NIR Spectroscopy based Non-Invasive Blood Glucose Concentration Measurement using BP Algorithm

Hyok Chol Song, Chol Jin Oh, Chol Hyon Sim, Chol Min Won

Abstract

Diabetes mellitus poses a significant challenge in clinical settings, necessitating frequent blood glucose measurements for insulin dosage determination. Conventional invasive methods, such as finger pricking, carry risks of infection and skin callusing. Non-invasive monitoring techniques offer a promising alternative for patients with hyperglycemia or hypoglycemia, enabling regular self-monitoring and advancing diabetes research. However, the accuracy and universality of most current non-invasive methods for measuring blood glucose concentration (BGC) remain inadequate. Achieving clinical credibility requires the elimination of individual discrepancies (IDs) in measurements. This study focuses on enhancing monitoring accuracy to a clinically acceptable level by mitigating the effects of IDs. We conducted a detailed analysis of factors influencing Near-Infrared spectroscopy measurements to reduce prediction error. An artificial neural network with a backpropagation algorithm was employed to predict BGC from the acquired spectral data. Experimental results confirm that our proposed BGC prediction model effectively leverages IDs, achieving performance that meets clinical standards.

Keywords: NBGM, BGC, NIR, Back-propagation

1. Introduction

Diabetes mellitus is a chronic disorder characterized by abnormal blood glucose concentration. The normal blood glucose level ranges from 60 to 110 mg/dL under pre-prandial conditions and up to 140 mg/dL postprandially. An individual is considered diabetic if their glucose level exceeds 126 mg/dL pre-prandially or 200 mg/dL postprandially. The primary danger of diabetes lies in its complications. Improper management can lead to serious health problems such as heart stroke, blindness, chronic kidney failure, ulcer formation, and blood vessel damage. Due to factors including poor diet and unhealthy lifestyles, the global diabetic population is rapidly increasing. The International Diabetes Federation estimates that the number of people with diabetes will reach nearly 592 million by 2035 [1].

Since diabetes is widely considered an incurable disease, stabilizing blood glucose concentration within the normal range is the primary strategy for preventing complications. For diabetics, determining insulin dosage and diet planning based on regular and frequent blood glucose measurements is essential for disease control and achieving a long life expectancy. Diabetic patients typically need to measure their glucose concentration up to five times daily to ensure precise management. However, with commercially available invasive blood glucose monitors, such frequent measurement is nearly impossible due to problems such as the risk of infection, skin calluses, and high cost.

Non-invasive blood glucose measuring techniques based on optical signal processing are widely investigated to develop affordable and accurate monitors, but most are not yet clinically acceptable. Several methods exist, including near-infrared (NIR), mid-infrared (MIR), Raman spectroscopy, photoacoustic spectroscopy, radio wave impedance, and fluorescence [2]. This paper proposes a method using NIR spectroscopy (700-2500 nm) for non-invasive glucose measurement. A wavelength of 940 nm was selected based on the characteristic glucose absorption peak in this range [3, 4].

Some previous studies have attempted to establish a direct relationship between analog-processed PPG signals and invasively measured glucose concentrations without accounting for individual discrepancies (IDs). Their blood glucose concentration (BGC) prediction models were based on linear or polynomial regression. Although their results were sometimes acceptable, the experimental data sets were often too small to produce a universal prediction model. For instance, Gayathri B et al. [5] developed linear and polynomial regression models using data from only 24 subjects aged 20 to 30. P. Daarani et al. [6] used 24 data samples to establish a second-degree polynomial relationship between measured and reference values. Kamol Sarkar et al. [7] collected 20 samples from 10 subjects (aged 22 to 48, BGC range 85–260 mg/dL) to derive a fourth-degree polynomial relationship.

