Biomedical Signal Processing and Control 106 (2025) 107750

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Infrared absorption spectroscopy-based non-invasive blood glucose monitoring technology: A comprehensive review

Taixiang Li [a], Quangui Wang [b], Ying An [a], Lin Guo [a], Linan Ren [a], Linghao Lei [a], Xianlai Chen [a,*]

[a] Big Data Institute, Central South University, Changsha, China
[b] Hunan Longgang Intelligent Technology Co., Ltd, Changsha, China

ARTICLE INFO

Keywords:
Non-invasive blood glucose monitoring
Infrared absorption spectroscopy
Machine learning
Multi-sensor
Photoplethysmography
Wavelength selection

ABSTRACT

Diabetes, characterized by hyperglycemia, is an incurable metabolic disorder with an alarmingly high prevalence rate. Self-monitoring of blood glucose holds exceptional significance in diabetes management. However, traditional invasive blood glucose monitoring devices have imposed inconvenience and discomfort on patients. This has propelled research in non-invasive blood glucose monitoring into the forefront, offering substantial clinical utility. In this survey, we reviewed the major technologies of non-invasive blood glucose monitoring based on absorption spectroscopy, including physical methodologies, signal and data processing techniques, and the progress in commercialization. This review can serve as an introduction to the modeling principles of non-invasive blood glucose monitoring, or as a collection of technical application methods of non-invasive glucose monitoring.

Contents

* Corresponding author.
E-mail address: chenxianlai@csu.edu.cn (X. Chen).

https://doi.org/10.1016/j.bspc.2025.107750
Received 5 February 2024; Received in revised form 20 January 2025; Accepted 15 February 2025
Available online 1 March 2025

1. Introduction

1.1. Overview of diabetes

According to the "IDF Diabetes Atlas" published in 2021, there are more than 530 million people in the world suffering from diabetes, and this number is expected to exceed 780 million by the end of 2045 [1]. In view of current medical restrictions, there is no way to completely eradicate diabetes, but there are still many factors that affect the way diabetes is controlled. To reduce the risk of diabetes-related complications, patients are often advised to follow the guidelines of healthcare providers on diet, exercise, self-monitoring, and medication regimens [2]. Among these measures, self-monitoring is crucial, as it allows patients to better understand their glycemic patterns and adjust their management strategies accordingly. In addition, early diabetes does not show obvious symptoms, making it difficult to distinguish and diagnose. Up to 70% of people with prediabetes eventually develop type 2 diabetes, which is one of the top 10 causes of death globally and is associated with comorbidities such as cardiovascular disease, kidney disease, neuropathy, and retinopathy [2]. Therefore, monitoring blood glucose concentration plays an important role in the treatment and early screening of diabetes.

1.2. Overview of invasive method

Currently, the methods used for glucose monitoring are mainly invasive. Depending on the type of sampling, these methods can be concluded as blood glucose monitoring (BGM) and continuous glucose monitoring (CGM). Among them, BGM aims to collect blood from veins or capillaries, and then apply a chemical-based method to get a relatively accurate blood glucose reading. Nonetheless, for each measure reading, it is accompanied by a blood draw, which can be distressing for individuals who require frequent monitoring. CGM, on the other hand, continuously tracks glucose levels using a sensor implanted just under the skin. It measures glucose in the interstitial fluid (ISF) rather than directly from the blood, which can result in a slight "lagging effect" between the sensor reading and real-time blood glucose levels [3]. Previously, CGM systems required regular calibration through fingerstick blood glucose readings to maintain accuracy. However, with the advent of factory-calibrated CGM devices, such as the Dexcom G6, calibration by the user is no longer necessary. Despite these advancements, CGM devices remain relatively expensive, and some users may experience skin irritation, contributing to psychological discomfort.

1.3. Overview of non-invasive method

The progress of sensor technology has progressively transformed the conventional method of blood detection into non-invasive technology, which does not require the use of instruments that penetrate the skin or

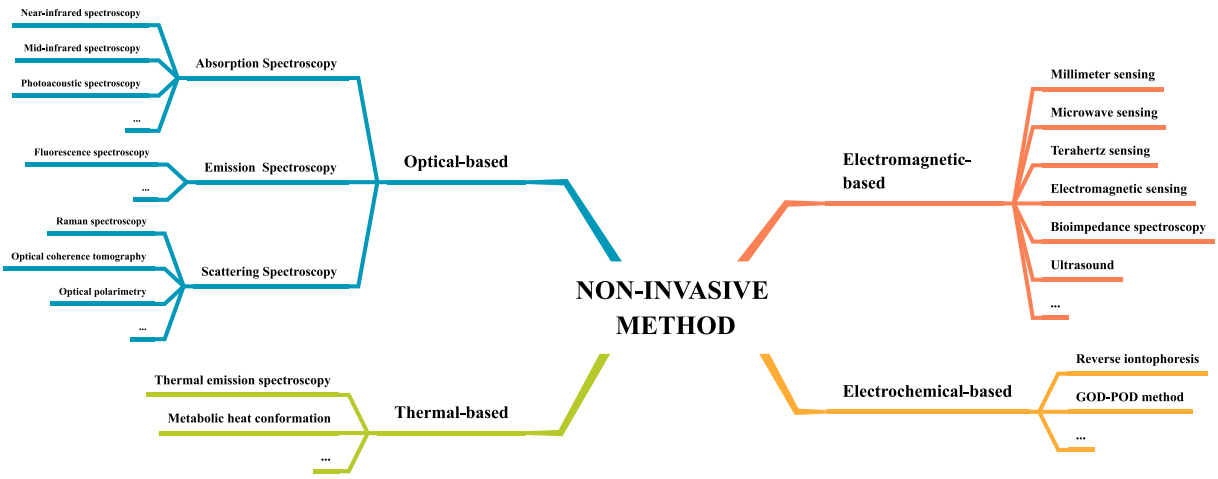

Fig. 1. Overview of sensor-based non-invasive glucose monitoring technologies.

physically enter the body. The investigation of blood glucose-specific sensors in subcutaneous blood glucose non-invasive extraction technology encompasses a range of techniques including optical, thermal, electromagnetic, and electrochemical methods (see Fig. 1).

The optical approach employs techniques that utilize the absorption, emission, and scattering properties of light as it passes through biological media in all spectral bands, including near-infrared (NIR) spectroscopy [4,5], mid-infrared (MIR) spectroscopy [6,7], optical coherence tomography (OCT) [8], fluorescence spectroscopy [9], Raman spectroscopy [10,11], optical polarimetry (OP) [12], photoacoustic spectroscopy [13]. The thermal metabolism method monitors glucose by detecting physiological indicators related to the generation of metabolic heat from the glucose molecule itself, including thermal emission spectroscopy [14,15] and the metabolic heat conformation (MHC) [16–18]. The electrical method utilizes the dielectric properties of glucose to detect blood glucose by emitting a small amount of electromagnetic radiation, electric current or ultrasonic waves, and correlating the dielectric constant of the tissue with the glucose concentration. Methods include millimeter sensing [19], microwave seining [20], terahertz sensing [21], electromagnetic sensing [22], bioimpedance spectroscopy [23], ultrasound [24]. Electrochemical methods mostly adopt the invasive subcutaneous implantation method for blood glucose monitoring, and some non-invasive research involves electrochemically analyzing the secretion of the human body (such as sweat, saliva, etc.), detecting the electrochemical signal of glucose in it, and then calculating the blood glucose concentration. Methods mainly include reverse iontophoresis (RI) [25], enzymatic method [26], non-enzymatic method [27], and others. Furthermore, there are studies aimed at predicting glucose levels by incorporating other physiological indicators, such as heart rate (HR) [28], oxygen saturation (SPO2), and electrocardiogram (ECG) [29,30].

Some of these methods have achieved phased results in the laboratory, but few methods can provide results comparable to invasive blood glucose monitoring in terms of measurement accuracy and measurement portability. As a research hotspot in the field of non-invasive detection, infrared (IR) spectroscopy technology shows great potential in the research of non-invasive blood glucose monitoring due to its strong penetrability, rapidity, non-destructive, and environmental protection.

1.4. Overview of non-invasive modeling method

From a measurement principle perspective, the conventional method of measuring blood glucose involves monitoring the electrical signal generated by the chemical reaction between glucose reactants (such as glucose oxidase) and glucose molecules in blood or ISF. Signal strength is typically strongly associated with blood glucose concentration, resulting in reasonably accurate measurements. Conversely, non-invasive detection primarily utilizes indirect physical signals to represent blood glucose levels, resulting in a lower signal-to-noise ratio (SNR), hence a weak correlation with blood glucose concentration. Therefore, it is necessary to construct a blood glucose signal extraction and prediction model based on a specific sensor.

The signal preprocessing procedure generally includes signal noise reduction, enhancement, and feature extraction. The specific preprocessing methods slightly vary for different signal sources. For noise reduction, algorithms such as Fourier transform, wavelet transform (WT), and adaptive filtering are frequently employed to mitigate noise introduced by the circuit system, myoelectric interference, and movement artifacts during data acquisition. The signal enhancement algorithm aims to maximize signal disinfection and minimize data differences arising from sensor calibration and sensitivity. Some studies employ algorithms such as Kalman Filtering for signal enhancement, where the methods differ slightly for different sensor types and usage environments. Feature extraction involves extracting signal features that exhibit a strong correlation with blood glucose signals from the original sensor data to reduce redundancy. This also forms a critical component of traditional machine learning algorithm modeling, where research methods include principal component analysis (PCA) and linear discriminant analysis (LDA) algorithms. Regression modeling is essential to accurately correlate sensor data or features with blood glucose values. Some studies employ traditional machine learning techniques such as partial least squares (PLS) [31] to develop blood glucose regression models using features extracted manually. In recent years, numerous studies have used deep learning algorithms for non-invasive blood glucose data modeling, and some progress has been made. The deep network can automatically extract data features, which somewhat resolves the challenge of extracting blood sugar signals from complex signals. However, the black-box nature of deep learning models decreases the model's interpretability, particularly for blood sugar monitoring, where interpretability directly relates to its clinical credibility for indicators closely linked to clinical diagnosis and treatment.

1.5. Structure of this paper

In this article, our objective is to summarize the key enabling technologies that allow a reliable IR-based detection of blood glucose levels. To achieve this, we begin by outlining the review methodology used to select and analyze relevant studies. Next, we present an overview of the invasive glucose monitoring technologies used for diabetes diagnosis and home monitoring, as well as the accuracy evaluation index for glucose monitoring devices. The following section of the review

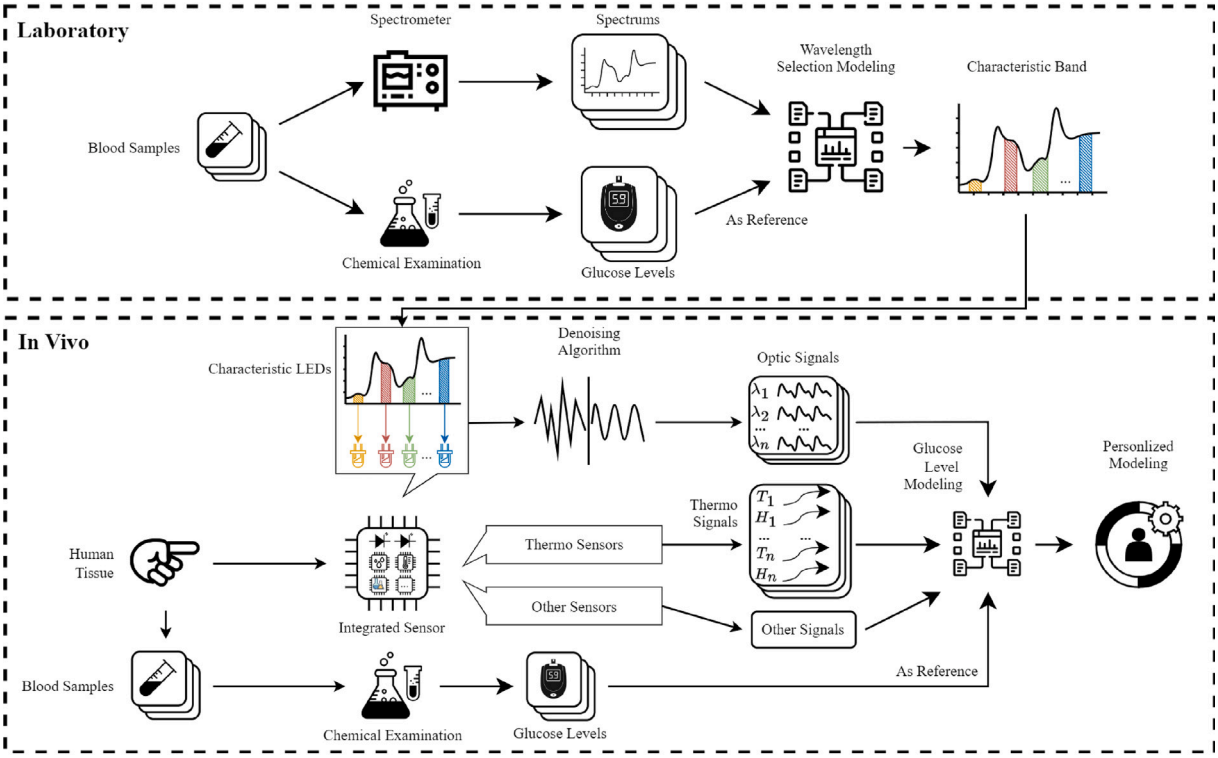

Fig. 2. Schematic overview of laboratory (upper) and in vivo (lower) processes in non-invasive blood glucose monitoring.

covers in detail the concept behind the approach of IR sensor-based technologies on non-invasive glucose monitoring, which includes technologies used in both laboratory and in-vivo scenarios. Subsequently, we focus on modeling-based technologies used in glucose monitoring, which include wavelength selection of IR spectrum, denoising techniques of Photoplethysmography (PPG), multi-sensor fusion process, sequential modeling technologies and personalized model calibration methods (see Fig. 2). After that, various current commercial non-invasive glucose monitoring devices were introduced along with device performance. Then, we discussed and analyzed the difficulties and challenges of existing technologies in non-invasive blood glucose monitoring. Finally, the conclusion part connects the previous sections and provides a further view into the future development of non-invasive glucose monitoring.

2. Review methodology

2.1. Literature search strategy

Eligible articles in the peer-reviewed literature were identified using several electronic databases, including Web of Science, IEEE Xplore, PubMed, and Google Scholar. The search strategy, which involved combinations of keywords such as 'non-invasive blood glucose monitoring', 'infrared absorption spectroscopy', 'machine learning' and 'Photoplethysmography', was adapted to each database to ensure a comprehensive retrieval of relevant studies.

Articles were included if they were published in English and fall into the categories of research or experimental reports, commentaries, narrative or descriptive articles, or book chapters. Studies were excluded if they were published in languages other than English, or if they were letters, editorials, media articles, or conference abstracts. Additionally, any study not directly related to non-invasive blood glucose monitoring or not utilizing infrared absorption spectroscopy was excluded. No limitations were set for the publication date to capture all relevant studies up to the knowledge cutoff date.

2.2. Selection criteria for commercial devices

Commercial non-invasive blood glucose monitoring devices utilizing infrared absorption spectroscopy were selected based on specific inclusion and exclusion criteria to ensure relevance and reliability. Devices were included if they were currently available on the market or in the late stages of development within the past five years, employed NIR or MIR absorption spectroscopy as the primary method for blood glucose measurement, and were capable of measuring blood glucose levels without requiring blood samples. In addition, selected devices must be supported by published clinical validation studies or substantial technical documentation. Devices were excluded if they lacked sufficient clinical validation or technical documentation, or if they did not employ NIR or MIR absorption spectroscopy as their primary measurement method.

The identification of relevant commercial devices began with extracting the names of companies from the reference lists of included academic papers and through a comprehensive search on the Internet. Subsequently, the official websites of these companies were reviewed to characterize the industry participants and their products in the non-invasive blood glucose monitoring field. To further ensure a thorough evaluation, publicly available patent information was searched across multiple patent offices, including the European Patent Office, the United States Patent and Trademark Office, and the World Intellectual Property Organization. This patent search aimed to identify additional devices and technologies that may not yet be prominently featured in commercial listings but have secured intellectual property protection, indicating their potential market presence.

2.3. Selection process

Fig. 3 illustrates the systematic flow of the selection process. Initially, a comprehensive search was conducted, yielding a total of 679 records, comprising 593 identified through the literature search and 86 identified through the commercial device search.

