB models achieved the lowest MAPE values, indicating their superior accuracy in predicting glucose levels. Further, examined the MSE in mg/dl^2. These values spanned from 9.16 to 509.3 and both the FFNN and LR models exhibited the lowest values, underscoring their higher precision in predictions. When considering the MAE in mg/dl, the range extended from 2.48 to 20.95. Here, the FFNN model attained the lowest value, reinforcing its accuracy in predicting glucose levels. Finally, the RMSE values ranged from 3.02 to 22.56 and the FFNN model boasted the lower RMSE values, which reflects its enhanced ability to capture differences between predicted and actual glucose levels. Based on the above metrics, both the FFNN and LR models consistently demonstrated well in terms of accuracy, precision, and prediction of glucose levels This robust evaluation underscores the effectiveness of our proposed FFNN approach and its alignment with the LR model's efficacy in handling this complex predictive task as shown in Table 2.

In addition, to demonstrate the strength of the proposed approach, a performance comparison between the current work and existing studies is provided in Table 3.

4.3. Relation between predicted glucose and input features

After the analysis of the predicted glucose, it was observed that there is a strong correlation between the input features and the glucose concentration in the subject's finger. Fig. 6, depict the relationship between the input features and the glucose concentration present in the subject's finger. As the voltage obtained by the detectors increases, the glucose molecules absorb low signal, resulting in a decrease in the predicted glucose concentration. Conversely, if the glucose concentration absorbs more signal, the obtained voltage decreases, indicating a higher predicted glucose concentration.in addition, the relationship between transmittance, absorbance, and reflections of the signal can be further explored in terms of predicted glucose concentration. As the predicted glucose concentration increases, the absorbance of the signal also increases, while the transmittance of the signal decreases with a decaying behavior depicted in Fig. 6. Based on the measured maximum and minimum glucose in both the reflective and transmission modes, respectively, they appear to be 400 mg/dl and 404 mg/dl, as well as 63 mg/dl and 64 mg/dl. Moreover, Table 4 presents the predicted glucose concentration ranges for minimum and maximum values based on absorbance, transmittance, and reflections at both wavelengths.

Furthermore, the measured predicted glucose with FFNN model and the reference glucose from reference device with respect to number of samples collected is shown in Fig. 7. The predicted glucose shows a linear relationship with the reference glucose. The

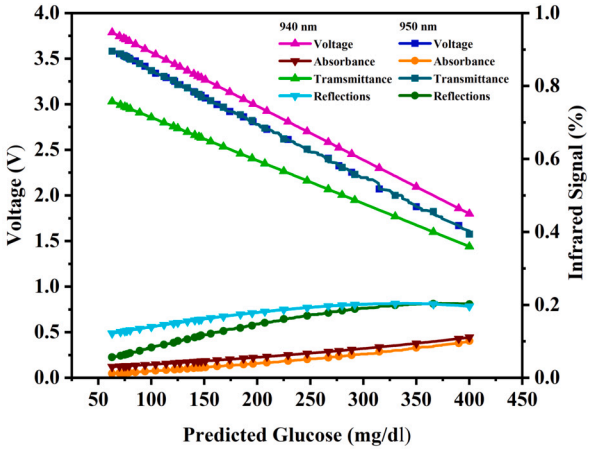


Fig. 6. Relation between predicted glucose and input features.

Table 4
Input features, minimum and maximum values compared to predicted glucose.

	950 nm Reflectance		940 nm Transmission	
	Min	Max	Min	Max
Glucose (mg/dl)	63	400	64	404
Absorbance (A)	0.12	0.44	0.04	0.40
Transmittance (T)	0.75	0.35	0.89	0.39
Reflections (R)	0.12	0.19	0.05	0.20

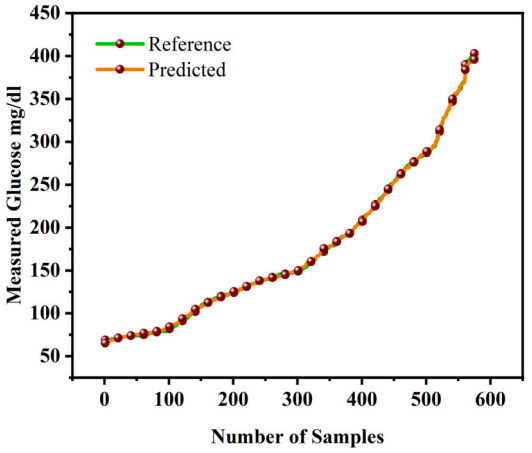


Fig. 7. Glucose comparison between reference and predicted.

subsequent section described the analysis of clinical accuracy and consistency of the glucose prediction model. To accomplish this, two widely used methods were employed, the SEG and Bland-Altman plot analyses.

4.4. Clinical accuracy and consistency analysis

The SEG and Bland-Altman plot are both commonly used methods for evaluating the accuracy of a measurement or prediction method, but they serve different purposes and are appropriate for different types of analyses.

By plotting the measured glucose values from a device against reference values of finger prick glucometer, the SEG method allows for a visual representation of the device's accuracy and clinical acceptability. The accuracy of the model for predicting glucose values was evaluated by undergoing a SEG analysis. To assess the accuracy of the model, the SEG analysis was conducted on both the train and test datasets. The grid's layout determines the zone for each data point, signifying the level of clinical risk linked to the measurement error. These zones range from A (green area) to E (dark red or black), where A is the most clinically acceptable and E signifies the highest risk. In the proposed system, the predicted glucose falls within zone A, which is clinically acceptable

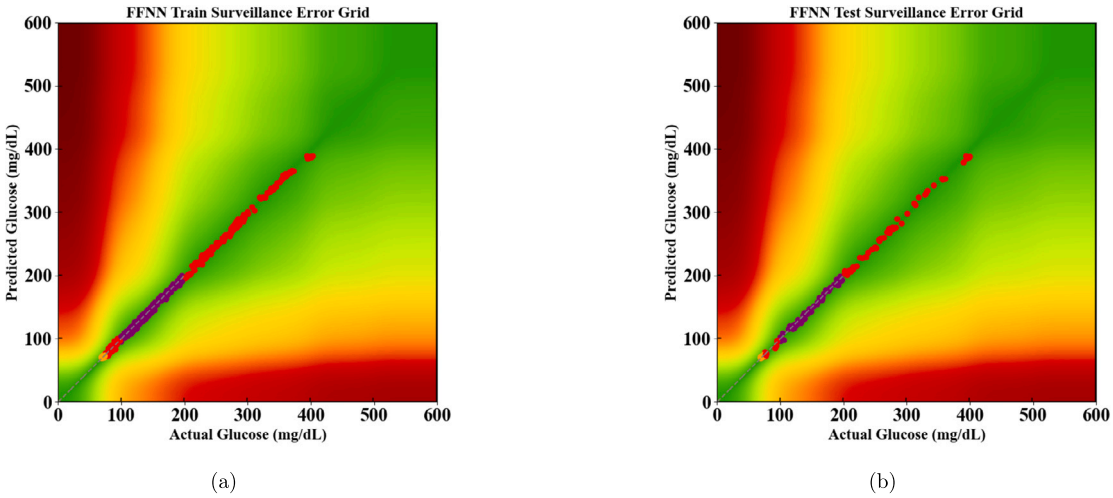


Fig. 8. SEG analysis: (a) Training dataset, (b) Test dataset.