Conversely, many researchers have recognized that individual factors influence non-invasive measurements and have attempted to use more advanced regression methods. Machine learning algorithms, such as Random Forest, have been proposed for BGC estimation. In other research, support vector regression (SVR) has been used to process multi-dimensional data. While these results have reached clinically acceptable levels, the databases used for this supervised learning were often small and narrow. The reference glucose values in these studies were typically within the normal range, or the number of patients was insufficient to confirm that the estimation models could effectively overcome IDs. For example, Kasun D. Pathirana et al. [3] used 315 data samples, but the glucose range was only 90–140 mg/dL. Zheng Li et al. [8] reported that EMD-based Random Forest modeling performed well within a BGC range of 70–150 mg/dL. Weijie Liu et al. [9] proposed a learning algorithm to overcome IDs and found that Independent Component Analysis (ICA) combined with Random Forest (RF) yielded the best performance compared to other methods; however, their device performance was evaluated on only six people. Yue Zhang et al. [10] identified 22 IDs as input parameters for an SVR model and collected 251 samples, but the BGC range was limited to 85–115 mg/dL.

Overcoming patient IDs across a wide range of BGC values is crucial for developing non-invasive blood glucose monitors (NBGM). Numerous individual factors in the NIR light path through a finger can distort the true reflection of glucose concentration. Several studies have contributed to identifying these factors to mitigate the effect of IDs [3, 9, 10, 11]. This paper proposes an artificial neural network with backpropagation (BP) to build a prediction model capable of overcoming IDs. Based on a detailed analysis of environmental parameters and IDs, we collected 155 learning data samples from a public clinic and used them as inputs for the neural network. The results are analyzed and evaluated using the Clarke error grid and correlation coefficient.

In the subsequent sections of this paper, Section 2 elaborates on the materials and methods used in the research. The study results are presented in Section 3, and the conclusion and directions for future development are provided in the final section.

2. MATERIAL AND METHODS

The overall workflow is as the following figure. In this section, the specification of hardware configuration is first covered and then theoretical basis and neural networking are mentioned in the second.

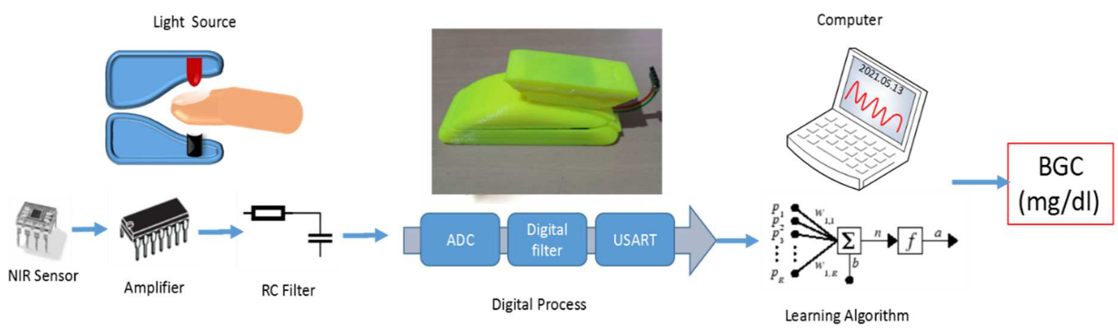

Fig. 1 Workflow of NIR glucose monitoring system

2.1 Hardware setting

2.1.1 Emitter and sensor design

A NIR transmitter and its detector are placed on opposite sides of fingers. The wavelength of the emitter is decided as 940nm because of its glucose absorption property. According to experiments conducted by Yang et al.(2018), it was demonstrated that NIR absorption by glucose molecules shows peak values at certain wavelength bands:near of 940nm. The research conducted by Yadav(2015) and his team confirmed the effectiveness of using 940nm. According to their result, 940nm wavelength also has the absorption peak by glucose. Furthermore, water spectra of under the wavelength shows approximately zero absorption meaning that water in human body would not affect the detection of glucose signal.

In some research, multi-wavelength technique is used to detect PPG signal. [4,7,9] However, single wavelength based technique is developed in this paper for its cost-effectiveness. TSAL5300 NIR LED was chosen as an emitter and OPT101(8-pin DIP) is used to detect 940nm wavelength. OPT101 shows the acceptable spectral responsivity between the wavelength of 900 and 1000nm according to the datasheet. In other researches, OPT101 was also selected for its high sensitivity of NIR spectra [1,12]. The following figure displays the circuit diagram of OPT101.

