A Smartphone-Based Non-Invasive Measurement System for Blood Constituents from Photoplethysmography (PPG) and Fingertip Videos Illuminated with the Near-Infrared LEDs

Md Hasanul Aziz
Marquette University

Follow this and additional works at: https://epublications.marquette.edu/dissertations_mu

Part of the Computer Sciences Commons

Recommended Citation
Aziz, Md Hasanul, "A Smartphone-Based Non-Invasive Measurement System for Blood Constituents from Photoplethysmography (PPG) and Fingertip Videos Illuminated with the Near-Infrared LEDs" (2022). Dissertations (1934 -). 1591.
https://epublications.marquette.edu/dissertations_mu/1591
A smartphone-based non-invasive measurement system for blood constituents from photoplethysmography (PPG) and fingertip videos illuminated with the near-infrared LEDs

by

Md Hasanul Aziz, M.Sc.

A Dissertation submitted to the Faculty of the Graduate School, Marquette University, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Milwaukee, Wisconsin
August 2022
A SMARTPHONE-BASED NON-INVASIVE MEASUREMENT SYSTEM FOR BLOOD CONSTITUENTS FROM PHOTOPLETHYSMOGRAPHY (PPG) AND FINGERTIP VIDEOS ILLUMINATED WITH THE NEAR-INFRARED LEDS

Md Hasanul Aziz
Department of Computer Science
Marquette University, 2022

At least two billion people are affected by hemoglobin (Hgb), diabetic-related, and other blood-related diseases. Regular clinical assessments of these problems are conducted by analyzing venipuncture-obtained blood samples in laboratories. A non-invasive, cheap, point-of-care, and accurate test is needed everywhere. We started with Hgb measurement, and after an extensive literature survey, we came up with a non-invasive solution with 10-second Smartphone videos of the index fingertips using custom hardware sets to illuminate the fingers.

We tested four lighting conditions with wavelengths in the near-infrared spectrum suggested by the absorption properties of two primary components of blood- oxygenated Hgb and plasma. We found a strong linear correlation between our measured and laboratory-measured Hgb levels in 167 patients with a mean absolute percentage error (MAPE) of 5%. In our initial analysis, critical tasks were performed manually. Now, using the same data, we have automated or modified all the steps. For all subjects, male subjects, and female subjects, we found a MAPE of 6.43%, 5.34%, and 4.85% and mean squared error (MSE) of 0.84, 0.5, and 0.49, respectively. The new analyses, however, have suggested inexplicable inconsistencies in our results, which we attribute to laboratory measurement errors reflected in a non-normative distribution of Hgb levels in our studied patients, as well as excess noise in the specific signals we measured in the videos.

To address these problems, we designed a customizable external attachment to the smartphone designed to limit the noise in this system. This attachment is a plastic box with a topside slot to accommodate the smartphone, and internally a 3-pronged- an electrical circuit, a holder box, and a pressure sensor for the fingertip pressing against the camera lens. The attachment is inexpensive, power-efficient, and portable, and the associated software is programmable to acquire optimal PPG signals. Measurement of blood constituents other than Hgb can also use this attachment with slight modifications, which we plan to do for glycated hemoglobin measurement. Towards that end, we have done exploratory work with an unbalanced dataset to detect hyperglycemic states and found an accuracy of 68% after oversampling with SMOTEENN.
ACKNOWLEDGEMENTS

Md Hasanul Aziz, M.S.

First and foremost, I would like to thank almighty Allah for giving me every opportunity to reach this stage and complete my Ph.D. Then I must acknowledge and express my utmost gratitude to my Ph.D. advisor Dr. Sheikh Iqbal Ahamed for his continuous support, motivation, and invaluable guidance throughout my graduate studies. He always provided the comfort I needed to reduce the stress of the stressful Graduate life, which increased my productivity manifold. He is truly a research Dad for me.

I would like to thank the members of my committee Dr. Praveen Madiraju and Rumi Ahmed Khan, M.D. for their time, comments, feedback, and encouragement. I would also like to extend my gratitude to Dr. Richard R. Love for his continuous mentorship in my research and for Dr. Kamrul Hasan for building the foundation of this work.

My late parents deserved all the credit for whatever I can accomplish today. It would never be enough to thank my Ammu (Mom) Ujjol and Abbu (Dad) Rontu for nurturing and raising me and giving me a wonderful childhood and teenage life, which is the building block of my life. I would like to thank my Apu (Sister) Shawrna, Ma (mother-in-law) Hiron Nahar, Baba (father-in-law) Mesbah Uddin, Vai (brother-in-law) Hira, and Bon (sister-in-law) Tamanna for always supporting, encouraging and loving me. They were a blessing, and I am lucky enough to have them in my life.

My high school Mirzapur Cadet College- a military school in Bangladesh pushed the habit of perseverance, discipline, and hard work in me. I want to acknowledge and express gratitude to my friend Rashed and Zobayer for their unconditional love and support. I am also thankful my childhood mentor Saad vai (brother) and my roommates in college life and post-college life. Furthermore, I would like to thank all my teachers, mentors, friends, relatives, and extended family back home who always wished me the best.

I could not complete my Ph.D. journey without my awesome group of friends in Milwaukee who became my second family. I would like to thank them for their friendship and support. This incredible group of people in Milwaukee was available whenever I needed them.

I would like to dedicate my work to my wife, Sumaiya Binte Mesbah whose love, encouragement, motivation, and support carried me through the journey. I am so thankful for her faith in me and immeasurable care for me. Whenever I had the slightest doubt about myself, she was there to assure me- "You are the best. If anyone can do it then that would be you".
# TABLE OF CONTENTS

ACKNOWLEDGEMENT ......................................................... i

LIST OF TABLES .............................................................. vi

LIST OF FIGURES ............................................................. vii

1 INTRODUCTION ............................................................. 1
  1.1 Significance .......................................................... 1
    1.1.1 Hemoglobin (Hgb) measurement .......................... 1
    1.1.2 Glycated Hemoglobin (HbA1C) measurement .......... 3
  1.2 Photoplethysmography (PPG) ..................................... 4
  1.3 Smartphone as a point-of-care tool ............................. 5
  1.4 Research questions ............................................... 6
  1.5 Dissertation organization ........................................ 6

2 LITERATURE REVIEW .................................................... 8
  2.1 Laboratory method ............................................... 8
  2.2 Minimal invasive techniques .................................... 8
  2.3 Non-invasive systems using spectrometers ................. 11
  2.4 Non-invasive systems using PPG signals .................... 14
  2.5 Non-invasive systems using smartphones ................... 16

3 METHODOLOGY ............................................................ 20
  3.1 Biological foundations: Absorption properties of blood in the near-infrared (NIR) spectrum ...................... 20
  3.2 Hardware setup and data collection .......................... 21
  3.3 Photoplethysmography (PPG) .................................... 22
3.4 Signal Processing ................................................. 26
  3.4.1 Fourier analysis ............................................. 26
  3.4.2 Savitzky-Golay smoothing filter ......................... 27
  3.4.3 Wavelet Transform(WT) .................................. 27
  3.4.4 Independent Component Analysis (ICA) ................. 27
  3.4.5 Butterworth Filter ......................................... 28
3.5 PPG cycle detection ............................................ 28
3.6 PPG selection and feature extraction ......................... 29
  3.6.1 Features from PPG ........................................ 32
  3.6.2 Features from 1st derivative PPG ....................... 34
  3.6.3 Features from 2nd derivative PPG ....................... 35
3.7 Feature selection and model generation ....................... 37
4  HEMOGLOBIN MEASUREMENT ................................. 40
4.1 Error Metrics ................................................... 40
  4.1.1 Bias ....................................................... 40
  4.1.2 Standard Deviation (SD) of the errors ................. 41
  4.1.3 Mean Square Error (MSE) .............................. 41
  4.1.4 Root Mean Square Error (RMSE) ....................... 42
  4.1.5 Root Mean Square Deviation (RMSD) ................... 42
  4.1.6 Mean Absolute Percentage Error (MAPE) ............... 42
  4.1.7 $R^2$ (R-squared)- Coefficient of determination .... 42
  4.1.8 Correlation Coefficient (R) ........................... 43
4.2 Results ......................................................... 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.1</td>
<td>Dataset: All subjects</td>
<td>46</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Dataset: Female subjects</td>
<td>48</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Dataset: Male subjects</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>Discussion</td>
<td>51</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Comparison with previous work</td>
<td>51</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Noise in signal</td>
<td>52</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Average line as a predictor</td>
<td>54</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Error metrics</td>
<td>54</td>
</tr>
<tr>
<td>4.3.5</td>
<td>Distribution of the Hgb levels in datasets</td>
<td>55</td>
</tr>
<tr>
<td>4.3.6</td>
<td>Separate machine learning model for men and women</td>
<td>56</td>
</tr>
<tr>
<td>4.3.7</td>
<td>Comparison with state-of-the-art</td>
<td>58</td>
</tr>
<tr>
<td>5</td>
<td>CUSTOMIZABLE DATA COLLECTION TOOL</td>
<td>60</td>
</tr>
<tr>
<td>5.1</td>
<td>Related works</td>
<td>60</td>
</tr>
<tr>
<td>5.2</td>
<td>Requirements</td>
<td>61</td>
</tr>
<tr>
<td>5.3</td>
<td>Description of complete system</td>
<td>62</td>
</tr>
<tr>
<td>5.3.1</td>
<td>The electrical circuit</td>
<td>62</td>
</tr>
<tr>
<td>5.3.2</td>
<td>3D designed box</td>
<td>66</td>
</tr>
<tr>
<td>5.3.3</td>
<td>The software system</td>
<td>69</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>75</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Technologies used in the software system</td>
<td>75</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Use of camera2 API over default camera</td>
<td>76</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Power Consumption</td>
<td>76</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Noise reduction</td>
<td>77</td>
</tr>
</tbody>
</table>
5.4.5 HIPAA compliance ........................................ 78
5.4.6 Future research ............................................ 78

6 HYPERGLYCEMIC STATE DETECTION ........................................ 79
6.1 Theory and Methodology ...................................... 79
6.2 Methodology and Results ...................................... 81
6.2.1 Feature set generation for machine learning model .......... 81
6.2.2 Initial Analysis ............................................. 82
6.3 Discussion ..................................................... 84

7 CONCLUSION AND FUTURE WORKS ............................... 85

BIBLIOGRAPHY .................................................. 87
LIST OF TABLES

1.1 Different Hgb levels (gm/dL) for mild, moderate and severe anemia [1] .................................................. 2

3.1 List of features extracted from each $S_{PPG}$ ................................................................. 30

4.1 Characteristics of the studied population and its laboratory hemoglobin levels. ........................................ 43

4.2 Sorted feature numbers for the models based on $f_{score}$ ...................................................... 46

4.3 Best result in Hgb measurement with all subjects without the outliers. .......................................... 48

4.4 Best results in models with female subjects. .................................................................................. 49

4.5 Best results in models with male subjects. .................................................................................... 59

4.6 Comparison with HemaApp and Masimo Pronto. ......................................................................... 59

6.1 Characteristics of the studied population and its laboratory HbA1c levels. ................................. 83

6.2 Classification result for HbA1c without using any sampling technique. Here, precision means ratio of true positive and predicted positives and recall means ratio of true positive and actual positives ........................................ 84

6.3 Classification result for HbA1C. Here, precision means ratio of true positive and predicted positives and recall means ratio of true positive and actual positives ........................................................................... 84
LIST OF FIGURES

1.1  Point of care (POC) tools for noninvasive Hb measurement a) Smart-Hb [2], b) Masimo SpHb [3], c) NBM 200 [4], and d) Astrim-Fit [5]. [6]  

2.1  Steps to measure Hgb with a Hemocue device.  

2.2  Color-based anemia test A) The cap and body of the device. A color scale sticker on the body enables user interpretation of the degree of anemia, if any. (B–E) Steps for using the POC color-based anemia test. (F) Ranging from blue to red, resultant solution colors correlate with different Hgb levels, as indicated.  

2.3  The response of hemoglobin, oxygenated hemoglobin, and water in the spectrum [7, 6].  

2.4  Hgb level measurement by transmittance mode using the NIR spectrophotometric system [8].  

2.5  HPPG sensor made of two LED of 660nm and 905nm to generate PPG signals from the fingertip [9].  

2.6  System overview of the HemaApp [10].  

3.1  System architecture (Here, ML- Machine Learning)  

3.2  The different hardware used for capturing the videos: A) A LED device in turned on condition, B) Look alike LED devices with different wavelength. C) A finger is placed on the LED device. D) Pixel2 smartphone is used for capturing a video while the finger is illuminated by LEDs.  

3.3  A PPG signal with some critical features.  

3.4  PPG signal of a block from 10x10 blocks in a video. The red dot in a block represent the average value of red pixels for that block. We get a PPG signal by taking the averages for the same block across all the frames.  

3.5  a) Averages of red pixels of block across all selected frame, b) Processed PPG signals from $S_u$.  

3.6  Detected PPG cycles in $S_{PPG}$.  

3.7  Selected PPG cycles are merged together.
3.8 A typical waveform of the PPG and its characteristic parameters, whereas
the amplitude of the systolic peaks is x while y is the amplitude of the
diastolic peak. [11]. ................................................................. 33

3.9 Original fingertip photoplethysmogram. A1 and A2 are the areas under
the whole PPG wave separated at the point of inflection. Thus, the inflec-
tion point area ratio can be calculated as the division of A2 by A1. [11]. . 34

3.10 Pulse interval for a PPG signal and peak to peak interval for two consecu-
tive peaks. ................................................................. 34

3.11 (a) Original fingertip photoplethysmogram (b) 1\textsuperscript{st} derivative wave of pho-
toplethysmogram [11]. ................................................................. 35

3.12 (a) Original fingertip photoplethysmogram (b) 2\textsuperscript{nd} derivative wave of pho-
toplethysmogram [11]. ................................................................. 36

3.13 Histogram of blocks by different systolic peak heights from all the S\textsubscript{f} for
a video. The systolic peak heights belong to the bins marked by green are
averaged to get systolic peak height feature for that video. ............... 37

4.1 Hemoglobin value distribution of all the subjects. .................. 44

4.2 Hemoglobin value distribution for male and female subjects. ...... 45

4.3 Performance of the models for increasing number of top features for all
subjects. ................................................................. 47

4.4 Performance of the models for increasing number of top features for fe-
male subjects. ................................................................. 50

4.5 Performance of the models for varying number of features for all male
subjects ................................................................. 51

4.6 Inconsistent PPG cycles in S\textsubscript{PPG}. ......................... 53

4.7 The difference relationships between venous and capillary (a surrogate
for arterial) hemoglobin levels in men and women. On the vertical axis
is the average difference between the measured values which decrease
with increasing hemoglobin levels, but at different rates between the gen-
ders. In usual hemoglobin level ranges (<15 gm/dL in men; <14 gm/dL
in women, women have lower arterial hemoglobin levels because the di-
ameters of their arterial vessels are smaller due to the Fahreus effect (see
text). From Murphy et al. [12]. ................................................................. 56
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Circuit design to keep applied pressure on the camera lens within a threshold.</td>
<td>63</td>
</tr>
<tr>
<td>5.2</td>
<td>Electrical circuit printed on a PCB board.</td>
<td>64</td>
</tr>
<tr>
<td>5.3</td>
<td>3D designed holder box for the smartphone and the electrical circuit.</td>
<td>66</td>
</tr>
<tr>
<td>5.4</td>
<td>Connection of all the components of the circuits in the box.</td>
<td>67</td>
</tr>
<tr>
<td>5.5</td>
<td>Placement of finger on the camera lens, and turning on of the LEDs with proper pressure.</td>
<td>68</td>
</tr>
<tr>
<td>5.6</td>
<td>Flow of decisions taken in the software part of the framework.</td>
<td>72</td>
</tr>
<tr>
<td>5.7</td>
<td>A PPG signal containing noise due to finger movement captured by the system keeping motion detection off.</td>
<td>77</td>
</tr>
<tr>
<td>5.8</td>
<td>A PPG signal that can be useful for further analysis generated by the system checking all the parameters in the runtime.</td>
<td>78</td>
</tr>
<tr>
<td>6.1</td>
<td>Absorption properties of HbA1c in NIR spectrum.</td>
<td>79</td>
</tr>
<tr>
<td>6.2</td>
<td>Absorption properties of HbA1c in visible and NIR spectrum.</td>
<td>80</td>
</tr>
<tr>
<td>6.3</td>
<td>Correlation among the features.</td>
<td>82</td>
</tr>
<tr>
<td>6.4</td>
<td>HbA1C value distribution for all subjects.</td>
<td>83</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

1.1 Significance

Globally, the population suffering from hemoglobin (Hgb) diseases-mostly anemia, is two billion individuals, from diabetes 463 million, and from kidney disease 1 million [13, 14, 15]. The standard blood testing systems for diagnosing and monitoring these diseases are to obtain and analyze venipuncture-obtained blood samples in laboratory settings. These systems have significant direct and indirect costs and are inconvenient for patients, particularly those who live in low- and middle-income countries. This process is very impractical for care of surgical and trauma patients where frequent and immediate results are needed. Non-invasive, point-of-care, rapid, patient-convenient, inexpensive, and accurate blood testing and measurement systems would benefit patients everywhere.

