Noninvasive Glucose Monitoring Using Polarized Light

Tianxing Li[†,#], Derek Bai[†], Temiloluwa Prioleau[†], Nam Bui[‡], Tam Vu[‡,§], and Xia Zhou[†]
[†]Department of Computer Science, Dartmouth College, Hanover, NH
[#]Department of Computer Science, UMass Amherst, Amherst, MA
[‡]Department of Computer Science, University of Colorado Boulder, Boulder, CO
[§]Department of Computer Science, Oxford University, UK
tianxingli@umass.edu,{derek.j.bai.21,temiloluwa.o.prioleau,xia.zhou}@dartmouth.edu
nam.bui@colorado.edu,tam.vu@cs.ox.ac.uk

ABSTRACT

We propose a compact noninvasive glucose monitoring system using polarized light, where a user simply needs to place her palm on the device for measuring her current glucose concentration level. The primary innovation of our system is the ability to minimize light scattering from the skin and extract weak changes in light polarization to estimate glucose concentration, all using low-cost hardware. Our system exploits multiple wavelengths and light intensity levels to mitigate the effect of user diversity and confounding factors (e.g., collagen and elastin in the dermis). It then infers glucose concentration using a generic learning model, thus no additional calibration is needed. We design and fabricate a compact (17 cm × 10 cm × 5 cm) and low-cost (i.e., <$250) prototype using off-the-shelf hardware. We evaluate our system with 41 diabetic patients and 9 healthy participants. In comparison to a continuous glucose monitor approved by U.S. Food and Drug Administration (FDA), 89% of our results are within zone A (clinically accurate) of the Clarke Error Grid. The absolute relative difference (ARD) is 10% . The r and p values of the Pearson correlation coefficients between our predicted glucose concentration and reference glucose concentration are 0.91 and 1.6×10^{-143}, respectively. These errors are comparable with FDA-approved glucose sensors, which achieve ≈90% clinical accuracy with a 10% mean ARD.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing**; • **Hardware** → **Sensor devices and platforms**.

KEYWORDS

Noninvasive glucose monitoring, light sensing

ACM Reference Format:
Tianxing Li[†,#], Derek Bai[†], Temiloluwa Prioleau[†], Nam Bui[‡], Tam Vu[‡,§], and Xia Zhou[†]. 2020. Noninvasive Glucose Monitoring Using Polarized Light. In *The 18th ACM Conference on Embedded Networked Sensor Systems (SenSys '20), November 16–19, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3384419.3430720

SenSys '20, November 16–19, 2020, Virtual Event, Japan

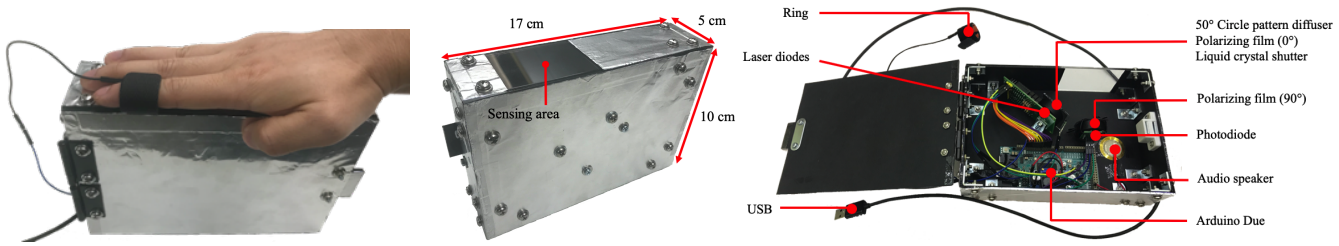
ACM ISBN 978-1-4503-7590-0/20/11…$15.00
https://doi.org/10.1145/3384419.3430720

1 INTRODUCTION

Diabetes is a major medical concern that affects more than 9% of the U.S. population [1]. The condition can affect body organs and increase the risk of heart disease, stroke, blindness, kidney failure, neuropathy, and congenital disabilities [67]. Currently, there is no cure for diabetes and patients can effectively manage their condition by monitoring the glucose concentration. A normal glucose concentration should be between 70 and 140 mg/dL (or 3.9 – 7.8 mmol/L) for non-diabetics. The glucose concentration is considered high if it is above 140 mg/dl after at least 8 hours without eating or drinking, or above 180 mg/dL after 2 hours without eating [45]. Clinical practice guidelines recommend that diabetic patients measure glucose concentration level at least three times a day (up to ten times for patients with Type 1 diabetes[1] [21]). Frequent glucose monitoring is crucial to maintaining their glucose concentration within the normal range and mitigating the risk of diabetes complications.

The mainstream glucose monitoring is currently performed invasively with enzymatic test strips measuring glucose concentration directly from a small sample of blood [9]. Although the invasive method is by far the most effective and accurate for monitoring blood glucose concentration level, it leverages automatic lancet devices to prick the fingertip and extract the blood sample. Using this process, it is painful to measure glucose level multiple times (i.e., >3 times) a day. Due to the discomfort, many diabetic patients rely on their symptoms and experiences to guess glucose concentration and insulin administration, which significantly increases the possibility of diabetic complications [71]. Alternatively, continuous glucose monitoring (CGM) systems can estimate glucose concentration in interstitial fluid (ISF) [31]. These minimally invasive devices rely on subcutaneously implanted biosensors to estimate real-time glucose concentration without the need of pricking the finger. However, CGM systems are still invasive, entailing risks of microbial infection, high financial cost of periodical replacement of biosensors [18], and the requirement of continuous calibration. Therefore, there is a significant need for a noninvasive, accurate, and low-cost glucose monitoring system.

[1]Type 1 diabetes is one of the two diabetes types, and it caused by genes and environmental factors, such as viruses.



(a) Prototype in use (b) Prototype design (c) Hardware components inside our prototype

Figure 1: An overview of the proposed system. The prototype contains a sensing box, an optics sensing unit, and a computation unit.

Existing noninvasive glucose monitoring systems still fall short in achieving these goals. Common approaches include optical, transdermal, and thermal techniques [31, 40, 49]. Optical techniques measure glucose concentration by analyzing the relationship between glucose concentration and optical signal changes such as optical rotation [13, 54] and light intensity [43, 63]). Due to the indirect and often weak relationship between the change of measured signal and glucose concentration, existing techniques are prone to low accuracy across diverse user groups, and affected by confounding factors and environmental changes like temperature. Transdermal techniques are based on the measurement of glucose molecules impedance [7, 70]. However, a limitation of this method is that low electric current passing through the skin can cause irritation [31]. Thermal techniques measure thermal generation, blood flow rate, hemoglobin, or oxyhemoglobin concentrations to estimate glucose levels, but the accuracy of existing systems is significantly affected by environmental factors like temperature and humidity.

In this paper, we propose a compact, low-cost, noninvasive glucose monitoring system with accuracy comparable to existing FDA-approved CGM sensors. Specifically, as shown in Figure 1(a), we propose a novel optical polarimetry that measures optical rotation related to the glucose concentration in ISF. At a high level, our system emits light towards the palm and then extracts optical rotation from the reflected light to estimate glucose concentration. Although optical polarimetry has been explored in prior literature, existing methods measure glucose concentration from either sample cells [10, 41] or transparent tissue in the body (eyes) [46, 54]. However, the former is still invasive, and the latter can cause photothermal damage to eyes. A recent method designs an optical coherence tomography to measure of glucose concentration on fingertips [13, 48, 51]. However, since these systems directly compute the amount of depolarization, they rely on bulky and expensive optical devices to measure small optical rotation. They also lack clinical evaluations on diabetic patients.

Challenges We face three main challenges to realize noninvasive blood glucose sensing using light. *First,* human skin is a highly absorbing and scattering medium. The majority of light is either absorbed or becomes depolarized due to tissue scattering, resulting in a very small portion of reflected light carrying the optical rotation associated with glucose concentration level. Thus, it is challenging to extract the weak optical rotation from reflected light. *Second*, other substances (e.g., collagen and elastin in the skin) in the body along the signal path also contribute to an observed optical rotation. These confounding factors can significantly reduce the accuracy of a light-based system for glucose monitoring. *Third*, user diversity presents an additional challenge. Our experiments show that skin characteristics (e.g., skin tone, skin thickness) affect the correlation between the observed signal and glucose concentration. Hence, a light-based glucose monitoring system and algorithm need to be robust against the aforementioned challenges.

Design We seek to address these challenges with three design considerations. *First*, to deal with depolarization from tissue scattering, we propose a novel methodology of *depolarization cancellation*, which cancels out depolarized light using low-cost, small hardware components. This sets a fundamental departure from prior works [13, 48, 51] that directly model depolarization. Specifically, our method leverages a low-cost LC and linear polarizer films to orient the light polarization to two orthogonal directions (parallel and perpendicular to scattering plane). Depolarization is canceled out via subtracting the measurements in these two directions. *Second*, to mitigate the impact of confounding factors, we exploit multiple wavelengths and intensity levels to control the light penetration depth beneath the skin. Since both the interstitial fluid and other confounding factors are not uniformly distributed in the body, probing skin with different light wavelengths and intensity levels allows the system to gather data at the penetration depth with higher linear correlations between observed signals and glucose concentration. Aggregating spatial features of reflected light under different wavelengths and light intensity levels also improves estimation of glucose level. *Finally*, to address the challenge of user diversity, we build a generic model based on boosted trees regression [75]. This generic model is trained with data from 50 participants. We then apply a resampling ensembles algorithm to handle imbalanced training data.

Evaluation To examine the feasibility of our approach, we built a compact (17 cm × 10 cm × 5 cm) and low-cost (<$250) prototype using off-the-shelf hardware components (Figure 1(b)). The prototype consists of three laser diodes (at 450 nm, 520 nm, and 658 nm), two linear polarizing films, a 50° circle pattern diffuser, an LC, and a photodiode (Figure 1(c)). We designed and fabricated two printed circuit boards (PCBs) to host the photodiode and three laser diodes. The photodiode, laser diodes, and LC are connected to an Arduino Due micro-controller. The micro-controller controls the orientation of the LC and light intensity of laser diodes, samples the recorded signal, generates features, and estimates glucose concentration in real time. Figure 1(a) shows a system overview.

We evaluated the prototype by collecting data from a local hospital and a university with 50 participants (31 male and 19 female, 18–81 years old), including 41 diabetic patients and 9 healthy subjects. We summarize our findings as below:

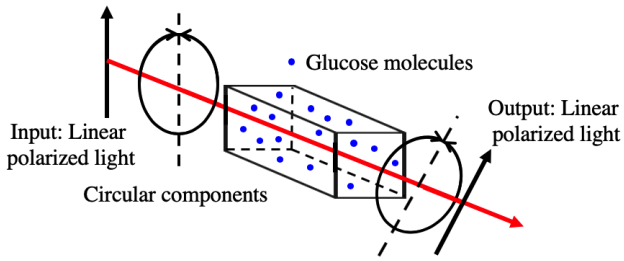


Figure 2: Glucose causes rotation of the plane of light polarization.