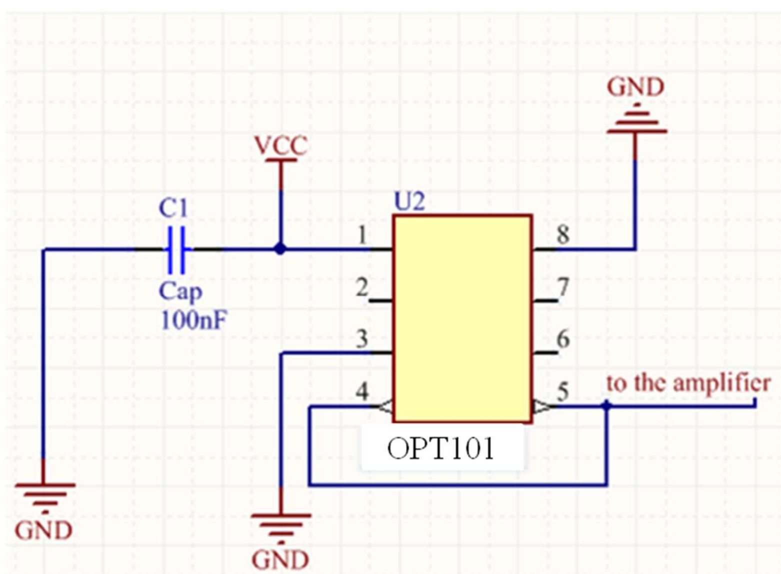

Fig. 2 Circuit diagram of OPT101

2.1.2 Analog rectifier and passband amplifier parameter decision

The PPG signal received by OPT101 is converted to voltage. The converted signal voltage range is from 0 to 20mV. It is needed to amplify the output signal and to condition the signal to eliminate noise and DC part. High-pass filter with the cutoff frequency 0.7Hz is implemented to remove lower frequency signals including DC part which is due to surrounding subjects. The formula for the cutoff frequency of the pass filter is as follows.

$$f = \frac{1}{2 \times \pi \times C \times R} \qquad (1)$$

The parameters of the high-pass filter were designed as C_{high}=1uF and R_{high}=230 $k\Omega$ according to this formula. Then, the output of the high-pass filter goes into a low-pass filter. The cutoff frequency of the low-pass filter should be greater than 2 since maximum human heart rate is 120bpm. The capacitance of the low-pass filter is 100nF and the resistance is 680 $k\Omega$.

The filtered signal is amplified about 202 times thorough two stages. The formula to calculate the gain is as follows.

$$A = 1 + \frac{R5}{R4} \qquad (2)$$

According to the formula, resistance is decided for R4, R5, R6 and R7 as in **Fig. 3** The output is then fed into a microcontroller analog pin to be processed. Fig.3 show

the circuit diagram of the implemented analog signal processing unit.

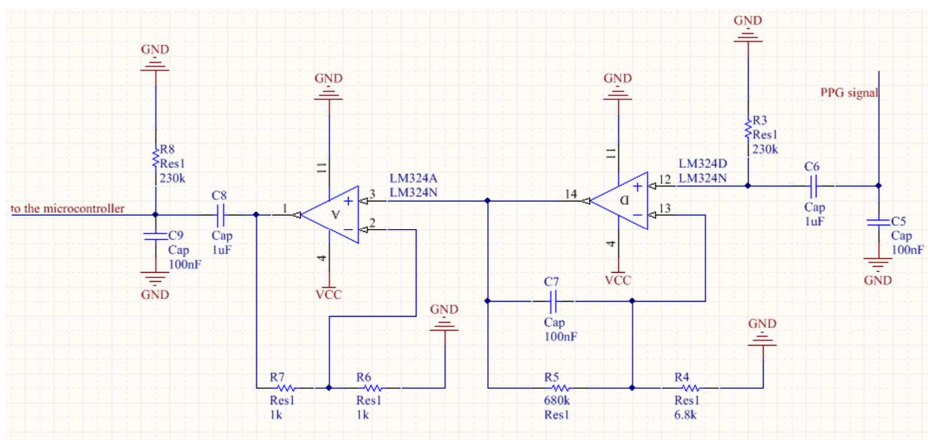

Fig. 3 Circuit diagram of analog amplifier and band pass filter

2.1.3 Skin temperature sensor

Skin temperature measurement was required to obtain the input parameter of learning algorithm. LM35-precision centigrade temperature sensor was chosen for its acceptable accuracy and proper measurement range for human body [13].