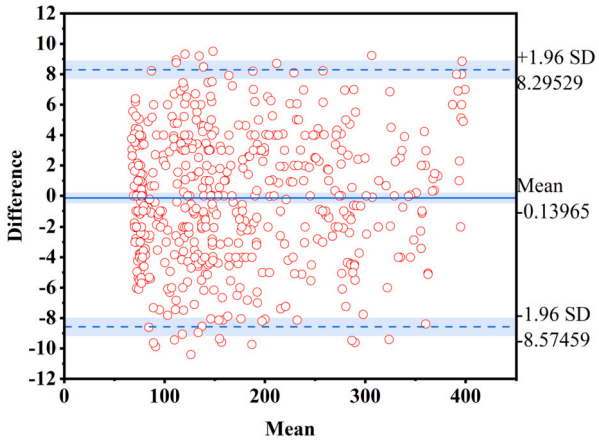


Fig. 9. Bland-Altman plot with mean difference for the proposed model.

in comparison to the reference device. This is demonstrated in Fig. 8(a) and Fig. 8(b). The results showed that the FFNN method achieved 100% linearity with the reference glucose values up to 430 mg/dl. This indicates that the model can accurately predict glucose values across a wide range of concentrations, which is crucial for effectively managing diabetes.

On the other hand, the Bland-Altman plot is a graphical method used to evaluate the level of agreement between two measurement methods. It is particularly useful for comparing the difference between the two methods with their average and plots them on a graph. This plot consists of a scatter plot of the difference between the two measurement methods on the y-axis and the average of the two methods on the x-axis. A horizontal line is also included to indicate the mean difference between the two methods, and two additional lines are included to show the upper and lower limits of agreement. In this study, the plot analysis shown in Fig. 9, revealed a mean difference between the predicted and actual glucose values of 0.16 mg/dl, which is relatively small and indicates good agreement between the predicted and actual glucose values. Moreover, the plot shows a cluster of data points around the mean difference line, which further supports the accuracy of the glucose prediction model. After analyzing the SEG and Bland-Altman plot results, the subsequent stage involves categorizing glucose levels into various groups based on predicted values, through a MLP and KNN classifiers. This process will be elucidated in the following section.

5. Multiclass glucose classification with MLP and KNN classifiers

In this section, we focus on the use of the MLP and KNN classifiers for the multiclass classification of glucose levels from the input features. The ML classification models are trained to accurately predict the class labels of new cases. The multiclass classification task is to classify the glucose levels into three classes - hypoglycemic (class 0), normal (class 1), and hyperglycemic (class 2) within the glucose concentration range of 63 mg/dL to 404 mg/dL.

Table 5
Optimized hyperparameters for MLP for different fold CV.

Evaluation Folds	Gridsearch CV Hyperparameters
5-Fold	activation = tanh, hiddenlayersize = 100, learningrate = constant, solver = adam
10-Fold	activation = tanh, hiddenlayersize = 100, learningrate = constant, solver = adam
15-Fold	activation = tanh hiddenlayersize = 100, learningrate = **constant**, solver = adam

*Constant learning rate: 0.001.

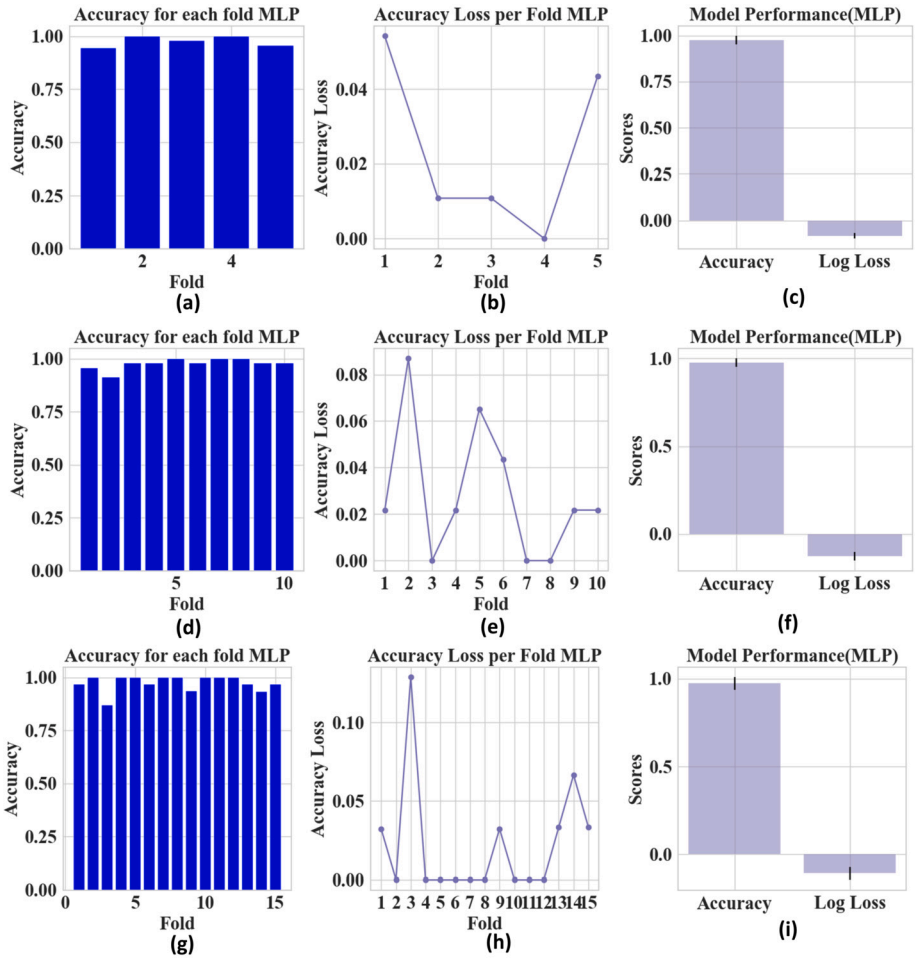


Fig. 10. MLP classifier performance (a) 5-fold accuracy, (b) 5-fold accuracy loss, (c) 5-fold model performance, (d) 10-fold accuracy, (e) 10-fold accuracy loss, (f) 10-fold model performance, (g) 15-fold accuracy, (h) 15-fold accuracy Loss, and (i) 15-fold model performance.

5.1. MLP classifier classification

The MLP classifier operates through training the network using a labeled data set, which includes a series of input features: $X_1, X_2, X_3, X_4, X_5, X_6, X_7$, and X_8. To divide the data, a stratified approach was employed. This involved splitting the data set into two parts: 80% was allocated for training, and the remaining 20% was reserved for testing. After training, the MLP classifier were used to predict the output class for new input. To enhance the classifiers' performance, hyperparameter tuning was performed using the grid search cross-validation (CV) method. During this process, a range of hyperparameters were examined and then selected the best combination from the initial parameters. The optimized hyperparameters for 5-fold, 10-fold, and 15-fold CV are described in Table 5. Fig. 10 illustrates the performance of the MLP model across various metrics, including accuracy (Fig. 10(a), Fig. 10(d), and Fig. 10(g)), loss (Fig. 10(b), Fig. 10(e), and Fig. 10(h)), and overall model performance (Fig. 10(c), Fig. 10(f), and Fig. 10(i)), utilizing three distinct CV techniques. The mean accuracy and neg log loss is given in Table 7.

Further, a multi-class classification analysis was conducted with three classes (0, 1, and 2) using 5, 10, and 15 CV folds. The results are presented in Fig. 11, Fig. 12, and Fig. 13, respectively. This analysis allowed us to observe both the correct and incorrect predictions made by the model for each class, utilizing a confusion matrix.