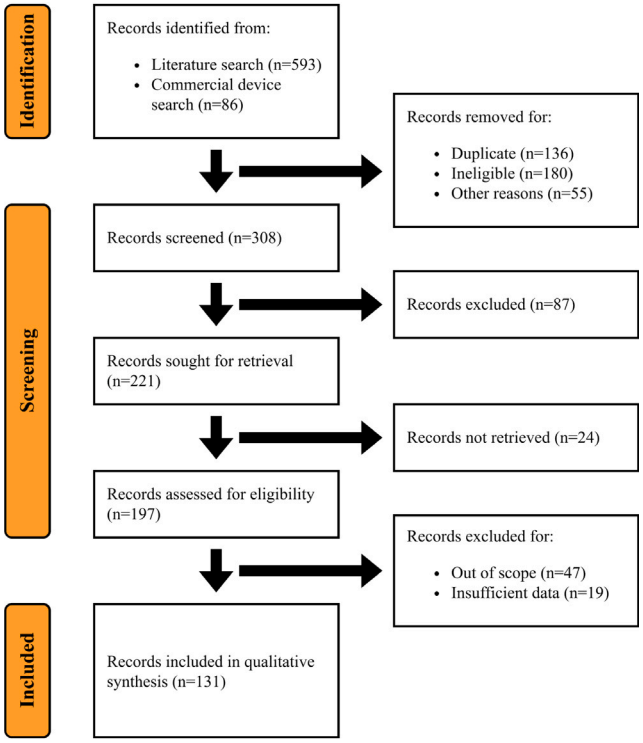

Fig. 3. Methodology flowchart of the search process.

In the first screening phase, one reviewer examined all titles and abstracts to determine if they met the basic inclusion and exclusion criteria. At this stage, 136 duplicates were removed along with 180 records deemed ineligible (e.g., not relevant to the research objectives) and an additional 55 records removed for other reasons (e.g., not published in English). This reduction left 308 records for further screening.

A subsequent screening narrowed this set to 221 records sought for retrieval. Of these, 24 could not be retrieved, leaving 197 records. The inclusion criteria required that studies address non-invasive blood glucose monitoring using infrared spectroscopy and be supported by relevant empirical findings or substantive technical documentation. Editorials, commentaries, and articles not directly aligned with the target research questions were excluded at this point. This step removed a total of 66 records, comprising those out of scope (n = 47) and those providing insufficient data (n = 19).

The remaining 131 records were then fully examined. Two reviewers independently assessed the full-text articles, evaluating study characteristics, methodologies, and relevance to the inclusion criteria. Discrepancies in study selection were resolved through discussion, and if necessary, consultation with a third reviewer to reach consensus. All included studies were organized and coded according to their level of relevance and quality. Specifically, studies were categorized as primary, secondary, or tertiary based on their relevance to the research questions, with primary studies being most directly relevant. This process helped prioritize studies for the qualitative synthesis.

The 131 records included in the qualitative synthesis refer specifically to studies that met the inclusion criteria of the systematic review. Two Refs. [1,2] are included in the total count for background purposes and are not directly relevant to the main findings of the review. Following this multi-stage review, a total of 133 records were included in this paper.

3. Overview of invasive glucose monitoring technologies

In the human organism, glucose plays a critical role as an energy source for cellular metabolism. The determination of glucose concentration can be performed on whole blood, plasma, or serum samples; however, the latter two are commonly preferred. This is because readings obtained from whole blood are often approximately 15% lower due to the presence of additional water content in the blood cells [32]. In addition to being present in the bloodstream, glucose is also widely distributed throughout various bodily fluids such as intracellular fluids (ICF), ISF, tears, saliva, and urine [33].

Currently, invasive blood monitoring technology has become the mainstream due to its convenience and practicality (see Table 1). Current techniques for determining glucose concentration are based on enzymatic and hexokinase methods. Both of these methods offer high levels of accuracy, specificity, and minimal cross-reaction. However, while both enzymatic and hexokinase methods are used in laboratory settings, the enzymatic method is preferred for point-of-care and home monitoring because of its simplicity and relative affordability.

3.1. Laboratory techniques

Enzymatic-amperometric and hexokinase techniques are the preferred methods for measuring blood glucose concentrations in laboratory settings. These methods offer high levels of specificity, sensitivity, and can detect a broad range of glucose concentrations, including under hypoglycemic and hyperglycemic conditions. Due to these features, they can serve as the gold standard for evaluating the performance of less accurate instruments, such as BGM, CGM, and potential non-invasive devices [34].

3.2. Blood glucose monitoring techniques

Conventional BGM devices read the glucose information directly from the bloodstream, and they are utilized in diabetes treatment include point-of-care testing (POCT) and self-monitoring of blood glucose (SMBG) devices [35], depending on the purpose of use.

3.2.1. Point-of-care testing

POCT refers to laboratory tests performed at the point of care, such as a doctor's office or a hospital. These tests are usually performed by healthcare professionals with a venous blood test and may include blood glucose, hemoglobin A1c (HbA1c), and other indicators [36]. POCT results are often available within minutes, directly providing the basis for clinical diagnosis and treatment of diabetes.

3.2.2. Self-monitoring of blood glucose

SMBG, on the other hand, is a self-testing technique commonly performed by individuals with diabetes. It involves the use of colorimetric reaction-based glucometers that measure blood glucose levels using a small drop of blood obtained through finger pricking [37]. SMBG glucometers are typically handheld devices that are compact and portable, making them convenient for use at home, work, or while traveling. This self-monitoring is typically performed multiple times a day, especially before and after meals, to aid people with diabetes in monitoring their blood glucose levels and adjusting their insulin doses as necessary. SMBG glucometers offer relatively quick and precise measurements of blood glucose levels, compared to POCT, and allow people with diabetes to adjust their diet, exercise, or medication as needed. Moreover, SMBG glucometers are relatively affordable and widely available. The most common brands of commercial SMBG devices currently available include Roche, LifeScan, Abbott, Bayer, Johnson & Johnson, Sinocare, and Sanofi. However, SMBG glucometers do have certain limitations. The user must prick their skin with a lancet, which can cause pain and skin irritation over time. Additionally, the user must carry test strips, lancets, and other supplies with them, which can be inconvenient [37].

Table 1
Comparison of invasive glucose monitoring technologies.

Characteristics	Laboratory	BGM		CGM	
		POCT	SMBG	FGM (iCGM)	RT-CGM
Accuracy	Very high	High	High	Medium	Medium
Sensitivity	Very high	High	High	Medium	Medium
Measurement time	Long	Quick	Quick	Quick	Quick
Professional training	Yes	No	No	No	No
Sample type	Blood, serum, plasma, urine	Blood	Capillary blood	ISF	ISF
Sampling method	Invasive	Invasive	Invasive	Interstitial	Interstitial
Physiological time delay	No	No	No	Yes	Yes
Testing cost	–	Low	Low	High	High
Diagnosis proof	Yes	Yes	No	No	No
Wearable	No	No	No	Yes	Yes
User calibration	–	No	No	No	Yes
Data display	–	On demand	On demand	On demand	Auto
Trend arrows	–	No	No	Yes	Yes
Alarms	–	No	No	No	Yes
Sensor duration	–	Disposable	Disposable	14 Days	Depending on the kind
Linkable to insulin pump	–	No	No	No	Yes
Auto insulin adjustment	–	No	No	No	Depending on the kind

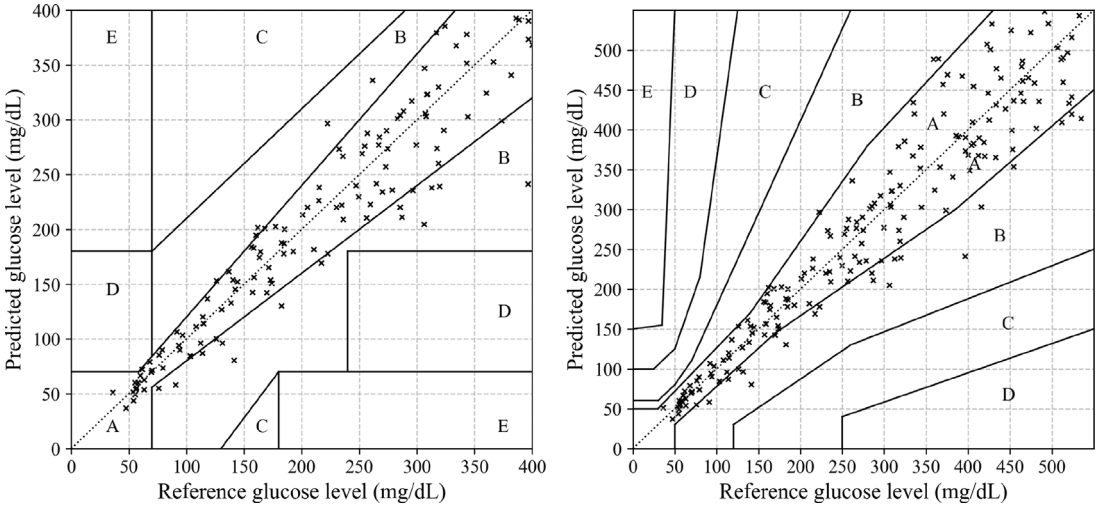

Fig. 4. Graphical representation of CEG (left) and PEG (right).

3.3. Continues glucose monitoring techniques

Unlike the conventional device, the CGM device is designed to measure glucose levels in the ISF, which is the fluid that surrounds and bathes cells in the body. This type of device is commonly employed by individuals diagnosed with diabetes to monitor their blood glucose levels. It comprises a small sensor that is inserted beneath the skin, typically in the abdominal area or the upper arm. The sensor functions by detecting glucose levels in the ISF and transmitting the data wirelessly to either a receiver or a smartphone application. The receiver or app provides real-time glucose level readings, enabling individuals to monitor their glucose levels throughout the day. Notably, a "lag time" exists between plasma and ISF, leading to interstitial glucose values that do not precisely correspond to blood glucose concentration. This can result in a loss of accuracy for both FGM and RT-CGM devices, particularly during periods of rapid glycemic excursions [3]. To ensure precise sensor glucose readings, some systems require daily capillary blood calibrations, usually performed twice a day at stable glucose values.

CGM devices can be classified into two types, namely flash glucose monitoring (FGM) and real-time CGM (RT-CGM), based on their scanning mode.

3.3.1. Flash glucose monitoring

FGM, which is also referred to as intermittent continuous glucose monitoring (iCGM), utilizes a wired glucose oxidase enzyme co-immobilized on an electrochemical sensor worn on either the arm or abdomen for a period of up to 14 days [38]. Although real-time interstitial glucose values are not continuously displayed, individuals can obtain them by placing a "reader" in proximity to the sensor. This system does not require daily fingerstick calibrations because the sensors are factory-calibrated.

3.3.2. Real-time glucose monitoring

In contrast, real-time glucose monitoring devices (RT-CGM) have the ability to show glucose levels in real-time, with values automatically displayed every 1–5 min, along with their rate of change and glucose trends. Moreover, some of these devices can be integrated with continuous subcutaneous insulin infusion (CSII), resulting in a sensor-augmented pump (SAP); in certain cases, this includes algorithms capable of interrupting insulin infusion when the glucose concentration reaches or is expected to reach a pre-defined level. Notably, RT-CGM systems require calibration to ensure accuracy, especially during initial usage. Calibration is typically performed by using fingerstick blood glucose measurements, and the frequency varies depending on the device.

Table 2
Percent error analysis.

Within ±5%	Within ±10%	Within ±15%	Within ±20%
X/Y (%)	X/Y (%)	X/Y (%)	X/Y (%)

3.4. Accuracy evaluation criteria

To assess the accuracy and effectiveness of glucose monitoring devices, there are several available tools. Percent error, mean absolute relative difference (MARD) and error grids are examples of metrics used to evaluate accuracy. They both provide a quantitative and qualitative assessment of their performance. These metrics help ensure that patients receive accurate and reliable glucose monitoring, which is essential in the management of diabetes and other conditions that require monitoring of blood glucose levels.

3.4.1. Percent error analysis

Percent error analysis is the most common method for evaluating blood glucose meters. This method evaluates the performance of a blood glucose meter by comparing the ratio of different error partitions between measured values and true values. The results of such an analysis are typically presented in Table 2, where X denotes the number of samples within the specified difference from the comparator method and Y denotes the total number of samples [39].

3.4.2. Clarke error grid analysis

Clarke error grid (CEG) was developed by David C. Clarke in 1987 as a tool to evaluate the clinical accuracy of glucose measurement devices [40]. It is based on comparing the glucometer measurements with the reference glucose measurements obtained from a laboratory test.

CEG analysis specifies the result of the measurement in five risk zones, each with a distinct level of clinical significance, with the goal of ensuring that most glucometer measurements fall within clinically acceptable zones (see Fig. 4). Zone A is the value within 20% of the reference reading. Zone B is designated as values that deviate >20% from the reference value but does not result in improper handling. Zone C indicates points that may lead to unnecessary treatment. Zone D represents the point at which hypoglycemia or hyperglycemia may not be detected. Zone E is a point of potential confusion for clinical treatment.

3.4.3. Parkes error grid analysis

Parkes error grid (PEG) is an extension of the CEG and is built on its framework [41], which categorizes measurements into five different risk levels that have specific clinical significance (see Fig. 4). The five risk levels, or zones, are differentiated by the impact they have on clinical action and outcome.

Zone A is the most stringent category, defined as a clinically accurate measurement that does not have an impact on clinical action. Zone B is more rigorously designated as an altered clinical effect with little or no effect on clinical outcome. On the other hand, zones C, D, and E are values that enable the risk to be assigned based on expertise, with each zone representing a different level of clinical significance.

It offers a more precise and comprehensive framework for assessing the accuracy and reliability of measurements, with the different risk zones indicating the potential impact on clinical action and outcome. This system is a valuable tool in ensuring that patients receive appropriate clinical treatment and that healthcare providers can make informed decisions based on the accuracy of the measurements.

Table 3
Approval criteria of invasive blood glucose meters.

Department	Glucose level	Metrics	Accuracy
FDA [39]	Entire range	Within 15%	95%
		Within 20%	99%
ISO [45]	<100 mg/dL	Within 15 mg/dL	95%
	≥100 mg/dL	Within 15%	95%

Table 4
iCGM special control performance requirements.

Department	Glucose level	Metrics	95% LCL
FDA [49]	<70 mg/dL	Within 15 mg/dL	>85
		Within 40 mg/dL	>98
	70–180 mg/dL	Within 15%	>70
		Within 40%	>99
	>180 mg/dL	Within 15%	>80
		Within 40%	>99
	Entire range	Within 20%	>87

LCL: Lower confidence limits.

3.4.4. Mean absolute relative difference

MARD is a widely used metric for measuring the accuracy of glucose monitoring devices [42], especially CGM, calculated as the mean percent of the absolute differences between the reference values $Reference$ and the meter readings $Measured$ (see Eq. (1)). The lower the MARD, the better the agreement between the device and the reference measurement.

$$MARD = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Reference_i - Measured_i}{Reference_i}\right| \times 100\% \quad (1)$$

Calculating the MARD for continuous CGM devices may seem like a straightforward task, and comparing the MARD values between different devices may seem uncomplicated. However, several complex factors (such as calibration, sensor duration, physiological time delay, etc.) [43] can make it difficult to interpret the MARD values provided by the manufacturers for their CGM systems. Consequently, MARD should not be considered an absolute indicator of accuracy without proper scrutiny, particularly in non-adjunctive use cases of such systems [44].

3.5. Approval criteria

The guidelines for the approval of invasive blood glucose meters vary between different countries, with most adopting standards based on ISO 15197:2013. However, in the United States, independent FDA standards are applied. Current standards often employ a proportion within a 15% error range as the evaluation criterion, with varying criteria for different ranges of blood glucose concentrations (see Table 3).

Although the guidelines for point accuracy of CGM systems can be relatively lenient compared to invasive meters (see Table 4), the MARD value is frequently used by manufacturers and clinicians to assess real-world performance, especially regarding the device's ability to reliably track glucose fluctuations and trends with fewer calibration requirements. For example, Dexcom G6 CGM and Abbott FreeStyle Libre have reported MARD less than 15% [46–48].

However, there are currently no specific evaluation guidelines for non-invasive detection products. As a result, non-invasive blood glucose products need to undergo certification based on the standards established for BGM and CGM devices.