2.1.4 Microcontroller configuration for the experiment

Microcontroller STM32f103 was chosen to convert analog signal to digital one and to implement a digital filter. The OPT101 signal and the temperature sensor signal are fed into different analog pins of STM32f103 and then converted into digital value in the range of 0 and 4091. Digital filter was designed and implemented to remove high frequency noise that is caused by circuit elements.**Fig.4**and **Fig.5**show the prototype of monitor and the system architecture we developed.

Fig. 4 Hardware setup

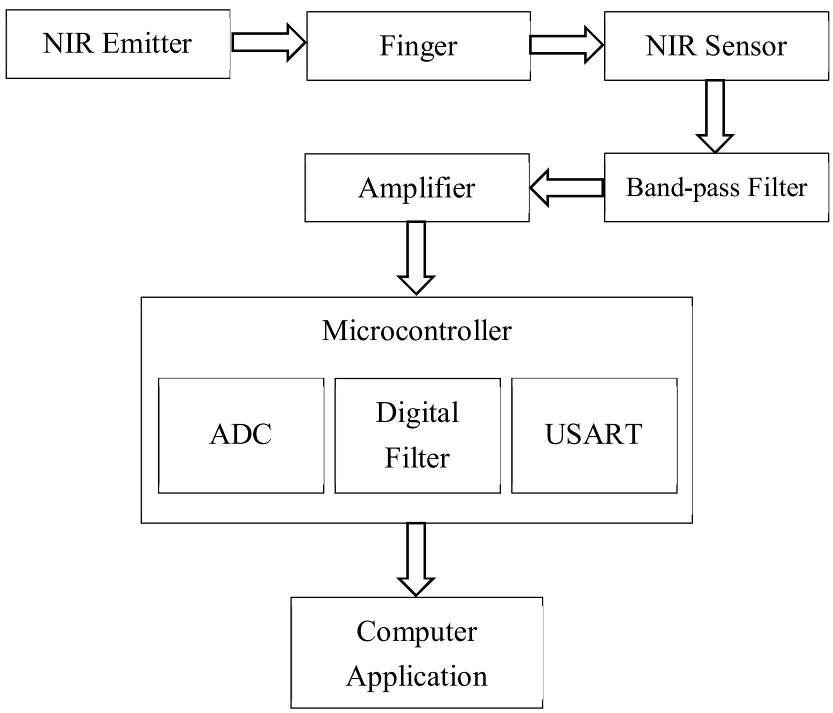

Fig. 5 System architecture

2.2METHODS

2.2.1 Main principle of NIR spectroscopy

When the light from the NIR emitter penetrates finger, it is absorbed and scattered. The scattered light ray is detected by the sensor. The scattered light reaches the sensor and drives output voltage signal of the sensor. The more ray arrives, the higher the output voltage of the sensor.

The development of NIR based glucose monitor has the Beer-Lambert Law as its theoretical basis. The Beer-Lambert Law states that the absorbance of a solution is directly proportional to the concentration of the absorbing material in the solution and the path length of the ray. It is represented as the following equation.

$$I = I_0 \cdot e^{-\mu_{eff} \bullet L} \qquad (3)$$

where, I is the transmitted light intensity, I_0 is the incident light intensity and L is the optical path length of NIR inside the tissue.

$$\mu_{eff} = [3\mu_s(\mu_s + \mu_s')]^{1/2} \qquad (4)$$

$$\mu_s = 2.303 \cdot \varepsilon \cdot C \qquad (5)$$

$$\mu_s' = \mu_s(1-g) \qquad (6)$$

Where ε is the molar extinction coefficient, C is the chromosome concentration and μ_s' indicates the reduced scattering coefficient. The absorption coefficient μ_{eff} is defined as equation (4) in terms of scattering coefficient μ_s. From these equations above, absorption coefficient is proportional to the scattering coefficient and the scattering coefficient is proportionally dependent on the chromosome concentration. Thus, transmitted light intensity exponentially decreases leading voltage of the sensor as glucose concentration increases and finger thickness increases.