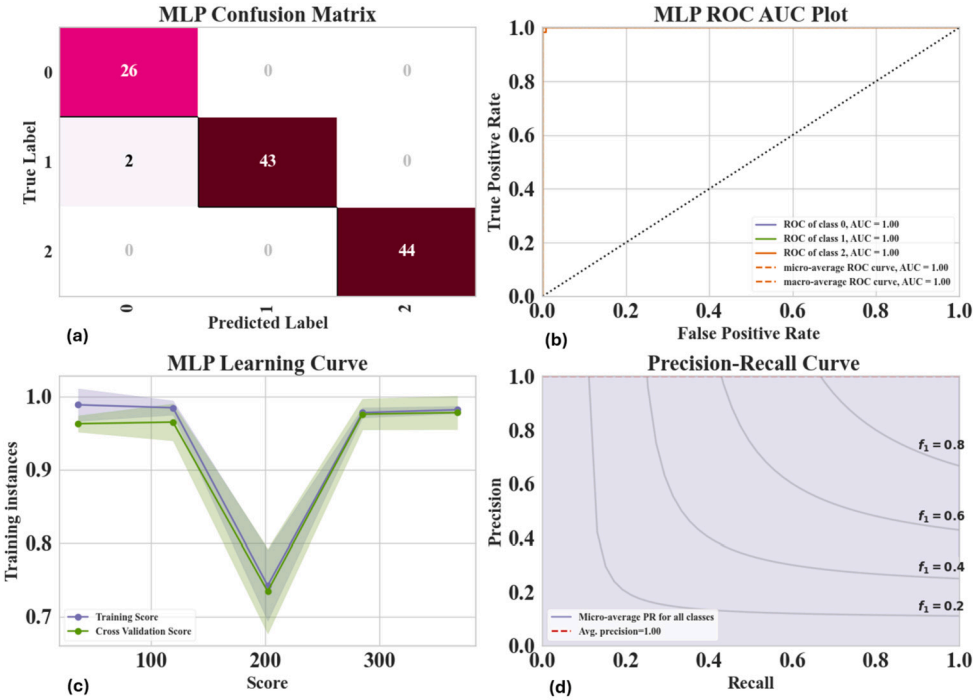


Fig. 11. MLP Classifier Metrics for 5-fold CV (a) Confusion matrix, (b) ROC AUC curve, (c) Learning curve, and (d) Precision-recall.

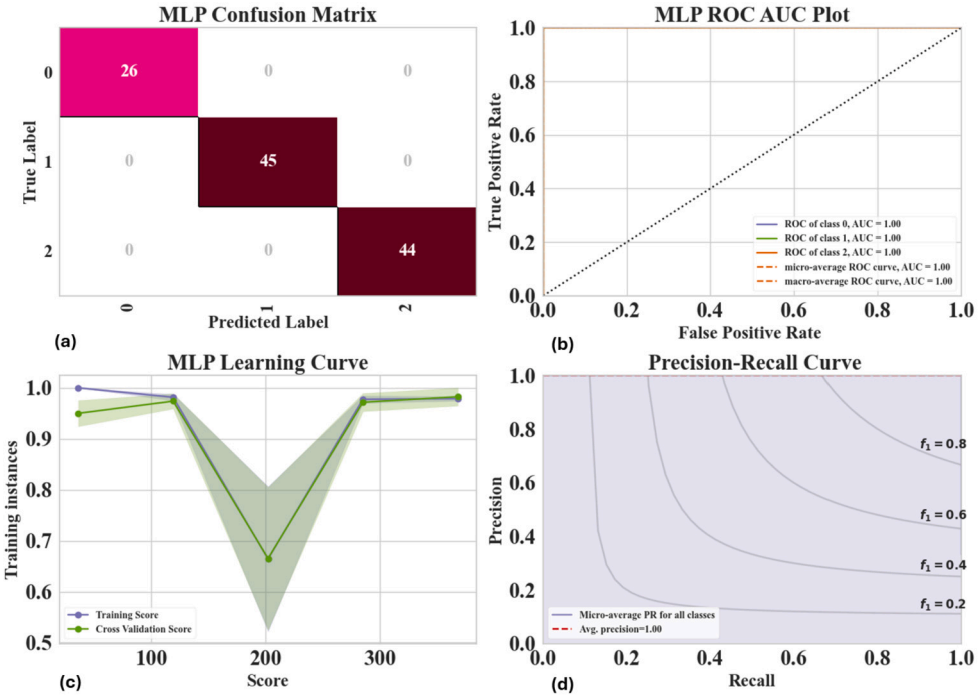


Fig. 12. MLP Classifier Metrics for 10-folds (a) Confusion matrix, (b) ROC AUC curve, (c) Learning curve, and (d) Precision-recall.

During the 5-fold validation process, the confusion matrix analysis (Fig. 11(a)) for Class 0 showed a precision of 93%, a recall of 100%, and an F1-score of 96% across 26 instances. In contrast, Class 1 demonstrated a precision of 100%, a recall of 96%, and an F1-score of 98% across 45 instances. For Class 2, both precision and recall were 100%, resulting in an F1-score of 100% across 44 instances. The model achieved an accuracy of 98% across 115 instances. The macro average for precision, recall, and F1-score was

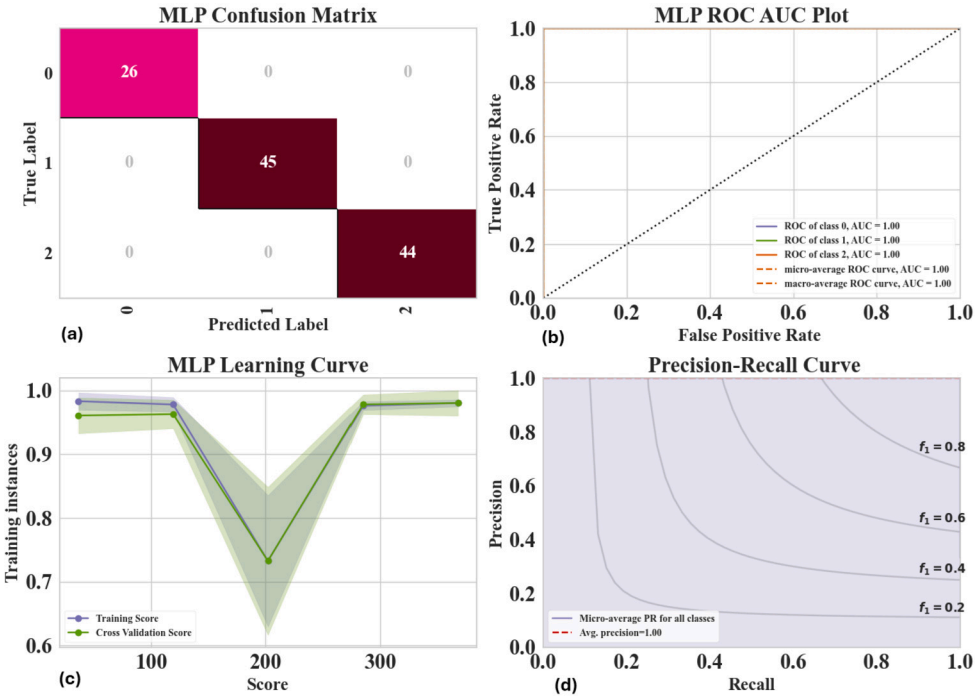


Fig. 13. MLP Classifier Metrics for 15-folds (a) Confusion matrix, (b) ROC AUC curve, (c) Learning curve, and (d) Precision-recall.

Table 6
Optimized hyperparameters of KNN for different fold CV.

Evaluation Folds	Gridsearch CV Hyperparameters
5-Fold	n-neighbors = 14
10-Fold	n-neighbors = 16
15-Fold	n-neighbors = 20

98%, 99%, and 98%, respectively. Similarly, the weighted average for precision, recall, and F1-score remained at 98% for the 5-fold CV in Fig. 11(d).

The ROC-AUC curve fitted well for class 0, class 1, and class 2 in the 5-fold MLP classifier (Fig. 11(b)). The learning curve maintained a balance from the initial to the final instances in the multiclass classification, as illustrated in Fig. 11(c).