4. Infrared absorption spectroscopy

4.1. Overview of infrared spectroscopy

IR spectroscopy deals with electromagnetic waves with longer wavelengths than visible (VIS) light, falling within the range of 770 nm

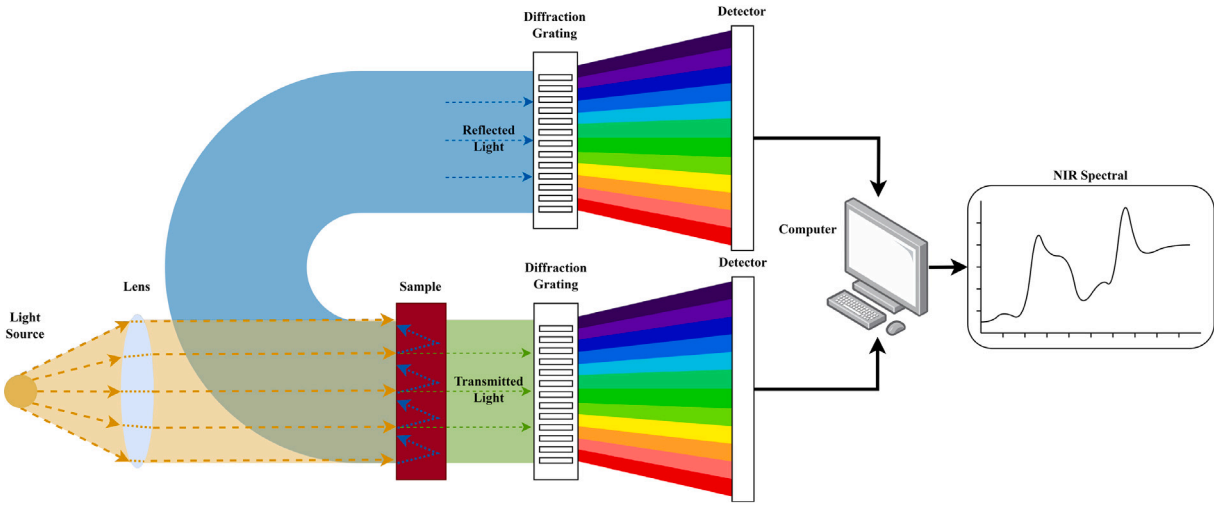

Fig. 5. Schematic representation of transmission mode (lower) and reflection mode (upper) of the IR spectroscopy.

to 1 mm. Conventionally, the IR spectrum is often categorized into three distinct regions: the NIR region (with wavelengths ranging from 780 nm to 2500 nm), the MIR region (with wavelengths spanning from 2500 nm to 25,000 nm), and the far-infrared (FIR) region (with wavelengths between 25 μm to 1000 μm).

When organic molecules with chemical bonds are exposed to infrared light, a resonance phenomenon occurs, which causes the molecules to vibrate and transition from the ground state to higher energy levels, resulting in the generation of a NIR spectrum. The main spectral information in the NIR region comes from the double and combined frequency absorptions of hydrogen-containing groups such as –CH, –NH and –OH, and the fundamental frequency absorption of these groups appears in the mid-infrared region. Therefore, the qualitative and quantitative analysis of substances can be performed by utilizing characteristics such as the position and intensity of the IR absorption spectrum of these substances.

Within the NIR range, light demonstrates a comparatively high capability to penetrate biofluids and soft tissues (>0.5 mm) and experiences lower levels of scattering than ultraviolet (UV) or VIS light [50]. Therefore, NIR sensing and measurement can be achieved through both reflection and transmission techniques. Conversely, MIR light possesses weaker penetration capabilities into soft tissues and is predominantly employed through reflection techniques. The configuration of IR spectroscopy primarily involves an IR light source with a specific wavelength, a sample that receives IR radiation, and a receiving photodiode (PD) that measures the attenuated light waves transmitted or reflected from the tissue sample to the detector (see Fig. 5).

4.2. Laboratory measurement

4.2.1. Transmission mode

For the transmission mode, the Beer–Lambert law provides a mathematical formulation to compute the absorbance of a sample based on its concentration and thickness:

$$A^{\lambda} = \lg\left(\frac{I_{in}^{\lambda}}{I^{\lambda}}\right) = \epsilon^{\lambda} l c \tag{2}$$

where A denotes the absorbance, I_{in} denotes the initial light intensity, I denotes the light intensity through the absorption medium, λ denotes the wavelength of the incident light, ϵ denotes the molecule absorption coefficient, which depends on λ the structure of the absorbing molecules, l denotes the absorption depth within the medium, c denotes the concentration of the substance to be tested. It can be expressed in advanced form as:

$$A^{\lambda} = \begin{bmatrix} A_{c_1}^{\lambda_1} & A_{c_2}^{\lambda_1} & \cdots & A_{c_n}^{\lambda_1} \\ A_{c_1}^{\lambda_2} & A_{c_2}^{\lambda_2} & \cdots & A_{c_n}^{\lambda_2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{c_1}^{\lambda_m} & A_{c_2}^{\lambda_m} & \cdots & A_{c_n}^{\lambda_m} \end{bmatrix} = \begin{bmatrix} \epsilon^{\lambda_1} l c_1 & \epsilon^{\lambda_1} l c_2 & \cdots & \epsilon^{\lambda_1} l c_n \\ \epsilon^{\lambda_2} l c_1 & \epsilon^{\lambda_2} l c_2 & \cdots & \epsilon^{\lambda_2} l c_n \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon^{\lambda_m} l c_1 & \epsilon^{\lambda_m} l c_2 & \cdots & \epsilon^{\lambda_m} l c_n \end{bmatrix} \tag{3}$$

where the horizontal axis denotes the absorbance array of samples within concentration $c_1 \sim c_n$ at wavelength λ_m, and the vertical axis denotes the absorbance array of the concentration c_n sample within wavelength $\lambda_1 \sim \lambda_m$.

The model assumes that light attenuation due to scattering is negligible compared to light absorption. Then the absorbance of transmitted or reflected light detected by PD depends on the concentration of absorbing molecules c, the thickness of the sample l, and the absorption coefficient of these molecules ϵ. For NIR range, the value of ϵ for glucose varies between 0 and 1 dL/(g cm) [51].

4.2.2. Reflection mode

The reflection mode is primarily employed when dealing with non-transparent samples. The presence of particles within the sample induces a scattering effect on light, leading to an uncertain optical path within the sample. Consequently, the correlation between the concentration of the analyte and the reflected light deviates from Beer–Lambert law, but it still reflects certain residual structural information of the analyte. Depending on the nature of the sample, the reflection mode can be categorized as either regular reflection or diffuse reflection. Regular reflection occurs under ideal conditions, where the angle of incidence matches the angle of reflection. However, in practical applications, rough surfaces constitute the majority of the objects being measured, resulting in predominantly diffuse reflection.

4.3. In vivo measurement

As noted, IR light exhibits a notably superior ability to penetrate biofluids and soft tissues, enabling the potential to determine glucose concentration through the analysis of light intensity transmitted or reflected across the skin tissue. However, in contrast to laboratory measurements, the complexity of skin tissues results in the superimposition of glucose absorption information with that of other molecules, and the measurements are further affected by light scattering within the skin tissue. Consequently, it becomes essential to diligently isolate and separate the pertinent glucose-related information to the maximum extent possible.

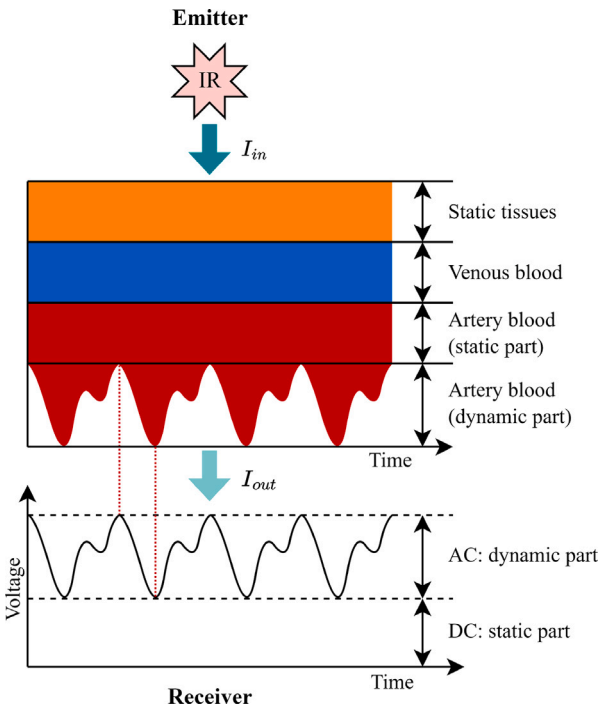

Fig. 6. Schematic diagram of PPG signal formation by IR light passing through tissue layers.

4.3.1. Characteristics of the skin tissue layers

The skin tissue comprises arterioles, venules, capillaries, and ISF. The stratification of skin tissue consists of the stratum corneum (10–20 μm), the epidermis (30–100 μm), the dermis (900–1500 μm), and the subcutaneous tissue (1000–5000 μm) [52]. The epidermal layer harbors an approximate ISF content of 15%–35%, yet lacks blood vessels. The dermal layer encompasses small arterioles, venules, and capillaries, accounting for approximately 40% of ISF. The subcutaneous tissue accommodates fat storage, a lesser extent of ISF (in comparison to the dermal layer), and vessels that establish a connection between the dermal component and systemic blood circulation [53]. Each distinct skin layer possesses inherent optical and dielectric characteristics, which may diverge due to variances in morphological features, thickness of skin layers, concentration of tissue/blood constituents (such as glucose), cutaneous blood perfusion, and other inter-individual disparities.

4.3.2. Photoplethysmography

PPG technology is an optical method employed to measure variations in blood volume throughout the cardiac cycle. Its ability to extract blood signals from background noise is of particular significance. PPG discerns alterations in blood volume within human tissue by means of changes in photoelectric reception.

In operation, a specific wavelength of light is directed onto the human body. As the emitted light traverses the tissue, it undergoes attenuation due to light-absorbing components within the tissue, before being detected by a PD at the opposite end. Since light absorption by tissues such as skin and bone remains relatively constant, signal variation is primarily driven by periodic changes in blood vessel volume and closure, synchronized with the cardiac cycle.

The PPG waveform, as shown in Fig. 6, is composed of two constituents: the direct current (DC) component and the alternating current (AC) component. The DC component corresponds to the tissue-transmitted or reflected light signal, depending on the tissue structure and the mean volume of arterial and venous blood. Typically, the DC component evolves gradually in response to respiration and activity, while the AC component, superimposed on the DC component, oscillates in accordance with fluctuations in blood volume between the systolic and diastolic phases of the cardiac cycle. The fundamental frequency of the AC component is dictated by the HR.

In more advanced scenarios, when complex wavelength light passes through human tissue, pulsatile fluctuations lead to alterations in transmittance at discrete wavelength points. This results in a dynamic three-dimensional spectral map (as shown in Fig. 7b). The wavelength and transmittance axes jointly illustrate the tissue transmission curve at specific wavelengths during the corresponding time intervals (as shown in Fig. 7a). Similarly, the time and transmittance axes show the temporal variation of the absorbance curve at the corresponding wavelengths (as demonstrated in Fig. 7c).

PPG, due to its ability to provide crucial cardiovascular insights [54], has gained significant traction beyond HR measurement within the realm of ubiquitous healthcare. Recent investigations have demonstrated its potential in the analysis of cardiac function [55,56] and the estimation of blood pressure [57–59], producing notable results. Given its non-invasiveness and cost-effectiveness attributes, PPG devices are progressively evolving towards wearable and portable configurations.

4.3.3. Infrared spectroscopy and absorption principles in glucose monitoring

Within human arterial blood, different substances exhibit varying absorption ratios across various bands of IR light. With reference to Lambert–Beer's law (Eq. (2)), the absorbance of a composite substance can be reformulated as follows:

$$A^{\lambda} = \lg\left(\frac{I_{in}^{\lambda}}{I^{\lambda}}\right) = E^{\lambda}\hat{L}C + G \tag{4}$$

where A denotes the total absorbance, I_{in} denotes the initial light intensity, I denotes the light intensity through the absorption medium, λ denotes the wavelength of the incident light, $E_{s_1,\ldots,s_n}$ denotes the molecule absorption coefficient matrix of absorbing substances $s_1 \sim s_n$, which depends on λ the structure of the absorbing molecules, $\hat{L}$ denotes the average path length of photon propagation, $C_{s_1,\ldots,s_n}$ denotes the concentration matrix of absorbing substances $s_1 \sim s_n$, G denotes the constant attenuation coefficient associated with the optical characteristics and geometric arrangement of superficial tissues (such as muscles and bones).

Hence the light intensity through the absorption medium I can be reformed as:

$$I^{\lambda} = I_{in}^{\lambda} \cdot e^{-(E^{\lambda}L^{*}C+G)} \tag{5}$$

Under the presumption of unchanging absorption within tissues beyond arterial blood, the mathematical representation for the intensity of transmitted light through the examined region is as follows:

$$I_{DC}^{\lambda} = I_{in}^{\lambda} \cdot e^{-E_c^{\lambda}\hat{L}C_c} \cdot e^{-E_v^{\lambda}\hat{L}C_v} \cdot e^{-G} \tag{6}$$

where E_c and C_c denotes the molecule absorption coefficient matrix and concentration matrix of the absorbing substance, which remains stable over short durations. E_v and C_v denote the molecule absorption coefficient matrix and the concentration matrix of the other variable substance.

In the presence of arterial pulsation, there are variations in the average length of the light absorption path within human tissue. Assuming an increase ΔL in length $\hat{L}$ leads to a change in the intensity of the transmitted light, denoted as I_{AC}^{λ}. This assumption presumes that the absorption of light by human tissue, excluding arterial blood, remains invariant. Accordingly, the transmitted light intensity during this scenario can be mathematically articulated as follows:

$$I_{DC}^{\lambda} - I_{AC}^{\lambda} = I_{in}^{\lambda} \cdot e^{-E_c^{\lambda}\hat{L}C_c} \cdot e^{-E_v^{\lambda}(\hat{L}+\Delta L)C_v} \cdot e^{-G} \tag{7}$$

By dividing Eqs. (6) and (7), the following formula can be derived:

$$\frac{I_{DC}^{\lambda} - I_{AC}^{\lambda}}{I_{DC}^{\lambda}} = e^{-E_v^{\lambda}\Delta LC_v} \tag{8}$$

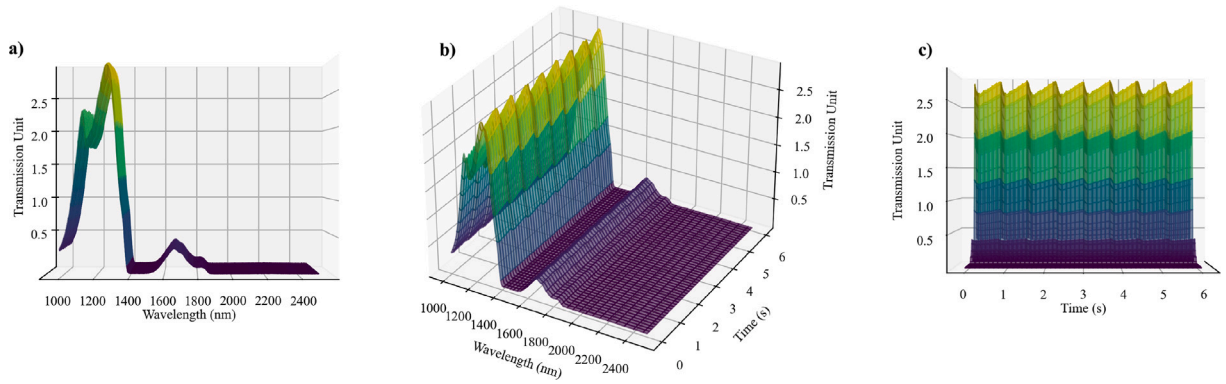

Fig. 7. 3D visualization of PPG and NIR interaction. (a) Wavelength-axis view; (b) Default view; (c) Time-axis view.