- 89% of results are within zone A (i.e., < 20% error) of the Clarke Error Grid [17], a common measure to examine glucose monitoring accuracy. The r and p values of the Pearson correlation coefficients between predicted and reference glucose concentration are 0.91 and 1.6×10^{-143}, respectively. The mean and median absolute relative differences are 10% and 9%, respectively.
- Glucose increasing and decreasing trends are detected with a precision of 88% and 87% and recall of 80% and 81% across all participants, respectively.
- The system is robust under various ambient lighting conditions (i.e., indoor and outdoor), body temperatures (e.g., $33^{\circ}C$ to $40^{\circ}C$), and hand pressure (5 N to 50 N).
- Our prototype received positive feedback from 41 diabetic patients for its noninvasive nature and low cost.

Contributions We make the following key contributions:

- We propose a novel optical polarimetry design capable of extracting weak optical rotation signals related to glucose concentration beneath the skin.
- We propose the concept of using multiple wavelengths and intensity levels to mitigate the impact of confounding factors and strengthen the correlation between observed optical rotation and glucose concentration.
- We build a generic learning model for estimating glucose concentration without the need for calibration. This model is robust across diverse users and more generalizable than personalized models.
- We design and implement a compact and low-cost prototype to estimate glucose concentration in real time using off-the-shelf hardware components.
- We evaluate our system with diverse users with and without diabetes and demonstrate system accuracy comparable to an existing FDA-approved CGM.

2 BACKGROUND AND CHALLENGES

Glucose is a monosaccharide (or simple sugar) with the molecular formula $C_6H_{12}O_6$. More importantly, glucose is an optically active molecule[2] because its molecule has four chiral centers, which gives 2^4 = 16 possible stereoisomers [3]. Based on the optical activity of glucose, polarimetric glucose sensing measures the linear polarization change to estimate glucose concentration. Linearly polarized light can be decomposed into the superposition of left and right circularly polarized light with equal amplitude. In an optically active medium, the velocities of left and right circularly polarized light are different [64]. Figure 2 shows that when linearly polarized light propagates through a medium with glucose molecules, the plane of light polarization at the output can be rotated [32]. In a clear medium, the amount of optical rotation α, which is linearly proportional to the glucose concentration C, can be computed with following equation [71]:

$$\alpha = R(\lambda, T) \cdot C \cdot L, \tag{1}$$

where R is the rotatory power of glucose molecule at light wavelength λ and temperature T, and L is the optical path in the medium that contains glucose.

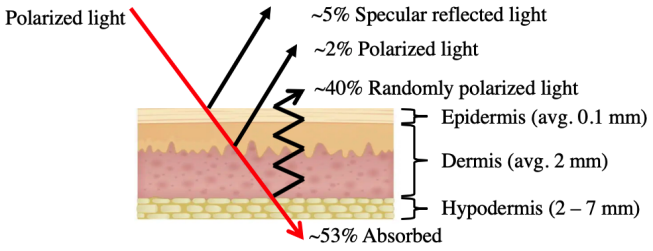


Figure 3: Impact of light absorption and scattering on the skin.

Based on this simple linear relationship between α and C, various polarimetric glucose sensors have been developed [42, 54]. However, existing methods require either extracting sample cells from the body or polarimetric sensing devices that can measure millidegree rotations to extract the weak polarized light. The former is painful and uncomfortable as patients with diabetes are requires to check their blood glucose multiple times (e.g. > 3 times) a day. The latter systems tend to be bulky and expensive and typically require high voltage for operation. Hence, there is still a critical need for a mobile, affordable, and accurate noninvasive glucose monitoring system.

A fundamental challenge is that human skin is highly absorbing and scatters incident light rays, thereby decreasing the signal-to-noise (SNR) of the reflected light and reducing the polarization change caused by the presence of glucose. The outermost layer of the skin is the epidermis (about 0.1 mm thick), which provides a waterproof barrier and creates a skin tone. Beneath the epidermis is the dermis (about 2 mm thick), which contains connective tissues and sweat glands. The deepest layer of the skin is the hypodermis (about 2–7 mm thick), which is made of fat tissue that stores nutrients and energy, insulates the body from cold temperatures, and provides shock absorption. As shown in Figure 3, when linearly polarized light illuminates skin surface, 53% of the light is absorbed by the skin, and roughly 5% is directly reflected from skin surface and therefore does not interact with glucose molecules below the skin [30]. The rest of the light penetrates the skin and may interact with glucose molecules before and while being reflected back. However, the majority (>95%) of such light is depolarized due to scattering from the tissue, which randomizes the orientation of polarization [52]. It leaves approximately only 4% of light retaining its polarization change because of its interaction with glucose molecules. Directly extracting polarized light reflected by the skin is not effective. Based on Mueller matrix decomposition, recent work has applied optical coherence tomography (OCT) to extract linear or circular polarized light in turbid media [27, 29, 39]. However, these methods require strict measurement of optical pathlength, detection depth, and ISF volume underneath the skin. These parameters

[2]Optical activity occurs if the molecules are one of two (or more) stereoisomers [66].

are hard to acquire in real-world settings. Besides, those methods require high-end hardware devices (e.g., high-speed lock-in amplifiers, non-polarizing beam splitters, compensation dispersers), which significantly increase both the size and cost of their systems.

Additionally, confounding molecules, as well as a chiral rotation from tissue birefringence, can also affect the optical rotation, and add noise to glucose estimations. For example, optically active chiral species in tissue (e.g., proteins in epidermis and dermis) and anisotropic tissue structures (e.g., collagen in the dermis) can contribute to the observed optical rotation, hindering access to specific glucose contribution.

Finally, user diversity presents an additional challenge for noninvasive glucose monitoring. Among different users, the optical path L in Eq. (1) is difficult to quantify and measure, thereby making it even more difficult to accurately estimate glucose concentration. Skin thickness and skin color both affect the ratio of absorbed/reflected light. Existing systems often require calibration (e.g., a fingerstick blood sample) to improve the estimation across different users. However, such solutions sacrifice user experience.

3 SYSTEM DESIGN

We address the aforementioned challenges via three design elements described below.

3.1 Depolarization Cancellation

Our first design element is a novel hardware design that exploits the use of LCs to cancel out randomly polarized light and extract weak polarized light associated with glucose concentration. Specifically, we apply different voltages on an LC to rotate the polarization of linearly polarized light and sense the reflected light in two orientations. For polarized light, if the orientation is parallel/perpendicular to a linear polarizing film, all/none of the light rays can pass through the linear polarizing film. For randomly polarized light, half the light rays can pass through a linear polarizing film, regardless of the orientation status of LC. Thus, the difference between these two measurements equals the intensity of polarized light.

Figure 4 shows the schematic of our system, which contains a laser diode, a photodiode, an optical diffuser, two linear polarizing films (crossed polarization), and an LC. After the optical diffuser and the first linear polarizing film (P_1), light emitted from the laser diode becomes linearly polarized light. It then reaches the skin surface after propagating through the LC. To exclude direct reflection from the skin surface, we place the laser diode on an axis that is 45° away from the photodiode's direction, so that glare from the skin surface reflects away from the photodiode. This allows both polarized (I_{pol}) and randomly polarized (I_{ran}) light to reach the second linear polarizing film (P_2) which is perpendicular to the plane of P_1. By changing the voltage on the LC, our system can orient the optical rotation of the incident light to either parallel or perpendicular to the orientation of P_2 and measure the two corresponding reflected light intensity values (I_{par} and I_{per}). These two values can be computed as below:

$$\begin{aligned} I_{per} &= I_{pol} \cdot \cos^2(\alpha + \frac{\pi}{2}) + \frac{I_{ran}}{2} = I_{pol} \cdot \sin^2(\alpha) + \frac{I_{ran}}{2} \\ I_{par} &= I_{pol} \cdot \cos^2(\alpha) + \frac{I_{ran}}{2} \end{aligned} \tag{2}$$

We compute the difference of these two measurements as the feature value $I_{feature}$. Thus we have

$$I_{feature} = I_{per} - I_{par} = -I_{pol} \cdot \cos(2\alpha). \tag{3}$$

Based on Eq. (1), $I_{feature}$ can be expanded as

$$I_{feature} = -I_{pol} \cdot \cos(2R(\lambda, T) \cdot C \cdot L). \tag{4}$$

Therefore, the glucose concentration C can be derived as follows:

$$C = \frac{\arccos(-\frac{I_{feature}}{I_{pol}})}{2R(\lambda, T) \cdot L}. \tag{5}$$

Since $I_{feature} < 0$ and $-I_{feature} < I_{pol}$, there is a positive correlation between C and $I_{feature}$. In Sec. 3.3, we will describe estimating C using $I_{feature}$.

Experimental validation We validate the above rationale using a bench-top experimental setup shown in Figure 5. The setup consists of a 658nm (red) laser diode (~ $60), a 50° circle pattern diffuser (~ $50), two linear polarizing films (~ $1), a low-cost LC (~ $1), and an OPT101 photodiode (~ $6). The hardware components were aligned properly on an optical breadboard to match the schematic of our system in Figure 4. We then evaluated the linear relationship between $I_{feature}$ and the glucose concentration.

In the first experiment, a male user placed his right hand (facing the photodiode) at the fixed location on the prototype for about 8 hours, and the system continuously sampled $I_{feature}$ every 5–15 minutes before and after the user drank 500 mL of sugar water and had some food. Ground truth data was collected using a FDA-approved Freestyle Libre CGM sensor [2] which also measures glucose concentration from ISF. The experiment was conducted in a dark room to minimize the impact of ambient light.

Figure 6 shows the correlation between the $I_{feature}$ and the glucose concentration (mg/dl). The black box shows the time window that the user drank sugar water and had a meal. The r and p values of the Pearson correlation coefficients are 0.86 and 0.0000012, respectively. The result demonstrates the correlation between $I_{feature}$ and the glucose concentration, and it aligns with the linear relationship in Eq. (4). Therefore, it is possible to estimate glucose concentration simply by measuring the change of $I_{feature}$.

The experiment was repeated on eight more users (6 males and 2 females, 18 to 33 years old), and we observed similar linear relationships across these users. However, the r value of the Pearson correlation coefficients can drop from 0.86 to 0.48 (last column in Table 1) due to various skin properties (e.g, color, thickness) and confounding factors. A low r value of the Pearson correlation coefficients means that the feature leads to a low signal-to-noise ratio (SNR), which can degrade the sensing performance significantly. This observation motivates our next two design elements.