Moving on to the 10-fold and 15-fold validation processes, the analysis of Class 0, Class 1, and Class 2 consistently showed a precision and recall of 100% in each fold, resulting in an F1-score of 100% across 26, 45, and 44 instances (Fig. 12(a) and Fig. 13(a)). Also, the ROC-AUC curve fitted well for all classes as shown in Fig. 12(b) and Fig. 13(b). Additionally, the learning curves maintained a balance from the initial to the final instances in multiclass classification, as illustrated in Fig. 12(c) and Fig. 13(c). Based on these results, the model achieved both macro average and weighted average accuracies of 100% for each metric, as depicted in Fig. 12(d) and Fig. 13(d) for the 10-fold and 15-fold CV, respectively.

5.2. KNN classifier classification

The KNN classifier is a type of ML method that is commonly for classification tasks. The k-nearest training data points to the test data point are considered, and the class that appears most frequently among these k-neighbors is assigned to the test data point. The closer a neighbor is to the test data point, the higher its weight. After training, the KNN classifier were used to predict the output class for new input. To enhance the classifiers' performance, hyperparameter tuning was performed using the grid search CV method. During this process, a range of hyperparameters were examined and then selected the best combination from the initial parameters. The optimized hyperparameters for 5-fold, 10-fold, and 15-fold CV are described in Table 6. Fig. 14 illustrates the performance of the KNN model across various metrics, including accuracy (Fig. 14(a), Fig. 14(d), and Fig. 14(g)), loss (Fig. 14(b), Fig. 14(e), and Fig. 14(h)), and overall model performance (Fig. 14(c), Fig. 14(f), and Fig. 14(i)), utilizing three distinct CV techniques. The mean accuracy and neg log loss is given in Table 7.

Further, a KNN multi-class classification analysis was conducted with three classes (0, 1, and 2) using 5, 10, and 15 CV folds. The results are presented in Fig. 15, Fig. 16, and Fig. 17, respectively. During CV process in each fold Class 0 confusion matrix analysis (Fig. 15(a), Fig. 16(a), and Fig. 17(a)) showed across 26 instances. In contrast, Class 1 demonstrated a precision of 98%, a recall of 100%, and an F1-score of 99% across 45 instances. For Class 2, both precision and recall were 100%, resulting in an F1-score of

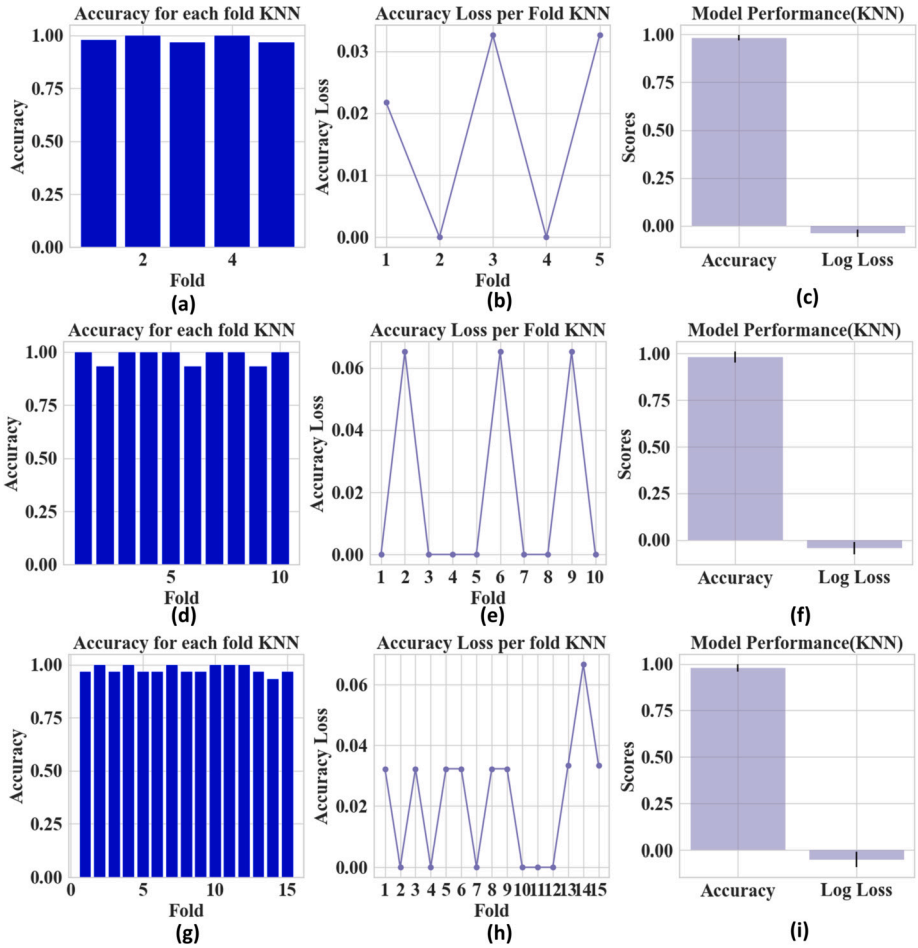


Fig. 14. KNN classifier performance (a) 5-fold accuracy, (b) 5-fold accuracy loss, (c) 5-fold model performance, (d) 10-fold accuracy, (e) 10-fold accuracy loss, (f) 10-fold model performance, (g) 15-fold accuracy, (h) 15-fold accuracy loss, and (i) 15-fold model performance.

100% across 44 instances. The model achieved an accuracy of 99% across 115 instances. The macro average for precision, recall, and F1-score was 99%, 99%, and 99%, respectively. Similarly, the weighted average for precision, recall, and F1-score remained at 99%, as shown in Fig. 15(d) for 5-fold CV, Fig. 16(d) for 10-fold CV, and Fig. 17(d) for 15-fold CV. The ROC-AUC curve fitted well for class 0, class 1, and class 2 in the 5-fold (Fig. 15(b)), 10-fold (Fig. 16(b), and 15-fold (Fig. 17(b)) KNN classifier. The learning curve maintained a balance from the initial to the final instances in all folds of KNN classifier as illustrated in Fig. 15(c), Fig. 16(c), and Fig. 17(c), respectively.

5.3. Classification performance evaluation

The two aforementioned classifier performance results were compared across various metrics such as precision, recall, and an F1-score. These comparative results have been presented in Table 7. Also, the classification metrics for 3 classes and two methods are shown in Fig. 18. The training and testing accuracy is depicted in Fig. 19, for 5-fold, 10-fold, and 15-fold. After evaluating the classifiers, it has been observed that both classifiers are effective in classifying the given data. The results indicate that the both the classifiers are comparatively higher ROC-AUC and precision-recall values, implying better overall performance.

However, to ascertain the superiority, an additional metric, known as negative log loss, was utilized to evaluate the models. This evaluation approach proves effective due to its capacity to provide insights into the alignment of a probabilistic model's predictions with actual outcomes. It takes into account the probabilities assigned to predicted outputs and assigns scores based on the magnitude of error. This metric can be achieved using equation (15).

$$neglogloss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log(y_{pred,ij}) \tag{15}$$

where, N is the total number of instances, C is the total number of classes, y_{ij} is a binary indicator (1 or 0) for whether class j is the correct classification for instance i $y_{pred,ij}$ is the predicted probability of instance i being classified as class j. Table 8 presents the