1.1.1 Hemoglobin (Hgb) measurement

Hemoglobin abnormality, both high and low levels of hemoglobin (Hgb), causes blood diseases. A high Hgb level can lead to fatal health issues, including heart attacks and strokes. A low level of Hgb results in iron deficiency, pregnancy-related anemia, and sickle cell anemia. Maintaining a standard level of Hgb is enormously essential for sound health since Hgb carries oxygen ($O_2$) from the lungs to tissues [16]. According to the World Health Organization (WHO), the cut-off Hgb value for males is 13 g/dl, and for females is 12 g/dl to define an anemic patient. The average Hgb levels for males are 13-16 g/dl, and for women are 12-15 g/dl. Based on these Hgb levels, the situation of anemia is categorized as mild, moderate, and severe [17]. Table 1.1 presents the Hgb ranges for mild, moderate, and severe anemia used in a study regarding the prevalence of moderate-severe anemia in the US Population (NHANES 2003-2012) [1].

Health care providers have a practical concern about the anemic people and
Table 1.1. Different Hgb levels (gm/dL) for mild, moderate and severe anemia [1].

<table>
<thead>
<tr>
<th>Population (age group)</th>
<th>Mild</th>
<th>Moderate-severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children (&lt;5 years)</td>
<td>&lt; 11.0</td>
<td>&lt; 10.0</td>
</tr>
<tr>
<td>Children (≥5 and &lt;12 years)</td>
<td>&lt; 11.5</td>
<td>&lt; 11.0</td>
</tr>
<tr>
<td>Children (≥12 and &lt;15 years)</td>
<td>&lt; 12.0</td>
<td>&lt; 11.0</td>
</tr>
<tr>
<td>Female (non-pregnant) (≥15 years)</td>
<td>&lt; 12.0</td>
<td>&lt; 11.0</td>
</tr>
<tr>
<td>Female (pregnant)</td>
<td>&lt; 11.0</td>
<td>&lt; 10.0</td>
</tr>
<tr>
<td>Male (≥15 years)</td>
<td>≥ 13.0</td>
<td>&lt; 11.0</td>
</tr>
</tbody>
</table>

intent to get the trend of Hgb levels regularly since anemia reduces the oxygen-carrying capacity of the blood, which causes weakness, fatigue, lightheaded, and other physiological problems [18]. In some cases, such as ill-health, less productivity, and premature deaths are treated as the cause of Hgb deficiency. In the United States, Hgb deficiency is observed frequently in infants, toddlers, adolescent females, and women of childbearing age [19]. Low Hgb concentration is caused by various reasons, such as decreased red blood cell (RBC) production, increased RBC destruction, iron deficiency, vitamin deficiency, or severe blood loss [19, 20]. Irrespective of what is causing the low or high level of Hgb in our blood, health service providers must know the Hgb level to provide better treatment to the patient before prescribing any medication.

Approximately 5.6% of the US population is anemic, and 1.5% of the population has moderate to severe anemia [1]. Sickle cell diseases (SCD) cost more than US $1.5 billion annually in the United States [21]. Globally, blood disorders and associated complications affect more than 5 million people. In Africa, approximately 250,000 babies are born with SCD every year [22] and 1.62 billion people are affected by Hgb-related abnormalities worldwide [23]. A reliable, affordable, and user-friendly solution is crucial to assessing the Hgb status of a large population. Clinical assessment of Hgb typically involves the cyan-methemoglobin method, which is considered to be reliable. However, this invasive process has several limitations, including that the diagnostic devices are not portable, results are not immediately available, and the
entire process is expensive. Thus, an Hgb disorder diagnosis based on an invasive method is not a perfect solution, especially for people in low- and middle-income countries [24, 25]. With available medical facilities, frequent invasive testing is also less convenient due to pain, anxiety, and infections [26]. A recent study estimated the cost for a complete blood count (CBC) test in Bari, Puglia, Italy, with approximately 1,000,000 inhabitants, to be US $3.14, resulting in a total cost of US $560,000 in 2018. Considering the entire national territory of Italy, the estimated cost will be more than US $20 million per year for public hospitals for outpatients. However, the laboratory costs for other cases, including hospitalized patients and private clinic patients, will be much higher than this previous estimation in Italy [27]. These multiple circumstances indicate the paramount importance of a non-invasive point-of-care (POC) method for Hgb measurement.

1.1.2 Glycated Hemoglobin (HbA1C) measurement

Another major public health issue like anemia is diabetes. International Diabetes Federation Diabetes Atlas estimated that globally diabetes was prevalent in 463 million people in 2019 where half of the people living with diabetes did not know that they have diabetes [15]. It is among the top 10 causes of death in adults [15]. In case of USA, 34.2 million people have diabetes, and 88 million people have prediabetes which is tantamount to 45% of total USA population [28].

World Health Organization (WHO) recommends Glycated hemoglobin (HbA1c) as a diagnostic test to detect diabetes condition. For diagnosing diabetes, a value of HbA1c greater than 6.5% indicates diabetes in the patients but values less than 6.5% does not exclude diabetes diagnosed using glucose test [29]. As HbA1c is a good index of long-term blood glucose, early diagnosis of diabetes is possible by measuring HbA1c [30]. Moreover, correlation between arterial stiffness can be used to prevent vascular complications from HbA1c value [31]. These multiple circumstances indicate that a noninvasive point-of-care (POC) method for HbA1c measurement is very
There are currently commercial devices that can measure Hgb non-invasively using near-infrared (NIR) spectroscopy and photoplethysmography, e.g. the Masimo SpHb measurement system, NBM 200, and TouchHb (Fig. 1.1). These devices are not highly accurate, measuring hemoglobin levels within ± 2 g/dl of laboratory-measured Hgb levels, and thus are in limited clinical use [32, 33, 4, 34]. Limited computational capabilities may account for the limited accuracy of these devices. Further liabilities of these devices are high expense, non-portability, and complicated set-ups.

Photoplethysmography (PPG) is an optical technique to observe the blood volume changes noninvasively. This process is a simple and cost-effective solution to see the blood volume changes in the microvascular bed of tissue [35]. To build a PPG system, we need a light source and a photodetector, where the light source illuminate the tissue area (e.g., finger) and the photodetector (e.g., smartphone camera) captures the variation of light intensity. The intensity variations are observed due to the systole and diastole effect from the heart. So, the PPG signal represents the variation of light intensity that changes with the peripheral pulse (synchronized to heartbeat). PPG works in a visible or a NIR region because the water in tissue absorbs ultraviolet (UV)
light as well as the longer infrared wavelengths (λ close to 1000 nm). Again, the light absorption of oxyhemoglobin and reduced hemoglobin (Hgb) has an isosbestic wavelength in this region [36]. Using features from PPG in various NIR regions, various studies determined Hgb, glucose, blood oxygen saturation, heart rate, blood pressure [37], cardiac output [38], respiration, arterial disease, vascular aging (for assessing hypertension) [39], viscoelastic properties of blood vessels [40], the degree and length of vasodilatation, lower limb chronic venous insufficiency (CVI) [41], vasospastic (cold sensitivity condition), and autonomic function (migraine patients).

Since a PPG signal reflects the blood movement from the heart to the fingertips through the blood vessels, the study of the characteristics parameter of a PPG signal may provide information on blood constituent levels. PPG features were used in several studied including Hematocrit, SpO2, pulse, and respiration [42, 9, 43, 44].

1.3 Smartphone as a point-of-care tool

Almost all of the invasive solutions suffers from laboratory dependency and their expensive nature. On the contrary, a major advantage of using a smartphone in non-invasive applications is its availability both in remote settings and for individual patients. Smartphone-based non-invasive solutions are becoming popular as a potential alternative to invasive clinical blood testing for their advantages in availability, user-friendliness, and easy attachability to different biosensing devices without adding little to no extra expenses.

2.6 billion people have smartphones around the world, and this availability make this tool attractive for medical diagnosis [45, 46, 47], and physiological parameter estimation [48, 49, 50, 51, 52]. Using a smartphone, Dantu et al. monitored blood glucose noninvasively by modifying and extending the Beer-Lambert law to accommodate multiple wavelengths [50]. They used the smartphone's camera as a photodetector to collect and analyze the spectrometric properties of the lights passed through a fingertip placed on the camera of the smartphone. Using this approach, the
smartphone has been used to measure different blood constituents, and in most cases, the data were collected from either fingertips or eyelids [53].

### 1.4 Research questions

After conducting a comprehensive literature survey, we summarized the following steps that can lead to improved non-invasive solution to facilitate the anemic and diabetic population taking better care of their health:

1. Measuring hemoglobin (Hgb) non-invasively with automated cardiac cycle analysis using the fingertip videos illuminated with NIR light.

2. Designing a robust and customizable data collection module to collect smartphone videos of the fingertip illuminated with NIR light.

3. Detecting hyperglycemic state from HbA1c values non-invasively using the PPG signals extracted from fingertip videos illuminated with NIR light.

### 1.5 Dissertation organization

This dissertation thesis will first present a comprehensive review of currently published literature on hemoglobin measurement—laboratory methods, minimal-invasive methods, and non-invasive methods and then in Chapter 3 it will describe the methodology to measure Hgb level non-invasively which can be extended to other blood constituent measurements. Chapter 3 also describes different signal processing techniques to reduce noise in the signal and features extracted from PPG signals to use with the machine learning models. Chapter 4 starts with different error metrics used to deem the success of the methods and which one is more suitable in our case. Then it describes performance of different models for different subset of our available dataset. Section 4.3 in this chapter discusses about different complexities in measuring Hgb, and how we can reduce the complexities.
Chapter 5 provides an in depth representation of a customizable data collection tool to reduce the complexity and increase the accuracy of blood constituent measurement systems. The next chapter explores the possibility of a non-invasive HbA1C measurement system using a similar approach. Finally, Chapter 7 describes the future works, and gives a conclusion to this dissertation.
Chapter 2
Literature Review

2.1 Laboratory method

In any circumstances, if people need to inform their Hgb level to the health service provider, then the usual way to know the Hgb level is cyanmethemoglobin method which is a very complicated process and performed in clinical settings only. In this method, the sample is mixed with Drabkin’s solution to form cyanmethemoglobin [54]. Then a photoelectric colorimeter measures the color of the cyanmethemoglobin at a wavelength of 540 nm, which in turn determines the Hgb level. The delayed response from this method may cause increased medical bills because of the different health parameter checking [25, 24]. Moreover, some patients may need continuous health monitoring when they are getting treatment for critical issues, like hypertension, diabetes, and other blood conditions [24, 55]. The patients in ICU, or operation room sometimes need continuous monitoring of Hgb level, which requires a repeated blood draw and catastrophic event may occur due to the lag time of the laboratory results. Therefore, point of care tools to measure hemoglobin level non-invasively is getting paramount importance in the research community.

2.2 Minimal invasive techniques

Apart from the invasive methods, the most frequently used minimal-invasive device is Hemocue® , which gives Hgb estimation within a minute after analyzing a droplet of blood from capillary or vein [56, 57]. In this minimal-invasive Hgb measurement system, a drop of blood is taken in a microcuvette covered with sodium deoxycholate and analyzed immediately by a photometer consists of two wavelengths 565nm and 880nm for the absorbance of reaction result with azidemethemoglobin (Fig. 2.1). Hemocue® estimated Hgb level with a sensitivity and specificity 94.1% and 95.2% respectively [58]. The fact Hemocue® is battery operated and does not need any
Figure 2.1: Steps to measure Hgb with a Hemocue device.

external power supply made it an ideal candidate to use in rural areas. However, some studies found that this system overestimated the Hb level compared with the clinical result [59, 54]. These studies reported that two factors might be responsible for this overestimation - the time interval between the micro-cuvette opening and their use, and hygroscopic property of Hemocue microcuvette in humid climatic conditions [59].

If Hgb could also be measured without using any photometer, its assessment would be more available everywhere. In 1995, the World Health Organisation (WHO) standardized a hemoglobin color scale (HCS) as an inexpensive, simple, and minimal-invasive tool to screen anemia in rural settings. The principle works behind this tool is that the color of a drop of blood differs in different hemoglobin level, and it can reliably detect anemia [60]. A drop of blood is taken on an absorbent test paper, and the color on the paper is matched against predetermined shades of red. These predetermined shades of red are printed on a small card comprises of 10 standard shades to represent hemoglobin levels at 3, 4, 5, 6, 7, 8, 9, 10, 12 and 14 g/dl [61]. Initial
Figure 2.2: Color-based anemia test A) The cap and body of the device. A color scale sticker on the body enables user interpretation of the degree of anemia, if any. (B–E) Steps for using the POC color-based anemia test. (F) Ranging from blue to red, resultant solution colors correlate with different Hgb levels, as indicated.

reports stated sensitivity of 95% and a specificity of 99.6% for the HCS. However, Darshana et al. found that the accuracy of HCS is not suitable for use in field settings; instead, we should use it only in clinical settings which contradict with the primary objective of HCS [62].

In HCS, the measured Hgb levels depend on human interpretation, which can differ for the same color shade from person to person. A smartphone can address this problem using its camera and computation capability. Tyburski et al. developed a solution based on HCS, which can use a smartphone to measure Hgb levels from a drop of blood mixed with reagents. Their built device consisted of 2 parts- a cap and a body (Fig. 2.2) [63]. The entire tube of the body get filled by capillary action and the cap prefilled with a reagent solution is placed into the body. After an hour, the solution of blood and reagent initiates a redox reaction and undergoes a stable color change. Then the Hgb level is estimated using a color scale sticker or an image of the solution through an optional smartphone app. They validated the platform on 238 patients and
found the sensitivity of the visual interpretation and smartphone analysis of the POC device was 90.2% and 91.1% respectively, where the specificity was 83.7% and 79.2% respectively. While this minimal-invasive approach can be an excellent alternative to traditional Hgb measurement, the main drawback lies within the limitation of reagent’s expiration date, quality of the cap and body of the device, quality, and resolution of the captured image, and the identification of an exact Hb level by a visual scale.

2.3 Non-invasive systems using spectrometers

Spectrometers can also replace the functions of the photometer for Hgb level measurement and omit the need for blood sampling. Researchers have explored the use of spectrometers in the NIR region to measure Hgb levels in invasive and non-invasive ways [8, 64, 65, 66].

Figure 2.3: The response of hemoglobin, oxygenated hemoglobin, and water in the spectrum [7, 6].

The complication of the other trivial non-invasive methods can be counteracted
by switching the measurement to the absorption properties of hemoglobin in the near-infrared spectrum instead, where the wavelength lies in between 780 nm and 2500 nm. Hemoglobin and water are the major absorbers of visible and infrared light among all the components of blood, respectively, have their lowest absorption coefficient in the NIR region around 650-900 nm [67]. Thus the penetration capability of the light in this region is much higher in the spectrum of light. However, as we see in figure 2.3, in the NIR region, around 650-840nm absorption coefficient of hemoglobin is much higher than that of water. So, we can use the NIR window of 650-840nm to measure hemoglobin level where absorption of light is directly correlated to the person’s hemoglobin level. Manufacturers considering these properties manufacture different PPG sensors which can be used to measure hemoglobin by analyzing the PPG signals captured by those sensors.