3.2 Exploiting Multiple Wavelengths & Light Intensity Levels

To boost our system's sensing robustness under various confounding factors and user diversity, our second design element exploits the diversity in light wavelengths and intensity levels. Specifically, it scans different skin layers using multiple wavelengths and intensity levels to strengthen the correlation between $I_{feature}$ and the glucose concentration in ISF. This sensing scheme has two benefits.

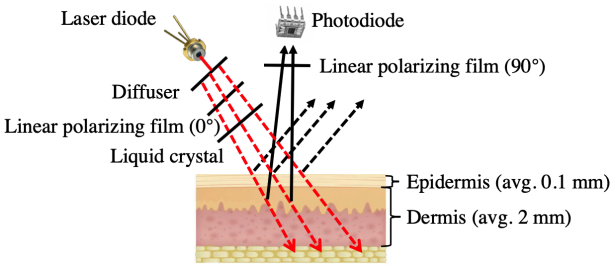


Figure 4: Schematic of our system. It leverages one liquid-crystal shutter and linearly polarizing film with proper alignment to extract weak polarized light reflected by ISF below the skin.

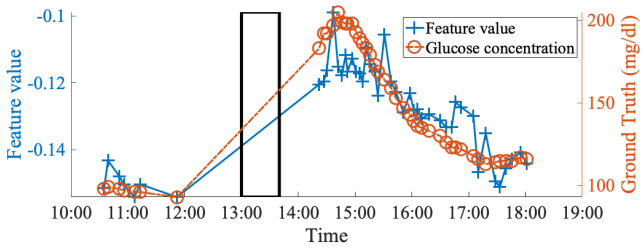


Figure 6: Linear relationship between our feature and the glucose concentration. The black box shows the time window that the user drank sugar water and had a meal.

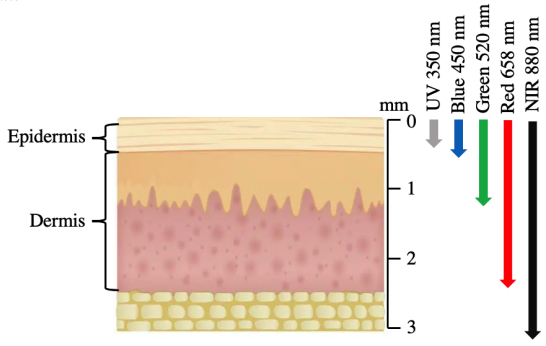


Figure 7: Effect of wavelength on penetration depth in skin tissue.

First, it can search skin layers to minimize the impact of confounding factors because ISF and confounding factors are not uniformly distributed underneath the skin [34]. Second, since wavelengths contribute to the optical rotation per Eq. (1), our system can generate multiple feature values at specific glucose concentration, which can further improve the sensing granularity.

Our system leverages M wavelengths within the visible light spectrum (400–700 nm) and N intensity levels (30–120 mW) to scan from 0.5 mm to 2 mm below the skin surface. In our implementation, $M = 3$ and $N = 7$. We select these wavelengths and intensity levels for two reasons. First, the wavelengths are safe for users. Below 400 nm, continuously illuminated by ultraviolet radiation can cause diseases like erythema, skin aging, and skin cancer [5]. Above 700 nm, infrared radiation can create heat on the skin and potentially increase skin temperature [16]. In our system, we select three candidates (450 nm, 520 nm, and 658 nm) within the visible light spectrum (400–700 nm) for continuous glucose sensing. Second, the intensity levels are appropriate for glucose sensing. ISF is mainly localized in the two outer layers of the skin, namely, epidermis and dermis. The volume of ISF in the epidermis and the

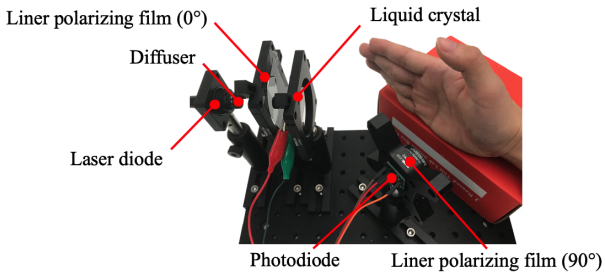


Figure 5: Experimental setup for validating the proposed depolarization cancellation methodology.

Age	Gender	Racial/ethnic category	Highest r value out of 21 features	r value (single wavelength)
31	Male	Asian	0.86 (658 nm, intensity level=5)	0.86
23	Male	Asian	0.83 (658 nm, intensity level=3)	0.64
18	Male	Asian	0.78 (658 nm, intensity level=3)	0.6
24	Male	Asian	0.81 (658 nm, intensity level=4)	0.63
23	Female	Asian	0.74 (520 nm, intensity level=4)	0.65
23	Male	White	0.75 (520 nm, intensity level=3)	0.71
25	Male	African American	0.78 (450 nm, intensity level=6)	0.46
24	Male	American Indian	0.74 (450 nm, intensity level=6)	0.51
33	Female	Asian	0.75 (520 nm, intensity level=4)	0.48

Table 1: Participant information and the r values of the Pearson correlation coefficients using multiple light and intensity levels and a single wavelength.

dermis are 15%–35% and 35%–45%, respectively [28]. The deeper layer (hypodermis) is made of fat and connective tissues (confounding factor), and the volume of ISF is less than 20%. Therefore, the dermis has higher ISF content and a lower ratio of confounding factors beneath the skin surface, therefore, we propose that it is ideal for glucose sensing. Figure 7 shows that visible light penetrates typical tissues to a depth of 0.5–3 mm in the skin [47]. We then leverage a Monte Carlo simulation, which is one of the most widely used stochastic methods for modeling light transport in human skin, to control the penetration depth of these light rays from 0.5–2 mm [6]. The penetration depth of these light rays can vary from 0.2-2.5 mm for different users/skin types (e.g., tone, the thickness of each layer). Therefore, our system was designed to scan layers in the dermis to improve the correlation between $I_{feature}$ and glucose concentration.

To validate this sensing approach, we designed and implemented a prototype, which is described in Sec. 4. We then asked nine participants (7 males and 2 females) to repeat the experiment in Sec. 3.1 using the new prototype. For each measurement, we scan their skin $M \cdot N \cdot 2$ times using M wavelengths and N intensity levels, which generates $M \cdot N$ features per Eq. (4). We then compute the r values of the Pearson correlation coefficients between the features and the glucose concentration.

Table 1 shows a summary of the highest r value of the Pearson correlation coefficients out of M by N features. Compared with the system that uses a single wavelength and intensity level (the last column in Table 1), the new design strengthens the correlation between $I_{feature}$ and the glucose concentration across users and significantly reduces the impact of confounding factors (the second last column in Table 1). We have two observations based on the experimental results. First, the light intensity should fit for the age. Skin thickness increases between 0 and 30 years of old and no significant variation in thickness between 30 and 65 years of

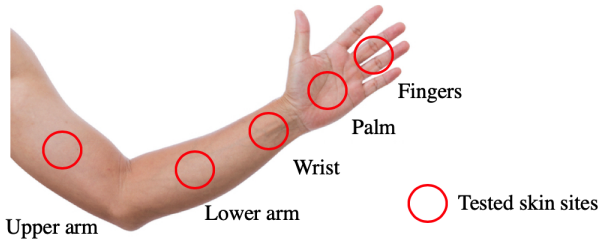


Figure 8: The five different measurement locations on the human arm.

old [69]. Thicker skin requires higher light intensity so that light rays can penetrate deeply underneath the skin to interact with more glucose molecules. However, overexposure can saturate the light sensor, which even degrades the sensing granularity. Second, people of different race/ethnic categories should select the proper light wavelength for the best SNR. Racial/ethnic categories decide skin color, which affects the light penetration depth underneath the skin. Most lightly pigmented skin types have approximately half as much epidermal melanin as the most darkly pigmented skin types [4, 57]. Since epidermal melanin absorbs the full spectrum range of visible light, the SNR of reflected light from dark skin is lower than that from light skin. Therefore, a higher light intensity is needed for dark skin users.

Figure 9 demonstrates the impact of wavelength on the r value of the Pearson correlation coefficients between the feature value ($I_{feature}$) and the glucose concentration for different racial/ethnic categories. The red lines represent the fitted linear models using linear regression. We observe that red (658 nm) or green (520 nm) light works well on Asian or African Americans (Figure 9(a)), while blue (450 nm) and red (658n m) light work well on White people and African Americans, respectively (Figure 9(c) and 9(b)). The epidermis thickness of African Americans is significantly higher compared with White people [56]. Therefore, for African Americans, blue light cannot penetrate deep enough to interact with glucose molecules within the dermis layers, introducing a low correlation between sensor data and glucose concentration. For White people, red light may penetrate too deep and interact with blood cells, introducing high noise due to confounding factors.

Identification of optimal skin sites We evaluated various skin sites for noninvasive glucose monitoring. As shown in Figure 8, we repeated the experiment above on five candidate sites (fingers, palm, wrist, lower arm, and upper arm). On each skin site, we collected three-hour of sensor data with ground truth before and after the user drank 500 ml sugar water. Although other sites like lips and ear may be relevant, we did not evaluate these sites due to the form factor of our prototype and potential discomfort to participants. Table 2 shows that both the palm and wrist are the most suitable sites for glucose monitoring using our prototype. The fingers were not considered further because the sensing area on the fingers are smaller than other sites. Lower arm and upper arm showed to be bad sites for our approach and system, we suspect this is due to the presence of a thick stratum corneum or other confounding factors.

3.3 Estimating Glucose Concentration

The last design element is to generate and select features and then leverage the features to estimate glucose concentration. To address user diversity, existing methods (e.g., Medtronic Guardian) require additional calibrations (e.g., a fingerstick blood sample) to refine their estimation, which sacrifices the user experience. We solve the problem with supervised learning to train an offline generic model mapping the sensing data to the glucose concentration. Specifically, we select gradient boosting decision trees, which capture the relationship between features and predicted values as a set of regression trees. This approach is beneficial for two reasons. First, the computation overhead of offline training and online prediction is low compared with feed-forward neural networks and SVM. Second, the accuracy can be high for low-dimensional sets of features and small datasets. After training the generic model, we use this model in our system to predict glucose concentration in real-time without need for calibration.

	Fingers	Palm	Wrist	Lower arm	Upper arm
Highest r value out of 21 features	0.46	0.8	0.74	0.12	0.09

Table 2: Comparison of r values of the Pearson correlation coefficients between the feature value ($I_{feature}$) and the glucose concentration on five skin sites.

Feature generation Our system generates two kinds of features (total = 27 features). Specifically, we extract $M \cdot N$ (i.e., 21 features based on our implementation) sensing features using Eq. (??). To reduce the impact of electrical noise (e.g., spike noise due to frequently turning the laser diodes on and off), we collect 1-min sensing data and calculate the median of this window to compute sensing features. Then, we include demographic factors like age, gender, race, skin tone, type of diabetes, and model of CGM as the features in our learning model to account for user diversity and variances among existing CGM devices. Since the mean absolute relative error of CGM devices varies from 6–12% [58], the model of CGM is only used to calibrate the labels in the learning model.