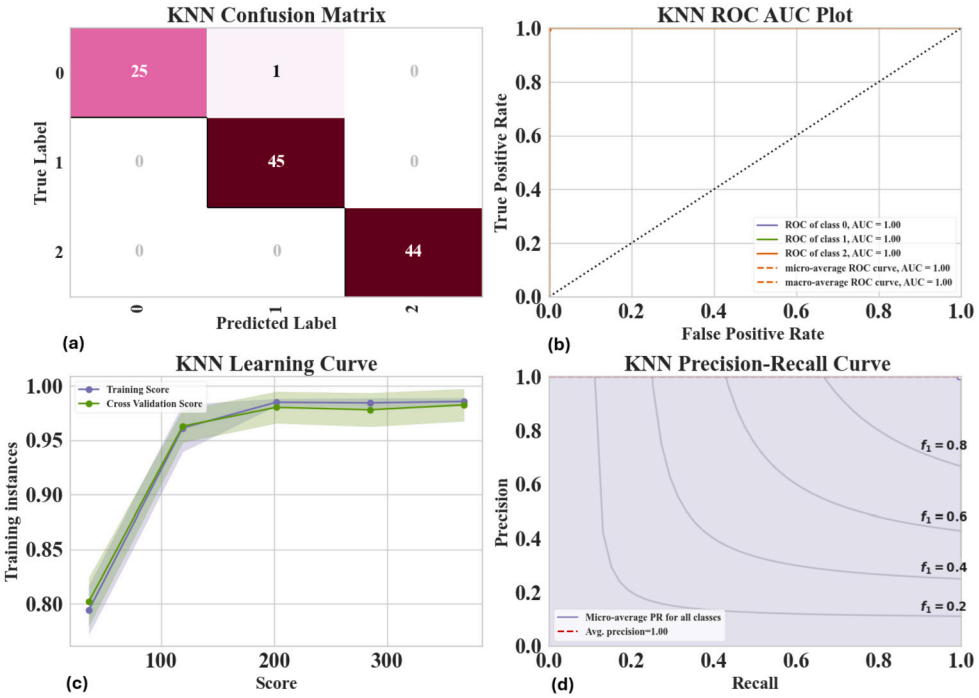


Fig. 15. KNN Classifier Metrics for 5-folds (a) Confusion matrix, (b) ROC AUC curve, (c) Learning curve, and (d) Precision-recall.

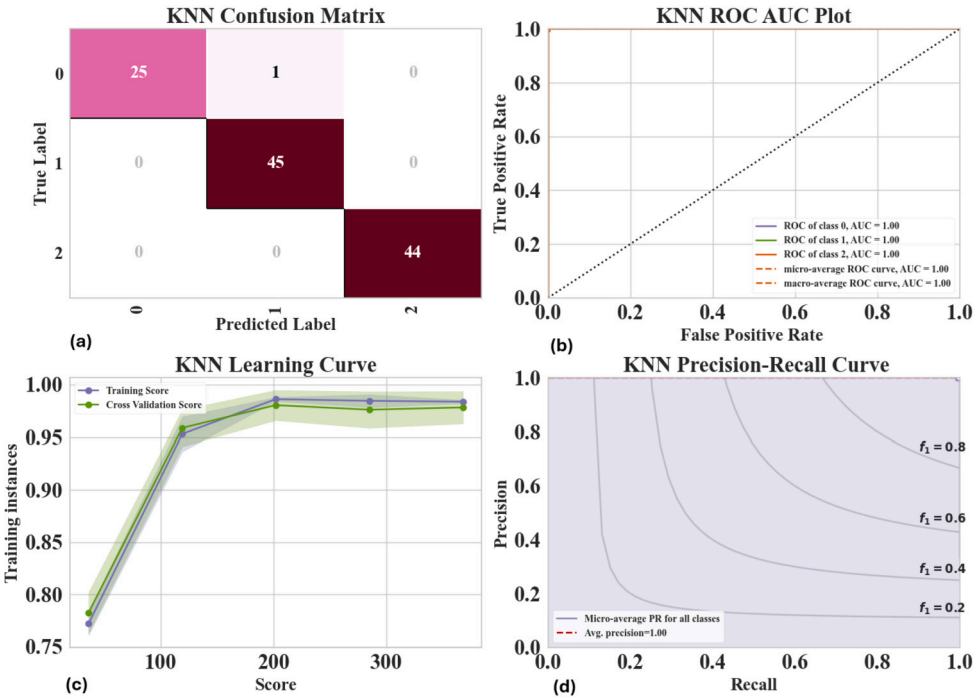


Fig. 16. KNN Classifier Metrics for 10-folds (a) Confusion matrix, (b) ROC AUC curve, (c) Learning curve, and (d) Precision-recall.

performance results of the model classifiers under different conditions of CV folds. The mean accuracy and mean negative log loss results collectively offer insights into the performance of both models.

The KNN model tends to achieve marginally higher mean accuracy and lower mean negative log loss values compared to the MLP model across various CV fold settings. These results suggest that the KNN model demonstrates better confidence in its predictions and a slightly superior average accuracy in classifying instances. From Table 9 we observed that the average training and CV

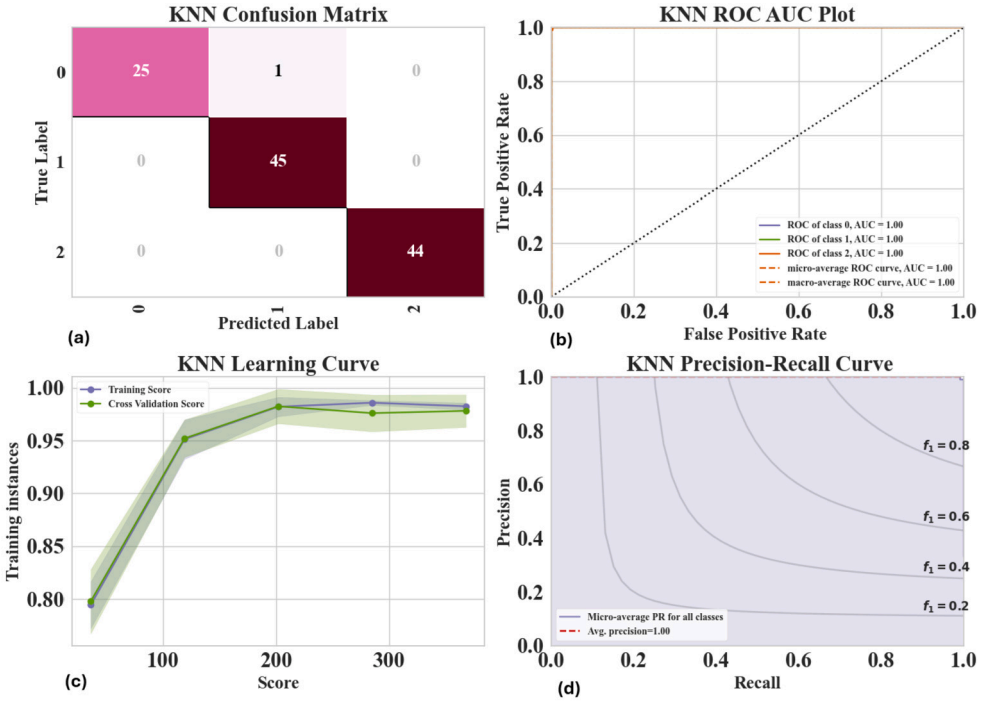


Fig. 17. KNN Classifier Metrics for 15-folds (a) Confusion matrix, (b) ROC AUC curve, (c) Learning curve, and (d) Precision-recall.

Table 7
Performance metrics for different folds and models.

		5-Fold		10-Fold		15-Fold	
		MLP	KNN	MLP	KNN	MLP	KNN
Class-0	Precision	0.93	1	1	1	1	1
	Recall	1	0.96	1	0.96	1	0.96
	F1-score	0.96	0.98	1	0.98	1	0.98
Class-1	Precision	1	0.98	1	0.98	1	0.98
	Recall	0.96	1	1	1	1	1
	F1-score	0.98	0.99	1	0.99	1	0.99
Class-2	Precision	1	1	1	1	1	1
	Recall	1	1	1	1	1	1
	F1-score	1	1	1	1	1	1

Table 8
Comparison for ML classification for mean accuracy and neg log loss.