Given that the DC signal, denoted as I^{λ}_{DC}, significantly outweighs the AC signal, denoted as I^{λ}_{AC}, within the observed detection signal, Eq. (8) can be simplified as:

$$\ln\left(1-\frac{I^{\lambda}_{AC}}{I^{\lambda}_{DC}}\right)\approx-\frac{I^{\lambda}_{AC}}{I^{\lambda}_{DC}}=-E^{\lambda}_{v}\Delta LC_{v} \tag{9}$$

Assuming the presence of n distinct light-absorbing substances within arterial blood, and considering the incident light spanning from λ_1 to λ_m, Eq. (9) can be extended in the following way:

$$\begin{bmatrix} I^{\lambda_1}_{AC}/I^{\lambda_1}_{DC} \\ I^{\lambda_2}_{AC}/I^{\lambda_2}_{DC} \\ I^{\lambda_3}_{AC}/I^{\lambda_3}_{DC} \\ \vdots \\ I^{\lambda_m}_{AC}/I^{\lambda_m}_{DC} \end{bmatrix} = \begin{bmatrix} \epsilon^{\lambda_1}_{s_1}\Delta Lc_{s_1}+\epsilon^{\lambda_1}_{s_2}\Delta Lc_{s_2}+\cdots+\epsilon^{\lambda_1}_{s_n}\Delta Lc_{s_n} \\ \epsilon^{\lambda_2}_{s_1}\Delta Lc_{s_1}+\epsilon^{\lambda_2}_{s_2}\Delta Lc_{s_2}+\cdots+\epsilon^{\lambda_2}_{s_n}\Delta Lc_{s_n} \\ \epsilon^{\lambda_3}_{s_1}\Delta Lc_{s_1}+\epsilon^{\lambda_3}_{s_2}\Delta Lc_{s_2}+\cdots+\epsilon^{\lambda_3}_{s_n}\Delta Lc_{s_n} \\ \vdots \\ \epsilon^{\lambda_m}_{s_1}\Delta Lc_{s_1}+\epsilon^{\lambda_m}_{s_2}\Delta Lc_{s_2}+\cdots+\epsilon^{\lambda_m}_{s_n}\Delta Lc_{s_n} \end{bmatrix} \tag{10}$$

Hence, in theory, the existence of a wavelength λ_i can be identified when the molar absorption coefficients of irrelevant substances, excluding the analyte s_g, approach zero. The relationship between the light intensity received and the concentration of the analyte can be established as follows:

$$c_g=\frac{I^{\lambda_i}_{AC}}{I^{\lambda_i}_{DC}}\cdot\frac{1}{\epsilon^{\lambda_i}_{s_g}\Delta L} \tag{11}$$

However, because the absorption characteristics of compounds are governed by hydrogen-containing groups, a significant degree of overlap is observed among the constitutive groups of various substances. Consequently, the identification of such an ideal wavelength becomes challenging. In addition, the absorbance of arterial pulse waves encompasses the cumulative absorption of all light-absorbing substances within the arterial blood, including the blood glucose concentration, which typically falls within the range of approximately 54 to 540 mg/dL (3 to 30 mmol/L). Particularly within the lower concentration intervals of blood glucose, the absorbance signal of glucose is challenging to capture because of signal weakness. Moreover, inter-individual variations in tissue characteristics can influence the acquisition of absorbance signals. Discrepancies that include factors such as skin thickness, skin color, muscle distribution, and skeletal structure can affect the DC component of the absorbance spectrum. Meanwhile, distinctions in variables such as vascular wall thickness, vascular wall elasticity, and vascular density can affect the AC component of the absorbance spectrum. Consequently, these variations contribute to measurement disparities.

5. Modeling technologies

5.1. Wavelength selection

The objective of identifying the characteristic wavelengths associated with glucose molecules is to determine an optimized set of wavelengths that can minimize the root mean square error (RMSE) between predicted values and reference values within the testing dataset.

$$\lambda=\underset{\lambda}{\operatorname{argmin}}\, f(\lambda)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(c_{ref}-g(\lambda))} \tag{12}$$

In practical operational procedures, IR spectra typically encompass thousands of wavelength points. For instance, the ThermoFisher Antaris™ II FT-NIR spectrometer employs a scanning resolution of 4 cm^{-1} (equivalent to 0.6 nm at 1250 nm) and can generate more than 3000 spectral points in the NIR range of 1000~2500 nm. These high-dimensional data present challenges posed by the "curse of dimensionality" that many statistical methods cannot address [60].

In addition, conducting infrared spectral analysis on organic compounds within complex mixtures often involves the issue of spectral feature overlap. This manifests as spectral overlap between the functional group spectra of the analyte molecule within mixed component solutions (such as blood) and the spectra of functional groups of other components. Furthermore, the spectral data exhibits substantial redundancy, characterized by narrow characteristic peaks of molecular groups and a low proportion of effective information. Moreover, IR spectral acquisition often necessitates specialized training, and the preparation, calibration, and data collection processes of samples can be intricate, resulting in a limited quantity of analyzable data samples.

Addressing the aforementioned issues, the predominant methodologies in current research include regularization, dimension reduction, and variable selection [61]. Regularization, as a means to address the problem of overfitting arising from high-dimensional data, involves introducing penalty terms into the objective function. The methods within this framework include the regression of the ridge [62], the elastic net [63], and the least absolute shrinkage and selection operator (LASSO) [64]. Dimension reduction techniques replace the original variables with latent variables or principal components with higher variance. These techniques include projection, principal component regression (PCR) [65], and partial least squares regression (PLSR) [66]. They aim to mitigate collinearity and spectral overlap stemming from irrelevant variables in the dataset while simultaneously reducing redundant noise. However, the latent variables obtained through these methods are scarcely interpretable, and their outputs are not easily transferable to subsequent tasks. Given that only a subset of spectral information correlates with the attributes of the analyte, variable selection is rooted in the assumption that enhancing model prediction performance requires only a limited number of selected variables. In comparison to other techniques, variable selection not only reduces computational complexity but also provides a more plausible chemical interpretation. For example, in infrared detection of moisture content,

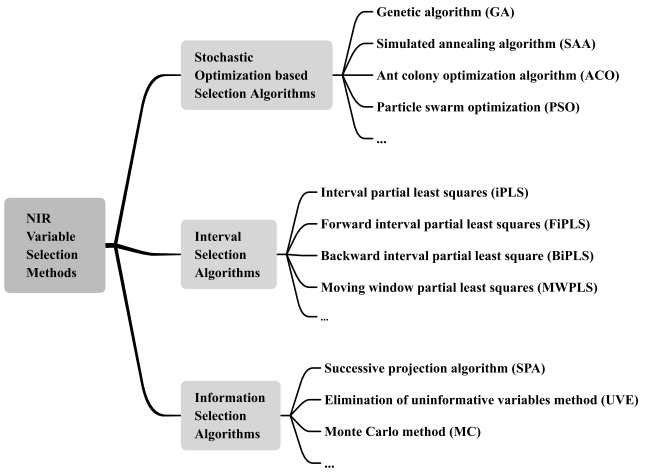

Fig. 8. Overview of NIR variable selection methods.

only variables related to the O–H bond can be considered effective, while others represent irrelevant or interfering variables, as moisture molecules only involve O–H bonds. Importantly, the variables selected through this approach possess more practical explanatory qualities, making them amenable to downstream applications. This includes guiding the development of feature-specific wavelength sensors aimed at achieving miniaturization and cost reduction of devices.

Diverse categories of feature selection emerge based on varying criteria and selection strategies. With respect to feature selection search strategies, they can be categorized as heuristic search, random search, and global search. Depending on the availability of class information within the samples, feature selection is classified as supervised or unsupervised. Based on feature selection criteria, it is categorized into distance measures, information measures, consistency measures, correlation measures, and classification error rate measures. In addition, its integration with subsequent classification algorithms classifies it as filtering, wrapper, or embedding methods. Furthermore, based on the form of variable selection, it can be categorized into univariate and multivariate selection. In the context of IR spectroscopy applications, feature selection methods can be broadly classified into three types: algorithms based on random optimization, algorithms based on variable intervals, and algorithms based on informative variables (see Fig. 8). Research indicates that the aforementioned methods are all employed in wavelength selection and model establishment, each possessing its own strengths and limitations.

5.1.1. Stochastic optimization based selection algorithms

The selection methods based on stochastic optimization involve constructing suitable objective functions to seek the optimal combination of target variables. Currently, there exist numerous common random optimization methods for variable selection. For instance, the genetic algorithm (GA), initially proposed by Holland in 1975 [67], draws inspiration from the principles of natural selection and genetic mechanisms in the biological realm. Through a continual process of genetic iteration that involves selection and exchange, variables with better values of objective function are preserved, while those with poorer values are eliminated, ultimately leading to optimal results. At present, GA has found widespread application in the field of analytical chemistry. Feng et al. [68] developed an optimization framework called GSMW-LPC-GA, based on GA, to enhance NIR quantitative calibration models. Within this framework, the grid search moving window (GSMW) strategy is employed to select information-rich NIR bands, which are then transformed into latent principal components (LPC) as input for GA optimization. Wang et al. [69] introduced a novel approach for constructing NIR spectroscopy multivariate calibration models, namely the genetic algorithm-wavelet transform-radial basis function network (WT-GA-RBFN). This method combines the advantages of wavelet transformation and GA. Furthermore, other optimization algorithms such as the simulated annealing algorithm (SAA), ant colony optimization (ACO), and particle swarm optimization (PSO) have also demonstrated promising results in the domain of NIR spectroscopy. Shi et al. [70] proposed a novel parallel heuristic SAA for variable selection in NIR spectroscopy. This algorithm employs a parallel mechanism to enhance search efficiency, generates high-quality candidate solutions through a heuristic mechanism, and employs the concept of the metropolis criterion to estimate the accuracy of candidate solutions. Huang et al. [71] introduced ant colony optimization-interval partial least squares (ACO-iPLS) and genetic algorithm-interval partial least squares (GA-iPLS) for measuring total anthocyanin content in flowering tea using NIR spectroscopy. Cao et al. [72] presented a wavelength selection method based on attractor-based random decision PSO for quantitative analysis using NIR spectroscopy.

5.1.2. Interval selection algorithms

Interval-based selection methods focus on studying variable intervals composed of multiple variables, rather than selecting individual variables. Typically, these methods exhibit greater efficacy and precision in variable selection. However, at times, defining the partitioning of spectral ranges (window width) and segments can be challenging. The interval partial least squares (iPLS) method divides the entire NIR spectrum into multiple equidistant intervals, performing PLS regression in each interval to obtain the minimum RMSE and assess the optimal interval. Subsequently, using this interval as a reference, the wavelength variables are expanded unidirectionally or bidirectionally to determine the optimal wavelength interval. The forward interval partial least squares (FiPLS) method builds upon iPLS by progressively incorporating subintervals with the best information or the least collinear variables, establishing optimal PLS models within the selected intervals until all intervals are used for modeling. In contrast, the backward interval partial least squares (BiPLS) method operates in the reverse manner as FiPLS. It progressively eliminates subintervals that contribute the least information or exhibit the highest collinearity, selecting multiple intervals associated with the lowest RMSE as the optimal combination. Zou et al. [73] proposed the BiPLS and FiPLS algorithms to effectively select wavelength regions to classify the soluble solid content (SSC) of apples. The moving window partial least squares (MWPLS) method is designed for each sub-wavelength region, selecting a spectral window of width w. Starting from the first wavelength point of the entire spectrum, this window is moved one wavelength point to the right until the end. Individual PLS models are established for each case, generating corresponding RMSE values for different principal components and thereby identifying one or more wavelength regions containing valuable information. Kasemsumran et al. [74] combined MWPLS with NIR spectroscopy based on DNA markers to distinguish turmeric varieties.

5.1.3. Information selection algorithms

In the context of information selection algorithms, information variables are defined as indicator vectors that can describe the magnitude of a variable's role in a model. NIR spectra consist of thousands of variables (wavelength points), with some variables having no positive impact on quantitative analysis of NIR spectra and are termed uninformative variables. These uninformative variables can easily decrease the predictive accuracy of the models. In recent years, researchers have introduced information-based variable selection methods, including the successive projections algorithm (SPA), uninformative variable elimination (UVE), and the Monte Carlo method (MC). SPA is a forward variable selection method that employs vector projection analysis to identify variable groups with the least redundant information. It effectively mitigates collinearity, singularity, and inconsistency among variables in the spectrum. The role of stability minimizes collinearity among vectors, reducing the number of variables used in modeling, and enhancing modeling efficiency. Tang et al. [75] introduced

the competitive adaptive reweighted sampling successive projection algorithm (CARS-SPA) to analyze nicotine in active tobacco ingredients within pesticide formulations. Initially, competitive adaptive reweighted (CAR) is used to select information variables, followed by SPA to optimize variables with minimal redundant information. UVE is based on PLS regression coefficients. The core concept is to use regression coefficients as a measure of the importance of the variables, eliminating wavelength variables that contribute less to the model. This reduces the number of model input variables and the complexity of the model. The combination of the Monte Carlo method and the PLS regression coefficients is used for wavelength variable selection. MC sampling randomly selects a portion of samples from the calibration set for PLS modeling. This process is repeated hundreds of times, and then wavelength variables corresponding to regression coefficients with significant impact are selected based on certain rules. Li et al. [76] employed the Monte Carlo uninformative variable elimination (MC-UVE) with nonlinear calibration methods to determine the content of gossypol in cotton seeds, demonstrating that MC-UVE can provide better and simpler calibration models compared to full spectra.

5.2. Signal denoising

As noted, PPG technology can measure dynamic photonic signals of blood volume, with the aim of minimizing the influence of static tissue signals within background noise. However, PPG signals, being a weak form of bio-signal, are susceptible to noise interference. PPG signals acquired in real-world settings can be influenced by various forms of noise due to factors such as sampling, system noise, and random noise [77], including optical scattering, motion artifacts, electromyographic interference, and poor circulation, affecting subsequent analyzes. Thus, signal denoising plays a significant role in the analysis of PPG signals.

PPG signal denoising methods often overlap with general bio-signal processing. However, it is challenging to determine whether important glucose-related information might be lost during noise filtering. Therefore, the choice of denoising techniques in non-invasive glucose monitoring largely depends on the stability of the device and the measurement environment. For wearable devices such as fitness trackers or smartwatches, real-time noise reduction is crucial due to frequent motion artifacts. These devices require robust filters to manage noise when tracking rapidly changing physiological parameters (Phy), such as HR.

5.2.1. On-device denoising methods

However, blood glucose levels exhibit slower fluctuations, often monitored at intervals of minutes or hours. Thus, measurements can be performed in calmer conditions where motion is minimal. In such scenarios, simple denoising techniques can effectively clean the signal without losing significant glucose-related data, while also reducing the demand on hardware computation resources. Using basic filters such as the Butterworth band-pass filter (BBPF) [78–81] or the moving average (MA) filter [82–84], devices can achieve effective noise reduction with relatively low computational complexity. This minimizes the processing power required, enabling more efficient operation on resource-constrained hardware, such as wearable devices, without sacrificing the quality of the glucose monitoring signals.

5.2.2. Advanced denoising

In dynamic and complex environments, where PPG signals are more likely to be affected by non-linear interactions and motion artifacts, advanced denoising methods become essential.

Techniques such as Kalman filtering are adept at handling time-series noise by predicting the state of the signal at each time step based on both the previous state and noise characteristics. This adaptability allows Kalman filters to compensate for various noise sources in real-time, making them suitable for wearable devices under continuous motion. Xu et al. [85] proposed a sophisticated approach that involves using WT, which decomposes signals into multiple frequency bands. This enables targeted denoising by filtering out frequency bands dominated by noise while preserving those with vital signal information. Unlike traditional filters, WT-based methods adapt to the non-stationary nature of the signal, effectively addressing motion artifacts while retaining critical signal details.

The slant transform-based bit plane method described by Guo et al. [86] represents a hybrid approach that combines spectral analysis and signal separation. PPG signals are transformed into the frequency domain using the Slant Transform, which is an orthogonal linear transformation, to identify the dominant signal components. By representing the signal in bit planes, lower-amplitude coefficients, often associated with noise, are discarded, while higher-amplitude coefficients representing the clean signal are retained.

Chen et al. [87] presented the multi-size weighted fitting (MSWF) algorithm, which introduces a polynomial-based smoothing approach to remove high-frequency noise and baseline drift from PPG signals. MSWF utilizes locally weighted fitting techniques with dynamic window sizes to create adaptive filtering based on signal characteristics. The algorithm iteratively updates the weights by minimizing the error between the measured and estimated signals, making it effective in enhancing the quality of the signal while preserving key physiological information. In non-invasive blood glucose monitoring tasks, MSWF demonstrated improvements in comparison to traditional BBPF and WT-based filtering models, showcasing its robustness in retaining critical glucose-related data while reducing noise interference.

5.3. Multi-sensor fusion

Non-invasive techniques, including IR spectroscopy analysis, predominantly employ indirect physical measurement methods. Considering the imperative of minimizing any harm to the human body, the power of the signal source is generally kept low. Consequently, the signal intensity received by sensors after traversing through human tissues is further attenuated. Furthermore, the received signal is compounded by various interference signals unrelated to the target signal, which weakens the inherent correlation between the acquired signal and the quantitative or qualitative analysis of the detected substance. Additionally, non-invasive blood glucose measurements are notably susceptible to environmental disturbances in real-world operational settings. This susceptibility contributes to the reduction of the SNR in measurement signals and consequently affects the integrity of the final detection results. Hence, the processing techniques applied to sensor signals, along with the amalgamation of multidimensional sensor data through fusion modeling methods, play a pivotal role in enhancing the precision of the ultimate blood glucose prediction model.

The application of multi-sensor fusion techniques in non-invasive glucose monitoring enables the combination of signals from different sensors to enhance measurement accuracy and reliability. For example, integrating PPG, IR, and electromagnetic data can help address the limitations posed by single-sensor noise and fluctuations in signal quality. In biological systems, sensor fusion has demonstrated its potential in areas such as human activity recognition and medical diagnostics [88]. This approach generally involves a workflow that includes data collection, signal preprocessing, feature extraction, and modeling, to combine complementary information from different sensors to improve the performance of predictive models (see Fig. 9). By employing sensor fusion in non-invasive glucose monitoring, it becomes possible to mitigate the impact of individual sensor noise and environmental influences, thereby enhancing the robustness of blood glucose estimation. In Table 5, we have provided a summary of recent studies on non-invasive glucose monitoring based on PPG technology, including dataset descriptions, input signal types, model structures, and performance metrics.