Feature selection We leverage boosted trees to select relevant features in the learning model [75]. A benefit of using boosted trees is that after the trees are constructed, it is straightforward to acquire the importance scores of each feature. For a single tree, we can compute the importance score of each feature by the amount that the feature improves the performance measure. We then average the score across all of the trees to compute the weighted score. The more a feature is used to make significant decisions within the model, the higher the weighted score becomes. In the experiment in Section 3.1, we observe that more than half of the sensing features do not have strong correlations with glucose concentration. Thus, we use the weighted score to rank the 21 sensing features and empirically select the top third of all sensing features, and then add the six demographic features to train a generic model and estimate glucose concentration.

Offline training We apply an Xgboost library [26] to construct boosted trees. To train a generic model, we collect data from 41 diabetes patients (18–81 years old) in a local hospital and 9 healthy participants (18–33 years old) in a university. Table 3 details the participants' information. Although users are relatively diverse, the distribution of their glucose concentration is highly non-uniform in the dataset. Most of the glucose concentration is within the normal range (70–140 mg/dl), and less than 22% of glucose concentration is high glucose (> 140 mg/dl). To handle the imbalanced dataset, we apply a resampling ensemble algorithm to train n (n = 5 in our

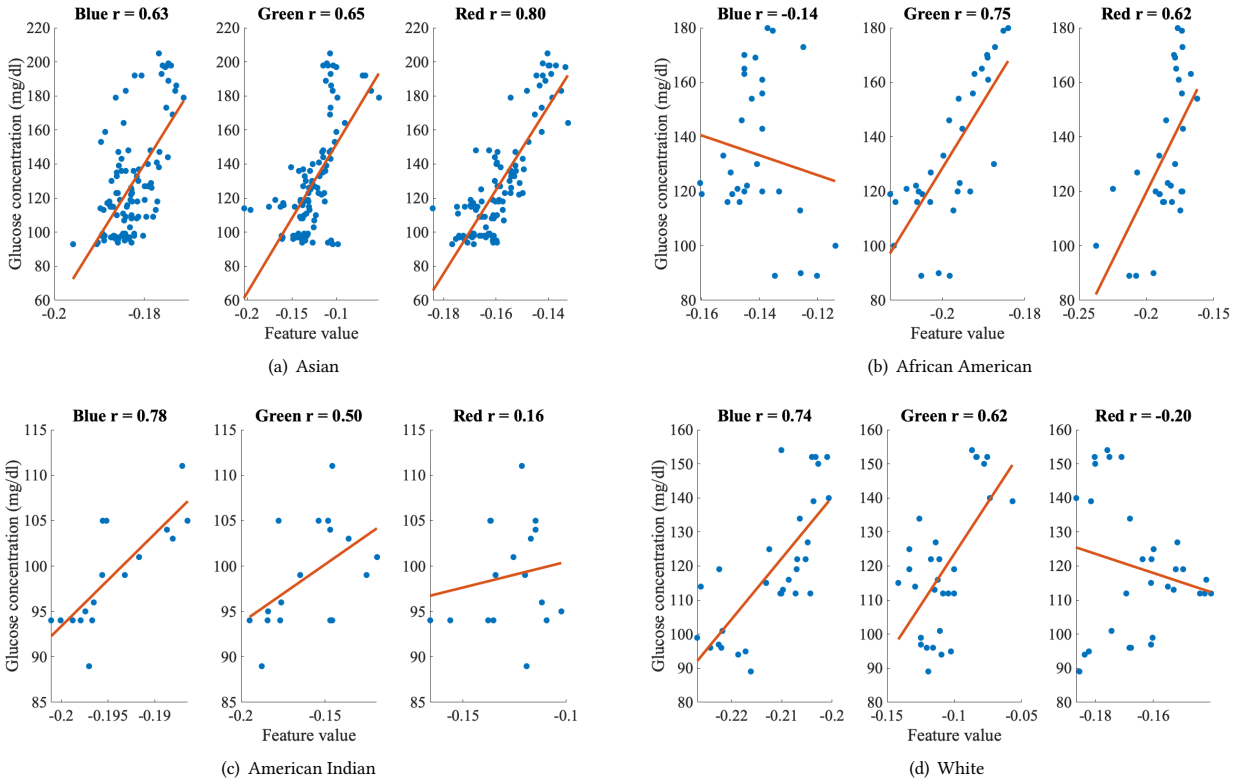


Figure 9: Impact of wavelength on the r value of the Pearson correlation coefficients for different racial/ethnic categories. The red lines represent the fitted linear model using linear regression.

implement) tree ensembles [55]. Let D denotes the training set that has X pairs of feature vector f and glucose concentration C. We split D into two datasets, namely, the samples that have high glucose concentration (D_{high}) and samples that have normal glucose concentration (D_{normal}). As shown in Figure 10, each ensemble is trained using all of the Y samples in D_{high} and $\frac{X-Y}{n}$ random samples in D_{normal}. Then, we leverage linear regression with additional training samples to linearly combine the n trained tree ensembles. Let C_i denotes the predicted glucose concentration from the i^{th} tree ensemble. The final estimation of glucose concentration C is derived by $C = \sum_{i=1}^{n} a_i \cdot C_i$. By using multiple tree ensembles and balanced datasets, the learning model effectively maps feature vector across a complete range of glucose concentration.

Online inference Users first need to input their age, gender, race, skin tone, types of diabetes to the system. The model of CGM is also used to calibrate the labels. No additional calibration is needed for online inference. As sensing data arrives on the fly, our system computes feature vectors and then leverage trained trees ensembles to select features and estimate glucose concentration. Each tree ensemble will generate a prediction score, and the linear combination of all scores is the final estimation of the glucose concentration.

4 PROTOTYPE

We designed and implemented a compact (17 cm × 10 cm × 5 cm) and low-cost (e.g., <$250) prototype (Figure 1) using off-the-shelf hardware components. The prototype consists of three main

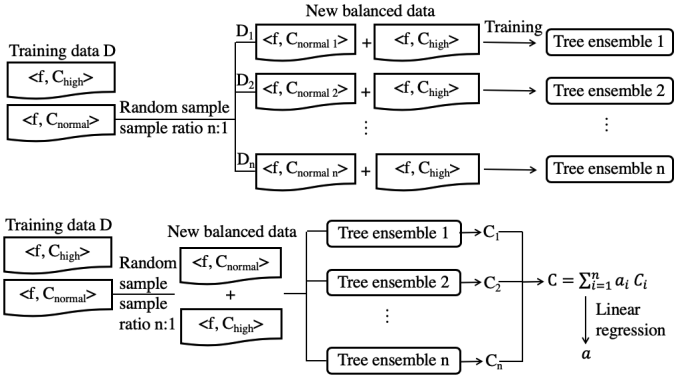
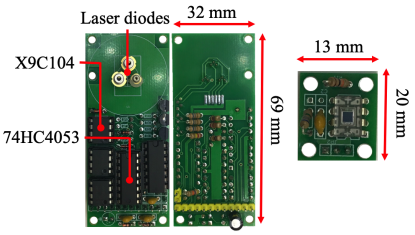


Figure 10: Dealing with imbalanced dataset, where n balanced datasets are generated by randomly sampling abundant dataset (D_{normal}) and samples of rare dataset (D_{high}). Then, the learning model leverages the n balanced datasets to train n tree ensembles. The linear combination of the n tree ensemble is derived by linear regression.

components, namely a sensing box, an optical sensing unit, and a computation unit.

Sensing box We cut acrylic plastic sheets (0.8 mm thickness) using a laser cutter to build the sensing box, which hosts the optical sensing unit and computation unit (Figure 1(b)). For debugging and maintenance purposes, the sensing box can be opened on the side (Figure 1(c)). Except for the sensing area (4.7 cm × 6.2 cm), we stick two layers of black paper on the inner side of the sensing box to block ambient light. We also place a small speaker inside



(a) Transmitter (front and back) (b) Receiver

Figure 11: Circuit boards for the optical sensing unit.

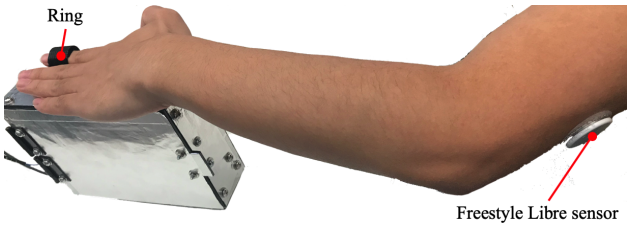


Figure 12: Experimental setup. A participant wears a ring and places the palm on the sensing area of our prototype. We leverage the readings from existing CGM sensors (e.g., Freestyle Libre, Dexcom G6, or Medtronic Guardian) as the ground truth.

the sensing box. Once the receiver detects ambient light due to the displacement of the palm, the speaker will play a sound to notify the user. Since users need to place their palms on top of the sensing box, we stick a layer of tin foil sheet on the outer side of the sensing box to share the ground between the sensing box and the palm.

Optical sensing unit We designed and fabricated two PCB boards (Figure 11) to host and control three (red, green, and blue) laser diodes as the transmitters and one photodiode (OPT101) as the receiver. On the receiver end, we increase the sensitivity of the photodiode by customizing DC gain and bandwidth to 8×10^6 V/A and 2 kHz, respectively (Figure 11(b)). The distance between the photodiode and the skin surface is 5 cm, which is the closest distance to the hand while not capturing any emitting light rays leaked from the laser diodes. In front of the photodiode, a linear polarizing film (90 ° polarization) is hosted on an optic mounting holder. On the transmitter end, we leverage an analog multiplexer (74HC4053) as switches (Figure 11(a)) to control the three laser diodes (at 450 nm, 520 nm, and 658 nm). For each laser diode, we leverage a digitally controlled potentiometer (X9C104) and a transistor (S9014) to control the power output from 30 mW to 120 mW (seven intensity levels). The pulse of the laser diode at each intensity level is about 250 μs. We then place the laser diodes at the axis 45 ° away from the photodiode axis to exclude specular reflection from the skin surface. Due to the size of our LC, the distance between the laser diodes and the skin surface is 7 cm to place the full transmitter board inside the sensing box. In front of the laser diodes, we place another linear polarizing film (0 ° polarization), a 50 ° circle pattern diffuser, and an LC shutter on an optic mounting holder. Applying a constant voltage on the LC can cause DC bias and damage the LC. To solve the problem, we leverage an AC wave between the two pins of the LC shutter. The two output pins are driven to opposite polarity from each other, which creates a true AC waveform with the Root Mean Square (RMS) voltage equal to the supply voltage. The frequency of the AC waveform is 2 kHz.