Model classifier	Grid Search CV Folds	Mean Accuracy	Mean neg logloss
MLP	5	0.976(+/-0.022)	-0.084(+/-0.015)
	10	0.976(+/-0.025)	-0.127(+/-0.025)
	15	0.974(+/-0.036)	-0.108(+/-0.030)
KNN	5	0.983(+/-0.015)	-0.040(+/-0.019)
	10	0.980(+/-0.030)	-0.044(+/-0.034)
	15	0.978(+/-0.020)	-0.048(+/-0.041)

performance using learning curves, the KNN model appears to be slightly better than the MLP model. It has higher training and CV scores, indicating better generalization to unseen data.

6. Discussion

The proposed dual wavelength NI blood glucose detection technique has been found to be superior to the technologies discussed in Section 2 and Section 3, with a high R^2 value and minimum MSE [57][61]. The R^2 value of 99%, the dual-wavelength model

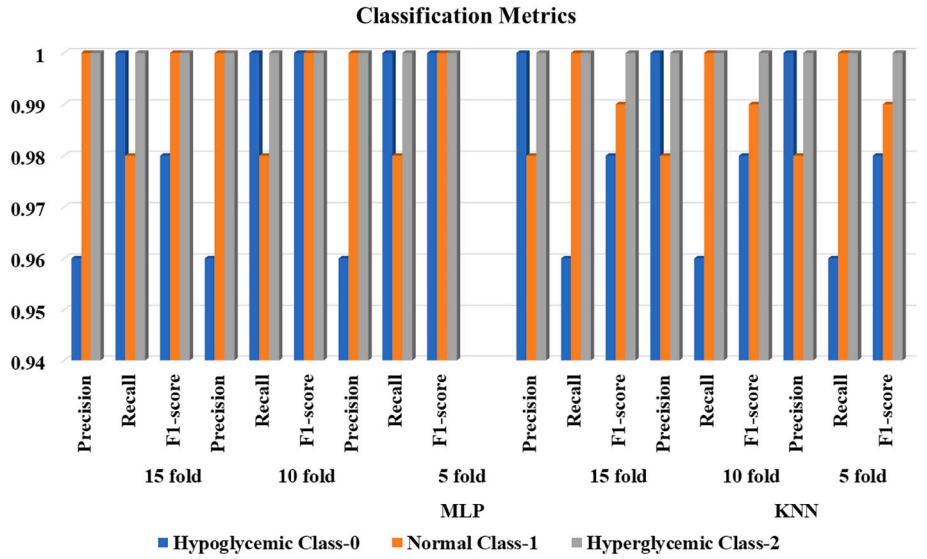


Fig. 18. Performance comparison for MLP and KNN classifiers.

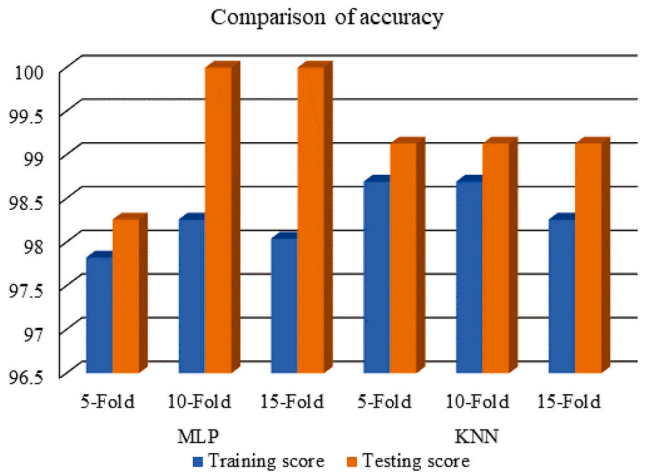


Fig. 19. Training and testing accuracy for MLP and KNN classifiers.

Table 9
Learning Curve (Average).

Model	Performance %	
	Training	Cross-validation
MLP	93.8	93.2
	93.0	91.8
	93.0	92.8
KNN	94.4	94.2
	93.4	93.0
	93.7	93.5

demonstrates strong predictive capabilities, effectively capturing the relationships between 8 input features and predicted glucose with respect to reference values. The utilization of an MLP and KNN classifier in the classification revealed that the KNN classifier demonstrated superior performance over the MLP classifier. The MLP classifier demonstrates stable training performance at around 93.8%, highlighting high accuracy on training data, with CV averaging around 92.6%, indicating good generalization. The KNN classifier shows slightly higher training performance (94.4%) and closely aligned CV (93.6%), reflecting effective learning and generalization capabilities. The study involved the collection of samples from 289 subjects, with hypo, normal and hyperglycaemia subjects (gender: male and female, age: 19-69 years; BMI at 29 ± 6, blood glucose level through Glucometer (BG ranging from 63 to

400 mg/dL). Also, the samples included 282 T2D and 7 T1D patients, overseen by a medical practitioner. The inclusion of samples from both patient groups is imperative for a more precise representation of real-world scenarios. Blood glucose levels were measured using the proposed glucose monitoring technique and a reference Dr Trust finger pick device. A total of 575 samples were collected, covering individuals with a history of hypoglycemia, normal, and hyperglycemia conditions. Categorizing blood glucose levels break down into hypoglycemic (< 70 mg/dL), normal (70 to 120 mg/dL), and hyperglycemic (> 120 mg/dL) ranges in fasting is crucial for both immediate clinical decision-making and long-term management. This categorization improves the interpretation of data obtained through the proposed system and the Dr Trust finger pick device, ensuring a comprehensive assessment of glycemic control. Informed consent was obtained from all participants, in compliance with the Helsinki guidelines.

The blood glucose levels were measured using a proposed glucose monitoring technique and a reference Dr Trust finger pick device, with the data obtained from a total of 575 samples collected from subjects with a history of hypoglycemia, normal, and hyperglycemia. The blood glucose levels obtained through the proposed method ranged from 63 to 404 mg/dL, with the levels categorized as hypoglycemic (BG level < 70 mg/dL), normal (69 > BG level < 180 mg/dL), and hyperglycemic (BG level > 180 mg/dL). The reference Dr Trust device uses the enzyme (FAD − GAD) on the principal Glucose dehydrogenase (GDH) flavin adenine dinucleotide (FAD), with a measuring range of 30-600 mg/dL, using 0.5 microlitre blood drop with the help of Dr Trust glucose strips. The device has been validated with HbA1C lab test of 7 patients, and the results were found to be 99% accurate with the reference device.

During the sample collection, it was observed that there was no significant correlation between finger width and finger length. The received light intensity for both hands pointing, middle and ring fingers was identical, and finger color did not affect blood glucose levels. With the prototype system setup, glucose measurement values were unaffected by changing light conditions. An analysis of the subject at age 38 was conducted to demonstrate this. Blood glucose levels were measured at 3-hour intervals throughout the day, and data was collected from various finger combinations. The proposed method was found to be suitable only for measuring fingertip tissue.

Despite these promising outcomes, the study acknowledges certain limitations for future improvements. The system has the potential to be enhanced into a continuous glucose monitoring system that can predict future glucose levels based on the patient's recent medical history. Moreover, furthermore, incorporating demographic information into the model, including factors like age and gender, holds promise for further enhancing measurements. Furthermore, ongoing investigation's primary focus has been on applying machine learning algorithms to optimize detector data for determining glucose levels. Future stages could explore integrating machine learning methodologies into the device design process itself to expand the horizons of our research.

7. Conclusion

In this study, a dual-wavelength system using 950 nm and 940 nm was proposed for non-invasive glucose level determination. The system was tested on 289 subjects, and a total of 575 real-time samples were collected and compared with measurements taken from a gold standard Dr. Trust glucose meter. To enhance accuracy, a FFNN regression model was used, resulting in an R^2 value of 99% and an MAE of 2.49 mg/dl, for real-time data analysis with the proposed method. The RMSE was found to be 3.02 mg/dl and the MSE was 9.16 $(mg/dl)^2$ compared with basic machine learning models. Two classification methods were used to predict multiclass, with 99% accuracy achieved using 2-classifiers. Based on these results, the proposed method appears to be more efficient for predicting glucose and glucose level classes. The scope of system can be further implemented for time series data with Type-I and Type-II diabetes sample for continuous glucose monitoring for future prediction of glucose based on previous history.