2.2.2 Regression model

■ Consideration of IDs

To determine BGC value in an acceptable level, different factors that influence actual glucose change and measurement should be considered. The thing to clarify is that the factors that can fake the measurement should only be identified and compensated. Other parameters that contribute glucose concentration change are not needed to be considered for the development of glucose monitor. Therefore, IDs that affect BGC were divided into two categories: ones that affect real BGC value itself and ones that have possibility to influence on the measurement. According to an article "42 factors that affect glucose" from http://diatribe.org, there are 42 factors that affect human BGC level. 11 factors were selected as effectors to measurement and the treatment methods of these factors in experiments were decided. Here, factors that can affect optical measurement were treated in two different ways. Some were selected as input parameters of neural network while others were decided as standard measurement condition. Table 1 is the list of IDs that can affect the measurement and their treatment methods.

Table1. The list of IDs and their treatment

No	Factors	Treatment
1	Age	Record and use it as an input to ANN.
2	Gender	Record and use it as an input to ANN.
3	Scar tissue	Avoid fingers with scar for measurement.
4	Exercise	Avoid exercise before measurement.
5	Prandial status	Record the status and set as an input to ANN.
6	Skin temperature	Record and use it as an input to ANN.
7	Skin humidity	Dry hands before measurement.
8	Skin color	Measure only yellowish skin subjects. (Korean is yellow race.)

- Consideration of environmental factors

The PPG signal detected by NIR sensor(TK11) contains DC component and AC component. DC component mainly represents scattered light from tissue and constant noise while AC part of the output voltage shows the blood volume change according to systolic and diastolic phases of the cardiac cycle.

Environmental factors can influence the output voltage of the photo detector. Ambient temperature, humidity and light affect NIR based measurement. First, ambient temperature and humidity influence on the behavior of photo detector – dark current and quantum efficiency [11]. Then, ambient light that penetrates finger can reach the photodiode and produce constant offset in the reading. The effect of these environmental factors are represented in DC part of the reading so high-pass filter could effectively eliminate it.

- BP(back-propagation) learning model

Back-propagation is the representative learning rule of multi-layer neural network. Multi-layer neural network is developed to overcome the limitation of single-layer architecture. The main drawback of single-layer neural network is that it can only be applicable in linear relationship problems. The hidden layers of multi-layer network are capable of solving non-linear problems which we meet in most practical cases. In BP algorithm, the input travels through input layer, hidden layers and output layer and then error starts from the output layer and propagate backwards towards the input layer to calculate delta. **Fig.5**shows the learning model of BGC with 11inputs and 1 hidden layer. Tangent function was used as an activation function.

