Published: 18 June 2025

RESEARCH-ARTICLE

DermaGlow: Objective Quantification of Melanin, Erythema and Skin-tone Using Wearable Optical Spectroscopy

TAREK HAMID, University of Virginia, Charlottesville, VA, United States

PATRICIA FLORES, University of Virginia, Charlottesville, VA, United States

JANE BYUN, University of Virginia, Charlottesville, VA, United States

ELIZABETH C. COURTNEY, University of Virginia, Charlottesville, VA, United States

KYLE C. QUINN, University of Virginia, Charlottesville, VA, United States

AMANDA WATSON, University of Virginia, Charlottesville, VA, United States

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 9, Issue 2 (June 2025)
https://doi.org/10.1145/3729474
EISSN: 2474-9567

DermaGlow: Objective Quantification of Melanin, Erythema and Skin-tone Using Wearable Optical Spectroscopy

TAREK HAMID, University of Virginia, USA
PATRICIA FLORES, University of Virginia, USA
JANE BYUN, University of Virginia, USA
ELIZABETH C. COURTNEY, University of Virginia, USA
KYLE C. QUINN, University of Virginia, USA
AMANDA WATSON, University of Virginia, USA

Accurate characterization of the skin is essential for optimizing diagnostic and therapeutic dermatological tools, as well as technologies like pulse oximetry that rely on skin perfusion. Traditionally, optical spectroscopy has been used for skin assessments through devices like commercial colorimeters, which are high-cost instruments that, while precise, only provide single measurements rather than continuous data. Additionally, medical wearable devices that use this technology often show variable accuracy based on skin tone. The limitations of existing devices demonstrate the need for a solution that can provide low-cost, accurate, and continuous skin monitoring across varying skin tones in a wearable form-factor. This paper introduces DermaGlow, a novel wearable optical spectroscopy framework designed for low-cost, non-invasive monitoring of melanin, erythema, and skin tone. DermaGlow utilizes an off-the-shelf multi-spectrum wearable device available in various configurations to enable real-time, personalized assessments across diverse skin conditions and skin tones. We assess the performance of the DermaGlow algorithm against a state-of-the-art colorimeter in a comprehensive user study involving a diverse group of 77 subjects, demonstrating a normalized mean absolute error (NMAE) of 5.33% (melanin) and 4.18% (erythema), and ΔE values less than 2.5 for CIE LAB measurements. Furthermore, we present an algorithm that utilizes DermaGlow outputs to correct for pulse oximeter inaccuracies typically found in those with darker skin pigmentation, resulting in an up to 75% decrease in mean absolute error (MAE) in hypoxic readings across skin tones relative to arterial blood measurements. Our findings highlight DermaGlow's potential for short and long-term skin monitoring and as a significant enhancement to existing wearable devices, particularly in improving the accuracy of pulse oximeter readings across different skin tones.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile devices**.

Additional Key Words and Phrases: spectroscopy, melanin, skin-tone, erythema, wearable technology

ACM Reference Format:
Tarek Hamid, Patricia Flores, Jane Byun, Elizabeth C. Courtney, Kyle C. Quinn, and Amanda Watson. 2025. DermaGlow: Objective Quantification of Melanin, Erythema and Skin-tone Using Wearable Optical Spectroscopy. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 2, Article 32 (June 2025), 28 pages. https://doi.org/10.1145/3729474

Authors' Contact Information: Tarek Hamid, pve8nt@virginia.edu, University of Virginia, USA; Patricia Flores, sjy7da@virginia.edu, University of Virginia, USA; Jane Byun, rjr5dn@virginia.edu, University of Virginia, USA; Elizabeth C. Courtney, xzy8ta@virginia.edu, University of Virginia, USA; Kyle C. Quinn, bak6sp@virginia.edu, University of Virginia, USA; Amanda Watson, aawatson@uva.edu, University of Virginia, USA.

ACM 2474-9567/2025/6-ART32
https://doi.org/10.1145/3729474

Fig. 1. DermaGlow is an advanced framework that utilizes wearable, multi-wavelength optical spectroscopy to deliver low-cost, continuous, and non-invasive measurements of melanin, erythema, and skin-color accurately across various skin tones and body locations.

1 Introduction

In the fields of dermatology and cosmetic science, precise analysis of skin characteristics serves as a key measure in diagnosing various skin conditions and tailoring cosmetic and medical interventions. Three of the most significant aspects in this analysis are melanin content, erythema level, and skin-tone, all of which not only characterize the overall appearance of the skin, but its physiological response to environmental factors and susceptibility to various dermatological conditions [1, 20, 58, 92]. These skin characteristics are also crucial factors in the performance of optical sensors used in wearable health monitoring devices, as they can significantly affect the accuracy of measurements like blood oxygen saturation [8, 69] and heart rate [44] by altering the absorption and scattering of light through tissue. Furthermore, these characteristics are not static; they can fluctuate significantly due to factors such as exposure to sunlight, hormonal changes, and/or skin injury [92], thus requiring the need for continuous monitoring or frequent periodic assessment to account for these dynamic changes.

Traditional practices for assessing these markers of skin have ranged from subjective visual assessments, such as the Fitzpatrick [28] or Monk [54] Skin Tone Scales, to more objective yet often invasive or sparse quantitative methods, such as biopsies [62] or portable spectrometers in the form of skin-tone pens [2, 12, 15, 19, 82]. These conventional methods, while valuable in specific contexts, each present notable limitations: subjective scales like the Fitzpatrick and Monk assessments inherently suffer from observer variability and lack quantitative precision; portable spectrometers, though objective, are typically expensive, require frequent calibration, and only provide discrete measurements rather than continuous monitoring; and biopsies, while providing detailed cellular information, are invasive procedures that cannot be performed frequently enough for ongoing monitoring of dynamic skin changes.