Ding et al. developed a high-performance spectrophotometric system by combining a broadband light source composed of 9 LEDs in the spectrum range of 600nm 1050nm with grating spectrograph and Silicon (Si) photodiode array with 16 elements (S4111-16R, Hamamatsu Photonics Co., Ltd) (Fig. 2.4) [8]. 109 volunteers participated in the study and provided spectra from their right index fingertip. The authors used the principal component analysis (PCA) to reduce the dimension of the spectral matrices collected by their system and provided the optimal principal components as input in back propagation artificial neural network (BP ANN). When the model predicted hemoglobin value, the correlation coefficient of prediction was 0.73, standard error of prediction (SEP) was 11.47g/L, and relative SEP was 7.41%.

By analyzing the response of NIR spectrometer from the blood flow of the forearm, Edwards et al. developed a non-invasive method of measuring Hb [64]. They used six laser diode having a wavelength of 797.5, 802.5, 831.2, 848.7, 866.5, and 907 nm to calculate the changes in deoxyhemoglobin and oxyhemoglobin concentrations by monitoring the variations in the absorption of NIR light in the arm.
Figure 2.4: Hgb level measurement by transmittance mode using the NIR spectrophotometric system [8].
A spectrometer can be used to separate and measure the spectral components of the NIR window in 650nm-1000nm after passing through the blood constituents. Wang et al. captured the spectral response of 60 subjects’ finger by GE65000 spectrometer (Ocean Optics Inc., USA) and applied a back-propagation neural network on the collected response to predict hemoglobin level [65]. They trained the model on 45 subjects and predicted on 15 subjects to get an R of 0.907 and root mean square error of 3.34 mmol/dl.

The main disadvantage of using a NIR-spectrometer is its expense, and the methods need separate computing devices and these make it infeasible to use as a point of care tool. While using the smartphone instead of a NIR-spectrometer can overcome these disadvantages, smartphones offer a less sophisticated and less accurate analysis of the spectrum than spectrometers.

2.4 Non-invasive systems using PPG signals

Using NIR-spectroscopy and photodetectors, photoplethysmograms (PPG)- a measure of the variations in the volume of blood with each cardiac cycle applying optical properties of blood, can be generated. Elgendi et al. and Allen et al. presented a thorough survey on the PPG signals captured from fingertips, its various features and applications in various biomedical fields [11, 68]. In the widespread clinical application of PPG signal, it is captured using a PPG probe, or sensors, or a pulse oximeter. The basic form of PPG capturing technology requires only a light source to illuminate the tissue and a photodetector to measure the variations in the intensity of light associated with changes in the volume of the blood [68].

The feasibility of measuring blood hemoglobin level from the ratio of changing attenuation of light to changing the length of the wave during arterial pulsations in a PPG signal was demonstrated by Aldrich et al. [69]. Among the characteristics of a PPG signal, Azarnoosh et al. found peak-to-peak value to have a correlation coefficient of -0.787 in Pearson method and -0.842 in Spearman method with hemoglobin [70]. They
have recorded PPG signal from 30 subjects’ fingertips at four different wavelengths, infrared (950 nm), red (660 nm), orange (590 nm) and green (560 nm) using special PPG probe. They used a ratio of peak-to-peak ratio at a wavelength to peak-to-peak ratio at another wavelength and found the ratio between orange (590 nm) wavelength and infrared (950 nm) had the highest correlation with hemoglobin.

There are two modes of capturing a PPG signal- transmitting mode and reflective mode [71]. In the transmitting mode, the light will pass through the network consists of tissue, blood, and bone and the intensity of the light not absorbed by the network will be measured by photodetector situated on the other side of the network. In the case of reflective mode, the position of the photodiode and the light source are adjacent at a certain distance for good signal readability, and the light beam not absorbed by the network after reflection will be captured at the photodiode. Applying the reflective method, Dewantoro et al. used 4 LEDs with a wavelength of 460 nm, 515 nm, 620–625 nm, and 940 nm and a photodiode to generate PPG signals [72]. Then they calculated the AC and DC ratio of the PPG signals at all the wavelengths and used the ratio of that at any two wavelengths to obtain the hemoglobin level by applying a linear regression method. They also developed an android app to display the result immediately and store the history for a subject. The system was validated on 30 subjects, the RMS error was 1.53 gm/dl with a maximum error of 3.7 gm/dl, and the standard deviation of precision was 0.24 gm/dl

Kavsaoğlu et al. used DCM03(APMKorea, Korea)- a PPG sensor made of two LED of 660nm and 905nm to generate PPG signals from the fingertip (Fig. 2.5) [9]. They used features from 40 characteristic properties of the generated PPG signal to predict hemoglobin level by applying machine learning techniques like Partial least squares regression (PLSR), Generalized regression neural network (GRNN), and Classification and regression trees (CART). They collected data from 33 people and split the data set into 70% training-30% testing to achieve a MAPE of 0.0791-lowest among all the
The main disadvantage of using PPG sensors to measure Hgb levels is the need for extra computation devices. In most reports micro-controllers are used, which have limited computation capability. In contrast, a smartphone can perform the functions of both the PPG sensor and micro-controller. So, the use of a smartphone might increase user-friendliness and accuracy in measuring Hgb levels.

2.5 Non-invasive systems using smartphones

Availability and affordability of smartphones make them a convenient computing platform for point of care (POC) solutions. 2.6 billion people have smartphones around the world, and this availability make this tool attractive for medical diagnosis [45, 46, 47], and physiological parameter estimation [48, 49, 50, 51, 52]. Using a smartphone, Dantu et al. monitored blood glucose noninvasively by modifying and extending the Beer-Lambert law to accommodate multiple wavelengths [50]. They used the smartphone's camera as a photodetector to collect and analyze the spectrometric properties of the lights passed
through a fingertip placed on the camera of the smartphone. Using this approach, the smartphone has been used to measure different blood constituents, and in most cases, the data were collected from either fingertips or eyelids [53].

Images captured by the cameras are a representation of the exposure of different wavelength's light on the lens of the camera. Based on this principle, researchers are trying to build non-invasive hemoglobin prediction models utilizing optical responses captured by smartphones under multiple wavelengths of light. However, the Smartphone camera cannot capture spectral data as detailed as those from a spectrometer or a digital infra-red camera. To overcome the limitation of the smartphone's camera in capturing the optical response of blood in the NIR spectrum, researchers usually use an extra attachment. Dimauro et al. enhanced the performance of the smartphone's camera to capture images of the eyelids by attaching an enclosed macro-lens to the smartphone camera [73]. The extra attachment helped to segment the region of interest more precisely. The difficulties in clinical operation however, suggest this is an impractical approach.

In order to compensate for the effect of the ambient light in the images, many researchers moved the data collection site to the fingertip. Mannino et al. developed an app for smartphones where users can take a photo of their fingernail bed by the smartphone camera and get an estimation of their hemoglobin level through the app [74]. In this app, when the user captures the image of their fingernails, the app prompts the user to select regions of interest. They have collected images of fingernail bed and the hemoglobin level from standard CBC test from 337 subjects to build a model applying multi-linear regression with a bi-square weighting algorithm. When users tap on the screen corresponding to their nail beds, the app starts relating the image parameter data from the selected region of interest to hemoglobin levels collected from the standard CBC test using their algorithm. Using this model, they predicted the hemoglobin level of 100 subjects within a range of ±2.4gm/dl and bias of
0.2gm/dl, where their sensitivity to detect anemia was 97%. However, this system may suffer if the nail-beds get discolored due to any external factors like jaundice and cyanosis.

Figure 2.6: System overview of the HemaApp [10].

Apart from the natural lighting source, we can use external hardware embodiment with a smartphone camera to get a better optical response in the images. Wang et al. built HemaApp- a smartphone-based POC tool, which is a smartphone application to measure blood hemoglobin level non-invasively, where they used external hardware embodiment [10, 55]. They experimented with three different hardware embodiment- (i) white flash and an infrared emitter. (ii) incandescent lamp with white flash and an infrared emitter. (iii) white flash and custom infrared LED array. The application captures video of fingertip under these light sources and extracts the red, green, and blue time series wave from each video. They have collected data from 31 subjects to build the SVM based regression prediction model to estimate the hemoglobin level. They found that the hardware embodiment with incandescent and LEDs provided the best accuracy of ±1.26 gm/dl with an R of 0.82. However, the presence of noise in the signals might be troublesome for peak detection. So, a noise detection strategy could make the system more robust.
The need for a spectrometer can also be avoided if the RGB image captured by the smartphones can be converted into a virtual hyperspectral image. Michelle et al. generated virtual spectral images from regular images and used those converted images to train a Hgb prediction model. They found an excellent $R^2$ value of 0.94. However, the dataset was too small to develop a reliable Hgb measurement model. Additionally, in capturing the image of the eyelid, it is challenging to minimize the effect of ambient light, and results may differ for different lighting conditions [52].
Chapter 3  
Methodology

The architecture of our system is shown in Figure 3.1. This system takes about 90 seconds, from capturing fingertip video to sending back the results to the smartphone. In the next sections, we will discuss it in depth.

![Figure 3.1: System architecture (Here, ML- Machine Learning)](image)

### 3.1 Biological foundations: Absorption properties of blood in the near-infrared (NIR) spectrum

When light passes through any medium, its absorption depends on the wavelength of the light. Lights in the NIR range, between 700-1000 nm, are least absorbed by Hgb (Figure 2.3) and this range is known as the "optical window" [7]. Utilization of the penetration capability of the NIR light in this region can enable determining the amount of light absorbed by Hgb, which will tell us the Hgb level in the blood by following the Beer-Lambert law. This law states that the absorption of light is directly proportional to the concentration of the medium and path length of the light through the medium, given by
\[ I = I_0 e^{-\alpha [C] d} \]  

(3.1)

where \( I \) is the intensity of light after absorption, \( I_0 \) is the intensity of the incident light, \( \alpha \) is the absorption coefficient, \([C]\) is the concentration of the medium, and \( d \) is the path length of the light. As \( I_0, \alpha, \) and \( d \) are constant for a subject and lighting condition in equation 3.1, we can derive that:

\[ [C] \propto -I \]  

(3.2)

where \([C]\) is the hemoglobin level.

### 3.2 Hardware setup and data collection

In work reported here, we have used the fingertip as our data source. While measuring Hgb levels, the fingertip needed to be illuminated with light having a wavelength (\( \lambda \)) in the optical window demonstrated in Figure 2.3 to get a better optical response. We selected LEDs with wavelengths 850nm, 940nm, and 1070nm in the optical window due to availability and price in the market. However, these LEDs cannot remove darkness, and the smartphone cannot operate appropriately in the dark and very susceptible to noises in such conditions. So, each hardware set comprised two white LEDs along with six NIR LEDs of the same \( \lambda \) connected with a printed circuit board and finger holder (Figure 3.2). The whole setup was operated with a 12-V DC battery. Each LEDs’ power dissipation was about 80mW, and for one-hour usage, the entire setup with eight LEDs spent around 0.001 electric unit. Such use of low power ensured no effect on the users’ regular power consumption.

213 participants of our study provided four videos with Pixel2 smartphone in four different lighting conditions- three with LEDs in optical window and one with default flash. The videos were captured at 60 frames per second (fps) for 10 seconds at full HD resolution (1920 × 1080), which gave us 600 frames or sample points for each
subject. The subjects also provided a blood sample at the same visit to measure Hgb level by standard laboratory methods.

In 213 subjects of southeast Asian origin, 100 male subjects and 113 female had similar skin color, and an age range from 14 to 70 years, the average (µ) age of 39.74, and a median age of 39 years. The data with informed consent were collected from a clinical care facility in Bangladesh after approvals from the Institutional Review Board (IRB) at Marquette University (project number: PRO00020317) and the Bangladesh Medical Research Council.

3.3 Photoplethysmography (PPG)

Photoplethysmography can measure blood volume changes in the micro-vascular bed of tissue using optical properties of blood [75]. The red or near infra-red (NIR) wavelength gives better PPG responses [76, 68]. Figure 3.3 shows some
important characteristics of a PPG signal.

Figure 3.3: A PPG signal with some critical features.

In a fingertip, the blood is circulated through non-uniformly distributed capillaries, and blood volume at the micro level is not equal across all locations. To get uniform optical responses across the regions of interest, we made our regions of interest smaller by dividing each frame into $10 \times 10$ blocks. As each frame had $1920 \times 1080$ pixels, we averaged the value of $192 \times 108$ red pixels for each block to get 100 PPG signals per video (Figure 3.4).

In our system, the camera captured the optical response of NIR light passing through the blood stream of a fingertip. As the blood's volume changes with the cardiac cycle, the change in optical response following the Beer-Lambert law was captured. The camera gave us these optical responses as pixel values of the frames, and the $\mu$ value of all the pixels in a block is proportional to the blood's volume in that instant. We got a PPG by plotting $\mu$ values of the same block across all the frames of a video as a time series.
Figure 3.4: PPG signal of a block from 10x10 blocks in a video. The red dot in a block represent the average value of red pixels for that block. We get a PPG signal by taking the averages for the same block across all the frames.

Pixels generally have three values—red, green, blue to represent all the visible colors. As NIR is not visible, the NIR response is represented as the nearest similar visible color. Moreover, as the wavelengths we used were very close to red, values of red pixels represent the optical response better than other pixel colors, and we used the $\mu$ of red pixels in a block to generate the PPGs.

In analyzing the frames of the videos, we found that the PPG signals in the videos had more noise at the beginning and end of the recordings due to power interference, finger movement, and placement issues. Consequently, we selected the frames from 51 to 550 of the videos for further PPG analysis. The steps in processing the PPGs signals came next. We designated unprocessed signals as $S_u$ (Figure 3.5a).

- Considering each frame as a sample point, the value of the Nyquist frequency in
Butterworth filter (BF) was half of the sampling rate $r$ which was determined by:

$$r = \frac{\text{number of frames in the video}}{\text{Duration of the video}}$$  \hspace{1cm} (3.3)

- BF does not cut-off the frequencies sharply around the passband. We used Fast Fourier Transformation (FFT) to remove the remaining residue frequencies, and that removed any frequency outside the limit - 0.5Hz to 5Hz. We denoted the output signal of this step as $S_c$.

- The output of BF was normalized relative values. From equation 3.2 we know that, Hgb level is negatively proportional to these values. So, in the end, we multiply $S_c$ with $-1$ to get $S_{PPG}$ (Figure 3.5b). Now, Hgb level will be directly proportional to the characteristics of $S_{PPG}$.  

Figure 3.5: a) Averages of red pixels of block across all selected frame, b) Processed PPG signals from $S_u$. 

3.4 Signal Processing

PPG signals are widely used for physiological and blood constituent assessment in wearable devices. The device generated PPG signal has motion artifacts, which is a critical problem. In this section, we discuss several signal processing strategies that are frequently used for PPG signal analysis to identify and remove the issues coming from finger movement, signal noise, and motion artifact and statistical tools.

3.4.1 Fourier analysis

The PPG signal is non-stationary and quasi-periodic where Fourier analysis applies to stationary periodic signals [71]. We can apply Fourier series on PPG signals on a cycle-by-cycle basis. To remove high-frequency noise of a PPG signal, we can filter the data using smoothing filters like Savitzky-Golay (SG) smoothing filter, Butterworth filter, Gaussian filter. Later, the cycle-by-cycle Fourier series (CFS) analysis can be carried out. Studies showed that this method reduced the measurement error of the PPG signal from 37% to 3% [77].

Figure 3.6: Detected PPG cycles in $S_{PPG}$
3.4.2 **Savitzky-Golay smoothing filter**

Savitzky-Golay developed a data smoothing filter using a least-squares polynomial approximation. They fitted a polynomial to a set of input data and evaluated the resulting polynomial at a single point, maintaining the shape and magnitude of the waveform peaks while smoothing the waveform. Building a low pass filter using this approach is known as Savitzky-Golay smoothing filter [78].