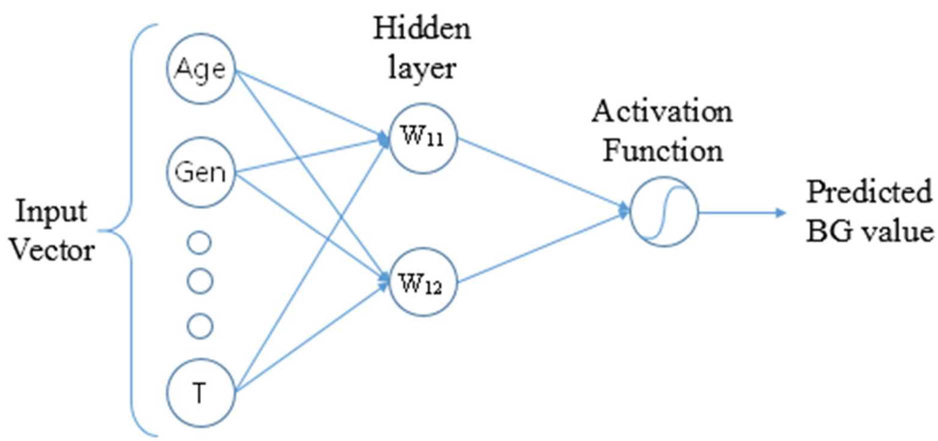

Fig. 6 The learning model of BGC

3. RESULTS

Data Collection

All the experiment was conducted in the morning and in the same room. 107 volunteers participated in the experiment and 155 valid records were collected from them. The data collection experiment was conducted two times in the morning; before breakfast and 2hours after breakfast. Among the volunteers, there were 52 males and 55 females. The range of their age was from 18 to 79 and BGC range was from 59 to 569 mg/dl. Each record of dataset had 10 attributes; gender, age, height, weight, finger temperature, finger thickness, blood pressure and fasting as IDs and monitor output and reference BGC as measured values.

Data modelling

Based on the above system configuration and data collection, we used back-propagation algorithm to o btain optimized prediction model. We configured a m ulti-layer neural network with one hidden layer and t hen repeated modelling by modifying several paramet ers of the network; learning rate, batch size, activatio n function and the number of epoch. To achieve opti mization, we tested input parameters whether each of them helps improve modelling. From the test, we c ould identify that blood pressure doesn't contribute to learning result. Therefore, this factor was eliminated. 90% of the data was used for training of neural net work and the rest 10% was used for validation.

Result analysis

Clarke Error Grid and correlation coefficient were employed to analyze the result. Clarke Error Grid analysis is used to evaluate clinical accuracy of the prediction model [8,9,10]. The finger-pricking BGC value (reference value) is represented by x-axis while y-axis shows predicted value by neural network. The grid is divided into 5 regions. Zone A represents that the non-invasive estimation of glucose concentration will lead correct decision and treatment of diabetes. Values falling within zone B would not be correct for accurate treatment and can lead no treatment. Zone C, D and E means "overcorrecting", "dangerous failure" and "erroneous treatment" zones respectively. In summary, zone A and B are clinically accurate and acceptable region, whereas the points in zone C, D and E would result serious mistake in treatment. Clarke Error Grid analysis was programmed in MATLAB referring the code from the site http://github.com . **Fig.7**shows the distribution of points in Clarke Error Grid. The grid analysis shows that 62.5% of test data is placed in region A and 37.5% of data placed in region B. No point is distributed in C, D, E regions where would lead incorrect clinical decision. The result represents that correlation coefficient(R^2) between reference value and predicted value is 0.87. This means that all the predicted BGC values are clinically correct and acceptable.

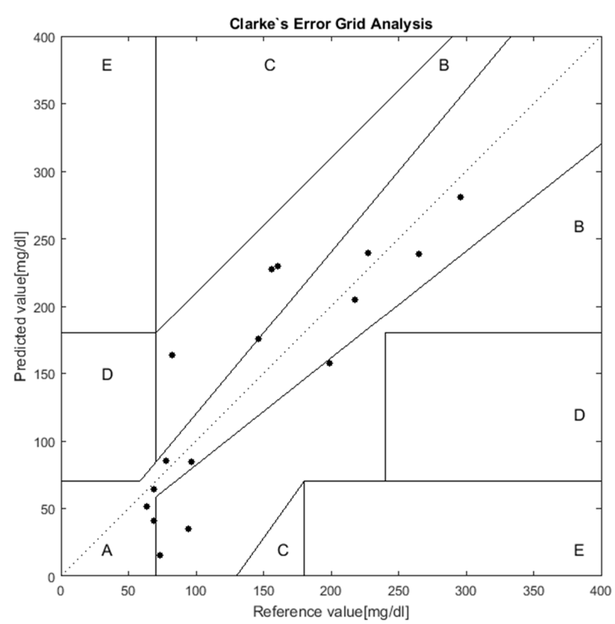

Fig. 7 The Clarke error grid analysis between reference BGC value and Predicted BGC value(R^2=0.87)