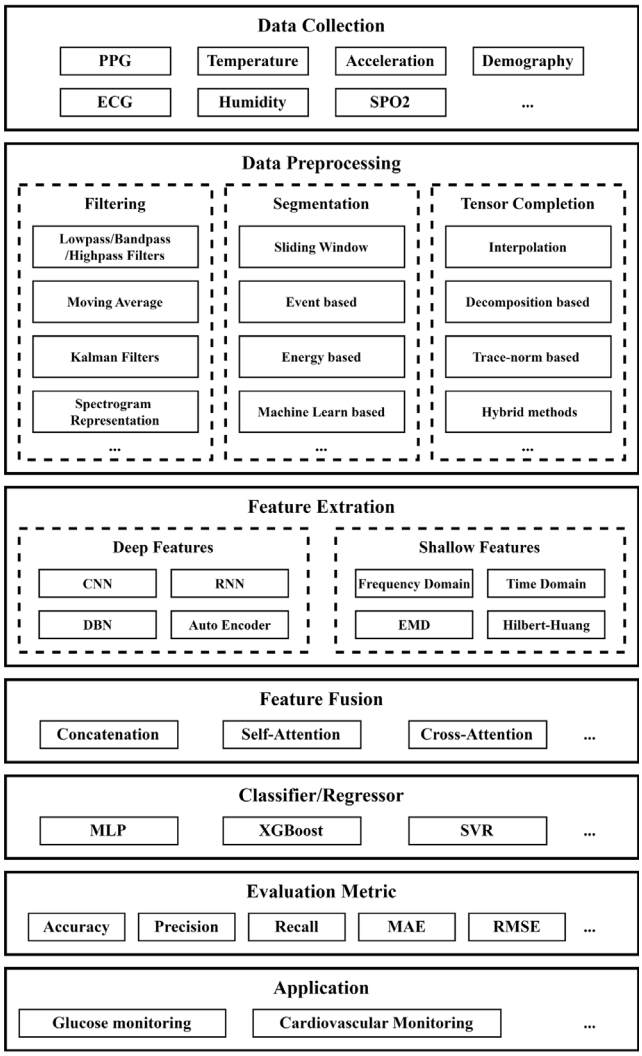

Fig. 9. Multi-sensor fusion framework for non-invasive glucose monitoring.

5.3.1. Data collection

As illustrated in Fig. 9, this endeavor starts with the collection of data from various sensors. Sensors are hardware components capable of capturing various types of signals, with miniaturized sensors incorporable into everyday devices such as smartphones, smart wristbands, and other wearable medical devices [89]. These devices are equipped with abundant sensor resources including motion sensors (accelerometers, gyroscopes, magnetometers, and motion cameras), environmental and physiological sensors (pressure sensors, pulse oximeters, electrocardiographs, thermometers, hygrometers, microphones, etc.), and GPS, collecting information at sampling rates ranging from 20 to 50 Hz according to activity types [90].

Recent studies have demonstrated the effectiveness of integrating PPG signals with other physiological and environmental sensor data to enhance non-invasive blood glucose monitoring. For instance, Lee et al. [79] proposed a method combining dual-channel PPG signals with ECG data to estimate blood glucose levels. Their results indicated that the combination of these two signals improved predictive accuracy due to the complementary information offered by the ECG on cardiovascular dynamics. Furthermore, Bent et al. [91] explored the integration of multiple physiological signals, including PPG, electrodermal activity (EDA), and skin temperature (ST), collected from wearable devices, with contextual information such as food logs (FL) and demographic inputs. Their approach demonstrated that combining physiological data with behavioral and lifestyle information enabled a more personalized prediction of glucose excursions, showing the potential of multi-sensor fusion for comprehensive glucose monitoring systems.

In addition, several researchers have investigated the combination of PPG with accelerometer and gyroscope data to address the influence of motion artifacts. Nabavi and Bhadra [92] proposed a robust fusion method to reduce motion artifacts in PPG signals by incorporating data from both a PPG sensor and an accelerometer.

5.3.2. Data preprocessing

Data preprocessing plays a crucial role in ensuring the integrity and quality of multimodal signals for non-invasive monitoring systems. For PPG signals specifically, various denoising methods have been extensively discussed in Section 5.2. When dealing with multimodal data, preprocessing must consider the distinct characteristics of each signal type. For example, motion sensors, accelerometers, and gyroscopes typically undergo noise reduction using low-pass or Kalman filters to handle sensor drift and motion artifacts effectively. Environmental sensors may require correction for baseline drifts, temperature calibration, and normalization techniques.

Data segmentation is another key step in preprocessing. Segmenting signals into manageable frames allows for effective feature extraction and computational efficiency. Common approaches include sliding windows, overlapping windows, and event-based segmentation to capture critical transitions or activities. Machine learning-based techniques are also employed to adaptively segment data based on signal patterns, as noted by Shoaib et al. [93], Banos et al. [94], and Londhe et al. [95].

In addition, dealing with missing or incomplete data is essential to maintain robustness. Techniques like tensor completion can address these gaps by employing interpolation, low-rank decomposition methods, and hybrid approaches that combine statistical and learning-based models to reconstruct missing segments with greater precision [96]. Such strategies are pivotal in preserving the reliability of multimodal datasets and enhancing the accuracy of downstream machine learning models.

5.3.3. Feature extraction

Feature extraction involves transforming raw data into meaningful characteristics or features that effectively represent the underlying patterns of the signal. This process is crucial for reducing noise, improving accuracy, and improving computational efficiency in machine learning models. The features can be categorized into shallow and deep features.

Shallow feature extraction techniques often involve manual methods based on domain knowledge. These include time-domain features (e.g. mean, standard deviation, skewness), frequency-domain features (e.g. spectral energy, dominant frequency), and more sophisticated signal decomposition techniques like Hilbert–Huang and empirical mode decomposition (EMD) [97,98]. When high-dimensional features are involved, dimensionality reduction techniques such as PCA, independent component analysis (ICA), and LDA are employed to retain essential information while reducing complexity. Hina et al. [84] utilized a Wrapper Algorithm to select six key features from the extracted set, which included time-domain features, spectral entropy (SE), logarithmic energy entropy (LogE), and Kaiser-Teager energy (KTE) features. By implementing these selected features in hardware, they achieved high-accuracy glucose estimation with minimal power consumption, making it suitable for wearable devices and real-time monitoring applications. Nie et al. [99] extracted 24 features from imaging photoplethysmography (IPPG) data, focusing on time-domain characteristics, KTE, and energy profile (EN). Using the Pearson Correlation Coefficient, they selected 12 key features for subsequent analysis. This feature selection approach demonstrated good consistency in their model, contributing to reliable performance in glucose monitoring. Zhang et al. [83] proposed a method that fits PPG signals using Gaussian functions and extracts 28 features from the time and frequency domains of the decomposed subwaves. This approach enabled effective

Table 5
Overview of recent studies on PPG-based non-invasive blood glucose monitoring techniques.

References	Dataset-subject	Dataset data pair	BG range (mg/dL)	Input	Reference device	Preprocessing methods	Feature extraction	Fusion methods	Regressor	ISO accuracy (%)	CEG Zone A (%)	MARD (%)
Chen et al. (2024) [87]	260	4684	72~216	PPG	BGM	MSWF	CNN	CAF	MLP	83.75	87.89	–
Moajjem et al. (2024) [100]	15	25,000	50~250	PPG EDA ST FL	CGM	Quality assessment	CNN	CAF	MLP	–	79.03	12.57
Nie et al. (2023) [99]	8	1280	–	IPPG	CGM	MA	Hand-crafted (12)	–	RFR	–	89.6	8.9
Lee et al. (2023) [79]	18	36	–	PPG (2-C) ECG,	BGM	BBPF	Hand-crafted (16)	–	LR	–	100	–
Prabha et al. (2022) [78]	217	7263	58~390	PPG PhyP	BGM	BBPF	MFCC (13)	–	XGBoost	–	98.97	–
Gupta et al. (2021) [101]	26	–	84~119	PPG (3-C) PhyP	BGM	FSW BBLF	Hand-crafted (17)	–	XGBoost	–	96.15	–
Alonso-Silverio et al. (2021) [102]	122	–	–	PPG	BGM	–	MFCC	–	AdaBoost	–	92	11.2
Bent et al. (2021) [91]	15	25,000	50~250	PPG EDA ST FL	CGM	–	Hand-crafted (69)	–	XGBoost	–	–	13.26
Islam et al. (2021) [103]	52	191	68~211	IPPG	BGM	GF ALS	Hand-crafted	–	PCR	–	–	–
Hina et al. (2020) [84]	200	–	–	PPG	BGM	MA	Hand-crafted (6)	–	SVR	–	95	7.62
Joshi et al. (2020) [104]	46	191	80~420	PPG	BGM	–	–	–	MPR	–	100	–

ALS: Asymmetric least squares; BBLF: Butterworth low-pass filter; CAF: Cross-attention fusion; FSW: fitting-based sliding window; GS: Gaussian filter; ISO accuracy: Percentage with relative errors within 15% (glucose > 100 mg/dL) or 15 mg/dl (glucose ≤ 100 mg/dL); LR: Linear regression; MPR: Multiple polynomial regression; PPG (n-C): Number of PPG channels used; RFR: Random forest regression.

correlation analysis of blood glucose levels and successfully separated samples into three distinct blood glucose level intervals, demonstrating its effectiveness in capturing variations across different glucose ranges. Prabha et al. [78] extracted 13 Mel-scale frequency cepstral coefficients (MFCC) from the frequency spectrum of PPG signals using discrete Fourier transform (DFT). These coefficients effectively captured the spectral characteristics of the PPG signals, enabling a compact and informative representation that improved the estimation of blood glucose level.

However, shallow features are highly dependent on expert knowledge and often require significant data annotation efforts. To address these limitations, deep learning-based approaches have become increasingly popular. Deep learning methods automatically extract features from raw sensor data using multi-layer neural networks, allowing for hierarchical representations from low-level to high-level abstractions. Common techniques include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and autoencoders, which have shown substantial improvements in non-invasive glucose monitoring tasks. Robert et al. [105] developed a CNN model based on a 39-layer residual neural network (ResNet) to identify diabetes using single-wave PPG signals recorded via smartphones. Through a cohort study involving more than 60,000 participants, they demonstrated that smartphone-captured PPG waveforms can effectively serve as non-invasive digital biomarkers for diabetes detection. Moajjem et al. [100] proposed the MMG-NET model, which designed multiple convolutional modules to extract the features of blood glucose. Yang et al. [106] proposed an improved CNN-based model called DCC-Net, designed to achieve end-to-end glucose prediction by capturing diverse temporal and spatial features within PPG signals.

5.3.4. Feature fusion

Feature Fusion focuses on combining features from different sensor sources to create a more comprehensive representation of the underlying physiological. By integrating data from multiple sensors, feature fusion aims to capture diverse attributes that are not discernible using a single sensor. Feature fusion plays an essential role in enhancing non-invasive blood glucose monitoring, where the integration of features from different physiological signals—such as PPG, accelerometers, and EDA—can help improve glucose estimation accuracy. Real-world signals often contain non-linearities due to varying environmental and physiological conditions, making it crucial to develop feature fusion strategies that can capture these complex relationships. Chen et al. [87] proposed a cross-view feature fusion mechanism (CVFF) to integrate complementary information and facilitate latent knowledge sharing among heterogeneous features from different physiological time-series data. This approach aims to capture the potential correlations between hemodynamic characteristics and blood glucose levels, thereby improving the accuracy and reliability of non-invasive glucose monitoring systems. Moajjem et al. [100] proposed a multi-layer neural network framework incorporating a cross-modal attention mechanism to fuse PPG, EDA, ST data. This approach demonstrated superior performance in non-invasive glucose prediction datasets, effectively using complementary information from different physiological signals.

5.3.5. Classification, regression, and evaluation

Classification, regression, and evaluation of learning algorithms encompass various tasks such as diabetes diagnosis, classification of blood glucose levels, and prediction of blood glucose.

In classification tasks, such as diagnosing diabetes or categorizing blood glucose levels into discrete ranges (e.g., hypoglycemia, normal, hyperglycemia), algorithms aim to assign samples to predefined classes based on extracted features. The performance of classification models is typically measured using metrics like accuracy, precision, recall, and the F1 score, which provide a comprehensive evaluation of the model's ability to distinguish between classes.

For regression tasks, where the objective is to predict continuous or single-point blood glucose values, models focus on accurately estimating glucose concentrations. The effectiveness of these models is evaluated using metrics such as MAE, RMSE and the coefficient of determination (R^2), which evaluate the precision of the model in capturing variations in blood glucose levels.

5.4. Sequence-based model

Differing from the single-point measurement approach employed by traditional BGM, the miniaturization and portability of non-invasive sensors enable continuous monitoring similar to CGM devices. These sensors offer complete non-invasiveness and eliminate the need for consumables, granting users the freedom to perform tests with peace of mind. Additionally, certain wearable devices are capable of continuously and uninterruptedly collecting monitoring data, thereby providing a substantial volume of historical test data for model optimization.

The glucose levels of different individuals can be influenced by variations in their physical characteristics, making it difficult to offer personalized descriptions due to the unobservability of these parameters. In contrast, empirical models driven by data offer a comparatively more accessible means of approximating the time-varying associations among variables [107]. Several calibration approaches grounded in non-linear data-driven architectures, utilizing historical readings from blood glucose monitors, have been investigated.

In recent years, as deep learning networks have advanced, an increasing number of researchers have initiated the application of these methodologies to the prediction of blood glucose levels. Kalita et al. [108] used the long-short-term memory network (LSTM) to provide an accurate glucose prediction output for closed-loop diabetes management devices. Dudukcu et al. [109] proposed a decision-level combinations of LSTM, WaveNet and gated recurrent unit (GRU) to predict the future blood glucose values of the patient using only the blood glucose values from the history of diabetes patients. Wang et al. [110] proposed a short-term blood glucose prediction model (VMD-IPSO-LSTM) that combines variational modal decomposition (VDM) and improved particle swarm optimization that optimizes the long short-term memory network (IPSO-LSTM) to improve glucose prediction accuracy. The results demonstrated superior predictive performance compared to the independent utilization of LSTM, VMD-LSTM, and VMD-PSO-LSTM methods. Mosquera et al. [111] introduced a robust glucose prediction model, denoted Gluco30, based on LSTM networks. This model was trained on an extensive dataset spanning 27,466 days, with the objective of forecasting glucose concentrations over a short-term horizon of 30 min. In particular, Gluco30 demonstrated superior predictive accuracy compared to alternative machine learning techniques. The authors attributed these favorable outcomes to several factors, including the ample size of the training dataset, the complexity of the network architecture, and the utilization of varying data volumes during the model training process. Consequently, it is plausible to say that deep learning networks tend to exhibit improved performance when provided with substantial training data.

5.5. Personalized calibration

As noted, the glucose levels of different individuals can be influenced by variations in their physical characteristics, making it difficult to offer personalized descriptions due to the unobservability of these parameters. Considering the significant differences in patient physiology and behavior, numerous research efforts have been undertaken to construct individualized deep learning models to enhance performance. Nevertheless, a frequently encountered scenario in clinical settings involves new patients who lack an adequate volume of data requisite to train a personalized model.

Transfer learning aims at applying knowledge from a specific domain (referred to as the source domain) to benefit an interconnected domain (referred to as the target domain), especially when there exists a scarcity of data specific to the target domain (see Fig. 10). Zhang et al. [112] introduced a State-by-State online transfer learning framework for the prediction of blood glucose. This framework incorporates incremental clustering to precisely categorize the states within patients' blood glucose time series. Subsequently, it constructs prediction models for each state and uses online transfer learning to train a new model when confronted with limited data. Deng et al. [113] devised deep learning techniques that leverage transfer learning and data augmentation to predict patient-specific blood glucose levels in different short-term time frames. They utilized patient-specific glucose measurements, recorded at 30-min intervals, to predict forthcoming glucose levels in the range of 5 min to 1 h. Yu et al. [114] applied TrAdaBoost to enhance the accuracy and stability of the glucose prediction model. Yu et al. [115] developed a predictive framework employing instance-based and network-based deep transfer learning for the purpose of cross-subject glucose prediction, utilizing segmented glucose time series data. To account for subject variability, the dynamic time warping (DTW) technique was employed to identify the most suitable source domain dataset, demonstrating the highest similarity to new subjects. Subsequently, a network-based deep transfer learning method was formulated, integrating cross-domain datasets to derive a personalized model with enhanced generalization capabilities. In a case study involving clinical datasets, it was observed that the proposed deep transfer learning framework, when supplemented with segmented datasets from other subjects, achieved markedly improved accuracy in predicting glucose levels for new subjects with type 2 diabetes.