3.4.3 **Wavelet Transform (WT)**

Biological signals are usually non-stationary by the property, and they tend to change substantially over time. For this reason, WT being a time-frequency method, works better with noise reduction and signal enhancement for PPG signals than other traditional methods [71]. Lee et al. used a stationary wavelet transform (SWT), and the wavelet transforms modulus maxima (WTMM) to reduce motion artifacts in PPG signals [79]. They have achieved an 87% reductions in HR estimation error, 76% in HRV estimation error, and 66% in instantaneous HR error. Moreover, Teng et al. used a continuous wavelet transform (CWT) to find the accurate position of peak and foot of the PPG signal [80]. Wu et al. and Bousefsaf et al. used CWT on PPG signal to measure HR [81] [82]. However, Foo et al. found that WT has limited capability in restoring corrupted PPG signal for both Heart Rate (HR) and Pulse Transmit Time (PTT) measurements [83]. They observed that WT’s inherent algorithm could induce unwarranted phase variability, which can compromise the clinical interpretation of HR and PTT. There are also some advanced technique like Synchrosqueezing Transform derived from WT has been used with PPG signal [84].

3.4.4 **Independent Component Analysis (ICA)**

Independent Component Analysis (ICA) can separate the additive non-Gaussian sub-components of a multivariate signal [85]. As motion artifacts in a PPG signal comes from an independent source, can be separated using ICA. It can also
be used to separate the effect of ambient light and other interference. Kim et al. used a combination of ICA and block interleaving with low-pass filtering to reduce motion artifact in PPG signals [86]. Holton et al. compared ICA with principal component analysis- another source separation technique for their effectiveness in PPG signal recovering from video recordings and found that ICA produces the most consistent result [87].

### 3.4.5 Butterworth Filter

Butterworth filter makes the frequency response of a signal as flat as possible in the passband. It is called a maximally flat filter because it does not have any ripple in the passband or the stopband [88]. It can be used as highpass, low-pass, or bandpass filter. Bonissi et al. used the Butterworth filter to preprocess the PPG signal on continuous authentication methods for PPG biometrics [89]. Luke et al. used an algorithm that consists of both Butterworth low pass filtering and wavelet transform to remove motion artifacts from PPG data [90]. While trying to monitor blood pressure using a PPG sensor, Riaz et al. used the Butterworth filter to preprocess the PPG signal [91]. However, Moreno et al. found that the Butterworth filter distorted the PPG signal to give a wrong number of peaks [92].

### 3.5 PPG cycle detection

A PPG cycle is analogous to a cardiac cycle that repeats in a PPG signal. In Figure 3.5(b), all the PPG cycles in $S_{PPG}$ do not have PPG cycles with usual critical features like diastolic peak or dicrotic notch (Figure 3.3). In our earlier work [6], we detected and selected these cycles manually by spotting them in the plots, which is not an optimal solution. To develop a generally usable system, we needed to detect these cycles automatically. In order to detect PPG cycles in $S_{PPG}$, we implemented the following logic-

1. In a PPG cycle, starting point, dicrotic notch, and ending point are consecutive
minima and systolic peak, and diastolic peak are consecutive maxima points.

2. Systolic peak height must be greater than the diastolic peak height.

3. The height of the dicrotic notch must be greater than that of the starting point and ending point.

4. The time elapsed for a PPG cycle must be within 20% of the time needed for a heartbeat of that subject which can be determined from FFT of the PPG signal. Here, we allowed a threshold of 20% error in determining the time needed for a heartbeat or PPG cycle.

![Figure 3.7: Selected PPG cycles are merged together.](image)

We reduced our search space by using `find_peak` from the `numpy` module of python to detect the peaks in the PPGs and then iterated through the peaks to detect the PPG cycles. Figure 3.6 depicts detected PPG cycles in a $S_{PPG}$.

### 3.6 PPG selection and feature extraction

As we need three PPG cycles to extract all the features, we selected three PPG cycles based on the following criteria from all the detected PPG cycles in each $S_{PPG}$.
• If at least three PPG cycles are detected, then select three PPG cycles based on descending systolic heights to select the PPG cycles with strong optical responses.

• If less than three PPG cycles are detected, then take all the available PPG cycles and replicate the PPG cycle with maximum systolic height to get three PPG cycles.

• If no PPG can be detected in $S_{PPG}$ for a block, then discard the block from further analyses.

The selected PPG cycles were merged together to get $S_f$ for feature extraction (Figure 3.7). While combining the PPG cycle can create discontinuity in the signal, the feature values are robust to these changes. For each block, we enhanced the analytic system of Elgendi et al. [11] to extract 45 features from $S_f$, as well as the 1st, and 2nd derivative of $S_f$ (Table 3.1).

<table>
<thead>
<tr>
<th>Feature Number</th>
<th>Feature Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0 : x$</td>
<td>Systolic peak height</td>
</tr>
<tr>
<td>$f_1 : y$</td>
<td>Diastolic peak height</td>
</tr>
<tr>
<td>$f_2 : z$</td>
<td>Dicrotic notch height</td>
</tr>
<tr>
<td>$f_3 : t_{pi}$</td>
<td>Pulse interval</td>
</tr>
<tr>
<td>$f_4 : y/x$</td>
<td>Augmentation index</td>
</tr>
<tr>
<td>$f_5 : (x - y)/x$</td>
<td>Relative augmentation index</td>
</tr>
<tr>
<td>$f_6 : z/x$</td>
<td>Ratio of $z$ and $x$</td>
</tr>
<tr>
<td>$f_7 : (y - z)/x$</td>
<td>Negative relative augmentation index</td>
</tr>
<tr>
<td>$f_8 : t1$</td>
<td>Systolic peak time</td>
</tr>
<tr>
<td>$f_9 : t2$</td>
<td>Dicrotic notch time</td>
</tr>
<tr>
<td>$f_{10} : t3$</td>
<td>Diastolic peak time</td>
</tr>
</tbody>
</table>
\( f_{11} : \Delta T = t_3 - t_1 \)  Time between systolic and diastolic peak
\( f_{12} : \frac{t_1}{2} \)  Time between half systolic peak points
\( f_{13} : \frac{A_2}{A_1} \)  Inflection point area ratio
\( f_{14} : \frac{t}{x} \)  Systolic peak rising slope
\( f_{15} : \frac{y}{t_{pi} - t_3} \)  Diastolic peak falling slope
\( f_{16} : \frac{t_1}{t_{pi}} \)  Ratio of \( t_1 \) and pulse interval time (\( t_{pi} \))
\( f_{17} : \frac{t_2}{t_{pi}} \)  Ratio of \( t_2 \) and pulse interval time (\( t_{pi} \))
\( f_{18} : \frac{t_3}{t_{pi}} \)  Ratio of \( t_3 \) and pulse interval time (\( t_{pi} \))
\( f_{19} : \frac{\Delta T}{t_{pi}} \)  Ratio of \( \Delta T \) and pulse interval time (\( t_{pi} \))
\( f_{20} : t_{a1} \)  Interval time from the first PPG cycle start point (\( l_1 \)) in the first derivative of PPG (\( S_f \)) to first maxima (\( a_1 \)) of \( S_f \).
\( f_{21} : t_{b1} \)  Interval time from point \( l_1 \) to first minima of the first PPG cycle (\( b_1 \)) in the \( S_f \).
\( f_{22} : t_{e1} \)  Interval time from point \( l_1 \) to the second maxima of the first PPG cycle (\( e_1 \)) in the \( S_f \).
\( f_{23} : t_{f1} \)  Interval time from point \( l_1 \) to the second minima of the first PPG cycle (\( f_1 \)) in the \( S_f \).
\( f_{24} : \frac{b_2}{a_2} \)  Ratio of first minima (\( b_2 \)) and first maxima (\( a_2 \)) in the second derivative of the PPG signal (\( S_f^{2} \)).
\( f_{25} : \frac{e_2}{a_2} \)  Ratio of the second maxima (\( e_2 \)) in \( S_f^{2} \) and \( a_2 \).
\( f_{26} : \frac{b_2 + e_2}{a_2} \)  Ratio of (\( b_2 + e_2 \)) and \( a_2 \) [40]
\( f_{27} : t_{a2} \)  Interval time from the second PPG cycle start point (\( l_2 \)) in the second derivative of PPG to \( a_2 \).
\( f_{28} : t_{b2} \)  Interval time from point \( l_2 \) to \( b_2 \).
\( f_{29} : \frac{t_{a1}}{t_{pi}} \)  Ratio of \( t_{a1} \) and \( t_{pi} \)
\( f_{30} : \frac{t_{b1}}{t_{pi}} \)  Ratio of \( t_{b1} \) and \( t_{pi} \)
\( f_{31} : \frac{t_{e1}}{t_{pi}} \)  Ratio of \( t_{e1} \) and \( t_{pi} \)
\( f_{32} : \frac{t_{f1}}{t_{pi}} \)  Ratio of time interval of \( l_1 \) (\( t_{f1} \)) and \( t_{pi} \)
\[ f_{33} : \frac{t_{a2}}{t_{pi}} \quad \text{Ratio of } t_{a2} \text{ and } t_{pi} \]
\[ f_{34} : \frac{t_{b2}}{t_{pi}} \quad \text{Ratio of } t_{b2} \text{ and } t_{pi} \]
\[ f_{35} : \frac{t_{a1} + t_{a2}}{t_{pi}} \quad \text{Ratio of } (t_{a1} + t_{a2}) \text{ and pulse interval } (t_{pi}) \]
\[ f_{36} : \frac{t_{b1} + t_{b2}}{t_{pi}} \quad \text{Ratio of } (t_{b1} + t_{b2}) \text{ and pulse interval } (t_{pi}) \]
\[ f_{37} : \frac{t_{e1} + t_{e2}}{t_{pi}} \quad \text{Ratio of } (t_{e1} + t_{e2}) \text{ and pulse interval } (t_{pi}) \]
\[ f_{38} : \frac{t_{f1} + t_{f3}}{t_{pi}} \quad \text{Ratio of } (t_{f1} + t_{f3}) \text{ and pulse interval } (t_{pi}) \]
\[ f_{39} : f_{\text{base}} \quad \text{Fundamental component frequency obtained from Fast Fourier Transformation (FFT)} \]
\[ f_{40} : |s_{\text{base}}| \quad \text{Fundamental component magnitude from FFT} \]
\[ f_{41} : f_{\text{2nd}} \quad \text{Second component frequency obtained from FFT. Such that, } f_{\text{base}} < f_{\text{2nd}} \]
\[ f_{42} : |s_{\text{2nd}}| \quad \text{Second component magnitude from FFT} \]
\[ f_{43} : f_{\text{3rd}} \quad \text{Third component frequency obtained from FFT. Such that, } f_{\text{base}} < f_{\text{2nd}} < f_{\text{3rd}} \]
\[ f_{44} : |s_{\text{3rd}}| \quad \text{Third component magnitude acquired from FFT} \]

### 3.6.1 Features from PPG

Noteworthy PPG features extracted from the original PPG signal are discussed next.

**Systolic peak height**

As shown Fig. 3.8, the systolic peak height (x) indicates the pulsatile changes in blood volume caused by arterial blood flow around the measurement site of the PPG [93, 94]. Systolic peak height has been related to stroke volume [95].

**Pulse area**

The as the total area under the PPG curve is used to measure pulse area. The PPG area response to skin incision is found to differ between movers and non-movers [96].
Figure 3.8: A typical waveform of the PPG and its characteristic parameters, whereas the amplitude of the systolic peaks is $x$ while $y$ is the amplitude of the diastolic peak. [11].

The pulse area can be divided into two areas at the dicrotic notch. Wang et al. found that the ratio of the two areas can be used as an indicator of total peripheral resistance (Fig. ??). This ratio is called the inflection point area ratio (IPA).

**Peak to peak interval**

The distance between two consecutive systolic peaks will be referred to as Peak-Peak interval (Fig. 3.10). The ECG signal's R-R interval is correlated with the Peak-Peak interval of the PPG signal as both represent a completed heart cycle. The Peak-Peak interval can be used to detect the heartbeat in PPG signals [97, 98].

**Pulse interval**

Pulse interval is the distance between the beginning and the end of the PPG waveform (Fig. 3.10). The Pulse interval is usually used instead of the peak to peak interval when the diastolic peaks are more clear and easier to detect compared to the systolic peak.

The ratio of Pulse interval to its systolic amplitude can provide an understanding of the properties of a person's cardiovascular system [99]. In 2008, Lu et
Figure 3.9: Original fingertip photoplethysmogram. A1 and A2 are the areas under the whole PPG wave separated at the point of inflection. Thus, the inflection point area ration can be calculated as the division of A2 by A1. [11].

Figure 3.10: Pulse interval for a PPG signal and peak to peak interval for two consecutive peaks.

al. [100] demonstrated that HRV in PPG and ECG signals are highly correlated by using the Pulse interval in PPG signals, and R-R intervals in ECG signals. They strongly suggested that PPG signals could be used as an alternative measurement of HRV.

3.6.2 Features from 1st derivative PPG

Diastolic point definition

Diastolic point can be defined as the point at which the 1st derivative of the waveform is closest to zero [101] (Fig. 3.11).
Figure 3.11: (a) Original fingertip photoplethysmogram (b) 1st derivative wave of photoplethysmogram [11].

**ΔT calculation**

ΔT is the peak-to-peak time which is related to the time taken for the pressure wave to propagate from the heart to periphery and back. The time between the systolic and diastolic peaks is ΔT. The definition of ΔT depends on the PPG waveform as its contour varies with subjects. When there is a second peak as in Fig. 3.11, ΔT is defined as time between the two maxima.

In other words, ΔT is the time between the two positive to negative zero-crossings of the derivative as in Fig. 3.11. However, in some PPG waveforms, there is no clear second peak. In this case, ΔT is defined as the time between the peak of the waveform and the inflection point on the down slope of the waveform which is a local maximum of the first derivative.

**3.6.3 Features from 2nd derivative PPG**

The second derivative is more commonly used than the first derivative. In literature, the second derivative of photoplethysmogram is also called the acceleration plethysmogram (APG) because it is an indicator of the acceleration of the blood in the finger.
As shown in Fig. 3.12, The waveform of the APG includes four systolic waves and one diastolic wave, namely a-wave (early systolic positive wave), b-wave (early systolic negative wave), c-wave (late systolic reincreasing wave), d-wave (late systolic redrecreasing wave) and e-wave (early diastolic positive wave). The e-wave represents the dicrotic notch as shown in Fig. 3.12. Its location corresponds to the closure of the aortic valve and subsequent retrograde blood flow, and can be used to monitor cardiac function [102].

The height of each wave was measured from the baseline, with the values above the baseline being positive and those under it negative. The ratios of the height of the each wave to that of the a-wave (b/a, c/a, d/a and e/a) are usually used for wave analyses [103]. The second derivative of the finger PPG waveform is used to stabilize the baseline and enable the individual features to be visualized and detected easily.
For $m$ available blocks for a video we obtained feature matrix $F_{m \times 45}$. The values in the columns $f_1, f_2, f_3, \ldots, f_{45} \in F$ may differ (Figure 3.13). So, to generate the final feature vector $\hat{f}$ for the subject, a column $f_i \in F$ underwent the following steps:-

- Created ten bins in between the minimum and maximum value of $f_i$.
- Put $f_{1,i}, f_{2,i}, \ldots, f_{m,i} \in f$ into the corresponding bin.
- Averaged all the members of largest bin and immediate adjacent bins.

We combined all $\hat{f}$ vectors of $n$ subject to get the final feature matrix $\hat{F}$ with a dimension of $n \times 45$. As we had four sets of videos of all the subjects under four different lighting settings, now we had four feature matrices:- $\hat{F}_{850}, \hat{F}_{940}, \hat{F}_{1070},$ and $\hat{F}_{Pixel2}$, one for each light setting.
In theory, Hgb level can be measured from equation 3.4 where $Hgb$ represents hemoglobin level, $m_{Hgb}$ represents mass of hemoglobin, and $V$ represents blood volume.

$$Hgb = \frac{m_{Hgb}}{V} \quad (3.4)$$

Blood is mostly comprised of Hgb and water. About 55% of the total volume of blood is plasma, which has 92% water. 40-45% of the total volume of blood is red-blood cells which have 95% of Hgb [104]. Now, equation 3.4 can be rewritten as:

$$Hgb = \frac{m_{Hgb}}{V_{Hgb} + V_w} \quad (3.5)$$

where, $V_{Hgb}$ is volume of hemoglobin and $V_w$ is the volume of water present in the blood.