Computation unit We use a micro-controller (Arduino Due) to power the optical sensing unit, control laser diodes and LC shutter, digitize analog signal from the photodiode, extract features, and estimate glucose concentration in real time. Inference results are stored in the micro-controller. They can also be transmitted to other devices through a USB cable. A ring with a layer of tin foil is connected to the ground of the micro-controller. Thus, when the user wears the ring and places the palm on the sensing box, the micro-controller, the palm, and the sensing box share the ground.

5 PROTOTYPE EXPERIMENTS

We conducted a three-month study in a local hospital and a university to evaluate our prototype.[3]

5.1 Study Setup

In our study, we enrolled 50 participants (18–81 years old), including 41 patients with diabetes (32 Type 1 diabetes and 9 Type 2 diabetes) and 9 healthy people. Table 3 summarizes participant information. We recruited more patients than healthy participants because diabetic patients are likely the major user group of the system. Furthermore, the glucose level of diabetic patients has a wider range and fluctuates more than that of healthy people[24] and thus it presents a more challenging scenario for glucose monitoring. We leverage FDA-approved CGM sensors (e.g., Freestyle Libre, Dexcom G6, and Medtronic Guardian) to collect the ground truth data since these sensors also measure glucose concentration in ISF. Figure 12 shows the experiment setup, where a participant wears a ring and places the palm on the sensing area. During the experiment, a participant sits on a chair with a relaxing posture. We collect data from each participant for roughly one to two minutes, which leads to approximately 300 measurements collected by the prototype. We compute the median of these measurements to compute feature values. If the palm did not cover the whole sensing area or the photodiode inside the sensing box detected ambient light, the speaker would notify the participant by playing a sound.

Data collection in a hospital We recruited 41 patients with diabetes who wore CGM sensors at Dartmouth-Hitchcock Medical Center (DHMC). Participants were not included if they were pregnant or planning pregnancy, or were considered by the doctors to be unsuitable to participate. Before and after an appointment with doctors,[4] we collected data from the patients, including the sensing data from our prototype, the readings from patients' CGM sensors, and some survey questions. On average, we collect three data points from each patients. About 60% of patients had meals or drinks before visiting the doctor. Before collecting data from patients, we cleaned our prototype using an alcohol pad and asked patients to wash hands with hand sanitizers. This cleaning protocol was required by the hospital to ensure that the device was cleaned between participants. For patients with diabetes, 69% of data is in

[3]We obtained the IRB approval to conduct the study in both institutions.
[4]Each doctor's appointment takes about 30–60 mins

	Gender		Racial and ethnic categories						Age			CGM devices			Diabetic patients		Healthy
	M	F	AI	A	AA	H/L	NH	W	18–30	30–50	>50	Freestyle Libre	Dexcom G5/G6	Guardian	Type 1	Type 2	participants
# of Users	31	19	2	14	4	3	3	24	16	15	19	37	10	3	32	9	9

Table 3: Participant information. AI: American Indian, A: Asian, AA: African American, H/L: Hispanic or Latino, NH: Native Hawaiian or Other Pacific Islander, W: White.

the normal glucose range (70–140 mg/dl) and 31% of data is in the high glucose range (>140 mg/dl).

Data collection in a university We enrolled one undergraduate student, eight graduate students, and one professor at Dartmouth College. Since all of the participants are healthy subjects, we gave them new Freestyle Libre sensors, which leverage adhesive patches to attach to the skin surface, to collect the ground truth. Participants were not included in the study if they had known allergies to medical-grade adhesives, pregnant, or planning pregnancy. We collect the sensing data from our device and readings from Freestyle Libre sensors under two glucose conditions, namely before and after a meal or a drink of 500 ml sugared water (e.g., Coke). The interval between two adjacent samples is 30–60 mins. To balance the training set, we collect at least 12 data points from each participants. No cleaning process was required in the study. For healthy participants, 95% of data is in the normal glucose range, 5% of data is in the high glucose range.

Data presentation and error metric Researchers and clinicians use Clarke error grid analysis (EGA) to assess the clinical accuracy of glucose monitor systems. As shown in Figure 13, EGA shows reference glucose values (ground truth) and the predicted glucose concentration on the x-axis and the y-axis, respectively. EGA contains five main zones, namely Zone A, B, C, D, and E. Zone A are those predicted values, which diverge as of the reference values by 20% or less, and is considered to be clinically accurate. Zone B are those estimated values differ as of the reference values by more than 20%, but are still clinically acceptable. However, Zone C leads to unnecessary treatment; Zones D fails to detect hypoglycemia or hyperglycemia; Zone E represents the erroneous results [17]. Therefore, researchers and clinicians look at two error metrics to quantify the performance of a glucose monitoring system: (a) the percentage of predicted glucose level which fall in Zone A or B and (b) the Pearson correlation coefficients between the predicted and reference glucose concentration [49]. In the experiment, we also look at the absolute relative differences (ARD) computed as below as an additional error metric to evaluate our system:

$$\text{ARD} = \frac{|\text{predicted glucose} - \text{reference glucose}|}{\text{reference glucose}}. \quad (6)$$

A predicted glucose concentration is considered as clinical accurate if the ARD is less 20%[49].

5.2 Overall Performance

We begin with examining the overall performance of our system with 50 participants (details in Table 3) by applying leave-one-participant-out cross-validation. Each time we train a model, we leave one user's data out as testing data and then rotate the dataset. Figure 13 shows the EGA. All predicted glucose concentration is in either of Zone A or Zone B. 89% of the predicted data is clinically accurate (Zone A), and 11% of the predicted data is clinically acceptable (Zone B). The r and p values of the Pearson correlation coefficients are 0.91 and 1.6×10^{-143}, respectively. The mean and median of ARD are 10% and 9%, respectively. Compared with recent work that leveraged optical coherence tomography to extract optical rotation [13], our system is low-cost and compact, and the system requires no calibration. Compared with the existing FDA approved CGM sensors (e.g., Freestyle Libre, Dexcom G6, or Medtronic Guardian), our system has no needles and does not require sensor replacement. We also evaluated system performance using a Support Vector Machine (SVM) [11] and a 10-layer feed-forward neural network in Matlab. The mean of ARD are 13%, and 21%, respectively. These models do not outperform boosted trees because our dataset is relatively small for neural networks. Running a large number of iteration can result in overfitting, which significantly degrades the performance. We plan to collect more data and re-evaluate these models in future work.

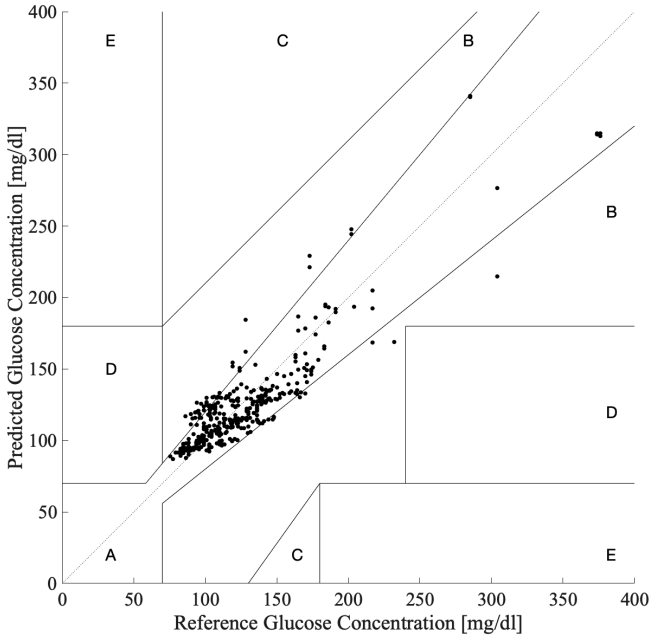


Figure 13: Clarke error grid analysis. All of the our predicted glucose levels are in either of Zone A and Zone B. 89% of the predicted data is clinical accurate (Zone A), and 11% the predicted data is clinical acceptable (Zone B).

Figure 14 shows the CDF of the ARD for overall, high, and normal glucose samples. The system can measure glucose concentration with 89th percentile error less than 20% (clinical accurate), regardless of its glucose concentration. Since the distribution of glucose concentration is highly non-uniform in our dataset, the results demonstrate the robustness of our resampling ensemble algorithm.

The trends of glucose concentration is a health indicator and correlate with physical activity. For example, in Type 1 diabetes, glucose trends can be used to monitor the risk of diseases like

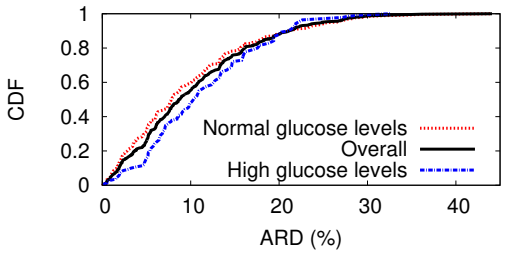


Figure 14: CDF of ARD for all, high, and normal glucose concentration sample. Regardless of glucose concentration, 89% of predicted glucose levels are clinical accurate, which demonstrates the robustness of our resampling ensemble algorithm.

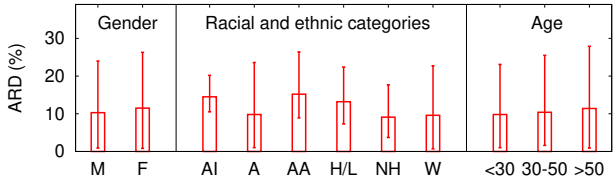


Figure 15: Impact of user diversity. AI: American Indian, A: Asian, AA: African American, H/L: Hispanic or Latino, NH: Native Hawaiian or Other Pacific Islander, W: White.

cardiovascular complications [74]. In the experiment, we define that an "increasing" or "decreasing" trend happens if the glucose concentration increases or decreases over 10% in 60 mins. Table 4 summarizes the results, where 77 "increasing trends" and 68 "decreasing trends" were collected from 50 participants in the study. Therefore the performance of our system is sufficient for home care monitoring devices and medicare applications.

	Precision	Recall
Increase trend	0.88	0.80
Decrease trend	0.87	0.81

Table 4: Precision and recall of detecting "increasing" and "decreasing" trends from 50 participants.

Impact of user diversity Figure 15 shows the mean ARD and 90% confidence intervals for different genders, race and ethic categories, and age groups. Overall, the mean ARD is similar across diverse participants. We make following observations. First, due to the demographic distribution in the local area, there are fewer participants from certain race and ethnic categories (e.g., AI, AA, and H/L, see Table 3). Limited training data degrades the sensing accuracy for these participants. For example, the largest error rates (e.g., ARD > 30%) come from either AA or AI group, which only have 4 and 2 participants, respectively. In future work, we plan to recruit more diverse subjects and collect more representative training data to mitigate this problem. Second, we observe some large errors come from participants with palm smaller than our sensing area. In this scenario, since users cannot place their palm at the same location on the device, their hand placement offset introduces noise to the system. To address this problem, we plan to reduce prototype size with more compact hardware components.