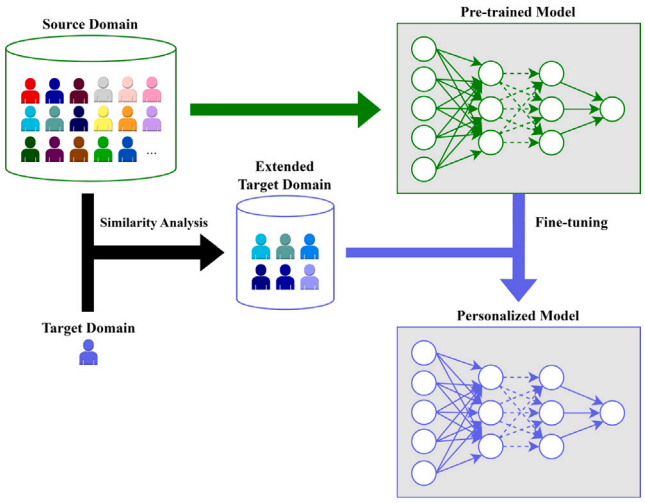

Fig. 10. Schematic diagram of transfer learning for blood glucose prediction in a new subject.

6. Commercial devices

From an economic perspective, non-invasive blood glucose monitoring technology holds significant market potential and commercial value. With the growing awareness of health issues and the increasing number of diabetes patients, non-invasive blood glucose monitoring technology is poised to become one of the mainstream products in the future diabetes management market. This will create substantial business opportunities for device manufacturers, healthcare providers, medical device suppliers, and others.

Following the review principles outlined in 2.2, this study selected representative commercial devices to exemplify and discuss the current state of non-invasive glucose monitoring technology. The selection process considered parameters such as the underlying technology, measurement principles, market readiness, and clinical validation. The specific devices and their detailed parameters are presented in Fig. 11 and Table 6.

6.1. D-POCKET

DiaMonTech is engaged in the development of a handheld non-invasive glucose monitoring device utilizing spectroscopy technology,

Table 6
Overview of recent commercial non-invasive glucose monitoring devices.

Device	Company	Technology	Target site	Accuracy	Data size	State	Ref
D-POCKET	DiaMonTech	MIR spectroscopy	Fingertip	PEG Zone A+B (98.8%)	100 subjects (41 with diabetes and 59 healthy) 2379 data pairs	CE approved	[116,117]
CoG - Hybrid Glucometer	CNOGA Medical	VIS-NIR spectroscopy	Fingertip	PEG Zone A (91.1%) Zone B (7.8%) MARD (18.1%)	100 subjects (all with diabetes)	CE approved	[118–121]
GlucoBeam	RSP Systems	Raman spectroscopy	Hand	PEG Zone A+B (98.8%) MARD (14.3%)	160 subjects (all with diabetes)	Under development	[122,123]
Wizmi	Wear2b	NIR spectroscopy	Wirst	CEG Zone A (93%) Zone B (7%) MARD (7.23%)	32 subjects (all healthy pregnant women)	Proof of concept	[124–126]
LTT	Light Touch Technology	NIR spectroscopy	Fingertip	N/A	N/A	Under development	[127,128]

Fig. 11. Representative commercial non-invasive blood glucose monitoring devices. (a) D-POCKET; (b) CoG-Hybrid Glucometer; (c) GlucoBeam; (d) Wizmi; (e) LTT.

referred to as "D-POCKET" (see Fig. 11a). This device emits a light beam through human figure tissue, captures the resultant temperature change, and subsequently calculates the glucose level [129]. In terms of regulatory approval, its prototype was announced to receive CE approval as a medical device in early 2019.

The D-POCKET device demonstrated promising accuracy in a clinical study, with most of the results falling within clinically acceptable zones according to PEG analysis. The study involved 100 participants, including 41 individuals with diabetes and 59 healthy subjects, covering a glucose range of 50 to 350 mg/dL. Invasive reference measurements using a clinical glucometer and non-invasive measurements on the finger were simultaneously recorded in 5-min intervals, starting from fasting glucose values for healthy subjects and low glucose values for diabetes patients, over a 2-h period. Specifically, 98.8% of the glucose measurements were within Zones A and B, increasing to 99.1% with an improved algorithm. Less than 1% of the data were in Zone C, with no data points in Zones D or E. The mean and median percent discrepancies between the non-invasive method and the reference invasive technique were 12.1% and 6.5%, respectively, for the full dataset, and improved to 11.3% and 6.4% with the enhanced algorithm [116].

However, it is important to note that this evaluation was conducted by the device's developers, which may introduce potential biases. The authors have disclosed conflicts of interest, as several key members of the study team are affiliated with DiaMonTech [116]. To ensure the robustness of these findings, further independent clinical trials are recommended, particularly those conducted by third-party researchers. These trials should involve a larger and more diverse participant pool and assess long-term stability and reproducibility in real-world settings. Additionally, the accuracy of the device in various patient populations and under different environmental conditions needs further validation.

6.2. CoG - Hybrid Glucometer

CNOGA developed a spectral-based non-invasive glucose monitoring device (see Fig. 11b). Within the finger compartment, a set of four LEDs emit light at specific wavelengths ranging from approximately 600 to 1000 nm, which is directed through the fingertip. As this light propagates through the tissue and blood capillaries, a portion of it is absorbed, leading to modifications in the light signal. Subsequently, the altered light is directed onto a color image sensor, essentially a camera. This specialized high-resolution camera captures real-time images of the light passing through the tissue of the fingertip for glucose calculation [118,119]. In terms of regulatory approval, CNOGA has received CE certification for monitors for vital and non-vital parameters (Certificate No.:6019GB410190910).

In clinical settings, the CoG device was evaluated for system accuracy according to ISO standards. The study involved 100 samples from participants with type 1 and type 2 diabetes, as well as healthy volunteers (43 females, 57 males; mean age: 53 ± 16 years). Glucose measurements were distributed according to ISO standards, and non-invasive tissue glucose readings were tested, followed by healthcare professionals compared to the YSI 2300 Stat Plus reference method. The CoG device met the ISO system accuracy criteria, with 99% raw data points falling within Zones A and B (91.1% and 7.8%, respectively) of CEG, although the MARD for the non-invasive component was relatively higher at 18.1% [121]. These results suggest that while the non-invasive component shows promising potential for pain-free glucose monitoring, further optimization is needed to improve its precision in glucose prediction.

It is also important to note that the study was supported by a research grant from CNOGA Medical Ltd., the manufacturer of the CoG device. This potential conflict of interest raises the possibility of bias in the design, data collection and analysis of the study. Given these circumstances, further independent clinical trials conducted by third-party researchers are recommended. Such trials should involve a larger and more diverse participant pool to confirm the repeatability of these findings and evaluate the long-term stability of the device in real-world settings. Independent studies would help ensure the objectivity and reliability of the device's performance across different patient populations and environmental conditions.

6.3. GlucoBeam

The GlucoBeam device (see Fig. 11c), developed by RSP Systems, employs Raman spectroscopy for glucose measurement in the ISF. This method involves directing a laser beam at the molecules present in the ISF, including glucose, and quantifying the frequency shift of the reflected light [122].

In a home-based clinical study involving 160 subjects with diabetes, 99.8% of the measurements fell within Zones A and B in PEG, with a MARD of 14.3% [123]. The GlucoBeam device demonstrated promising results in terms of accuracy, with the vast majority of measurements falling within clinically acceptable zones. However, the MARD of 14.3% suggests that while the device performs well, there is still room for improvement in terms of precision.

Same as the above devices, several authors of the clinical study are affiliated with RSP Systems, the manufacturer of the GlucoBeam device, including inventors on patent applications related to the device. This potential conflict of interest should be considered when interpreting the study results. To ensure unbiased results, future trials conducted by independent parties are recommended [123].

6.4. Wizmi

Wizmi (see Fig. 11d), developed by Wear2b, utilizes NIR light to capture absorption data through the skin and blood vessels, enabling the calculation of glucose levels. The sensor set is located as a wrist bracelet, in direct contact with the skin surface [124,125]. The device is still under development.

Throughout the clinical trial period, a total of 32 female participants were involved, resulting in a dataset comprising 224 data pairs. Within these 224 paired measurements, all data points fell within the clinically appropriate zones according to the CEG analysis, with 208 data pairs (93%) located in Zone A and 16 data pairs (7%) within Zone B. The MARD between the non-invasive glucose measurements obtained using Wizmi and the laboratory plasma glucose reference values was found to be 7.23% or 9.66 mg/dl [126].

However, since the device is still in the research stage and has not obtained any medical device certification, its accuracy and performance under various conditions remain questionable.

6.5. LTT

Light Touch Technology Inc. is currently engaged in the development of a non-invasive glucose monitoring device that employs MIR spectroscopy technology, as shown in Fig. 11e. At the present time, the device/sensor remains in the conceptual stage of development. To date, there is a dearth of clinical trial data available in connection with this particular device.

7. Discussion

Non-invasive glucose monitoring represents a highly promising avenue for enhancing the management of diabetes. However, the successful creation of reliable non-invasive glucose monitoring devices has proven to be a challenging objective within the realm of diabetes management.

7.1. Integration of multiple technologies

Unlike laboratory measurements, the generation, storage, and transformation of glucose molecules within the human body constitute an exceedingly complex process, manifesting in various physical information facets. Individual non-invasive measurement methods focus solely on one aspect of glucose. NIR spectroscopy emphasizes the optical transmitted absorption characteristics of glucose molecules [5], while MIR spectroscopy concentrates on the optical reflected absorption characteristics [6]. Thermal-based methodologies focus on the thermodynamic consequences arising from the metabolic decomposition of glucose [16]. Furthermore, non-optical methods consider specific glucose properties. Electromagnetic methods are mainly concerned with the electrical conductivity of glucose molecules within the human body [23], while electrochemical methods prioritize monitoring the concentration of glucose as it permeates from bodily fluids to the external surface of the skin [130]. By employing a methodology that amalgamates various techniques, non-invasive glucose information can be obtained from various dimensions, theoretically compensating for the issue of low accuracy associated with a single method. However, such devices that integrate multiple methods pose challenges in terms of the miniaturization and integration of sensors. Moreover, integrated methods often result in increased device costs and restricted usage conditions, issues that require resolution in future research efforts.

7.2. Device calibration

The indirect methodology of non-invasive measurements requires calibration against contemporaneous blood glucose measurements, thus offering an estimate of glucose concentration. This calibration procedure is undertaken prior to device utilization, with the aim of mitigating the influence of individual quasi-stable variables, such as tissue thickness and structure. Typically, this calibration process involves multiple sets of paired invasive and non-invasive measurements, with the frequency of these pairs varying based on the specific device and technology employed. Given that calibration has been identified as a potential source of discomfort [131], the adoption of a simple and brief calibration procedure is anticipated to improve the usability and user satisfaction of the device. Nonetheless, the majority of non-invasive devices have traditionally imposed a lengthy and intricate calibration process. For example, the CoG - Hybrid Glucometer requires at least 25 invasive calibration measurements performed at different times of the day, such as wake-up, before meal, 1–2 h after meal, and before sleep [118]. At present, the calibration of non-invasive devices appears to be an essential step to guarantee high levels of accuracy. Research in the field of non-invasive blood glucose devices is increasingly focusing on the duration, complexity, and validity period of the calibration process, with efforts aimed at rendering the calibration procedure more user-friendly and suitable for home use.

7.3. Suitability for various people

As noted, the primary constraint associated with optical techniques lies in the influence of human factors, particularly the characteristics of the skin. This is due to the fact that the transmission of light at various wavelengths is contingent upon factors such as skin thickness, color, and structure, as well as the presence of bone, blood, and other materials through which the light must traverse. Although the use of PPG can mitigate the impact of these variables by extracting pulsed blood flow, there are instances where the PPG waveform remains undetectable, primarily due to a low perfusion index (PI). This is especially pertinent in conditions characterized by low blood pressure and inadequate peripheral circulation [132]. Furthermore, the efficacy of non-invasive devices can be subject to the user's health status, including factors such as cardiac health, blood pressure levels and insulin injection. A study by Segman et al. [119] evaluated the precision and reliability of the measurements obtained from the CoG - Hybrid Glucometer in individuals with type 1 diabetes, type 2 diabetes, as well as healthy participants and individuals who had undergone cardiac surgery. However, there is still a limited body of research that elucidates the relationship between the accuracy of these devices and the health conditions of their users.

7.4. Model interpretability

At the current stage, non-invasive glucose monitoring technologies rely predominantly on machine learning algorithms to achieve exceptional predictive performance. Nevertheless, a significant portion of these algorithms are characterized as "black-box" models, lacking the capacity to provide explanatory insights into their predictions. The demand for interpretability in machine learning models within the healthcare domain is paramount, as clinicians require well-founded

decisions supported by reliable results accompanied by comprehensive explanations [133]. Although non-invasive sensors can provide estimations of blood glucose levels based on sensor readings, predictions lacking an understandable rationale may lead to patient confusion. This, in turn, can undermine the dependability of the device, especially when patients are accustomed to conventional invasive methods. Future research efforts should evaluate interpretability methodologies, such as signal correlation analysis, and integrate them into proposed solutions.

7.5. Medical applicability

A significant constraint observed in the existing body of research is the relatively small sample sizes in the majority of studies. This limitation has clear implications for the clinical utility of the findings and poses challenges in terms of the generalizability of the results. Furthermore, some studies did not include trials involving patients with diabetes, which restricted the relevance of proposed solutions to the intended end users. Furthermore, most of these studies were conducted in controlled environments and involved subjects of a younger demographic. It is imperative that future research endeavors involve larger and more diverse participant cohorts, encompassing a broader age spectrum, and specifically encompassing patients with both type 1 and type 2 diabetes.

7.6. Regulatory challenges and accuracy evaluation

The development of non-invasive glucose monitoring devices faces significant regulatory hurdles due to their current inability to meet the precision standards achieved by BGM. As mentioned earlier, the accuracy of most non-invasive devices has not been fully validated by independent third-party testing agencies, and the reported accuracy in the literature remains questionable. Despite advances in technology, non-invasive methods are still seen as less accurate. Unlike CGM, which can be assessed using MARD as a standalone metric, non-invasive devices often lack such a clear, universal accuracy benchmark. This absence of a standardized accuracy evaluation framework complicates the assessment of non-invasive devices, making it difficult for them to gain regulatory approval. As a result, further technological improvements and the development of reliable metrics for accuracy evaluation are crucial to enabling non-invasive glucose monitors to meet regulatory and clinical expectations.

8. Conclusion

This review paper presents a comprehensive perspective on non-invasive blood glucose measurement methods based on infrared absorption spectroscopy. Its objectives encompass not only the engineering disciplines associated with this technology, but also pertinent physiological and biochemical theories, along with the modeling techniques underpinning it. Subsequently, we endeavor to showcase the relevant non-invasive blood glucose monitoring devices and their current state of development.

A prominent constraint of present non-invasive devices lies in their inability to completely supplant conventional glucose meters. As such, ongoing enhancements are imperative, directed towards augmenting performance through substantial refinements in algorithms, software, and device attributes. Furthermore, additional clinical investigations are imperative to determine whether the use of non-invasive devices genuinely results in improved glycemic control among patients using this technology. It is also important to acknowledge that non-invasive glucose monitoring devices offer economic advantages to consumers by eliminating the requirement for lancets or test strips. However, it is imperative to recognize that certain technologies under consideration still involve significant expenses. Given the emergence of novel technologies and the ongoing enhancement of existing ones, we maintain the belief that the advent of a truly non-invasive glucose monitoring solution is merely a question of time.

Funding

This work was financially supported by grants from National Key Research and Development Program of China (2023YFC3604601) and Natural Science Foundation of Hunan Province of China (2023JJ70019).

CRediT authorship contribution statement

Taixiang Li: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Quangui Wang:** Supervision, Resources, Funding acquisition, Conceptualization. **Ying An:** Writing – review & editing, Visualization, Resources. **Lin Guo:** Investigation, Formal analysis. **Linan Ren:** Visualization, Data curation. **Linghao Lei:** Resources. **Xianlai Chen:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Xianlai Chen reports financial support was provided by National Key Research and Development Program of China (2023YFC3604601). Quangui Wang reports financial support was provided by Natural Science Foundation of Hunan Province of China (2023JJ70019). If there are other authors, they 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

No data was used for the research described in the article.

References

[1] Hong Sun, Pouya Saeedi, Suvi Karuranga, Moritz Pinkepank, Katherine Ogurtsova, Bruce B Duncan, Caroline Stein, Abdul Basit, Juliana CN Chan, Jean Claude Mbanya, et al., IDF diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045, Diabetes Res. Clin. Pract. 183 (2022) 109119.

[2] Adam G Tabák, Christian Herder, Wolfgang Rathmann, Eric J Brunner, Mika Kivimäki, Prediabetes: A high-risk state for developing diabetes, Lancet 379 (9833) (2012) 2279.

[3] Udo Hoss, Erwin Satrya Budiman, Factory-calibrated continuous glucose sensors: The science behind the technology, Diabetes Technol. Ther. 19 (S2) (2017) S–44.