From equation 3.2 we know that the intensity of the light received after absorption in the medium is negatively proportional to the concentration or mass of the medium. Figure 2.3 shows that light of 850 nm and 1070 nm wavelength is very much responsive to hemoglobin and water, respectively. As the volume of water in the tissues is constant over time, the effect will be neutralized when Butterworth Filter normalized the PPG signal, and $I_{1070}$ will depend on water in blood only. Thus, equation 3.5 becomes:

$$Hgb = \frac{I_{850}}{I_{850} + I_{1070}} \quad (3.6)$$

where, $I_{850}$ and $I_{1070}$ are the intensities of light after passing through the finger at exposure wavelengths 850 nm and 1070 nm respectively.

Based on equation 3.6, we derived a new feature set $\hat{F}_D$ from $\hat{F}_{850}$, and $\hat{F}_{1070}$ by applying $\frac{\hat{F}_{850}}{\hat{F}_{850} + \hat{F}_{1070}}$. We fed feature sets $\hat{F}_{850}, \hat{F}_{940}, \hat{F}_{1070}, \hat{F}_{Pixel2},$ and $\hat{F}_D$ into Support Vector Regression (SVR) to build the models $M_{850}, M_{940}, M_{1070}, M_{Pixel2},$ and $M_D$, which
will give hemoglobin level measurements.
Chapter 4
Hemoglobin Measurement

4.1 Error Metrics

There are multiple metrics or measurement techniques which have been used to assess accuracy of laboratory tests such as hemoglobin measurement. The commonest have been bias and standard deviation (S.D.) [105, 106]. Bland-Altman plots, mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), root mean square deviation (RMSD), correlation coefficient (R), r-squared (R2), mean average percentage error (MAPE) are also regularly used. Succinctly, the advantages and disadvantages of each of these are as follows:

4.1.1 Bias

Bias is the mean of differences between an estimated measurement and actual measurement, which can be expressed by the following formula:

$$
\bar{d} = \frac{1}{n} \sum_{k=1}^{n} (X_k - Y_k)
$$

where, n is the number of samples, X is the estimated and Y is the actual measurement. Bias is used with Bland-Altman plot to describe the agreement between the measurements. Having a low bias means both the measurements are in great agreement and a high bias means the opposite. For estimations, a low bias can mean that the estimations are remarkably close to the actual value. A low bias is a desirable outcome but sometimes it can be misleading. A positive difference can minimize the effect of a negative difference in the bias calculation. Moreover, comparing the performance of two methods using bias can be misleading too when the methods are working with different datasets of the same type. The model dealing with a dataset having larger numbers or bigger average can have more bias than that of the other model, and still can be a better model in terms of overall performance if we consider the smaller numbers only. So, we need to observe other error metrics along with bias to compare performance of different models.
4.1.2 **Standard Deviation (SD) of the errors**

Standard Deviation (SD) is used with bias to determine the limits of agreement which gives us insight about the spread of the errors. SD of the errors for an estimated measurement is calculated by the following formula:

\[ d = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} (X_k - Y_k)^2} \]

where, \( \bar{d} \) is the bias and \( X \) is the estimated and \( Y \) is the actual measurement.

A high SD indicates that the model has large errors for some samples and a low SD indicates similar error for all the samples. SD always needs bias to describe the performance of the models and vice versa. A combination of low bias and SD is always desirable. However, the combination of bias and SD may not work while comparing the performance of two methods working on different datasets of the same type for the similar reasons described earlier.

4.1.3 **Mean Square Error (MSE)**

MSE is a version of bias where the positive and negative differences cannot cancel each other. We can measure MSE with the following formula:

\[ \text{MSE} = \frac{1}{n} \sum_{k=1}^{n} (X_k - Y_k)^2 \]

where, \( X \) is the estimated and \( Y \) is the actual measurement.

While MSE neutralize the trouble of positive and negative error, it amplifies the problem with the magnitude of the samples in the dataset. As it squares each error, errors greater than one will be less than the squared error and less than one will be greater than the squared error. By estimating the mean for all the samples, a model working with a normally distributed dataset having standard deviation less than 1 can have a lower MSE than a model working with a dataset having standard deviation greater than 1 which is estimating different numbers for different samples. So, when comparing the performance of two methods working on different datasets of the same type, MSE may fail to distinguish the better performing model.
4.1.4 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is square root of MSE. It has similar advantages and disadvantages to MSE.

4.1.5 Root Mean Square Deviation (RMSD)

Root Mean Square Deviation is exactly same to RMSE.

4.1.6 Mean Absolute Percentage Error (MAPE)

MAPE is used as a loss function and in model evaluation. It returns the percentage. It usually expresses the accuracy as a ratio defined by the formula:

\[ M = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{X_t - Y_t}{Y_t} \right| \]

where X is the estimated and Y is the actual measurement. The MAPE is also sometimes reported as a percentage, which is the above equation multiplied by 100. As MAPE uses absolute value of the errors, MAPE does not have any issue with positive and negative errors. Moreover, MAPE normalizes the error by the original values which offsets the largeness or smallness of the original values. It can differentiate the accuracy of the different system having different dataset. For these reasons, MAPE can be considered as a better metric to compare performances among different systems, and we will focus on MAPE.

4.1.7 \( R^2 \) (R-squared) - Coefficient of determination

Any regression model needs to evaluate to know that how good it fits the dataset. \( R^2 \) can determine how well the model is performing against the baseline model which always predicts the mean of the dataset. \( R^2 \) can be measured with the following formulas:

\[ \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \quad SS_{tot} = \sum_i (y_i - \bar{y})^2, \quad SS_{res} = \sum_i (y_i - \hat{y}_i)^2, \quad R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \]

where, x is the estimated and y is the actual measurement. The highest and best value of \( R^2 \) can be 1 which means the estimated values are same as the actual values. A low value of \( R^2 \) may indicate the bad performance of the model which is not always true. A good model can have a low \( R^2 \) value due to low variance in the actual data. So, while a high \( R^2 \) value is desirable, we cannot discard a model for having a low \( R^2 \) value.
4.1.8 Correlation Coefficient (R)

The sample correlation coefficient (R) is a measure of the closeness of association of the points in a scatter plot to a linear regression line based on those points. Possible values of the correlation coefficient range from -1 to +1, with -1 indicating a perfectly linear negative, i.e., inverse, correlation (sloping downward) and +1 indicating a perfectly linear positive correlation (sloping upward). When we estimate any value with model, we try to evaluate the model by determining the correlation coefficient between the measured value, x and actual value, y. While its necessary to have at least a moderate high R value to substantiate the stability of the model, having a high value of R does not guarantee high accuracy or low accuracy. It tells us how closely x follows y when y increases or decreases.

We have elected to use MAPE and correlation coefficient as our primary endpoint metrics for the aforementioned reasons.

Table 4.1. Characteristics of the studied population and its laboratory hemoglobin levels.

<table>
<thead>
<tr>
<th>Dataset type</th>
<th>Count</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Maximum (g/dl)</th>
<th>Minimum (g/dl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>213</td>
<td>11.28</td>
<td>1.08</td>
<td>14.4</td>
<td>6.7</td>
</tr>
<tr>
<td>Male</td>
<td>100</td>
<td>11.85</td>
<td>0.85</td>
<td>14.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Female</td>
<td>113</td>
<td>10.76</td>
<td>0.93</td>
<td>12.1</td>
<td>6.7</td>
</tr>
</tbody>
</table>

4.2 Results

We collected four sets of videos from 213 subjects by following the procedures described in section 3.2. Table 4.1 data describe the standard information of our subjects by gender. Distribution of the Hgb levels of all subjects are shown in Figure 4.1, the most prevalent Hgb levels are 11.2 g/dl and 12.1 g/dl. The by-gender distribution of Hgb levels is shown in Figure 4.2; the Hgb levels of female subjects cover a wider range than those of male subjects.
Distribution of the Hgb levels of all subjects are shown in Figure 4.1, the most prevalent Hgb levels are 11.2 g/dl and 12.1 g/dl. The by-gender distribution of Hgb levels is shown in Figure 4.2; the Hgb levels of female subjects cover a wider range than those of male subjects.

We extracted features for each of the models from $S_f$. These model features were ranked by $f_{score}$ against the lab measured Hgb level using SelectKBest method from scikit-learn. For $j$th feature in $\hat{F}$, if $X_{1,j}, X_{2,j}, X_{3,j}, \ldots, X_{n,j} \in X_{:,j}$ are the feature values and $y_1, y_2, y_3, \ldots, y_n \in y$ are the actual Hgb levels then the $f_{score}(fs)$ can be determined by the equation 4.1.

$$f_{sj} = \sum_{i=1}^{n} \frac{(X_{i,j} - \bar{X}_{:,j})(y_i - \bar{y})}{SD(X_{:,j})SD(y)}$$

where,

$$\bar{X}_{:,j} = \frac{1}{n} \sum_{i=1}^{n} X_{i,j},$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i,$$

$$SD(X_{:,j}) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_{i,j} - \bar{X}_{:,j})^2},$$
In equation 4.1, the numerator determines the discrimination between the feature and target, and the denominator determines the one within the feature and target. The larger the \( f_{score} \) is, the stronger the correlation between the feature and target, and the feature is better for the model. Based on \( f_{score} \), we selected \( k \in \{3, 4, \ldots, 45\} \) features from the feature set \( \hat{F} \) for each of the available models and tried to find out which collection of features gave the best result. Table 4.2 shows the ranking of the features in descending order and the numbers in that table represent the features in table 3.1. For example, the top 3 features with highest \( f_{score} \) for M850 are \( f_{25}, f_{24}, \) and \( f_{15} \). From table 3.1, we can see that these are ratio of \( e_2 \) and \( a_2 \), ratio of \( b_2 \) and \( a_2 \), and diastolic peak falling slope. This ranking was used to select top \( k \) features while building the model.

We used the SVR method from scikit-learn module of python [107], where the performance depends on the value of parameters. Towards that end, for each set of features, we found the best set of parameters based on the lowest MSE in the whole
dataset from a possible set of parameters using the GridSearchCV method. Using those parameters with SVR, models were built to measure Hgb level. The performance of the models was then evaluated by 10-fold cross-validation using scikit-learn. The datasets were randomly shuffled and split into ten unique sets where iteratively, each set was used as a test set and the rest of the data as the training set. The results for a dataset were compiled by merging these 10 test sets’ results. We compare the results of all models as well as the average line considered as a model, which we will refer to as $M_A$. In $M_A$, all the subjects have the same prediction value, which is the dataset’s average Hgb value.

### 4.2.1 Dataset: All subjects

213 subjects are studied in this dataset, 100 male and 113 female (Table 4.1). $S_f$ for only 206 subjects were available in all four light settings. Since we wished to compare the result of all models, the results here are based on these 206 subjects. For this 206 subject set, the Hgb level had an $\mu$ of 11.28 g/dl and a standard deviation ($\sigma$) of 1.046 g/dl.

---

**Table 4.2. Sorted feature numbers for the models based on $f_{score}$**

<table>
<thead>
<tr>
<th>Model</th>
<th>Feature numbers from Table 3.1 in descending order of $f_{score}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{850}$</td>
<td>26, 25, 16, 8, 4, 5, 7, 36, 6, 19, 1, 11, 35, 27, 41, 13, 33, 24, 40, 2, 28, 20, 15, 30, 32, 34, 17, 23, 38, 43, 0, 39, 21, 10, 42, 44, 37, 12, 31, 18, 22, 29, 14, 9, 3</td>
</tr>
<tr>
<td>$M_{940}$</td>
<td>7, 23, 33, 40, 19, 16, 9, 5, 13, 41, 34, 25, 26, 4, 31, 3, 20, 11, 38, 10, 1, 8, 6, 44, 39, 36, 17, 22, 42, 35, 0, 12, 1, 15, 30, 2, 21, 29, 24, 28, 43, 8, 32, 27, 37</td>
</tr>
<tr>
<td>$M_{1070}$</td>
<td>13, 37, 11, 38, 19, 10, 31, 9, 18, 42, 17, 44, 4, 5, 22, 20, 28, 6, 39, 30, 34, 21, 12, 3, 43, 33, 27, 16, 0, 41, 14, 26, 2, 40, 25, 32, 24, 7, 35, 1, 23, 8, 29, 36, 15</td>
</tr>
<tr>
<td>$M_{Pixel2}$</td>
<td>29, 5, 42, 22, 20, 4, 23, 30, 44, 21, 28, 6, 43, 1, 13, 39, 37, 31, 2, 15, 25, 34, 41, 32, 38, 11, 26, 3, 8, 40, 9, 16, 7, 14, 0, 19, 18, 35, 12, 24, 17, 10, 27, 36, 33</td>
</tr>
<tr>
<td>$M_D$</td>
<td>27, 8, 42, 37, 31, 13, 33, 30, 28, 22, 38, 17, 10, 34, 9, 36, 32, 24, 18, 43, 41, 21, 7, 16, 26, 35, 1, 39, 3, 40, 25, 11, 5, 23, 2, 20, 4, 19, 6, 15, 14, 12, 29, 0, 44</td>
</tr>
</tbody>
</table>
We built the models by selecting best $k \in \{3, 4, \ldots, 45\}$ features following the serial of the features given in Table 4.2 for $M_{850}$, $M_{940}$, $M_{1070}$, $M_{\text{Pixel2}}$, and $M_D$. We did 10-fold cross validation to find which set of the features gave the best results, and the results are depicted in Figure 4.3. In this figure, Figure 4.4, and 4.5, values of Mean square error (MSE), Mean absolute percentage error (MAPE), and Correlation coefficient (R) are plotted against increasing count of features. From these figures, we can understand the impact of various combination of features on the models. With an essentially flat line, the varying count of features do not provide new information to the models, and the models cannot measure Hgb accurately.

![Graphs showing performance of models for increasing number of top features](image)

Figure 4.3: Performance of the models for increasing number of top features for all subjects.

The performance of the models, except $M_{940}$, did not vary much with the increasing number of top features. The MSE was $\sim 1.1 \text{g}^2/\text{dl}^2$, and MAPE was $\sim 7.32\%$ in most cases, but $M_{940}$ had the lowest MSE of 1.01 for top 32 features, and the lowest MAPE of 7.12\% with top 6 features. Moreover, in most cases, the $\sigma$ in the measured Hgb levels were very low which means that the model predicted around the $\mu$ of the original
Table 4.3 Best result in Hgb measurement with all subjects without the outliers.

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE ↓</th>
<th>MAPE ↓</th>
<th>R ↑</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>M850</td>
<td>0.838 (5)</td>
<td>6.58 (4)</td>
<td>0.169 (4)</td>
<td>0.02 (4)</td>
</tr>
<tr>
<td>M940</td>
<td>0.809 (24)</td>
<td>6.53 (24)</td>
<td>0.297 (24)</td>
<td>0.179 (24)</td>
</tr>
<tr>
<td>M1070</td>
<td>0.84 (4)</td>
<td>6.547 (4)</td>
<td>0.2 (4)</td>
<td>0.004 (4)</td>
</tr>
<tr>
<td>MPixel2</td>
<td>0.812 (20)</td>
<td>6.43 (20)</td>
<td>0.248 (20)</td>
<td>0.083 (20)</td>
</tr>
<tr>
<td>MD</td>
<td>0.86 (-)</td>
<td>6.67 (-)</td>
<td>0.12 (12)</td>
<td>0.001 (-)</td>
</tr>
<tr>
<td>MA</td>
<td>0.86 (N/A)</td>
<td>6.65 (N/A)</td>
<td>N/A</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Hgb values only and the model is very underfitted.

Table 4.3 displays the best results for all the subjects after removing the outliers for all the models. In tables 4.3-4.5, we used the number in parenthesis for a result to indicate the number of features used to get that result. Inside the parenthesis, a "-" means that the result is the same for all number of top features. An up arrow beside a performance metric means that a higher value of that performance metric indicates a better result. Similarly, the down arrow means the lower, the better.

We identified four subjects as outliers (Figure 4.1) as their value is far away from the average Hgb value with respect to σ having one occurrence. Hgb levels of the remaining 202 subjects without the outliers had an μ of 11.29 g/dl and a σ of 0.929 g/dl. In this scenario, M940 had lowest MSE of 0.809 gm²/dl² with top 24 features, MPixel2 had lowest MAPE of 6.43% with top 20 features, and M940 had highest R-value of 0.297 with top 24 features. As for the full dataset, the (σ) in the measured Hgb levels was very low, and the performance of M850, M1070, and MD did not change with varying number of features in the models.