5.3 Generic Model v.s. Personalized Model

We conducted a one-week experiment to demonstrate the effectiveness of our generic model over the personalized model. In the experiment, we collected data from two healthy participant every 30 mins before and after lunch/dinner. The new dataset contains 205 samples. Then, we leverage a different number of samples to train personalized models. And the rest of the dataset is used as the testing data for both personalized and generic models. Figure 16 shows that our generic model is better than personalized models if the training data of personalized models is less than half of the dataset. Once we build the generic model, the system does not require sampling data (calibration) from users. Therefore, the generic model significantly improves the system's usability and robustness. In future work, we plan to recruit patients to further demonstrate the effectiveness of our generic model over the personalized model.

5.4 Practical Considerations

Wavelengths and light intensity levels We downsampled the features to evaluate the impact of wavelengths and light intensity levels. Figure 17(a) and 17(b) plots the mean accuracy using different wavelengths and light intensity levels. We also include error bars covering 90% confidence intervals. In Figure 17(a), we observe that the combination of red, green, and blue light outperforms any of the individual wavelengths. In future work, we plan to add more wavelengths within the visible light spectrum to further explore the impact of wavelengths. In Figure 17(b), since the performance converges with more than six light intensity levels, the seven light intensity levels are optimal in our system.

Ambient light Our prototype can block the ambient light by covering two layers of black paper except for the sensing area. In the experiment, we turned all laser diodes off and collected the photodiode reading under three ambient light conditions, namely, 0 lux (dark room), 600 lux (indoor), and 5000 lux (outdoor). For all ambient light conditions, the photodiode readings were zero if the user properly placed the palm on the sensing area. For inappropriate placement of the palm, our system will notify the user by playing sounds. We also built a wristband to evaluate the impact of ambient light under wearable scenarios. A photodiode was placed in the center of the wristband facing the skin's surface. In this scenario, the photodiode readings were still zero in all ambient light conditions.

Pressure and temperature We attached a 10 cm^2 force sensing resistor and a non-contact infrared temperature sensor to the prototype to evaluate the impact of pressure and temperature on a healthy subject. The subject was asked to gradually increase the pressure on the sensing area from 5 to 50 N. To vary the temperature of the palm, the subject held either of a cup of hot or ice water before placing the palm on the sensing area. By changing the duration of holding the hot/ice cup, the temperature of the palm varied from 33°C and 40°C. Figure 18 summarizes the results. We observed that neither pressure nor temperature introduced a visible impact on the analog signal. Specifically, the difference in the analog signal change between 5 N and 50 N pressure is less than 1%, and the difference between 33°C and 40°C is less than 0.8%. Both differences are much lower than electrical noise (e.g., spike noise).

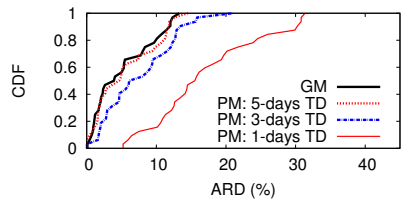
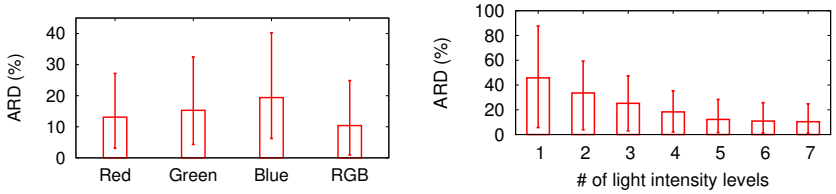
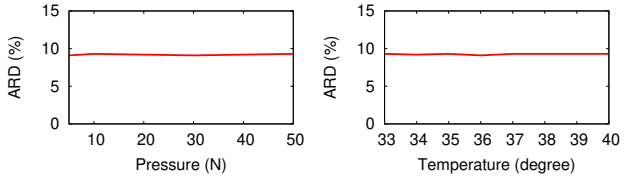


Figure 16: Generic model (GM) v.s. Personalized model (PM) with training data (TD) in different size.



(a) Wavelengths (b) Intensity levels

Figure 17: The impact of wavelengths and light intensity levels to our system.



(a) Impact of pressure (b) Impact of temperature

Figure 18: Impact of pressure and temperature

5.5 User Experience Study

We conducted a user experience study with 41 patients with diabetes who had worn CGM devices in a hospital. In the study, we asked the participants to list the pros and cons of their sensors and ours. Except for one patient, all participants agreed that the existing CGM sensors were much more convenient than the typical finger-prick tests, and the sensors provided good accuracy to help them monitor and manage diabetes. However, a third of the patients complained about the unpleasant wearing experience of the sensors since each sensor still had invasive needles. All of the participants disliked the requirement of calibrating some sensors (e.g., Dexcom G5 and Medtronic Guardian). Four participants reported that their sensors suffered from significant (more 20%) accuracy drops without calibrations. A third of the patients complained about the high price of the sensors ($60 to $400 for each sensor) since these sensors need to be replaced every 7 to 14 days. Compared with the existing CGM sensors, all of the patients liked the low-cost, no calibration, and the noninvasive benefits of our system. However, six patients complained about the size of our prototype, which might be heavy for them to carry around. In Section 6, we will discuss potential miniaturization of the system.

6 DISCUSSION

We discuss the limitations of our study, the lessons we learned from our study and implementation, and plans for future work.

Ground truth To acquire the ground truth of glucose concentration, we leverage existing FDA-approved CGM devices (e.g., Freestyle Libre, Dexcom), which measure glucose concentration in ISF rather than directly in the bloodstream. Since glucose moves from blood vessels and capillaries first and then into ISF, there is a 5 to 10 minutes delay in ISF glucose response to changes in blood glucose [8]. Recent work shows that the translation between ISF and blood glucose concentration is not just a shift in time [19]. Also, in our dataset, few patients have Type 2 diabetes because patients with Type 2 diabetes take insulin less than 3 times a day and therefore would not meet criteria per their insurance company to qualify for using a CGM sensor. Also, the mean absolute relative error of CGM devices varies from 6–12%, higher than invasive methods like finger-prick tests [58]. Therefore, we will consider using finger-prick tests to acquire more accurate blood glucose concentration as the ground truth and recruiting more patients with Type 2 diabetes.

Accuracy improvement The system's current glucose monitoring accuracy is constrained by three factors. First, the training data set is still relatively small. The number of patients limits the performance of our prototype. In our study, we only collected data from patients who already had worn CGM devices and brought their readers to the hospital because their CGM sensors can only be read from their own readers. This requirement limits the number of participants in our study. Second, due to the demographic distribution in the local area, few patients with diabetes had dark skin, which lowers the sensing performance of our prototype. Moving forward, we plan to collect data from a more diverse set of participants over a longer term. We also plan to conduct a long-term user study includes a quantitative survey carried on both patients and doctors and investigate potential issues when the system is used over a long period. Finally, our current learning model is based on a coarse discretization of skin color (e.g., white, black, yellow) and yet human skin tone is at a continuous scale. We plan to leverage the Fitzpatrick scale [61] to investigate the impact of skin tone and consider a real-number representation of skin tone as the feature for the learning model.

User activities Participants reported to us that activities (e.g., running) could significantly affect the accuracy of their CGM sensors. Due to the limitations of our study, we only evaluate our system when users sit on a chair without doing any activities. In our future work, we plan to design a separate study to evaluate the performance of the prototype under various user activities.

Cost and size Compared with commercial CGM sensors, our system does not require sensor replacement, much lower usage cost in long-term usage. Also, it is possible to significantly reduce the total cost of our prototype from $250 to less than $50. The main cost of our prototype comes from three laser diodes ($150) and the engineered diffuser ($50). We can replace laser diodes and the engineered diffuser with LEDs (< $3) and a regular diffuser (< $10), respectively. However, the downside is that the light intensity distribution on the palm will be highly non-uniform (e.g., Gaussian-like distribution). Thus, it is difficult to control the light penetration depth at specific wavelengths within the sensing area. To address

the problem, we can leverage a narrowband optical filter to filter out the irrelevant spectrum. In addition, the size of our current prototype limits the body locations we can test. We will explore miniaturization of the system. Current bottleneck of miniaturization is the size of LC. With smaller LCs available, the whole system can be more compact, possibly integrated with wearable devices to examine a wider set of body locations.

7 RELATED WORK

In this section, we summarize the main technologies used for glucose monitoring, including invasive, minimally invasive, and non-invasive techniques.

Invasive techniques Hospitals employ invasive electrochemical biosensors to measure blood glucose concentration. These biosensors leverage automatic lancet devices to prick the finger for extracting blood samples. By leveraging Faraday rotators [10], electro-optic modulator [46], or liquid crystal [41], existing methods can modulate the polarization state of the light beam to infer the glucose concentration from sample blood. Although recent work has made efforts to reduced blood sample requirement to less than 1 μL [68], these methods are still painful since patients with diabetes have to frequently measure (> 3 times) blood glucose every day. Our work differs in that our system is noninvasive.

Minimally invasive techniques Minimally invasive approaches leverage subcutaneous sensors to measure glucose concentration in body fluids other than blood, e.g., sweat, saliva, and interstitial fluid. These methods, known as continuous glucose monitoring (CGM), can monitor glucose continuously and automatically. CGM sensors leverage tiny electrode (invasive needles), which functionalised with an enzyme film using an electropolymerisation method [62], to measure glucose concentration in ISF. The overall measurement error is approximately ±10% [60]. However, CGM sensors have needles, and they require periodic replacement of sensors. Some CGM sensors (Medtronic Guardian and Dexcom G5) require calibration (e.g., fingerstick blood samples) for optimal sensor accuracy. Our work differs in that our system has no needles, and it requires neither of periodic sensors replacement nor calibration.

Non-invasive techniques The technologies used for noninvasive glucose monitoring include optical, transdermal, and thermal techniques. Comprehensive reviews of recent noninvasive glucose monitoring techniques can be found in [31, 40, 49, 52, 67].

Polarimetry and spectroscopy are the two typical optical techniques for non-invasive glucose monitoring. Polarimetry is based on the rotation of the linear polarization vector of light by the glucose concentration [13, 46, 54]. However, existing methods measure glucose concentration from transparent tissue in the body (i.e., eyes) [46, 54], which can cause photo-thermal damage to eyes [22]. A recent method designs an optical coherence tomography to measure glucose concentration on fingertips [13]. However, their systems requires expensive and bulky optical devices to measure small optical rotation, and the sensing performance suffers from interference of confounding optical rotations. Spectroscopy is the study of objects based on their wavelength spectrum when they emit or absorb light [42]. Four spectroscopic techniques that have been applied to non-invasive glucose monitoring: near-infrared spectroscopy [37, 73], mid-infrared spectroscopy [12, 35, 36, 38], Raman spectroscopy [23, 25, 43, 63], and photoacoustic spectroscopy [50, 59, 63]. However, environmental changes like humidity, atmospheric pressure, and temperature can affect the measured glucose values [20]. More importantly, existing systems lack of comprehensive clinical evaluations on diabetic patients. Small LEDs and photodiodes have been used to build compact prototypes [44, 72]. However, all users are non-diabetic and young. Our work differs in that our system entails a much simpler set up with a low-cost liquid crystal and a photodiode, and we evaluate our system on 50 participants, including 41 diabetic patients.