[4] H. Michael Heise, Sven Delbeck, Ralf Marbach, Noninvasive monitoring of glucose using near-infrared reflection spectroscopy of skin—constraints and effective novel strategy in multivariate calibration, Biosensors 11 (3) (2021) 64.

[5] Aminah Hina, Hamza Nadeem, Wala Saadeh, A single led photoplethysmography-based noninvasive glucose monitoring prototype system, in: 2019 IEEE International Symposium on Circuits and Systems, ISCAS, IEEE, 2019, pp. 1–5.

[6] Ryosuke Kasahara, Saiko Kino, Shunsuke Soyama, Yuji Matsuura, Noninvasive glucose monitoring using mid-infrared absorption spectroscopy based on a few wavenumbers, Biomed. Opt. Express 9 (1) (2018) 289–302.

[7] Takuya Koyama, Naoto Shibata, Saiko Kino, Atsushi Sugiyama, Naota Akikusa, Yuji Matsuura, A compact mid-infrared spectroscopy system for healthcare applications based on a wavelength-swept, pulsed quantum cascade laser, Sensors 20 (12) (2020) 3438.

[8] YT Lan, YP Kuang, LP Zhou, GY Wu, PC Gu, HJ Wei, K Chen, Noninvasive monitoring of blood glucose concentration in diabetic patients with optical coherence tomography, Laser Phys. Lett. 14 (3) (2017) 035603.

[9] Jian Zhang, Yongjun Zheng, Jimmy Lee, Alex Hoover, Sarah Ann King, Lifeng Chen, Jing Zhao, Qiuning Lin, Cunjiang Yu, Linyong Zhu, et al., Continuous glucose monitoring enabled by fluorescent nanodiamond boronic hydrogel, Adv. Sci. (2023) 2203943.

[10] Jeon Woong Kang, Yun Sang Park, Hojun Chang, Woochang Lee, Surya Pratap Singh, Wonjun Choi, Luis H Galindo, Ramachandra R Dasari, Sung Hyun Nam, Jongae Park, et al., Direct observation of glucose fingerprint using in vivo Raman spectroscopy, Sci. Adv. 6 (4) (2020) eaay5206.

[11] Yun Sang Park, Sungmo Ahn, Hojun Chang, Woochang Lee, Sung Hyun Nam, Influence of Raman spectrometer collection efficiency on performance of noninvasive blood glucose detection for device miniaturization, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society, EMBC, IEEE, 2020, pp. 6139–6142.
[12] Zhen Fang Yu, Casey W. Pirnstill, Gerard L. Coté, Dual-modulation, dual-wavelength, optical polarimetry system for glucose monitoring, J. Biomed. Opt. 21 (8) (2016) 087001–087001.
[13] Yujiro Tanaka, Takuro Tajima, Michiko Seyama, Kayo Waki, Differential continuous wave photoacoustic spectroscopy for non-invasive glucose monitoring, IEEE Sens. J. 20 (8) (2019) 4453–4458.
[14] Marcel Mueller, Michael Grunze, Edward H Leiter, Peter C Reifsnyder, Ulrike Klueh, Don Kreutzer, Non-invasive glucose measurements in mice using mid-infrared emission spectroscopy, Sensors Actuators B: Chem. 142 (2) (2009) 502–508.
[15] Tomoya Kitazaki, Yusuke Morimoto, So Yamashita, Daichi Anabuki, Shiori Tahara, Akira Nishiyama, Kenji Wada, Ichiro Ishimaru, Glucose emission spectra through mid-infrared passive spectroscopic imaging of the wrist for non-invasive glucose sensing, Sci. Rep. 12 (1) (2022) 20558.
[16] Yang Zhang, Jian-ming Zhu, Yong-bo Liang, Hong-bo Chen, Shi-min Yin, Zhen-cheng Chen, Non-invasive blood glucose detection system based on conservation of energy method, Physiol. Meas. 38 (2) (2017) 325.
[17] Ok Kyung Cho, Yoon Ok Kim, Hiroshi Mitsumaki, Katsuhiko Kuwa, Noninvasive measurement of glucose by metabolic heat conformation method, Clin. Chem. 50 (10) (2004) 1894–1898.
[18] Fei Tang, Xiaohao Wang, Dongsheng Wang, Junfeng Li, Non-invasive glucose measurement by use of metabolic heat conformation method, Sensors 8 (5) (2008) 3335–3344.
[19] George Shaker, Karly Smith, Ala Eldin Omer, Shuo Liu, Clement Csech, Udeshaya Wadhwa, Safieddin Safavi-Naeini, Richard Hughson, Non-invasive monitoring of glucose level changes utilizing a mm-wave radar system, Int. J. Mob. Hum. Comput. Interact. (IJMHCI) 10 (3) (2018) 10–29.
[20] Amyrul Azuan Mohd Bahar, Z Zakaria, AAM Isa, Rammah A Alahnomi, Norhanani Abd Rahman, Complex permittivity measurement based on planar microfluidic resonator sensor, in: 2018 18th International Symposium on Antenna Technology and Applied Electromagnetics, ANTEM, IEEE, 2018, pp. 1–5.
[21] Olga Cherkasova, Maxim Nazarov, Alexander Shkurinov, Noninvasive blood glucose monitoring in the terahertz frequency range, Opt. Quantum Electron. 48 (2016) 1–12.
[22] Helena Cano-Garcia, Rohit Kshirsagar, Roberto Pricci, Ahmed Teyeb, Fergus O'brien, Shimul Saha, Panagiotis Kosmas, Efthymios Kallos, Enhancing the accuracy of non-invasive glucose sensing in aqueous solutions using combined millimeter wave and near infrared transmission, Sensors 21 (9) (2021) 3275.
[23] Jan-Hugo Andersen, Olav Bjerke, Fatos Blakaj, Vilde Moe Flugsrud, Fredrik Alstad Jacobsen, Marius Jonsson, Eirik Nobuki Kosaka, Petter André Langstrand, Øyvind Grannes Martinsen, Alexander Stene Moen, et al., Bioimpedance and NIR for non-invasive assessment of blood glucose, J. Electr. Bioimpedance 10 (1) (2019) 133.
[24] Svenja Meyhöfer, Britta Wilms, Flavia Ihling, Anne Windjäger, Hannes Kalscheuer, Andrej Augustinov, Vera Herrmann, Hendrik Lehnert, Sebastian M Schmid, Evaluation of a near-infrared light ultrasound system as a non-invasive blood glucose monitoring device, Diabetes Obes. Metab. 22 (4) (2020) 694–698.
[25] Yao Yao, Jingyao Chen, Yuhan Guo, Tian Lv, Zilin Chen, Ning Li, Shaokui Cao, Bingdi Chen, Tao Chen, Integration of interstitial fluid extraction and glucose detection in one device for wearable non-invasive blood glucose sensors, Biosens. Bioelectron. 179 (2021) 113078.
[26] Yanxiang Cheng, Xia Gong, Jian Yang, Guizhou Zheng, Ying Zheng, Yanjun Li, Yunsheng Xu, Gang Nie, Xi Xie, Meiwan Chen, et al., A touch-actuated glucose sensor fully integrated with microneedle array and reverse iontophoresis for diabetes monitoring, Biosens. Bioelectron. 203 (2022) 114026.
[27] Jia Zhu, Shangbin Liu, Zhihui Hu, Xianzhe Zhang, Ning Yi, Kairui Tang, Michael Gregory Dexheimer, Xiaojun Lian, Qing Wang, Jian Yang, et al., Laser-induced graphene non-enzymatic glucose sens. for on-body measurements, Biosens. Bioelectron. 193 (2021) 113606.
[28] Marjan Gusev, Lidija Poposka, Gjoko Spasevski, Magdalena Kostoska, Bojana Koteska, Monika Simjanoska, Nevena Ackovska, Aleksandar Stojmenski, Jurij Tasic, Janez Trontelj, et al., Noninvasive glucose measurement using machine learning and neural network methods and correlation with heart rate variability, J. Sens. 2020 (2020).
[29] Jingzhen Li, Igbe Tobore, Yuhang Liu, Abhishek Kandwal, Lei Wang, Zedong Nie, Non-invasive monitoring of three glucose ranges based on ECG by using DBSCAN-CNN, IEEE J. Biomed. Heal. Inform. 25 (9) (2021) 3340–3350.
[30] Mihaela Porumb, Corbin Griffen, John Hattersley, Leandro Pecchia, Nocturnal low glucose detection in healthy elderly from one-lead ECG using convolutional denoising autoencoders, Biomed. Signal Process. Control. 62 (2020) 102054.
[31] Osamah Alrezj, Mohammed Benaissa, Saleh A. Alshebeili, Digital bandstop filtering in the quantitative analysis of glucose from near-infrared and midinfrared spectra, J. Chemom. 34 (3) (2020) e3206.
[32] J. Michael McMillin, Blood glucose, Clinical Methods: The History, Physical, and Laboratory Examinations, 3rd ed., Butterworths, 1990.
[33] Danielle Bruen, Colm Delaney, Larisa Florea, Dermot Diamond, Glucose sensing for diabetes monitoring: Recent developments, Sensors 17 (8) (2017) 1866.
[34] Wilbert Villena Gonzales, Ahmed Toaha Mobashsher, Amin Abbosh, The progress of glucose monitoring—A review of invasive to minimally and non-invasive techniques, devices and sensors, Sensors 19 (4) (2019) 800.
[35] Shizuki Sai, Mitsunori Urata, Iwao Ogawa, Evaluation of linearity and interference effect on SMBG and POCT devices, showing drastic high values, low values, or error messages, J. Diabetes Sci. Technol. 13 (4) (2019) 734–743.
[36] Vladimir Gubala, Leanne F Harris, Antonio J Ricco, Ming X Tan, David E Williams, Point of care diagnostics: Status and future, Anal. Chem. 84 (2) (2012) 487–515.
[37] Hon-Ke Sia, Chew-Teng Kor, Shih-Te Tu, Pei-Yung Liao, Jiun-Yi Wang, Self-monitoring of blood glucose in association with glycemic control in newly diagnosed non-insulin-treated diabetes patients: A retrospective cohort study, Sci. Rep. 11 (1) (2021) 1–9.
[38] Lorena Alarcon-Casas Wright, Irl B. Hirsch, Metrics beyond hemoglobin A1C in diabetes management: Time in range, hypoglycemia, and other parameters, Diabetes Technol. Ther. 19 (S2) (2017) S–16.
[39] U.S. Food and Drug Administration, Self-monitoring blood glucose test systems for over-the-counter use: Guidance for industry and food and drug administration staff, 2020, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/self-monitoring-blood-glucose-test-systems-over-counter-use.
[40] William L. Clarke, The original clarke error grid analysis (EGA), Diabetes Technol. Ther. 7 (5) (2005) 776–779.
[41] Joan L Parkes, Stephen L Slatin, Scott Pardo, Barry H Ginsberg, A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose, Diabetes Care 23 (8) (2000) 1143–1148.
[42] Thomas Danne, Revital Nimri, Tadej Battelino, Richard M Bergenstal, Kelly L Close, J Hans DeVries, Satish Garg, Lutz Heinemann, Irl Hirsch, Stephanie A Amiel, et al., International consensus on use of continuous glucose monitoring, Diabetes Care 40 (12) (2017) 1631–1640.
[43] Lutz Heinemann, Michael Schoemaker, Günther Schmelzeisen-Redecker, Rolf Hinzmann, Adham Kassab, Guido Freckmann, Florian Reiterer, Luigi Del Re, Benefits and limitations of MARD as a performance parameter for continuous glucose monitoring in the interstitial space, J. Diabetes Sci. Technol. 14 (1) (2020) 135–150.
[44] Florian Reiterer, Philipp Polterauer, Michael Schoemaker, Guenther Schmelzeisen-Redecker, Guido Freckmann, Lutz Heinemann, Luigi Del Re, Significance and reliability of MARD for the accuracy of CGM systems, J. Diabetes Sci. Technol. 11 (1) (2017) 59–67.
[45] International Organization for Standardization, In vitro diagnostic test systems — Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus (ISO 15197), 2013, https://www.iso.org/standard/54976.html.
[46] Georgia M Davis, Elias K Spanakis, Alexandra L Migdal, Lakshmi G Singh, Bonnie Albury, Maria Agustina Urrutia, K Walkiria Zamudio-Coronado, William H Scott, Rebecca Doerfler, Sergio Lizama, et al., Accuracy of dexcom G6 continuous glucose monitoring in non–critically ill hospitalized patients with diabetes, Diabetes Care 44 (7) (2021) 1641–1646.
[47] Kevin Hanson, Mark Kipnes, Hien Tran, Comparison of point accuracy between two widely used continuous glucose monitoring systems, J. Diabetes Sci. Technol. 18 (3) (2024) 598–607.
[48] Shridhara Alva, Timothy Bailey, Ronald Brazg, Erwin S Budiman, Kristin Castorino, Mark P Christiansen, Gregory Forlenza, Mark Kipnes, David R Liljenquist, Hanqing Liu, Accuracy of a 14-day factory-calibrated continuous glucose monitoring system with advanced algorithm in pediatric and adult population with diabetes, J. Diabetes Sci. Technol. 16 (1) (2022) 70–77.
[49] U.S. Food and Drug Administration, Integrated continuous glucose monitoring system - 21 CFR 862.1355, 2022, https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=862.1355.
[50] Akikazu Sakudo, Near-infrared spectroscopy for medical applications: Current status and future perspectives, Clin. Chim. Acta 455 (2016) 181–188.
[51] Michael F.G. Wood, Daniel Côté, I. Alex Vitkin, Combined optical intensity and polarization methodology for analyte concentration determination in simulated optically clear and turbid biological media, J. Biomed. Opt. 13 (4) (2008) 044037–044037.
[52] Andreas Caduff, Mark S. Talary, Pavel Zakharov, Cutaneous blood perfusion as a perturbing factor for noninvasive glucose monitoring, Diabetes Technol. Ther. 12 (1) (2010) 1–9.
[53] Paul A.J. Kolarsick, Maria Ann Kolarsick, Carolyn Goodwin, Anatomy and physiology of the skin, J. Dermatol. Nurses' Assoc. 3 (4) (2011) 203–213.
[54] John Allen, Photoplethysmography and its application in clinical physiological measurement, Physiol. Meas. 28 (3) (2007) R1.
[55] Pak-Hei Chan, Chun-Ka Wong, Yukkee C Poh, Louise Pun, Wangie Wan-Chiu Leung, Yu-Fai Wong, Michelle Man-Ying Wong, Ming-Zher Poh, Daniel Wai-Sing Chu, Chung-Wah Siu, Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting, J. Am. Hear. Assoc. 5 (7) (2016) e003428.

[56] Jinseok Lee, Bersain A Reyes, David D McManus, Oscar Maitas, Ki H Chon, Atrial fibrillation detection using an iphone 4S, IEEE Trans. Biomed. Eng. 60 (1) (2012) 203–206.

[57] Joonnyong Lee, JangJay Sohn, Jonghyun Park, SeungMan Yang, Saram Lee, Hee Chan Kim, Novel blood pressure and pulse pressure estimation based on pulse transit time and stroke volume approximation, Biomed. Eng. Online 17 (2018) 1–20.

[58] Enric Monte-Moreno, Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques, Artif. Intell. Med. 53 (2) (2011) 127–138.

[59] Ludi Wang, Wei Zhou, Ying Xing, Xiaoguang Zhou, et al., A novel neural network model for blood pressure estimation using photoplethesmography without electrocardiogram, J. Heal. Eng. 2018 (2018).

[60] Iain M. Johnstone, D. Michael Titterington, Statistical challenges of high-dimensional data, 2009.

[61] Yong-Huan Yun, Hong-Dong Li, Bai-Chuan Deng, Dong-Sheng Cao, An overview of variable selection methods in multivariate analysis of near-infrared spectra, TRAC Trends Anal. Chem. 113 (2019) 102–115.

[62] Zhang Man, Liu Xu-hua, He Xiong-kui, Zhang Lu-da, Zhao Long-lian, Li Jun-hui, Study on the application of ridge regression to near-infrared spectroscopy quantitative analysis and optimum wavelength selection, Spectrosc. Spectr. Anal. 30 (5) (2010) 1214–1217.

[63] Xin Huang, Yi-Ping Luo, Qing-Song Xu, Yi-Zeng Liang, Elastic net wavelength interval selection based on iterative rank PLS regression coefficient screening, Anal. Methods 9 (4) (2017) 672–679.

[64] Xiaoli Luan, Jin Liu, Fei Liu, Multilevel LASSO-based NIR temperature-correction modeling for viscosity measurement of bisphenol-A, ISA Trans. 107 (2020) 206–213.

[65] Xueguang Shao, Wei Wang, Zhenyu Hou, Wensheng Cai, A new regression method based on independent component analysis, Talanta 69 (3) (2006) 676–680.