4.2.2 Dataset: Female subjects

113 female subjects were studied separately, which is a subset of 4.2.1; the number of female subjects that had $S_f$ available in all four light settings was 109. As in 4.2.1, we compared the results of all models based on these 109 common subjects. The Hgb level of these 109 had an μ of 10.76 g/dl and a σ of 0.93 g/dl. The distribution of the
Hgb levels of this subset is given in Figure 4.2.

Table 4.4. Best results in models with female subjects.

<table>
<thead>
<tr>
<th>Model</th>
<th>With or without outlier</th>
<th>MSE ↓</th>
<th>MAPE ↓</th>
<th>R ↑</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{850}$</td>
<td>Outlier</td>
<td>0.86 (8)</td>
<td>6.80 (8)</td>
<td>0.25 (5)</td>
<td>0.02 (0.26)</td>
</tr>
<tr>
<td>$M_{1070}$</td>
<td>Outlier</td>
<td>0.85 (7)</td>
<td>6.95 (7)</td>
<td>0.19 (7)</td>
<td>0.06 (7)</td>
</tr>
<tr>
<td>$M_{Pixel2}$</td>
<td>Outlier</td>
<td>0.85 (9)</td>
<td>6.76 (9)</td>
<td>0.25 (3)</td>
<td>0.06 (9)</td>
</tr>
<tr>
<td>$M_{940}$</td>
<td>Outlier</td>
<td>0.75 (25)</td>
<td>6.26 (25)</td>
<td>0.52 (21)</td>
<td>0.73 (21)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.49 (29)</td>
<td>5.34 (29)</td>
<td>0.46 (29)</td>
<td>0.42 (29)</td>
</tr>
<tr>
<td>$M_{D}$</td>
<td>Outlier</td>
<td>0.68 (13)</td>
<td>6.12 (17)</td>
<td>0.46 (13)</td>
<td>0.31 (17)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.483 (30)</td>
<td>5.48 (30)</td>
<td>0.49 (30)</td>
<td>0.487 (30)</td>
</tr>
<tr>
<td>$M_{A}$</td>
<td>Outlier</td>
<td>0.86 (N/A)</td>
<td>6.96 (N/A)</td>
<td>N/A</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.62 (N/A)</td>
<td>6.03 (N/A)</td>
<td>N/A</td>
<td>0.79 (N/A)</td>
</tr>
</tbody>
</table>

Following same procedures described in section 4.2.1, the models were generated again for these 106 female subjects (Figure 4.4). Here, the performance of the models is modestly better. The lowest MSE was ~0.68 gm$^2$/dl$^2$ for $M_{D}$ with top 13 features, the lowest MAPE was ~6.12% for $M_{D}$ with top 17 features, and highest R was ~0.52 for $M_{940}$ with 21 features. Moreover, $\sigma$ in the measured Hgb levels had improved compared with the earlier results.

In the case of female subjects, we found two outliers following same argumentation used for the subjects in section 4.2.1 (Figure 4.2). As we see in Figure 4.4, performance of $M_{940}$ and $M_{D}$ is noticeably better than the other models. So, at this stage, we compared the performance of $M_{940}$ and $M_{D}$ only and presented in Table 4.4 where other models are omitted. Without the outliers, 111 common female subjects having $S_f$ for at least one block had an average Hgb level of 10.84 g/dl and a $\sigma$ of 0.79 g/dl. $M_{D}$ had lowest MSE of 0.48 g$^2$/dl$^2$ with top 30 features, $M_{940}$ had lowest MAPE of 5.34% with top 29 features, and $M_{D}$ had highest R-value of 0.49 with top 30 features. In contrast to dataset 4.2.1, here, the $\sigma$ in the measured Hgb levels and the performance of $M_{940}$, and $M_{D}$ are much better than $M_{A}$. 

Figure 4.4: Performance of the models for increasing number of top features for female subjects.

4.2.3 Dataset: Male subjects

100 male subjects, a subset of the dataset in 4.2.1 was studied; 97 subjects had $S_f$ in all four light settings. As for the datasets in 4.2.1 and 4.2.2, we compared the performance of all the models on these 97 subjects. Their Hgb levels had an $\mu$ of 11.85 g/dl and a $\sigma$ of 0.853 g/dl. The distribution of the Hgb levels for male subset is given in Figure 4.2.

The result for this analysis in Figure 4.5 shows us that the performance of $M_{850}$ and $M_{940}$ did not vary much with increasing number of top features. $M_D$ gave us the lowest MSE of 0.66 gm$^2$/dl$^2$ with top 6 features, $M_{Pixel2}$ gave us lowest MAPE of 5.39% with top 5 features, and $M_D$ gave the highest R of 0.3 with top 6 features. Also, compared to dataset 4.2.1, the $\sigma$ of the measured Hgb levels got better but it is worse than that for dataset 4.2.2.

As in the earlier subgroup analyses, we removed four subjects as outliers (Figure 4.1). Without the outliers, we had 93 male subjects, and their Hgb levels had an $\mu$ of 11.85 g/dl and a $\sigma$ of 0.73 g/dl. In this case, $M_D$ gave us the lowest MSE of 0.823 g$^2$/dl$^2$
Figure 4.5: Performance of the models for varying number of features for all male subjects

with top 26 features, and lowest MAPE of 4.85% with top 3 features. $M_D$ also had the highest R-value of 0.27 with the top 33 features. Also, the $\sigma$ of the measured Hgb levels in these 93 male subjects were poor.

4.3 Discussion

The foundation of this Hgb measurement system laid on the characteristics of the PPG signals. Value of the features extracted from the PPG signals depended on the intensity of light captured by the smartphone camera which is proportional to the concentration of Hgb (Equation 3.2). We used machine learning techniques to establish this relation among PPG signals, intensity of light, and Hgb. So, to get a proper measure, each step related to analysis of PPG signals was needed to be accurate. Similarly, improvement in any of these steps may increase the accuracy of the whole system.

4.3.1 Comparison with previous work

The problem with a Hgb measurement system with manual PPG cycle selection is that the human intervention component allows selection bias. Manual selection also increases the time required for measuring Hgb, and chances for human error.
Removing these uncertainties from the system has been our principal goal in creating an automated system reported here. However, with automation, the system may lose the opportunity to select more informative PPG cycles. Thus, selecting criteria for the best PPG cycles from those available is a major challenge. Here we have selected PPG cycles with greater height, and averaged the features in median range in order to get most relevant information from the PPG signals. Further work in determining of the quality of a signal and PPG cycle, and finding more robust methods to determine which PPG cycles hold more relevant information based on their quality, is needed.

In initial analytic activities [6], we manually selected optimal PPG cycles, which were only available in 167 of the 213 cases. Splitting these 167 cases into 80-20% for training and testing respectively, in the testing set we found a MAPE of 5.45%, an R=0.84, and a Bland Altman plot where the bias and precision obtained were -0.114 and 0.475 respectively. These metrics figures suggest good accuracy, high correlation with the gold standard laboratory measures, and limited bias and tight limits of agreement.

As note, in these initial analyses, 167 subjects had PPG cycles available for all four light settings and the analyses were done on these subjects only. To compare the performance of the current automated model with our previous work, we also built our model with the same 167 subjects and did 10-fold cross validations. We achieved an MSE of 0.86 in our current approach as opposed to 0.88 in the old approach which confirms that automation in the present work can replace previous manual PPG cycle selection.

4.3.2 Noise in signal

The accuracy of the Hgb measurement system depends on the values of the features extracted from the PPG cycles. If noises are present in the signals, then the feature values will be corrupted, and error will occur in the Hgb measurement. Moreover, PPG cycles may become unrecognizable due to acute noise in the PPG signals. Limiting noise in the signals is desirable for accurate Hgb measurement.
In detecting PPG cycles in $S_{PPG}$, we found that many $S_{PPG}$ did not have consistently similar PPG cycles (Figure 4.6). These characteristics also differ from block to block for the same subject. For this reason, we applied majority voting in the feature value selection before averaging the feature values for a subject so that most common feature values across all the blocks of a subject are used for our model. In this way, we ensured the minimum effect of noise on the feature values and measured Hgb values.

However, these inconsistencies mainly occurred due to motion artifacts or improper finger pressure while recording the videos. We can reject the videos having motion artifacts or inappropriate finger pressure at the time of recording the videos and requiring the recording of only noiseless videos. Moreover, a robust hardware set consisting of LED boards and smartphone holder for capturing the videos should also help to reduce motion artifacts. Towards that end, for future study, we are building a hardware set and mobile application to try to ensure less noise in the videos. Moreover, we are trying to make our PPG cycle detection more robust like Bui et al. [108].
4.3.3 **Average line as a predictor**

$M_A$ of all the datasets had better MSE and RMSE than that of the state-of-the-art stated in Table 4.6 where $M_A$ is not a valid estimator. If $M_A$ has low MSE or RMSE then estimating around the average value with low $\sigma$ will also have low MSE and RMSE. This demonstrates how important it is to consider the distribution of the estimations, and how much the accuracy improved compared to $M_A$.

4.3.4 **Error metrics**

Based on the results from table 4.3-4.5, it is evident that we get different best models based on best result from different error metrics. So, it is very important which error metric we choose. We believe that we should select models based on lowest MAPE for the following reasons.

The value of MSE depends on the magnitude of the original values. It is normal to have large MSE if we have large original values. Two systems having different MSEs does not necessarily mean that the system having lower MSE is better, unless both the systems are working with the same data. So, we should not compare MSE of one system with that of another system working on different dataset.

The value of R tells us the correlation between the estimated values, $y$ and original values, $x$. Having a high value of R does not guarantee high accuracy or better accuracy. It tells us how close $\Delta x$ and $\Delta y$ is if the original and estimated values increase or decrease $\Delta x$ and $\Delta y$ are.

On the other hand, MAPE offsets the largeness or smallness of the original values with normalization by the original values. It tells us the accuracy of the system. For these reasons, we suggest that MAPE is a better metric to compare performances among different systems.
4.3.5 Distribution of the Hgb levels in datasets

The distribution of the laboratory measured Hgb levels in our data set is shown in Figure 4.1 and 4.2. Noteworthy are the following:

- The Hgb values ranges for both men and women in our subjects are lower than usual normal ranges by gender. Normal ranges are for men, 13.5 to 17.5 g/dl, and for women, 12.0 to 15.5 g/dl [109] where we had 3 males and 12 females in this normal range. So, it is not possible to tell how our model would perform for people with Hgb levels across the entire normal range.

- More importantly, in normal population the distributions of Hgb levels are continuous. With a reasonably sized patient population of over 200 cases, we had laboratory-determined Hgb levels clustered at specific values. Such findings suggest recurrent laboratory measurement errors and/or rounding off of laboratory results. Because accurate gold standard measures cannot be interpreted to characterize our studied population, it was impossible for us to demonstrate accurate Hgb measurements with our non-invasive system.

- Specifically, out of the total 16 of specific Hgb levels, number of subjects were dense at 7 levels in the dataset used in subsection 4.2.1 where number of subjects at two Hgb levels were more than double than that of any other Hgb levels. But the dataset used in subsection 4.2.2 had 11 different Hgb levels, and among those levels, number of the subjects were dense at 6 levels. The Hgb level distribution of the dataset used in subsection 4.2.3 closely followed that of the dataset used in subsection 4.2.1. In this case, the number of subjects were dense at only 3 Hgb levels, where 12 Hgb level were present.

- These circumstances and data show that better distribution of Hgb levels in analyzed data sets gave us greater accuracy and \( \sigma \) in our measured Hgb levels.
From the above observations, we hypothesize that if we can test our model with a dataset having better distribution and covering the whole range of usual Hgb levels, then we should be able to demonstrate greater accuracy and generalizability which is also backed by Goodwin et al. [110].

Figure 4.7: The difference relationships between venous and capillary (a surrogate for arterial) hemoglobin levels in men and women. On the vertical axis is the average difference between the measured values which decrease with increasing hemoglobin levels, but at different rates between the genders. In usual hemoglobin level ranges (<15 gm/dL in men; <14 gm/dL in women, women have lower arterial hemoglobin levels because the diameters of their arterial vessels are smaller due to the Fahreus effect (see text). From Murphy et al. [12].

4.3.6 Separate machine learning model for men and women

When our datasets are divided into subsets based on gender, the results and performance of the models differed. Our models were built on features extracted from the PPG signals which is made of arterial pulse wave in the fingertip. However, the laboratory measured Hgb values were venous Hgb values. Arterial fingertip Hgb values
differ from venous Hgb values [111]. Further, Murphy et al. found that the differences between venous and arterial Hgb values for males and females differ across the middle of the spectrum of usual clinical hemoglobin values (The population studied was of blood donors that were required to have normal levels of hemoglobin) (Figure 4.7) [12]. The Fahraeus effect explains this phenomenon- the average concentration of red blood cells decreases as the diameter of the tube in which the blood is flowing decreases. Mean arterial diameters are higher in men than women, and the differences between mean arterial diameters between the genders decrease as arteries become arterioles and become smaller in diameter (Figure 4.7) [112]. The impact of estrogen is considered to be the cause of these decreased diameters of blood vessels in women [113].

Men and women with comparable Hgb levels in the fingertip arterioles/capillaries can have significantly different Hgb levels in their arm venous bloodstreams. Other authors have developed similar findings to those of Murphy: Cable et al. found female donors had, on average, venous Hgb levels 0.5 to 0.8 g/dL lower than male donors with the same fingertip Hgb levels [114]. Bell et al. also found differences between capillary and venous Hgb levels that varied by gender [115]. The studied sample sizes in the Murphy et al. and Cable et al. reports were 36,258 and 21,840 patients respectively, which strongly suggest that these findings are robust and accurate. Arterial and capillary hemoglobin levels are reportedly very similar [116, 117].

Machine learning models relating fingertip Hgb levels from the fingertip videos with the laboratory-measured Hgb levels from arm venous samples cannot estimate two different laboratory-measured Hgb levels for men and women from the same fingertip Hgb levels. However, when differences between venous and arterial Hgb values, can be specifically described, as for example in the figures presented (Figure 4.7), machine learning models can learn these relationships, and use them predict fingertip Hgb values from venous Hgb values. Based on these observations, we propose
to develop separate models for men and women.

4.3.7 Comparison with state-of-the-art

HemaApp [10] also collected Hgb measurements using the FDA cleared Masimo Pronto 7 to compare their measured Hgb level, and Masimo Pronto measured Hgb levels with standard laboratory-measured Hgb levels. HemaApp provided only Root Mean squared Error (RMSE) of their measured Hgb levels. We converted RMSE into MSE to compare their results with ours. The conversion follows the traits described in equation 4.2. From Table 4.6, we can see that our model performed better than both HemaApp and Masimo as assessed by MSE and RMSE, but poorer as assessed by R which is caused by the poor distribution in our dataset [110]. Here, the female dataset without outliers is used, which gave us the best result in term of MSE.

\[
\begin{align*}
\text{MSE} &> \text{RMSE}, \text{ if } \text{RMSE} > 1 \\
\text{MSE} &< \text{RMSE}, \text{ if } \text{RMSE} < 1
\end{align*}
\]

(4.2)

Moreover, our algorithm brought the following improvements from the state-of-the-art methods- 1. Minimization of the presence and effect of the signal's noises. 2. Techniques to detect and select PPG cycles.
Table 4.5. Best results in models with male subjects.