Funding statement

This research received no external funding.

Institutional review board statement

The data collection and measurement method were approved by Institutional Ethical Committee (IEC) VIT-AP IEC/20220803 and collected samples as per guidelines with the World Medical Association (WMA). Declaration of Helsinki. Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Informed consent

The Informed written consent was obtained from all subjects involved in the study.

CRediT authorship contribution statement

M. Naresh: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **V. Siva Nagaraju:** Writing – review & editing, Visualization, Data curation. **Sreedhar Kollem:** Writing – original draft, Investigation, Data curation. **Jayendra Kumar:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Samineni Peddakrishna:** Writing – review & editing, Supervision, Project administration, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Chat-GPT in order to improve the language proficiency. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgements

The authors would like to thank Health center, VIT-AP University, and Embedded systems Lab for testing and calibration. The referenced blood glucose concentration values are taken from human blood. These values have been collected using Dr Trust blood glucometer from the health center. In this work, all financial and material support have been done by VIT-AP University, Amaravati, Andhra Pradesh, India.

References

[1] W. H. Organization, Noncommunicable diseases, https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases, 2022. (Accessed 3 June 2022).
[2] IDF Diabetes Atlas, tenth edition, available online: https://diabetesatlas.org/atlas/tenth-edition, 2022.
[3] A.D. Association, Diagnosis and classification of diabetes mellitus, Diabetes Care 30 (2007) S42–S47.
[4] A. Heller, B. Feldman, Electrochemical glucose sensors and their applications in diabetes management, Chem. Rev. 108 (2008) 2482–2505, https://doi.org/10.1021/cr068123a.
[5] L. Tang, Shih-Jie Chang, C.-J. Chen, Jyh-Jong Liu, Non-invasive blood glucose monitoring technology: a review, Sensors 20 (2020) 6925.
[6] W. Villena Gonzales, A.T. Mobashsher, A. Abbosh, The progress of glucose monitoring—a review of invasive to minimally and non-invasive techniques, devices and sensors, Sensors 19 (4) (2019) 800.
[7] T. Wu, X. You, Z. Chen, Hollow microneedles on a paper fabricated by standard photolithography for the screening test of prediabetes, Sensors 22 (11) (2022) 4253.
[8] C. Chen, X.-L. Zhao, Z.-H. Li, Z.-G. Zhu, S.-H. Qian, A.J. Flewitt, Current and emerging technology for continuous glucose monitoring, Sensors 17 (1) (2017) 182.
[9] J. Smith, The pursuit of noninvasive glucose: hunting the deceitful Turkey, 2015, revised and expanded.
[10] P. Jain, A. Joshi, S. Mohanty, iglu: an intelligent device for accurate noninvasive blood glucose-level monitoring in smart healthcare, IEEE Consum. Electron. Mag. 9 (1) (2019) 35–42, https://doi.org/10.1109/MCE.2018.2883500.
[11] A. Hina, W. Saadeh, Noninvasive blood glucose monitoring systems using near-infrared technology—a review, Sensors 22 (13) (2022) 4855.
[12] J. Ikbal, G.S. Lim, Z. Gao, The hybridization chain reaction in the development of ultrasensitive nucleic acid assays, TrAC, Trends Anal. Chem. 64 (2015) 86–99.
[13] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to Linear Regression Analysis, John Wiley & Sons, 2021.
[14] Y. Ever, K. Dimililer, B. Sekeroglu, Comparison of machine learning techniques for prediction problems, Adv. Intell. Syst. Comput. 927 (2019) 713–723.
[15] M. Solihin, Y. Shameem, T. Htut, C. Ang, M. Hidayab, Non-invasive blood glucose estimation using handheld near infra-red device, Int. J. Recent Technol. Eng. 8 (2019) 16–19.
[16] E. Alpaydin, Introduction to Machine Learning, MIT Press, 2020.
[17] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 2020.
[18] Y. Freund, R. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, in: European Conference on Computational Learning Theory, Springer, 1995, pp. 23–37.
[19] H. Drucker, Improving regressors using boosting techniques, in: Proceedings of the IC, vol. 97, 1997, pp. 107–115.
[20] J. Ferreira, I. Pires, G. Marques, N. Garcia, E. Zdravevski, P. Lameski, F. Flórez-Revuelta, S. Spinsante, Identification of daily activities and environments based on the adaboost method using mobile device data: a systematic review, Electronics 9 (2020) 192.
[21] A. Rajkomar, J. Dean, I. Kohane, Machine learning in medicine, N. Engl. J. Med. 380 (14) (2019) 1347–1358.
[22] S. Kumar, S. Gadadhar, A random forest classifier based on genetic algorithm for cardiovascular diseases diagnosis (research note), Int. J. Eng. 30 (2017) 1723–1729.
[23] O. Mujahid, I. Contreras, J. Vehi, Machine learning techniques for hypoglycemia prediction: trends and challenges, Sensors 21 (2021) 546.
[24] T.T. Islam, M.S. Ahmed, M. Hassanuzzaman, S.A. Bin Amir, T. Rahman, Blood glucose level regression for smartphone ppg signals using machine learning, Appl. Sci. 11 (2) (2021) 618.
[25] M. Lee, T.M. Gatton, K.K. Lee, A monitoring and advisory system for diabetes patient management using a rule-based method and knn, Sensors 10 (2010) 3934–3953.
[26] G. Battineni, G.G. Sagaro, C. Nalini, F. Amenta, S.K. Tayebati, Comparative machine-learning approach: a follow-up study on type 2 diabetes predictions by cross-validation methods, Machines 7 (2019) 74.
[27] I.I. Abiodun, A.H. Jantan, A. Omolara, K. Dada, N. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: a survey, Heliyon 4 (6) (2018) e00938.
[28] Z. Car, S. Baressi, N. Anđelić, I. Lorencin, V. Mrzljak, Modeling the spread of Covid-19 infection using a multi-layer perceptron, Comput. Math. Methods Med. 2020 (2020) 5714714.
[29] M. Zhao, P.S. Leung, Revisiting the use of biological fluids for non-invasive glucose detection, Future Med. Chem. 12 (2020) 645–647.
[30] M. Hassan, C. Vyas, B. Grieve, P. Bartolo, Recent advances in enzymatic and non-enzymatic electrochemical glucose sensing, Sensors 21 (13) (2021) 4672.
[31] R. Zhang, Noninvasive electromagnetic wave sensing of glucose, Sensors 19 (5) (2019) 1151.
[32] A. Caduff, M.S. Talary, M. Mueller, F. Dewarrat, J. Klisic, M. Donath, L. Heinemann, W.A. Stahel, Non-invasive glucose monitoring in patients with type 1 diabetes: a multisensor system combining sensors for dielectric and optical characterisation of skin, Biosens. Bioelectron. 24 (9) (2009) 2778–2784.