[66] V. Gaydou, J. Kister, Nathalie Dupuy, Evaluation of multiblock NIR/MIR PLS predictive models to detect adulteration of diesel/biodiesel blends by vegetal oil, Chemometr. Intell. Lab. Syst. 106 (2) (2011) 190–197.

[67] John H. Holland, Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, MIT Press, 1992.

[68] Quanxi Feng, Huazhou Chen, Hai Xie, Ken Cai, Bin Lin, Lili Xu, A novel genetic algorithm-based optimization framework for the improvement of near-infrared quantitative calibration models, Comput. Intell. Neurosci. 2020 (2020).

[69] Yan Wang, Bingren Xiang, Radial basis function network calibration model for near-infrared spectra in wavelet domain using a genetic algorithm, Anal. Chim. Acta 602 (1) (2007) 55–65.

[70] Jiyong Shi, Xuetao Hu, Xiaobo Zou, Jiewen Zhao, Wen Zhang, Xiaowei Huang, Yaodi Zhu, Zhihua Li, Yiwei Xu, A heuristic and parallel simulated annealing algorithm for variable selection in near-infrared spectroscopy analysis, J. Chemom. 30 (8) (2016) 442–450.

[71] Huang Xiaowei, Zou Xiaobo, Zhao Jiewen, Shi Jiyong, Zhang Xiaolei, Mel Holmes, Measurement of total anthocyanins content in flowering tea using near infrared spectroscopy combined with ant colony optimization models, Food Chem. 164 (2014) 536–543.

[72] Hui Cao, Yanxia Wang, Sanchun Yang, Yan Zhou, A wavelength selection method based on random decision particle swarm optimization with attractor for near-infrared spectral quantitative analysis, J. Chemom. 29 (5) (2015) 289–299.

[73] Xiaobo Zou, Jiewen Zhao, Yanxiao Li, Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models, Vib. Spectrosc. 44 (2) (2007) 220–227.

[74] Sumaporn Kasemsumran, Nattapom Suttiwuitpukdee, Vichein Keeratinuakal, Rapid classification of turmeric based on DNA fingerprint by near-infrared spectroscopy combined with moving window partial least squares-discrimination analysis, Anal. Sci. 33 (1) (2017) 111–115.

[75] Guo Tang, Yue Huang, Kuangda Tian, Xiangzhong Song, Hong Yan, Jing Hu, Yanmei Xiong, Shungeng Min, A new spectral variable selection pattern using competitive adaptive reweighted sampling combined with successive projections algorithm, Analyst 139 (19) (2014) 4894–4902.

[76] Cheng Li, Tianlun Zhao, Cong Li, Lei Mei, En Yu, Yating Dong, Jinhong Chen, Shuijin Zhu, Determination of gossypol content in cottonseeds by near infrared spectroscopy based on Monte Carlo uninformative variable elimination and nonlinear calibration methods, Food Chem. 221 (2017) 990–996.

[77] Mohamed Elgendi, Optimal signal quality index for photoplethysmogram signals, Bioengineering 3 (4) (2016) 21.

[78] Anju Prabha, Jyoti Yadav, Asha Rani, Vijander Singh, Intelligent estimation of blood glucose level using wristband PPG signal and physiological parameters, Biomed. Signal Process. Control. 78 (2022) 103876.

[79] Po-Lei Lee, Kuo-Wei Wang, Chen-Yuan Hsiao, A non-invasive blood glucose estimation system using dual-channel PPGs and pulse-arrival velocity, IEEE Sens. J. (2023).

[80] Erick Javier Argüello-Prada, Sonia María Bolaños, On the role of perfusion index for estimating blood glucose levels with ultrasound-assisted and conventional finger photoplethysmography in the near-infrared wavelength range, Biomed. Signal Process. Control. 86 (2023) 105338.

[81] Jingzhen Li, Jingjing Ma, Olatunji Mumini Omisore, Yuhang Liu, Huajie Tang, Pengfei Ao, Yan Yan, Lei Wang, Zedong Nie, Noninvasive blood glucose monitoring using spatiotemporal ECG and ppg feature fusion and weight-based choquet integral multimodel approach, IEEE Trans. Neural Netw. Learn. Syst. (2023).

[82] Chia-Yu Liao, Wai-Chi Fang, LRCN-based noninvasive blood glucose level estimation, in: 2023 IEEE International Symposium on Circuits and Systems, ISCAS, IEEE, 2023, pp. 1–5.

[83] Gaobo Zhang, Zhen Mei, Yuan Zhang, Xuesheng Ma, Benny Lo, Dongyi Chen, Yuanting Zhang, A noninvasive blood glucose monitoring system based on smartphone PPG signal processing and machine learning, IEEE Trans. Ind. Inform. 16 (11) (2020) 7209–7218.

[84] Aminah Hina, Wala Saadeh, A noninvasive glucose monitoring SoC based on single wavelength photoplethysmography, IEEE Trans. Biomed. Circuits Syst. 14 (3) (2020) 504–515.

[85] Ke Xu, Xinyu Jiang, Sijie Lin, Chenyun Dai, Wei Chen, Stochastic modeling based nonlinear Bayesian filtering for photoplethysmography denoising in wearable devices, IEEE Trans. Ind. Inform. 16 (11) (2020) 7219–7230.

[86] Weizhi Guo, Bingo Wing-Kuen Ling, Yiting Wei, Denoising of photoplethysmograms for non-invasive blood glucose estimation via slant transform based bit plane method, in: 2023 IEEE International Conference on Industrial Technology, ICIT, IEEE, 2023, pp. 1–4.

[87] Shisen Chen, Fen Qin, Xuesheng Ma, Jie Wei, Yuan-Ting Zhang, Yuan Zhang, Emil Jovanov, Multi-view cross-fusion transformer based on kinetic features for non-invasive blood glucose measurement using PPG signal, IEEE J. Biomed. Heal. Inform. (2024).

[88] Henry Friday Nweke, Ying Wah Teh, Ghulam Mujtaba, Mohammed Ali Al-Garadi, Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions, Inf. Fusion 46 (2019) 147–170.

[89] Ivan Miguel Pires, Nuno M Garcia, Nuno Pombo, Francisco Flórez-Revuelta, From data acquisition to data fusion: A comprehensive review and a roadmap for the identification of activities of daily living using mobile devices, Sensors 16 (2) (2016) 184.

[90] Guoxin Zhang, Zengcai Wang, Lei Zhao, Yazhou Qi, Jinshan Wang, et al., Coal-rock recognition in top coal caving using bimodal deep learning and Hilbert-Huang transform, Shock. Vib. 2017 (2017).

[91] Brinnae Bent, Peter J Cho, Maria Henriquez, April Wittmann, Connie Thacker, Mark Feinglos, Matthew J Crowley, Jessilyn P Dunn, Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches, Npj Digit. Med. 4 (1) (2021) 89.

[92] Seyedfakhreddin Nabavi, Sharmistha Bhadra, A robust fusion method for motion artifacts reduction in photoplethysmography signal, IEEE Trans. Instrum. Meas. 69 (12) (2020) 9599–9608.

[93] Muhammad Shoaib, Stephan Bosch, Ozlem Durmaz Incel, Hans Scholten, Paul JM Havinga, Complex human activity recognition using smartphone and wrist-worn motion sensors, Sensors 16 (4) (2016) 426.

[94] Oresti Banos, Juan-Manuel Galvez, Miguel Damas, Hector Pomares, Ignacio Rojas, Window size impact in human activity recognition, Sensors 14 (4) (2014) 6474–6499.

[95] Aboli N. Londhe, Mithilesh Atulkar, Semantic segmentation of ECG waves using hybrid channel-mix convolutional and bidirectional LSTM, Biomed. Signal Process. Control. 63 (2021) 102162.

[96] Qingquan Song, Hancheng Ge, James Caverlee, Xia Hu, Tensor completion algorithms in big data analytics, ACM Trans. Knowl. Discov. from Data (TKDD) 13 (1) (2019) 1–48.

[97] Andreas Bulling, Ulf Blanke, Bernt Schiele, A tutorial on human activity recognition using body-worn inertial sensors, ACM Comput. Surv. 46 (3) (2014) 1–33.

[98] Huile Xu, Jinyi Liu, Haibo Hu, Yi Zhang, Wearable sensor-based human activity recognition method with multi-features extracted from Hilbert-Huang transform, Sensors 16 (12) (2016) 2048.

[99] Zihan Nie, Meng Rong, Kaiyang Li, Blood glucose prediction based on imaging-photoplethysmography in combination with machine learning, Biomed. Signal Process. Control. 79 (2023) 104179.

[100] Moajjem Hossain Chowdhury, Muhammad EH Chowdhury, Abdulrahman Alqahtani, MMG-net: Multi modal approach to estimate blood glucose using multi-stream and cross modality attention, Biomed. Signal Process. Control. 92 (2024) 105975.

[101] Shantanu Sen Gupta, Tae-Ho Kwon, Shifat Hossain, Ki-Doo Kim, Towards non-invasive blood glucose measurement using machine learning: An all-purpose PPG system design, Biomed. Signal Process. Control. 68 (2021) 102706.

[102] Gustavo A Alonso-Silverio, Víctor Francisco-García, Iris P Guzmán-Guzmán, Elías Ventura-Molina, Antonio Alarcón-Paredes, Toward non-invasive estimation of blood glucose concentration: A comparative performance, Mathematics 9 (20) (2021) 2529.

[103] Tanvir Tazul Islam, Md Sajid Ahmed, Md Hassanuzzaman, Syed Athar Bin Amir, Tanzilur Rahman, Blood glucose level regression for smartphone ppg signals using machine learning, Appl. Sci. 11 (2) (2021) 618.
[104] Amit M Joshi, Prateek Jain, Saraju P Mohanty, Navneet Agrawal, iGLU 2.0: A new wearable for accurate non-invasive continuous serum glucose measurement in IoMT framework, IEEE Trans. Consum. Electron. 66 (4) (2020) 327–335.
[105] Robert Avram, Jeffrey E Olgin, Peter Kuhar, J Weston Hughes, Gregory M Marcus, Mark J Pletcher, Kirstin Aschbacher, Geoffrey H Tison, A digital biomarker of diabetes from smartphone-based vascular signals, Nature Med. 26 (10) (2020) 1576–1582.
[106] Yucen Yang, Jie Chen, Jie Wei, Zhikang Wang, Jiangning Song, Yuanting Zhang, Yuan Zhang, Jichao Zhao, Non-invasive blood glucose detection system with infrared pulse sensor and hybrid feature neural network, IEEE Sens. J. (2024).
[107] Ashenafi Zebene Woldaregay, Eirik Årsand, Ståle Walderhaug, David Albers, Lena Mamykina, Taxiarchis Botsis, Gunnar Hartvigsen, Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes, Artif. Intell. Med. 98 (2019) 109–134.
[108] Deepjyoti Kalita, Khalid B. Mirza, Glucose prediction using wide-deep LSTM network for accurate insulin dosing in artificial pancreas, in: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society, EMBC, IEEE, 2022, pp. 4426–4429.
[109] Hatice Vildan Dudukcu, Murat Taskiran, Tulay Yildirim, Blood glucose prediction with deep neural networks using weighted decision level fusion, Biocybern. Biomed. Eng. 41 (3) (2021) 1208–1223.
[110] Wenbo Wang, Meng Tong, Min Yu, Blood glucose prediction with VMD and LSTM optimized by improved particle swarm optimization, IEEE Access 8 (2020) 217908–217916.
[111] Clara Mosquera-Lopez, Robert Dodier, Nichole Tyler, Navid Resalat, Peter Jacobs, Leveraging a big dataset to develop a recurrent neural network to predict adverse glycemic events in type 1 diabetes, IEEE J. Biomed. Heal. Inform. (2019).
[112] Xinyu Zhang, Xia Yu, Zhanhu Zhang, Hongru Li, Jingyi Lu, Jian Zhou, A state-by-state online transfer learning framework with incremental clustering for blood glucose prediction, Biomed. Signal Process. Control. 86 (2023) 105200.
[113] Yixiang Deng, Lu Lu, Laura Aponte, Angeliki M Angelidi, Vera Novak, George Em Karniadakis, Christos S Mantzoros, Deep transfer learning and data augmentation improve glucose levels prediction in type 2 diabetes patients, NPJ Digit. Med. 4 (1) (2021) 109.
[114] Yan Yu, Jipeng Huang, Juan Zhu, Shili Liang, An accurate noninvasive blood glucose measurement system using portable near-infrared spectrometer and transfer learning framework, IEEE Sens. J. 21 (3) (2020) 3506–3519.
[115] Xia Yu, Tao Yang, Jingyi Lu, Yun Shen, Wei Lu, Wei Zhu, Yuqian Bao, Hongru Li, Jian Zhou, Deep transfer learning: A novel glucose prediction framework for new subjects with type 2 diabetes, Complex & Intell. Syst. (2021) 1–13.
[116] Thorsten Lubinski, Bartosz Plotka, Sergius Janik, Luca Canini, Werner Mäntele, Evaluation of a novel noninvasive blood glucose monitor based on mid-infrared quantum cascade laser technology and photothermal detection, J. Diabetes Sci. Technol. 15 (1) (2021) 6–10.
[117] Werner Mäntele, Miguel Angel Pleitez Rafael, Tobias Lieblein, Otto Hertzberg, Alexander Bauer, Hermann Von Lilienfeld-Toal, Arne Küderle, Tabea Pfuhl, Non-invasive substance analysis, 2021, US Patent 10, 883, 933.
[118] Cog Hybrid Glucometer, https://www.cnogacare.co/cog-hybrid-glucometer. (Accessed 10 July 2023).
[119] Yosef Segman, Device and method for noninvasive glucose assessment, J. Diabetes Sci. Technol. 12 (6) (2018) 1159–1168.
[120] Yosef Segman, Method and apparatus for non-invasive glucose measurement, 2020, US Patent 10, 687, 739.
[121] Andreas Pfützner, Filiz Demircik, Johannes Pfützner, Kim Kessler, Stephanie Strobl, Jan Spatz, Anke H Pfützner, Alexander Lier, System accuracy assessment of a combined invasive and noninvasive glucometer, J. Diabetes Sci. Technol. 14 (3) (2020) 575–581.
[122] GlucoBeam, https://rspsystems.com/glucobeam/. (Accessed 10 July 2023).
[123] Anders Pors, Kaspar G Rasmussen, Rune Inglev, Nina Jendrike, Amalie Philipps, Ajenthen G Ranjan, Vibe Vestergaard, Jan E Henriksen, Kirsten Nørgaard, Guido Freckmann, et al., Accurate post-calibration predictions for noninvasive glucose measurements in people using confocal Raman spectroscopy, ACS Sens. 8 (3) (2023) 1272–1279.
[124] Ohad Bashan, Oded Bashan, Aharon Klein, Ben Zion Dekel, Device, system and method for non-invasive monitoring of physiological measurements, 2021, US Patent 11, 129, 556.
[125] Ohad Bashan, BAR-SAKAI Giora, Oded Bashan, Adjustable non-invasive wearable monitoring device, 2022, US Patent App. 17/614, 600.
[126] Eran Hadar, Rony Chen, Yoel Toledano, Kinneret Tenenbaum-Gavish, Yuval Atzmon, Moshe Hod, Noninvasive, continuous, real-time glucose measurements compared to reference laboratory venous plasma glucose values, J. Matern.-Fetal Neonatal Med. 32 (20) (2019) 3393–3400.
[127] LTT, https://www.light-tt.co.jp/product?lang=en. (Accessed 10 July 2023).
[128] Kanade OGAWA Koichi Yamakawa, Substance-in-blood concentration measurement device and substance-in-blood concentration measurement method, 2021, Publication of EP4224147A1.
[129] DIAMONTECH D-POCKET, https://www.diamontech.de/en/solutions/d-pocket. (Accessed 10 July 2023).
[130] Juliane R Sempionatto, Muyang Lin, Lu Yin, Ernesto De la Paz, Kexin Pei, Thitaporn Sonsa-Ard, Andre N de Loyola Silva, Ahmed A Khorshed, Fangyu Zhang, Nicholas Tostado, et al., An epidermal patch for the simultaneous monitoring of haemodynamic and metabolic biomarkers, Nat. Biomed. Eng. 5 (7) (2021) 737–748.
[131] Giada Acciaroli, Martina Vettoretti, Andrea Facchinetti, Giovanni Sparacino, Claudio Cobelli, From two to one per day calibration of dexcom G4 platinum by a time-varying day-specific Bayesian prior, Diabetes Technol. Ther. 18 (8) (2016) 472–479.
[132] P. Sivaprasath, Raja Mookka Gounder, B. Mythili, Prediction of shock by peripheral perfusion index, Indian J. Pediatr. 86 (2019) 903–908.
[133] Radwa ElShawi, Youssef Sherif, Mouaz Al-Mallah, Sherif Sakr, Interpretability in healthcare: A comparative study of local machine learning interpretability techniques, Comput. Intell. 37 (4) (2021) 1633–1650.