Transdermal techniques include reverse iontophoresis [7, 53, 65], and impedance spectroscopy [70]. Reverse iontophoresis transports glucose outward from the skin. The uncharged glucose molecules present in the interstitial fluid are carried along with the ions across the skin and collected at the cathode [7, 53, 65]. Impedance spectroscopy is based on the measurement of the impedance of a tissue using alternating currents of known intensity [70]. However, the low electric current passing through the skin can cause irritation [31]. Compare with transdermal techniques, our system leverage visible light to estimate glucose concentration, which will not cause skin irritations.

Thermal techniques involve measurements of thermal generation, blood flow rate, hemoglobin, and oxyhemoglobin concentrations, which correspond to the blood glucose levels [14, 15, 33]. However, this technique has a strong probability of interfering with the environmental conditions, such as temperature and pH. Our work differs in that our system is based on the rotation of the linear polarization vector of light by glucose concentration, which is unaffected by temperature and pH fluctuations.

8 CONCLUSION

We presented a noninvasive glucose monitoring system using polarized light. We tackled the challenge of skin scattering by designing a new liquid-crystal-based depolarization cancellation methodology. We then leveraged multiple wavelengths and light intensity levels and develop a generic learning model to address the challenge of user diversity and confounding factors. Therefore, no calibration or retraining is required in our proposed model. We built a compact (17 cm × 10 cm × 5 cm) and low-cost (e.g., <$250) prototype using off-the-shelf hardware components, and we evaluate the system through a study with 41 patients with diabetes and 9 healthy subjects. Experimental results show that our system achieves 89% clinical accuracy, and the mean absolute relative differences is 10%. We plan to expand our training data with larger-scale user studies across the hospitals in the U.S. to further improve system accuracy.

9 ACKNOWLEDGMENTS

We sincerely thank reviewers for their insightful feedback. Special thanks to Dr. Andrew Crawford, Dr. Stephanie Kassels, Dr. Erika Kline, and Dr. Richard Comi at DHMC for helping us conduct experiments with diabetic patients. This work is supported in part by National Science Foundation (CNS-1552924 and SenSE-2037267).

REFERENCES

[1] 2019. Diabetes Statistics. https://www.diabetesresearch.org/diabetes-statistics. (2019).
[2] 2019. FreeStyle Libre CGM sensors. https://www.freestylelibre.us/. (2019).
[3] 2019. Stereochemistry of Sugars: Diastereorners. http://www.chem.uiuc.edu/organic/. (2019).
[4] Simon Alaluf, Alan Heath, NIK Carter, Derek Atkins, Harish Mahalingam, Karen Barrett, RIA Kolb, and Nico Smit. 2001. Variation in melanin content and composition in type V and VI photoexposed and photoprotected human skin: the dominant role of DHI. *Pigment Cell Research* 14, 5 (2001), 337–347.
[5] Alexandra Amaro-Ortiz, Betty Yan, and John A D'Orazio. 2014. Ultraviolet radiation, aging and the skin: prevention of damage by topical cAMP manipulation. *Molecules* 19, 5 (2014), 6202–6219.
[6] Caerwyn Ash, Michael Dubec, Kelvin Donne, and Tim Bashford. 2017. Effect of wavelength and beam width on penetration in light-tissue interaction using computational methods. *Lasers in medical science* 32, 8 (2017), 1909–1918.
[7] Amay J Bandodkar, Wenzhao Jia, Ceren Yardımcı, Xuan Wang, Julian Ramirez, and Joseph Wang. 2014. Tattoo-based noninvasive glucose monitoring: a proof-of-concept study. *Analytical chemistry* 87, 1 (2014), 394–398.
[8] Ananda Basu, Simmi Dube, Michael Slama, Isabel Errazuriz, Jose Carlos Amezcua, Yogish C Kudva, Thomas Peyser, Rickey E Carter, Claudio Cobelli, and Rita Basu. 2013. Time lag of glucose from intravascular to interstitial compartment in humans. *Diabetes* 62, 12 (2013), 4083–4087.
[9] Alexander Bauer, Otto Hertzberg, Arne Küderle, Dominik Strobel, Miguel A Pleitez, and Werner Mäntele. 2018. IR-spectroscopy of skin in vivo: Optimal skin sites and properties for non-invasive glucose measurement by photoacoustic and photothermal spectroscopy. *Journal of biophotonics* 11, 1 (2018), e201600261.
[10] Brent D Cameron and Gerard L Cote. 1997. Noninvasive glucose sensing utilizing a digital closed-loop polarimetric approach. *IEEE Transactions on Biomedical Engineering* 44, 12 (1997), 1221–1227.
[11] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)* 2, 3 (2011), 1–27.
[12] Jason Yuanzhe Chen, Qi Zhou, Gu Xu, Ryan Taoran Wang, Edward Guangqing Tai, Longhan Xie, Qianzhi Zhang, Yanyan Guan, and Xiaochun Huang. 2019. Non-invasive blood glucose measurement of 95% certainty by pressure regulated Mid-IR. *Talanta* 197 (2019), 211–217.
[13] Tseng-Lin Chen, Yu-Lung Lo, Chia-Chi Liao, and Quoc-Hung Phan. 2018. Non-invasive measurement of glucose concentration on human fingertip by optical coherence tomography. *Journal of biomedical optics* 23, 4 (2018), 047001.
[14] Zhen-cheng Chen, Xing-liang Jin, Jian-ming Zhu, Di-ya Wang, and Ting-ting Zhang. 2009. Non-invasive glucose measuring apparatus based on conservation of energy method. *Journal of Central South University of Technology* 16, 6 (2009), 982.
[15] Ok Kyung Cho, Yoon Ok Kim, Hiroshi Mitsumaki, and Katsuhiko Kuwa. 2004. Noninvasive measurement of glucose by metabolic heat conformation method. *Clinical chemistry* 50, 10 (2004), 1894–1898.
[16] Soyun Cho, Mi Hee Shin, Yeon Kyung Kim, Jo-Eun Seo, Young Mee Lee, Chi-Hyun Park, and Jin Ho Chung. 2009. Effects of infrared radiation and heat on human skin aging in vivo. In *Journal of Investigative Dermatology Symposium Proceedings*, Vol. 14. Elsevier, 15–19.
[17] Md Koushik Chowdhury, S Anuj, S Neeraj, and S Shiru. 2015. Error Grid Analysis of Reference and Predicted Blood Glucose Level Values as Obtained from The Normal and Prediabetic Human Volunteer. *American Journal of Biomedical Engineering* 5, 1 (2015), 6–14.
[18] Mark Christiansen, Timothy Bailey, Elaine Watkins, David Liljenquist, David Price, Katherine Nakamura, Robert Boock, and Thomas Peyser. 2013. A new-generation continuous glucose monitoring system: improved accuracy and reliability compared with a previous-generation system. *Diabetes technology & therapeutics* 15, 10 (2013), 881–888.
[19] Claudio Cobelli, Michele Schiavon, Chiara Dalla Man, Ananda Basu, and Rita Basu. 2016. Interstitial fluid glucose is not just a shifted-in-time but a distorted mirror of blood glucose: insight from an in silico study. *Diabetes technology & therapeutics* 18, 8 (2016), 505–511.
[20] Gerard L Cote. 2001. Noninvasive and minimally-invasive optical monitoring technologies. *The Journal of nutrition* 131, 5 (2001), 1596S–1604S.
[21] Leszek Czupryniak, László Barkai, Svetlana Bolgarska, Agata Bronisz, Jan Broz, Katarzyna Cypryk, Marek Honka, Andrej Janez, Mladen Krnic, Nebojsa Lalic, and others. 2014. Self-monitoring of blood glucose in diabetes: from evidence to clinical reality in Central and Eastern Europe—recommendations from the international Central-Eastern European expert group. *Diabetes technology & therapeutics* 16, 7 (2014), 460–475.
[22] Carlos Eduardo Ferrante do Amaral and Benhard Wolf. 2008. Current development in non-invasive glucose monitoring. *Medical engineering & physics* 30, 5 (2008), 541–549.
[23] Annika MK Enejder, Thomas G Scecina, Jeankun Oh, Martin Hunter, WeiChuan Shih, Slobodan Sasic, Gary L Horowitz, and Michael S Feld. 2005. Raman spectroscopy for noninvasive glucose measurements. *Journal of Biomedical Optics* 10, 3 (2005), 031114.
[24] Michael M Engelgau, KM Narayan, and William H Herman. 2000. Screening for type 2 diabetes. *Diabetes care* 23, 10 (2000), 1563–1580.
[25] A Ergin, MJ Vilaboy, A Tchouassi, R Greene, and GA Thomas. 2003. Detection and analysis of glucose at metabolic concentration using Raman spectroscopy. In *2003 IEEE 29th Annual Proceedings of Bioengineering Conference*. IEEE, 337–338.
[26] Jerome H Friedman. 2002. Stochastic gradient boosting. *Computational statistics & data analysis* 38, 4 (2002), 367–378.
[27] Nirmalya Ghosh, Michael FG Wood, Shu-hong Li, Richard D Weisel, Brian C Wilson, Ren-Ke Li, and I Alex Vitkin. 2009. Mueller matrix decomposition for polarized light assessment of biological tissues. *Journal of biophotonics* 2, 3 (2009), 145–156.
[28] Willemijn Groenendaal, Golo Von Basum, Kristiane A Schmidt, Peter AJ Hilbers, and Natal AW van Riel. 2010. Quantifying the composition of human skin for glucose sensor development. (2010).
[29] Xinxin Guo, Michael FG Wood, Nirmalya Ghosh, and I Alex Vitkin. 2010. Depolarization of light in turbid media: a scattering event resolved Monte Carlo study. *Applied optics* 49, 2 (2010), 153–162.
[30] Steven L Jacques, Kenneth Lee, and Jessica C Ramella-Roman. 2000. Scattering of polarized light by biological tissues. In *Saratov Fall Meeting'99: Optical Technologies in Biophysics and Medicine*, Vol. 4001. International Society for Optics and Photonics, 14–28.
[31] Jayoung Kim, Alan S Campbell, and Joseph Wang. 2018. Wearable non-invasive epidermal glucose sensors: A review. *Talanta* 177 (2018), 163–170.
[32] DF Kimball, D Budker, DS English, C-H Li, A-T Nguyen, SM Rochester, A Sushkov, VV Yashchuk, and M Zolotorev. 2001. Progress towards fundamental symmetry tests with nonlinear optical rotation. In *AIP Conference Proceedings*, Vol. 596. AIP, 84–107.
[33] Jae B Ko, Ok K Cho, Yoon O Kim, and Kazuo Yasuda. 2004. Body metabolism provides a foundation for noninvasive blood glucose monitoring. *Diabetes care* 27, 5 (2004), 1211–1212.
[34] Paul AJ Kolarsick, Maria Ann Kolarsick, and Carolyn Goodwin. 2011. Anatomy and physiology of the skin. *Journal of the Dermatology Nurses' Association* 3, 4 (2011), 203–213.
[35] Jonas Kottmann, Julien Rey, and Markus Sigrist. 2016. Mid-Infrared photoacoustic detection of glucose in human skin: towards non-invasive diagnostics. *Sensors* 16, 10 (2016), 1663.
[36] Jonas Kottmann, Julien M Rey, Joachim Luginbühl, Ernst Reichmann, and Markus W Sigrist. 2012. Glucose sensing in human epidermis using mid-infrared photoacoustic detection. *Biomedical optics express* 3, 4 (2012), 667–680.
[37] Chak-hing Lam. 2009. *Clinical evaluation of non-invasive blood glucose measurement by using near infrared spectroscopy via inter-and intra-subject analysis.* Ph.D. Dissertation. The Hong Kong Polytechnic University.
[38] Sabbir Liakat, Kevin A Bors, Laura Xu, Callie M Woods, Jessica Doyle, and Claire F Gmachl. 2014. Noninvasive in vivo glucose sensing on human subjects using mid-infrared light. *Biomedical optics express* 5, 7 (2014), 2397–2404.
[39] Chia-Chi Liao and Yu-Lung Lo. 2015. Extraction of linear anisotropic parameters using optical coherence tomography and hybrid Mueller matrix formalism. *Optics Express* 23, 8 (2015), 10653–10667.
[40] Tamar Lin, Avner Gal, Yulia Mayzel, Keren Horman, and Karnit Bahartan. 2017. Non-invasive glucose monitoring: a review of challenges and recent advances. *Curr. Trends Biomed. Eng. Biosci* 6 (2017), 1–8.
[41] Yu-Lung Lo and Tsung-Chih Yu. 2006. A polarimetric glucose sensor using a liquid-crystal polarization modulator driven by a sinusoidal signal. *Optics communications* 259, 1 (2006), 40–48.
[42] A Losoya-Leal, S Camacho-León, G Dieck-Assad, and SO Martínez-Chapa. 2012. State of the art and new perspectives in non-invasive glucose sensors. *Revista Mexicana De Ingeniería Biomédica* 33, 1 (2012), 41–52.
[43] Signe M Lundsgaard-Nielsen, Anders Pors, Stefan O Banke, Jan E Henriksen, Dietrich K Hepp, and Anders Weber. 2018. Critical-depth Raman spectroscopy enables home-use non-invasive glucose monitoring. *PloS one* 13, 5 (2018), e0197134.
[44] Mon Arjay Malbog and Noel Linsangan. 2018. Non-invasive glucose meter for android-based devices. In *Proceedings of the 2018 10th International Conference on Computer and Automation Engineering*. 161–165.
[45] Roger Mazze, Yariv Yogev, and Oded Langer. 2012. Measuring glucose exposure and variability using continuous glucose monitoring in normal and abnormal glucose metabolism in pregnancy. *The Journal of Maternal-Fetal & Neonatal Medicine* 25, 7 (2012), 1171–1175.
[46] Roger J McNichols and Gerard L Cote. 2000. Optical glucose sensing in biological fluids: an overview. *Journal of biomedical optics* 5, 1 (2000), 5–17.
[47] Mehrdad Mehdizadeh. 2010. Microwave/RF Applicators and Probes. (2010).
[48] Pradipta Mukherjee, Nathan Hagen, and Yukitoshi Otani. 2019. Glucose sensing in the presence of scattering by analyzing a partial Mueller matrix. *Optik* 180 (2019), 775–781.
[49] Asmat Nawaz, Per Øhlckers, Steinar Sælid, Morten Jacobsen, and M Nadeem Akram. 2016. Non-invasive continuous blood glucose measurement techniques. *Journal of Bioinformatics and Diabetes* 1, 3 (2016), 01.