4. *DISCUSSION*

This paper has provided the study of NIR spectroscopy BGC prediction modelling by artificial neural network. First, the hardware that emits and receives NIR spectra and processes analog and digital signal is configured. Next, PP(physical parameter) and EP(environmental parameter)analysis has been done for deciding input vector of ANN learning. Then, experiments are conducted in public clinic to collect data samples with the device. Finally, artificial intelligence learning using back-propagation algorithm has been implemented to find non-linear model of BGC prediction from processed signal and IDs.

We confirmed that the BGC prediction with suggested model produces clinically acceptable result. Clarke error grid and correlation coefficient proves this hypothesis. Even though the limitation of the number of sample data caused some inaccuracy, this study provides the feasibility of developing NIR based non-invasive BGC monitor which is clinically available.

Our future research will be focused on increasing sample data with wide glucose and age range. Since our experiment was conducted in summer, temperature range was not broad. Therefore, we will also focus on improving hardware structure to confirm reproducibility of the device regardless of season and measuring time.

Compliance with Ethical Standards

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent: Informed consent was obtained from all individual participants included in the study.

References

[1] Yadav J, Rani A, Singh V, Muran BM. Near-infrared LED based non-invasive blood glucose sensor. In: Proceedings of the International Conference on Signal Processing and Integrated Networks; 2014; Noida, India. p. 123-127.

[2] Lin T, Gal A, Mayzel Y, Horman K, Bahartan K. Non-invasive glucose monitoring: a review of challenges and recent advances. J Biomed Sci Eng. 2017;10(6):1-15.

[3] Pathirage KD, Roopasinghe P, et al. Removing subject dependencies on non-invasive blood glucose measurement using hybrid techniques. In: Proceedings of the 2019 IEEE Conference; 2019; p. 7197-200.

[4] Jain P, Maddila R. A precise non-invasive blood glucose measurement using NIR spectroscopy and Huber's regression model. Opt Quantum Electron. 2019;51(1):1-15.

[5] Gayathri B, Sruthi K, Menon KAU. Non-invasive blood glucose monitoring using near infrared spectroscopy. In: Proceedings of the 2017 IEEE International Conference on Communication and Signal Processing; 2017; Chennai, India. p. 1139-42.

[6] Daarani P, Kavithamani A. Blood glucose level monitoring by non-invasive method using near infrared sensor. Int J Latest Trends Eng Technol. 2017;8(3):141-7.

[7] Sarkar K, Ahmad D, et al. Design and implementation of a noninvasive blood glucose monitoring device. In: Proceedings of the 21st International Conference on Computer and Information Technology (ICCIT); 2018 Dec 21-23; Dhaka, Bangladesh.

[8] Li Z, et al. Wearable non-invasive blood glucose estimation via empirical mode decomposition based

hierarchical multiresolution analysis and random forest. In: Proceedings of the 2016 IEEE Conference; 2016.

[9] Liu W, et al. Overcoming individual discrepancies: a learning model for non-invasive blood glucose measurement. Appl Sci (Basel). 2019 Jan;9(15):3043.

[10] Zhang Y, Wang Z. Non-invasive blood glucose estimation using near-infrared spectroscopy based on SVR. In: Proceedings of the 2017 IEEE Conference; 2017; p. 594-8.

[11] Pathirage KD, et al. Removing subject dependencies on non-invasive blood glucose measurement using hybrid techniques. In: Proceedings of the 2019 IEEE Conference; 2019; p. 7197-200.

[12] Truong VA, Hien P. Non-invasive glucose monitoring system utilizing near-infrared technology. In: Vasant P, et al., editors. Intelligent Computing and Optimization. Singapore: Springer; 2020. p. 120-9.

[13] Liu W, et al. Transfer and collaborative learning method for personalized noninvasive blood glucose measurement modeling. In: Proceedings of the 2019 IEEE Conference; 2019; p. 1179-83.