<table>
<thead>
<tr>
<th>Model</th>
<th>With or without outlier</th>
<th>MSE ↓</th>
<th>MAPE ↓</th>
<th>R ↑</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{850}$</td>
<td>Outlier</td>
<td>0.71 (5)</td>
<td>5.67 (-)</td>
<td>0.16 (5)</td>
<td>0.02 (4)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.56 (-)</td>
<td>5.07 (-)</td>
<td>0.149 (4)</td>
<td>0.076 (4)</td>
</tr>
<tr>
<td>$M_{1070}$</td>
<td>Outlier</td>
<td>0.71 (7)</td>
<td>5.59 (7)</td>
<td>0.11 (33)</td>
<td>0.22 (7)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.53 (23)</td>
<td>5.09 (23)</td>
<td>0.18 (23)</td>
<td>0.197 (23)</td>
</tr>
<tr>
<td>$M_{Pixel2}$</td>
<td>Outlier</td>
<td>0.68 (43)</td>
<td>5.39 (5)</td>
<td>0.24 (43)</td>
<td>0.34 (5)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.514 (15)</td>
<td>4.99 (14)</td>
<td>0.22 (15)</td>
<td>0.119 (15)</td>
</tr>
<tr>
<td>$M_{940}$</td>
<td>Outlier</td>
<td>0.71 (13)</td>
<td>5.67 (-)</td>
<td>0.21 (13)</td>
<td>0.32 (13)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.71 (13)</td>
<td>5.67 (-)</td>
<td>0.07 (8)</td>
<td>0.003 (8)</td>
</tr>
<tr>
<td>$M_D$</td>
<td>Outlier</td>
<td>0.66 (6)</td>
<td>5.42 (11)</td>
<td>0.305 (6)</td>
<td>0.34 (6)</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.502 (26)</td>
<td>4.85 (3)</td>
<td>0.27 (33)</td>
<td>0.187 (26)</td>
</tr>
<tr>
<td>$M_A$</td>
<td>Outlier</td>
<td>0.73 (N/A)</td>
<td>5.77 (N/A)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>No outlier</td>
<td>0.53 (N/A)</td>
<td>5.17 (N/A)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 4.6. Comparison with HemaApp and Masimo Pronto.

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset size</th>
<th>σ of the Dataset</th>
<th>Cross Validation</th>
<th>MSE ↓</th>
<th>RMSE ↓</th>
<th>R ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ours</td>
<td>111</td>
<td>0.787</td>
<td>10</td>
<td>0.48</td>
<td>0.69</td>
<td>0.49</td>
</tr>
<tr>
<td>HemaApp</td>
<td>31</td>
<td>2.2</td>
<td>n-1</td>
<td>1.58</td>
<td>1.26</td>
<td>0.82</td>
</tr>
<tr>
<td>Masimo Pronto</td>
<td>31</td>
<td>2.2</td>
<td>N/A</td>
<td>1.63</td>
<td>1.28</td>
<td>0.81</td>
</tr>
</tbody>
</table>
Chapter 5  
Customizable Data Collection Tool

5.1 Related works

Most methods measuring blood constituents including smartphone-based systems, use the spectral responses of blood to near-infrared light. To capture such spectral responses in great quality, smartphones need extra attachments as sources of near-infrared lights in both the hardware, and software form, and the analysis of these responses requires storage and transfer of these responses to an external computational device. A smartphone is an excellent tool for storage and transfer of data and options for displaying records and results. [118, 119, 120, 55, 121]. These reported methods do not exercise a smartphone's full potential in using other sensors to measure the blood constituents.

Using smartphones' different sensors, measuring blood constituents is achievable with or without adding external attachments. The details of external attachments to measuring systems depend on their targets. Vennemann et al. used the smartphone to power up an implant to measure the blood flow for remote monitoring of prosthetic heart valve function [122]. These investigators utilized the smartphone's RF field to harvest energy and transferred the energy between the smartphone and the implant through the NFC sensor.

Another system consisting of four LED lights of different wavelengths, an amplifier, and a photodiode captured a finger’s photoplethysmography (PPG) signal facilitated with a micro-controller unit, calculated the ratio between AC and DC signal of the PPG, and transferred the estimations of hemoglobin levels to a smartphone through Bluetooth [72]. This system used the smartphone as a storage and displaying device only.

A smartphone can replace attachment microcontrollers as a computing component. Matsumura et al. came up with iPhysioMeter- an iOS app that calculates
heart rate (HR), and normalized pulse volume (NPV) deriving PPG using a smartphone camera only [123]. iPhysioMeter had a bias of -0.12 and SD of 0.36, and a bias of 0.01 and SD of 0.43 in measuring HR and NPV respectively.

Heart rate variability (HRV) is an important marker of automatic nervous system function and this rate is a direct consequence of blood flow. Huynh et al.’s VitaMon use the smartphone’s front camera to calculate HRV with an average error of 14.26 ms. [124]. VitaMon used features of the extracted (PPG) signal from the recorded video of the face of the users and measured the user’s heart rate and HRV without any extra attachments.

Adding an attachment to the smartphone, Ilyas et al. detected and monitored sickle cell disease using micro-invasively obtained drops of patient’s blood [121]. They provided an external lens attached to the smartphone’s rear camera, a light source powered with DC batteries, and a smartphone holder. They analyzed images captured by the smartphone and determined the presence of sickle cells in the blood drop.

Wang et al. used external hardware with a smartphone application to measure hemoglobin from fingertip videos captured with the smartphone's front camera [10]. With multiple LEDs, the hardware package was placed beside the smartphone camera, and the smartphone was placed inside a box. Although the phone app was programmed to be resistant to noise, it could not detect motion artifacts or proper finger pressure and placement.

In summary, what is characteristic of these several currently available measurement systems using smartphones, is absence of components to reduce noise artifacts.

5.2 Requirements

We have been working to develop combined hardware and software systems to measure blood parameters using fingertip cell phone obtained videos [125, 126]. Our initial efforts have been in measuring blood hemoglobin (hgb) levels. The theoretical
basis for our systems lies in the Beer-Lambert law which states that the absorption of light is directly proportional to the concentration of the medium and path length of the light through the medium, which can enable blood constituents’ measurement. A smartphone’s camera is a good tool for correlating the light absorption with blood constituents’ levels using photoplethysmography [127].

Photoplethysmography uses optical properties of blood to measure volume change of the blood in the micro-vascular bed of tissue [75]. Figure ?? depicts an ideal PPG signal annotating some important features. The change in blood’s volume with the cardiac cycle changes the optical response following the Beer-Lambert law and these changes also occur in a video recorded with a smartphone camera. Pixel values of the frames in the video are proportional to these optical responses, and the blood volume in an instant is proportional to the average value of all the pixels in a frame at that instant. A PPG signal is generated by plotting the average values pixels across all the frames of a video as a time series [126]. What is critical is creating the conditions for obtaining high quality, meaning minimally noise-affected and consistent PPG signals, at low cost. Here we report the tools we have used to create such conditions for our hemoglobin measurement package which can be used for any other blood constituents’ measurement package using PPG features from smartphone fingertip video, and near infra-red (NIR) lights.

5.3 Description of complete system

5.3.1 The electrical circuit

The principal functionality of our designed electrical circuit is controlling the NIR lights to illuminate the fingertip for the videos and ensuring proper finger pressure on the camera lens while capturing the fingertip video (Figure 1). The key features and architecture of this electrical circuit are:
Figure 5.1: Circuit design to keep applied pressure on the camera lens within a threshold.

**Power Source:**

We use the smartphone as the power source for the tool—everything in the electrical circuit is designed to be powered by the smartphone itself. The circuit design has a USB C port which draws the power from the smartphone's USB port through an appropriate USB cable. By doing this, an extra power source becomes unnecessary. The overall power consumption from the smartphone is very minimal which ensures a negligible effect on users' regular power consumption.

While capturing the videos, the values of the pixels in the frames of the video change with the fingertip pressure applied on the camera lens. PPGs derived from the pixel values of the videos of the subjects with similar Hgb levels should have features with high co-relation. However, if they exert pressure on different level, then such characteristics will be missing from their PPGs. For all subjects, keeping the pressure on the camera lens in a certain range can ensure consistent feature values. A force-sensitive sensor (FSR) provides variable voltage depending on how much pressure is applied to it. Comparing the voltage from the FSR with a fixed voltage, we
can ensure the applied pressure is within an optimal range.

**Maintaining proper pressure:**

An operational amplifier (opAmp) works as a comparator circuit as it takes two input voltages and produces an output voltage based on the comparison between the input voltages. So, we made comparator circuits using the output of the FSR, and a fixed voltage as input of the opAmp. We added two such comparator circuits by connecting the FSR in the positive end of opAmp in one, and in the negative end of the opAmp in the other one (Fig. 5.1).

Here, one comparator unit is used to ensure the applied force is greater than the lower threshold, and another comparator unit is used to ensure the applied force is less than the upper threshold. These thresholds have been chosen by a trial-and-error method to define optimal pressure. The output of the two comparators is used as input of an AND gate, and if both conditions are met, then the AND gate will give a 5V output which is the output of this part of the electrical circuit. This output is used as VCC or positive end of the LEDs.
**Customizable Pressure Threshold:**

We used variable resistors in-between the pressure sensor and the opAmps to control the pressure thresholds. Changing the value of these resistors, the minimum, and maximum allowable pressure can be changed, which will be helpful in the exploratory phase.

**Customizable LEDs:**

Customizing LEDs allows for ease in changing the LEDs of the circuit. This feature has been useful in discovery phase with our developing systems because

1. Different blood constituents are responsive to different wavelengths of light, and having the option of using any type of LED can enable us to use the framework to determine various blood constituents.

2. Even if there is a resistor between the LEDs and power input, LEDs being delicate in nature, can stop working when the environment is not stable. This can also happen when the quality of LEDs are poor.

3. LEDs with a different wavelength are needed. 2*1 header pins for each of the LEDs in the circuit enable an option of customizable LEDs.

With customizable LEDs, both time and money can be saved, as other circuit components remain the same.

**Two arrays of LEDs:**

The electrical circuit for our hemoglobin system has two arrays of three LEDs and these two arrays are enclosed by two LEDs. In two arrays of LEDs, two different kinds of LEDs or same kind of LEDs can be used. They have separate control switch so that they can be turned on or off separately. Two outer LEDs do not have a switch, they turn on with an appropriate pressure on the FSR. Usually these two LEDs are used to illuminate the finger, but any other type of LEDs can be used here.
Printed circuit board (PCB):

The designed electrical circuit is printed on a board. A PCB keeps all the components very stable and the price of the whole setup low. We drew the digital circuit for our hemoglobin system on EasyEda - a free online tool [128] (Figure 5.1), and generated the required file for printing and assembling the PCB boards. Bill of materials (BOM) was one of the required files, which provides an extensive list of components and their brief description, quantity, and source. We had to modified that file to change the information of a few components due to unavailability from EasyEda’s default source. Using the files from EasyEda, our PCB boards were manufactured from PCBWay, at less than $10 per PCB unit including the materials, components, and assembly cost [129] (Figure 5.2).

5.3.2 3D designed box

The primary function of the box is to keep the smartphone and the other necessary components immobile while capturing the videos from the fingertip (Figure
2). If we hold the smartphone with our hands to capture the videos, then the videos demonstrate noise from hand motion.

The box is designed to work with a Pixel2 smartphone. However, to keep it unrestricted to a specific smartphone model, we used SolidWorks for the design. With SolidWorks, the dimensions of the box can be changed. This process makes a similar hardware setup for any smartphone model easily constructible.

Having multiple separable parts, assembly of this box is straightforward using slider mechanisms (Fig. 5.3). The box parts are-

1. **Base for smartphone with camera aperture, \(b_s\)** - the aperture is positioned from the sides to align with the smartphone camera,

2. **Side cover with fingertip aperture** - this aperture is horizontally aligned with the camera aperture, and vertically aligned with the \(b_s\),

3. **Base box** - this is the main part of the box which can be considered as the
foundation of the box. All other parts are inserted into this using the sliders. It also has placeholders for the switches. The electrical circuit lies on the ground of this base box. This base box also has a slider for the smartphone.

4. **Side cover with aperture for power cable passing through** - the aperture is adjacent to the base box. The power cable delivers the power from the smartphone to the electrical circuit, and it passes through the aperture to come outside from the inside of the box.

Figure 5.4 depicts all the components of the box and the circuit connected at the proper position.
**Power Source:**

The components can be dismantled quickly. A 3D printer can print all the parts of this 3D-designed box with inexpensive plastic. As manual labor is not involved, the overall cost to produce this box is minimal. We ordered the 3D printing of the boxes from a third-party manufacturing company named 3DHubs [130] where it cost $100 for four sets of boxes. The cost could be significantly less if the boxes were manufactured in bulk quantity from a wholesaler.

If this box is placed on a table or any even surface, then the smartphone and the NIR lights will be immobile and stable. The patient’s arm rests on a table surface to allow comfortable maintenance of fingertip position and pressure (Fig. 5.5).

5.3.3 **The software system**

A two-part software system with an android application and a backend server has been designed to record high-quality PPG signals from the fingertip videos and analyze these signals.

The android application serves as a front-end mobile application for video capture and system output display. To capture the fingertip video accurately, our data collection tool uses this mobile application for optimal placement of the fingertip relative to the lights and camera lens, and for achieving optimal reproducible pressure of the fingertip on the camera lens as described above.

This application will prompt for demographic information from the subject and assign a subject id. Then it can start video recording and stop the recording after the required time for successful video capturing. However, if the application detects any motion artifacts in real-time, then it stops video recording immediately and prompts for repeat recording.

In real time motion artifact detection summarized in Algorithm 1, the application ensures two things-
Algorithm 1 Real Time Image Analysis

Require: $image \neq \text{Null}$
Ensure: $\text{timeElapsed} \leq 10$

global variables
$\text{firstImage}$
end global variables

procedure IMAGEANALYSIS
$\text{timeStarted} \leftarrow \text{getCurrentTimeInSeconds}()$
$image \leftarrow \text{getNextImage}()$
$\text{firstImage} \leftarrow image$
$\text{framecount} \leftarrow 0$
while nextImageAvailable() do
$\text{timeElapsed} \leftarrow \text{getCurrentTimeInSeconds}()$
if $\text{timeElapsed} - \text{timeStarted} == 2$ then
if adjustFingerPosition(image) == True then
stopRecording()
endif
endif
if $\text{framecount \mod 10} == 0$ then
$\text{prevImg} \leftarrow \text{image}$
endif
$\text{detect} \leftarrow \text{isDifferent}(\text{image}, \text{prevImg})$
if $\text{detect} == \text{True}$ then
stopRecording()
endif
$image \leftarrow \text{getNextImage}()$
$\text{framecount} ++$
end while
end procedure

1. Proper finger placement on the camera lens.

2. Minimal motion artifacts.

Proper finger placement on the camera lens

The system uses Algorithm 2 to ensure that the finger is placed properly on the lens. To do this, the software system process the red, green, and blue pixel values from the RGB array of the frame. The system takes the first frame and calculates the average of pixel values. As the response captured by the camera is mostly comprised of blood,
Algorithm 2 Finger Placement Modification

Require: image \neq \text{Null}

procedure \text{ADJUSTFINGERPOSITION}(image)
    \hspace{1em} r \leftarrow \text{getRedAverage()}
    \hspace{1em} g \leftarrow \text{getGreenAverage()}
    \hspace{1em} b \leftarrow \text{getBlueAverage()}
    \hspace{1em} \text{if } r \leq 80 \text{ then}
        \hspace{1.5em} \text{return } \text{True}
    \hspace{1em} \text{end if}
    \hspace{1em} \text{if } b \geq 70 \text{ or } g \geq 70 \text{ then}
        \hspace{1.5em} \text{return } \text{True}
    \hspace{1em} \text{end if}
    \hspace{1em} \text{return } \text{False}
\end{procedure}

the frames should be bright red with the proper placement of the finger. To ensure this, the system takes the following steps-

i. Divides the frame into a smaller region of interest by creating a grid of $10 \times 10$.
   Thus, the system has a total of 100 blocks, and with proper placement of the finger, the average of red pixels of a block should be greater than that of the green pixels, and blue pixels across all these blocks. Let the average of red, green, and blue pixels be $r$, $g$, and $b$.

ii. Maintains $r > r_t$ where $r_t$ is the minimum threshold for the average value of the red pixels in a block where the maximum value of $r$ can be 255. The system tries to keep the value of $r$ for all the blocks in the video capture session above a certain threshold value so that the proper finger placement on the camera lens is ensured.

iii. Maintains $b < b_t$ and $g < g_t$ where $b_t$ and $g_t$ is maximum thresholds for blue and green pixels respectively where the maximum value of $b$, and $g$ can be 255. In the RGB system, for bright red color, the value of red pixels will be much higher (closer to 255) than that of the green and blue pixels. The system keeps track of
the value of the $g$ and $b$ so that it doesn’t exceed a certain threshold value. It ensures that the finger is not partially placed or not placed at all.