[50] Praful P Pai, Pradyut Kumar Sanki, Arijit De, and Swapna Banerjee. 2015. NIR photoacoustic spectroscopy for non-invasive glucose measurement. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 7978–7981.

[51] Quoc-Hung Phan, Tzu-Hsiang Jian, Yu-Ru Huang, You-Rui Lai, Wei-Zhe Xiao, and Shin-Wei Chen. 2020. Combination of surface plasmon resonance and differential Mueller matrix formalism for noninvasive glucose sensing. *Optics and Lasers in Engineering* 134 (2020), 106268.

[52] Raju Poddar, Joseph Thomas Andrews, Pratyoosh Shukla, and Pratima Sen. 2008. Non-invasive glucose monitoring techniques: A review and current trends. *arXiv preprint arXiv:0810.5755* (2008).

[53] Russell O Potts, Janet A. Tamada, and Michael J. Tierney. 2002. Glucose monitoring by reverse iontophoresis. *Diabetes/metabolism research and reviews* 18, S1 (2002), S49–S53.

[54] Georgeanne Purvinis, Brent D Cameron, and Douglas M Altrogge. 2011. Non-invasive polarimetric-based glucose monitoring: an in vivo study. *Journal of diabetes science and technology* 5, 2 (2011), 380–387.

[55] Yun Qian, Yanchun Liang, Mu Li, Guoxiang Feng, and Xiaohu Shi. 2014. A resampling ensemble algorithm for classification of imbalance problems. *Neurocomputing* 143 (2014), 57–67.

[56] Bernard Querleux, Thérèse Baldeweck, Stéphane Diridollou, Jean De Rigal, Etienne Huguet, Frédéric Leroy, and Victoria Holloway Barbosa. 2009. Skin from various ethnic origins and aging: an in vivo cross-sectional multimodality imaging study. *Skin Research and Technology* 15, 3 (2009), 306–313.

[57] Anthony V Rawlings. 2006. Ethnic skin types: are there differences in skin structure and function? 1. *International journal of cosmetic science* 28, 2 (2006), 79–93.

[58] Florian Reiterer, Philipp Polterauer, Michael Schoemaker, Guenther Schmelzeisen-Redecker, Guido Freckmann, Lutz Heinemann, and Luigi Del Re. 2017. Significance and reliability of MARD for the accuracy of CGM systems. *Journal of diabetes science and technology* 11, 1 (2017), 59–67.

[59] Zhong Ren, Guodong Liu, and Zhen Huang. 2014. Noninvasive detection of glucose level based on tunable pulsed laser induced photoacoustic technique. In *International Symposium on Optoelectronic Technology and Application 2014: Laser and Optical Measurement Technology; and Fiber Optic Sensors*, Vol. 9297. International Society for Optics and Photonics, 929709.

[60] David Rodbard. 2017. Continuous glucose monitoring: a review of recent studies demonstrating improved glycemic outcomes. *Diabetes technology & therapeutics* 19, S3 (2017), S–25.

[61] Silonie Sachdeva and others. 2009. Fitzpatrick skin typing: Applications in dermatology. *Indian Journal of Dermatology, Venereology, and Leprology* 75, 1 (2009), 93.

[62] Sanjiv Sharma, Zhenyi Huang, Michelle Rogers, Martyn Boutelle, and Anthony EG Cass. 2016. Evaluation of a minimally invasive glucose biosensor for continuous tissue monitoring. *Analytical and bioanalytical chemistry* 408, 29 (2016), 8427–8435.

[63] Joo Yong Sim, Chang-Geun Ahn, Eun-Ju Jeong, and Bong Kyu Kim. 2018. In vivo microscopic photoacoustic spectroscopy for non-invasive glucose monitoring invulnerable to skin secretion products. *Scientific reports* 8, 1 (2018), 1059.

[64] Jitendra Solanki, Om Prakash Choudhary, Pratima Sen, and Joseph Thomas Andrews. 2013. Polarization sensitive optical low-coherence reflectometry for blood glucose monitoring in human subjects. *Review of Scientific Instruments* 84, 7 (2013), 073114.

[65] MJ Tierney, HL Kim, MD Burns, JA Tamada, and RO Potts. 2000. Electroanalysis of glucose in transcutaneously extracted samples. *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis* 12, 9 (2000), 666–671.

[66] GT Tucker and MS Lennard. 1990. Enantiomer specific pharmacokinetics. *Pharmacology & therapeutics* 45, 3 (1990), 309–329.

[67] Sandeep Kumar Vashist. 2012. Non-invasive glucose monitoring technology in diabetes management: A review. *Analytica chimica acta* 750 (2012), 16–27.

[68] Sandeep Kumar Vashist, Dan Zheng, Khalid Al-Rubeaan, John HT Luong, and Fwu-Shan Sheu. 2011. Advances in carbon nanotube based electrochemical sensors for bioanalytical applications. *Biotechnology advances* 29, 2 (2011), 169–188.

[69] Jeanette M Waller and Howard I Maibach. 2005. Age and skin structure and function, a quantitative approach (I): blood flow, pH, thickness, and ultrasound echogenicity. *Skin research and technology* 11, 4 (2005), 221–235.

[70] Stuart Alan Weinzimer. 2004. Analysis: PENDRA: The Once and Future Noninvasive Continuous Glucose Monitoring Device? *Diabetes technology & therapeutics* 6, 4 (2004), 442–444.

[71] Michael FG Wood, Nirmalya Ghosh, Xinxin Guo, and I Alex Vitkin. 2008. Towards noninvasive glucose sensing using polarization analysis of multiply scattered light. *Handbook of optical sensing of glucose in biological fluids and tissues* 12 (2008).

[72] Jyoti Yadav, Asha Rani, Vijander Singh, and Bhaskar Mohan Murari. 2014. Near-infrared LED based non-invasive blood glucose sensor. In *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 591–594.

[73] Kamal Youcef-Toumi and Vidi A Saptari. 1999. Noninvasive blood glucose analysis using near infrared absorption spectroscopy. *The Home Automation and Healthcare Consortium* (1999), 2–3.

[74] Chiara Zecchin, Andrea Facchinetti, Giovanni Sparacino, Chiara Dalla Man, Chinmay Manohar, James A Levine, Ananda Basu, Yogish C Kudva, and Claudio Cobelli. 2013. Physical activity measured by physical activity monitoring system correlates with glucose trends reconstructed from continuous glucose monitoring. *Diabetes technology & therapeutics* 15, 10 (2013), 836–844.

[75] Huiting Zheng, Jiabin Yuan, and Long Chen. 2017. Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation. *Energies* 10, 8 (2017), 1168.