[33] M. Nakamura, M. Taki, Y. Kuwahara, T. Iwasaki, S. Shoji, T. Mizutani, M. Nishizawa, H. Sato, Selectivity-enhanced glucose measurement in multicomponent aqueous solution by broadband dielectric spectroscopy, in: 2016 IEEE MTT-S International Microwave Symposium (IMS), IEEE, 2016, pp. 1–4.
[34] M. Hofmann, R. Weigel, U. Helbig, A. Kölpin, A. Hagelauer, Microwave-based noninvasive concentration measurements for biomedical applications, IEEE Trans. Microw. Theory Tech. 61 (5) (2013) 2195–2204.
[35] M. Abdolrazzaghi, V. Nayyeri, F. Martin, Techniques to improve the performance of planar microwave sensors a review and recent developments, Sensors 22 (12) (2022) 6946.
[36] Y. Jin, Y. Yin, C. Li, H. Liu, J. Shi, Non-invasive monitoring of human health by photoacoustic spectroscopy, Sensors 22 (3) (2022) 1155.
[37] O. Amir, D. Weinstein, S. Zilberman, M. Less, D. Perl-Treves, H. Primack, A. Weinstein, E. Gabis, B. Fikhte, A. Karasik, Continuous noninvasive glucose monitoring technology based on “occlusion spectroscopy”, J. Diabetes Sci. Technol. 1 (2007) 463–469.
[38] A. Fercher, Optical coherence tomography principles and applications, Diabetes Technol. Ther. (2003).
[39] R.L. Shelton, W. Jung, S.I. Sayegh, D.T. McCormick, J. Kim, A. Quach, G.L. Horowitz, B.-H. Park, S.-H. Yun, Optical coherence tomography for advanced screening in the primary care office, J. Biophotonics 7 (2014) 525–533.
[40] B.H. Malik, G.L. Coté, Real-time, closed-loop dual-wavelength optical polarimetry for glucose monitoring, J. Biomed. Opt. 15 (2010) 017002.
[41] J.-M. Buchert, Thermal emission spectroscopy as a tool for noninvasive blood glucose measurements, Proc. SPIE 5566 (2004) 100–111.
[42] J.L. Lambert, J.M. Morookian, S.J. Sirk, M.S. Borchert, Measurement of aqueous glucose in a model anterior chamber using Raman spectroscopy, J. Raman Spectrosc. 33 (7) (2002) 524–529.
[43] N. Kuhar, S. Sil, T. Verma, S. Umapathy, Challenges in application of Raman spectroscopy to biology and materials, RSC Adv. 8 (2018) 25888–25908.
[44] A.M. Enejder, T.G. Scecina, J.-H. Oh, M. Hunter, W.-C. Shih, S. Sasic, G.L. Horowitz, M.S. Feld, Raman spectroscopy for noninvasive glucose measurements, J. Biomed. Opt. 10 (3) (2005).
[45] M. Naresh, A. G, S. Peddakrishna, P. Kumar, A. Singh, Design and implementation of low-cost smart contactless thermometer with polynomial regression model, 2022, pp. 303–305.
[46] O.S. Khalil, Non-invasive glucose measurement technologies: an update from 1999 to the dawn of the new millennium, Diabetes Technol. Ther. 6 (5) (2004) 660–697.
[47] R.A. Shaw, H.H. Mantsch, Infrared spectroscopy in clinical and diagnostic analysis, in: Encyclopedia of Analytical Chemistry, John Wiley & Sons, 2006, pp. 83–102.
[48] M. Goodarzi, W. Saeys, Selection of the most informative near infrared spectroscopy wavebands for continuous glucose monitoring in human serum, Talanta 146 (2016) 155–165.
[49] A. Aloraynan, S. Rassel, C. Xu, D. Ban, A single wavelength mid-infrared photoacoustic spectroscopy for noninvasive glucose detection using machine learning, Biosensors 12 (3) (2022) 166.
[50] S. Gebhart, M. Faupel, R. Fowler, C. Kapsner, D. Lincoln, V. McGee, J. Pasqua, L. Steed, M. Wangsness, F. Xu, Glucose sensing in transdermal body fluid collected under continuous vacuum pressure via micropores in the stratum corneum, Diabetes Technol. Ther. 5 (2) (2003) 159–166.
[51] A. Hina, W. Saadeh, Non-invasive blood glucose monitoring systems using near-infrared technology—a review, Sensors 22 (2022) 4855.
[52] W. Yang, N. Liao, H. Cheng, Y. Li, X. Bai, C. Deng, Determination of nir informative wavebands for transmission non-invasive blood glucose measurement using a Fourier transform spectrometer, AIP Adv. 8 (3) (2018) 035216.
[53] N.C. Dingari, I. Barman, G.P. Singh, J.W. Kang, R.R. Dasari, M.S. Feld, Investigation of the specificity of Raman spectroscopy in non-invasive blood glucose measurements, Anal. Bioanal. Chem. 400 (1) (2011) 287–299.
[54] Y. Uwadaira, A. Ikehata, A. Momose, M. Miura, Identification of informative bands in the short-wavelength nir region for non-invasive blood glucose measurement, Biomed. Opt. Express 7 (2016) 2729–2737.
[55] Y. Uwadaira, N. Adachi, A. Ikehata, S.-i. Kawano, Factors affecting the accuracy of non-invasive blood glucose measurement by short wavelength near infrared spectroscopy in the determination of the glycaemic index of foods, J. Near Infrared Spectrosc. 18 (5) (2010) 291–300.
[56] S. Haxha, J. Jhoja, Optical based noninvasive glucose monitoring sensor prototype, IEEE Photonics J. 8 (2016) 1–11.
[57] P. Jain, R. Madilla, A. Joshi, A precise non-invasive blood glucose measurement system using nir spectroscopy and Huber’s regression model, Opt. Quantum Electron. 51 (2019) 1–15.
[58] S.H. Lee, M.S. Kim, O.-K. Kim, H.-H. Baik, J.-H. Kim, Near-infrared light emitting diode based non-invasive glucose detection system, J. Nanosci. Nanotechnol. 19 (10) (2019) 6187–6191.
[59] M. Naresh, S. Peddakrishna, M. Thirupathi, Non-invasive glucose measurement with 940 nm sensor using short wave nir technique, in: Intelligent Computing and Optimization, Springer Nature Switzerland, Cham, 2023, pp. 95–103.
[60] J. Yadav, A. Rani, V. Singh, B.M. Murari, Near-infrared led based non-invasive blood glucose sensor, in: 2014 International Conference on Signal Processing and Integrated Networks (SPIN), 2014, pp. 591–594.
[61] P. Anupongongarch, Y. U-Pratya, N. Damrongwiriyanupap, W. Kositanurit, Development of a non-invasive blood glucose sensor, Int. J. Appl. 12 (2019).
[62] M. Naresh, S. Peddakrishna, Non-invasive near-infrared-based optical glucose detection system for accurate prediction and multi-class classification, Int. J. Exp. Res. Rev. 31 (Spl Volume) (2023) 119–130, https://doi.org/10.52756/10.52756/ijerr.2023.v31spl.012
.
[63] M. Naresh, S. Peddakrishna, Non-invasive glucose measurement using 950 nm reflective short wave nir technique, Res. Biomed. Eng. 39 (2023) 747–757, https://doi.org/10.1007/s42600-023-00308-5.
[64] K.V. Larin, M.S. Eledrisi, M. Motamedi, R.O. Esenaliev, Noninvasive blood glucose monitoring with optical coherence tomography: a pilot study in human subjects, Diabetes Care 25 (12) (2002) 2263–2267.
[65] P.P. Pai, A. De, S. Banerjee, Accuracy enhancement for noninvasive glucose estimation using dual-wavelength photoacoustic measurements and kernel-based calibration, IEEE Trans. Instrum. Meas. 67 (1) (2017) 126–136, https://doi.org/10.1109/TIM.2017.2763880.
[66] H. Ali, F. Bensaali, F. Jaber, Novel approach to non-invasive blood glucose monitoring based on transmittance and refraction of visible laser light, IEEE Access 5 (2017) 9163–9174.
[67] J. Zhu, Y. Zhou, J. Huang, A. Zhou, Z. Chen, Noninvasive blood glucose concentration measurement based on conservation of energy metabolism and machine learning, Sensors 21 (21) (2021) 6989.