**Minimal motion artifacts**

While capturing the video, the finger on the lens should not move. This can be assessed by keeping track of the pixel values of the frames from the captured videos. Pixel values in a frame are directly related to the amount of blood flowing through the fingertip in that instance. With the rhythmic pumping from the heart, the blood flows through the fingertip in a rhythm, and the pixel values should also reflect that. A sudden change in the pixel values in the same position of consecutive frames indicates an interruption in the rhythmic flow of pixel values and indicates that motion artifacts are being generated.
Algorithm 3 Image Difference

Require: \( \text{img} \neq \text{Null} \)
Ensure: \( \text{prevImage} \neq \text{Null} \)

procedure isDifferent(img, prevImg)
    localPixels ← 0
    globalPixels ← 0
    for each imgPixel ∈ img, prevImgPixel ∈ prevImg, initImgPixel ∈ firstImg do
        diffPixels ← abs(imgPixel – prevImgPixel)
        if diffPixels ≥ localPixelThresh then
            localPixels++
        end if
        globalDiffPixels ← abs(imgPixel – initImgPixel)
        if globalDiffPixels ≥ globalPixelThresh then
            globalPixels++
        end if
        if localPixels ≥ localThreshold or globalPixels ≥ globalThreshold then
            return different ← True
        end if
    end for
    return different ← False
end procedure
The RGB values of every frames are compared in real-time from the beginning of the video, and two counts—global and local count of how many pixels are differed significantly in current frame with compared to the reference frame—are kept. The sudden change in motion is usually reflected in a number of consecutive frames rather than two consecutive frames, so we decided to update our reference frame on every 0.3s- at 30 fps which is tantamount to 10 frames. While the local count only considers the current frame and the reference frame in the computation, the global count keeps track of the total number of pixels that differ significantly from the beginning of the video. If the different number of pixels in the two consecutive frames reach the maximum threshold either locally or globally, the system detects a motion artifact, and stops recording the video for that session. The options to set the margin of difference between the values of two pixels to be considered as significant, and the thresholds for the maximum number of pixels can differ in two consecutive frames locally and globally before the system detects motion artifacts are available in the settings. Such flexibility will help determining the optimal value for the margins and thresholds. This whole process is summarized in Algorithm 3.

These image processing techniques ensures the proper finger placement and stability of finger placement on the camera lens. The value of $r_t$, $b_t$, and $g_t$ was determined after multiple trial run with proper, and improper finger placement. With the apparent successful recording of the video identified by a subject id, the video is sent to the backend server, where the captured video is analyzed frame by frame in real-time.

**Video analysis in the backend:**

After the video is uploaded to the backend server, the following steps are operated following Aziz et el. [126] to measure Hgb level, which can be different for other blood constituents-

1. Pixel values are extracted from the video frames.
2. Each frame is divided into 100 blocks, and the pixels are averaged for each block to produce raw PPG signals.

3. Raw PPG signals are smoothed with signal processing techniques, and features are extracted from the smoothened signals.

4. Features are used with a pre-generated machine learning model to estimate the blood hemoglobin level.

A Flow of decisions taken in the software system is depicted in Figure 5.6.

5.4 Discussion

5.4.1 Technologies used in the software system

The heart of the software system lies in a backend server program accessed through an Application Programming Interface (API). The server software system will generate PPG signals from the video sent by mobile devices. The raw frames are first subdivided into blocks. The red pixel values of each block are then averaged so that all the red pixel values of each block collectively generate PPG signals. The PPG signals may have some noise. To reduce such noise in these signals, we use a Butterworth bandpass filter and Fast Fourier Transformation. The three best PPG cycles from a block based on the standard features of a PPG cycle are selected. If the system is not able to identify at least three cycles, it replicates one cycle to make three. After identifying the three best signals these are fed into a machine learning model for feature extraction and model generation for hemoglobin level prediction [126].

Following are the technologies we can use for the software system:


2. Front end Mobile Application: Android
5.4.2 Use of camera2 API over default camera

Our first approach was to use the built-in Android camera app but the caveat was that the built-in Android camera app does not provide any support for real-time image analysis for video recording. We can get the fully recorded video after the recording is finished. But this hinders two crucial features of our system which are system-aided finger placement and motion artifact detection.

After extensive research and review, we found the Camera2 API available in `android.hardware.camera2` packages which fits best according to our system requirements. Camera2 API has several advantages for processing captured images in real-time. Camera2 API also has the functionality to operate and get feedback from each of the camera lenses separately. The captured images in case of a video recording is a stream of images. A capture session needs to be created to process the incoming stream of images. To process the real-time stream of images, an ImageReader object is required. ImageReader provides an `OnAvailableImageListener`, a callback function that handles each image and provides the functionality to process the image. All of these functionalities combined enabled us to implement the features that we envisioned for our system.

5.4.3 Power Consumption

The sole power source of our tool is the smartphone, and the power consumption would vary depending on the make and model of the smartphone which is a Google Pixel2 in our case. Theoretically, only 0.0135 unit of electricity is used to fully charge a Pixel 2 smartphone's 2700mAh battery from 0% to 100% which costs 0.19 cents considering the average cost of per unit electricity across all the states in the US is 14.11 cents [131]. The original cost would depend on the power used by the adapter and power transfer efficiency of the adapter. However, the practical cost would not deviate much from the theoretical cost, and the minimal theoretical cost keeps the practical cost minimal.
The data collection tool dropped the smartphone's battery percentage by 31% in 45 minutes continuous usage session on average and 8-10 successful readings were taken with a couple of failed attempts due to motion in this time period. So, it costs less than 0.01 cents while conducting each Hgb measurement.

5.4.4 Noise reduction

This system reduces noises in the captured data by detecting motion artifacts, and ensuring proper placement, and optimal pressure of the finger on the camera lens. Finger movements and flinching usually causes the motion artifacts. Figure 5.7 depicts a PPG signal captured by the system after disabling the check for motion artifact and it shows how finger movements can affect the PPG signal. In such cases, a PPG signal can become inconsistent, and not provide all the feature values (Fig. ??). For calculating heart rate, or Hgb, many methods need to determine the number of systolic peaks present in the signal which would give erroneous result with this kind of PPG signals.

Figure 5.8 depicts a PPG signal captured by the system keeping the detection for motion artifact on.
Figure 5.8: A PPG signal that can be useful for further analysis generated by the system checking all the parameters in the runtime.

5.4.5 HIPAA compliance

Using this system, data will be collected from identifiable individuals with some demographic, and physiological information. This system maintains all the aspect of Health Insurance Portability and Accountability Act (HIPAA). All the information is stored in the server maintaining all the privacy and security protocols of HIPAA.

5.4.6 Future research

This system broadens the scope of the quality check for devices made for PPG signal acquisition using smartphones. Further work on determining the threshold values to detect proper finger placement can increase the performance of this system. Moreover, along with run-time motion detection, a feedback-based motion detection module can be added to the system.
Chapter 6  
Hyperglycemic state detection

6.1 Theory and Methodology

Detection of hyperglycemic state from HbA1c does not require to calculate the exact value of HbA1c, rather it needs us to measure the HbA1c value in a range. This requires us to divide the HbA1c values into some classes, and finding out the HbA1c value is in which class. However, we need to know how HbA1c is measured in other methods.

![Figure 6.1: Absorption properties of HbA1c in NIR spectrum.](image)

The standard way of calculating HbA1c is to determining the ratio of HbA1c to the total Hgb concentration. Laboratory-based gold reference method of HbA1c needs sophisticated equipment and professional personnel to operate the system. Apart from the laboratory method, some Point-of-care (POC) HbA1c devices are available, but their accuracy needs improvement too. These POC devices needs blood droplet too which brings some of the complications of the invasive solutions. Moreover, most of the methods uses extra chemicals or electrodes to measure HbA1c and Hgb.
concentration [132]. Dependency on the chemicals and sophisticated techniques makes POC tools complicated, and inconvenient for personal use.

A study tried to quantify HbA1c through near-infrared spectroscopy and found that HbA1C is more responsive in the waveband of 1492nm to 1858nm (Fig. 6.1) [133]. On the other hand, in a recent study, HbA1c is found more responsive under green light (Fig. 6.2 [134], where Hgb’s absorbance is very low under green light [135]. The authors used chemical solutions equivalent to HbA1c and Hgb. So, further experiment is required to establish this theory.

However, based on this theory, following the non-invasive solution to measure Hgb described earlier, we can try green pixel values of the frames of the videos to determine the class of the HbA1c value. For this method, we can use the same dataset used for Hgb measurement as in that dataset, we have video recordings with the smartphone only without any NIR lights.
6.2 Methodology and Results

Our methodology for HbA1c class detection is same to the methodology described for Hgb measurement except-

1. using green pixel values instead of red.

2. using features from videos with Pixel2 only equivalent for features of HbA1C.

3. using features from videos under a wavelength of 940nm equivalent for features of hemoglobin.

4. using a support vector classifier instead of support vector regressor.

6.2.1 Feature set generation for machine learning model

As HbA1C is more responsive to green light (Fig. 6.2), following Beer-Lambert law (Eqn. 3.1), and the formulation of the features in Section 3.7, the ideal formula to generate feature set for a machine learning model to detect hyperglycemic state would be-

\[ HbA1C = \frac{I_{550}}{I_{940}} \]  

where, 550nm corresponds to wavelength of the green light to provide the absorption information of HbA1c, and 940nm corresponds to a peak of Hgb absorption to provide the absorption information of Hgb.

However, we do not have any dataset that contains videos under green light, and collecting new data is time consuming, and lengthy process. For the exploratory purpose, we used the videos with Pixel2’s camera only and considered the value of green pixels from the videos analogous to green light. Based on equation 6.1, we generated the feature set from \( \frac{F_{Pixel2}}{F_{940}} \) where \( F_{Pixel2} \) is the features from green pixel values of videos captured with Pixel2 only, and \( F_{940} \) is the features from red pixel values of videos captured with Pixel2 smartphone and 940nm wavelength of NIR light. In the
videos captured with pixel2 only, the value of the green pixels is less than that of the red
pixels, and around ten subjects did not have any valid PPG cycles across all the blocks.
As we are dividing the feature value of two sets of videos, we considered only the
subjects having at least one PPG cycle across all the blocks in both the sets, and there
were 149 such subjects.

![Figure 6.3: Correlation among the features.](image)

Figure 6.3 depicts the correlation between the features, and some of the features
are highly correlated with other features. To reduce the variance of the model, and
make the model faster, we normalized the feature set and applied principal component
analysis to reduce the dimensions of the feature set. PCA needed 35 components to
explain the overall variance of the feature set.

### 6.2.2 Initial Analysis

Our dataset consists of 200 subjects is densely populated in 4.5-5.5 (Fig. 6.4). The standard deviation, maximum, and minimum of the dataset is reported in Table
Table 6.1. Characteristics of the studied population and its laboratory HbA1c levels.

<table>
<thead>
<tr>
<th>Dataset type</th>
<th>Count</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Maximum (g/dl)</th>
<th>Minimum (g/dl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>200</td>
<td>5.07</td>
<td>0.63</td>
<td>9.5</td>
<td>4.5</td>
</tr>
<tr>
<td>Subjects (4.5-5.5)</td>
<td>170</td>
<td>4.84</td>
<td>0.18</td>
<td>5.5</td>
<td>4.5</td>
</tr>
</tbody>
</table>

6.1. Among these 200 subjects, 176 subjects had normal HbA1c value (less than 6%) and rest are prediabetic and diabetic. We labeled the subjects having HbA1C less than 6% as class 0 and rest of the subjects as class 1. Although we have a great imbalance between class 0 and class 1, and sampling is the to-go method to work with imbalanced dataset, at first, we used SVC without any sampling to get a general understanding. Whenever we built a support vector classifier, we selected its parameters using GridSearchCV of scikit-learn module and used k-fold cross-validation technique with a value of 5 for k to determine the accuracy of the model.

Table 6.2 represents the performance of the model in detecting hyperglycemic
state without using any sampling method. Although this model has moderate accuracy of 68%, we cannot accept this model for having a low precision and recall for class 1 which translates to "this model fails to detect hyperglycemic subjects".

Table 6.2. Classification result for HbA1c without using any sampling technique. Here, precision means ratio of true positive and predicted positives and recall means ratio of true positive and actual positives

<table>
<thead>
<tr>
<th>Class</th>
<th>#Subjects</th>
<th>Precision</th>
<th>Recall</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 0</td>
<td>127</td>
<td>0.83</td>
<td>0.79</td>
<td>68</td>
</tr>
<tr>
<td>Class 1</td>
<td>22</td>
<td>0.07</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

We used SMOTEENN oversampling method combined with random undersampling of the majority class to balance the ratio of subjects from class 0 and 1 to 2:1. Following this steps, we got an accuracy of 68%, but the precision and recall increased. The classification report is provided in Table 6.3.

Table 6.3. Classification result for HbA1C. Here, precision means ratio of true positive and predicted positives and recall means ratio of true positive and actual positives

<table>
<thead>
<tr>
<th>Class</th>
<th>#Subjects</th>
<th>Precision</th>
<th>Recall</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 0</td>
<td>127</td>
<td>0.88</td>
<td>0.73</td>
<td>68</td>
</tr>
<tr>
<td>Class 1</td>
<td>22</td>
<td>0.21</td>
<td>0.41</td>
<td></td>
</tr>
</tbody>
</table>

6.3 Discussion

Although we did not get a very good accuracy to detect hyperglycemic state, the results can guide us to build a substantial system to determine the hyperglycemic state. From the results, it is evident that with a balanced dataset, the model built with information from green pixels of the smartphone videos can detect hyperglycemic state with a moderate accuracy. So, if the model can be built with information from the smartphone videos under green light directly, then it should detect hyperglycemic state with a better accuracy.
Chapter 7  
Conclusion and Future Works

Measuring hemoglobin (Hgb) and other blood constituents in a non-invasive manner could facilitate health services around the world, especially now, when at-home contactless services are a necessity during the pandemic. Current non-invasive systems being expensive and not accurate enough are not ready to use as a point-of-care (POC) tool yet. Our motivation was to research a cheap and reliable non-invasive POC tool to measure blood constituents to ameliorate the current situation.

This dissertation proposes a novel solution for a POC tool to measure hemoglobin and other blood constituents in a non-invasive manner. As shown in the different chapters, this dissertation contributes to build such an inexpensive, ubiquitous, and reliable tool.

The initial works in this dissertation explored non-invasive Hgb measurement using machine learning techniques where the features are extracted from Photoplethysmography (PPG) signal generated from smartphone videos and achieved a great accuracy. In the process, we automated the process of detecting and selecting PPG cycles and figured out the biological reasoning to use separate machine learning model for male and female. We also looked into the issues of our analysis, and built a framework- both hardware- and software-based to overcome these drawbacks. Then we explored the possibility of hyperglycemic state detection using the framework.

While further refinements in PPG signal analysis to improve the accuracy of the system we have developed are certainly needed, we propose that with the hardware and software framework we have built, at this time a large validation study of our system is appropriate. Key issues in the pursuit of such a study will be recruitment of a large study population with a full range of low and high hemoglobin levels, standard laboratory hemoglobin measurements by a single large volume Clinical Laboratory Improvement Act-accredited facility, locked-down data collection and analysis
sequence systems, and rigorous data monitoring procedures [136]. While collecting the data in the future, we need to follow a strict data collection protocol which would ensure some crucial factors like data security, storing the data following HIPPA compliance, and female patients do not have acrylic paints on their fingernails. If we can collect diversified data, we will be able to analyze the racial and skin-color effect on our analysis. Apart from measuring Hgb levels, this whole system can be modified and extended to do white blood cell count, creatinine level measurement, and hyperglycemic state detection.

In summary, we have developed a hardware and software system for the noninvasive assessment of hemoglobin levels using fingertip videos obtained by a smartphone, which provides results beginning to approximate accuracy levels within ±1g/dl, and it can be considered for general clinical use. With data collection refinements suggested by our results, we suggest that more accurate and clinically applicable measurements will be possible with our system.


colour strip (hcs-hll), a digital haemoglobinometer (truehb) and a non-invasive device (touchhb) for screening patients with anaemia,” *Journal of clinical pathology*, vol. 69, no. 2, pp. 164–170, 2016.


[128] “Easyeda.”

[129] “Pcbway.”

[130] “3d hubs.”

[131] “Average price of electricity to ultimate customers by end-use sector.”