The advent of wearable technology offers a promising solution for non-invasive monitoring of health parameters. Within this context, wearable optical spectroscopy is a particularly useful approach for objective, long-term skin-tone analysis. These devices present a cost-effective alternative to the aforementioned skin-tone pens, which

cost several thousand dollars. Current wearable devices already use optical spectroscopy to measure biomarkers such as heart rate and blood oxygen, which can be extended to potentially measure skin composition. Developing a wearable device capable of measuring melanin, erythema, and skin-tone continuously and non-invasively can be beneficial in several ways:

(1) **Addressing the pulse oximeter problem** - the ability to measure melanin and erythema in a wearable form-factor can improve the accuracy of pulse oximeters, which often show biased readings in individuals with darker skin [27, 36, 43, 71, 79], as well as enhance the reliability of other optical biosensors that operate under similar principles.
(2) **Diverse Skin-Tone Inclusivity** - incorporating comprehensive data on melanin and skin-tone across diverse skin-tones can enhance representation and fairness, especially in AI model deployment in healthcare or remote patient monitoring/recruitment.
(3) **Personalized Skin Care** - understanding individual skin composition can aid in personalized treatments, especially in dermatology where personal skin properties can have dramatic influences on treatment effectiveness.
(4) **Sun Exposure Management** - monitoring changes in melanin and erythema content can provide personalized recommendations for sun protection, potentially reducing the risk of skin cancer by allowing for timely and tailored interventions based on an individual's unique skin responses to UV exposure.

In this paper, we address the following research questions:

RQ1: How can we non-invasively measure melanin, erythema, and skin-tone content accurately using wearable optical spectroscopy?
RQ2: How can we use these skin measures to correct for inaccuracies seen in pulse oximetry measurements in darker skin-tones?

We introduce DermaGlow, a novel optical spectroscopic framework that enables non-invasive, continuous analysis and reporting of melanin, erythema, and overall skin-tone. The DermaGlow framework uses a commercially available device that utilizes reflectance spectroscopy with wavelengths across the visible spectrum in multiple wearable form factors. The device leverages a multi-spectral sensor paradigm to enhance visible light resolution beyond that of current commercial wearables, allowing for greater data density. We leverage this data to differentiate between subtle changes in skin properties, and make a quantitative assessment of the underlying structure of the skin.

We developed an algorithm that processes multi-spectral signals obtained from reflectance spectroscopy and returns melanin, erythema, and skin-tone values. The algorithm processes raw three-dimensional spectral data and employs a machine learning model to optically differentiate between skin constituents across various skin-tones. Our approach achieves normalized mean absolute error (NMAE) values of 5.33% for melanin and 4.18% for erythema, with ΔE values below 2.5 for CIE L*, a*, and b* measurements - a standardized color space used for skin-tone where L* indicates lightness/darkness, a* represents green-red, and b* represents blue-yellow - respectively, across skin-tones, with consistent results regardless of overall skin-tone and measurement location. We evaluate our algorithm on its accuracy, mean absolute error, and generalizability relative to a skin-tone pen used and trusted in dermatological settings.

Furthermore, we developed an algorithm tested against an open-source database that incorporates skin-tone measurements into blood oxygen measurements to correct inaccuracies found in pulse oximeter measurements, which is amplified in subjects with darker skin pigmentation. Overall, we demonstrate up to a 75% decrease in mean absolute error (MAE) and a decrease in overall variability across skin-tones simply by incorporating these skin-tone values in blood oxygen prediction.

Our contributions can be summarized as follows:

(1) An end-to-end framework that employs commercially available technology for non-invasive, continuous skin-tone monitoring.
(2) A model for transforming wearable multi-spectrum data to melanin, erythema, and skin-tone measurements.
(3) A comprehensive user study demonstrating the accuracy of this system across a variety of skin-tones and body locations.
(4) An algorithm to correct pulse oximeter inaccuracies in darker skin pigmentation by incorporating skin measurement in blood oxygen calculation.

The remainder of this work is structured as follows: Section 2 provides a background of melanin, erythema, and skin-tone, and the methods that have been used to attempt to quantify them. Section 3 details the DermaGlow framework, specifically the off-the-shelf device that is used for system set-up, the gold standard device that is used for comparison, and the algorithm developed for reporting of melanin, erythema, and skin-tone, as well as correcting for pulse oximeters. In Section 4, we evaluate our device against the gold standard measurements in a user study and evaluate our algorithm to correct pulse oximeter inaccuracies in dark skin participants. Lastly, we discuss our results and plans for future work in Section 5, and conclude in Section 6.

2 Background and Related Work

There has been substantial fundamental research on melanin and overall skin-tone constituents. Melanin, the pigment primarily responsible for skin color, plays a critical role not only in cosmetic appearance but also in protecting the skin from UV radiation [45, 59, 94]. Erythema, often manifested as redness of the skin, can indicate inflammation, allergic reactions, or other skin conditions [76]. Understanding the relation of these two skin constituents, their interaction with one another, and traditional measurement practices for these biomarkers is important in accurately assessing skin health. We will begin this section with an overview of melanin and erythema, two components of skin-tone that dramatically impact its visual appearance as well as its response to environmental and clinical factors. Then, we will discuss the traditional visual approaches of skin-tone quantification and classification. Lastly, we will discuss quantitative methods that have been developed to enhance these approaches.

2.1 Skin Composition and Evaluation

The visual appearance of skin is largely dictated by its chromophores, a term coined to describe specific molecules that absorb light radiation and emit color as a result. The main chromophores affecting skin color are melanin, which imparts a range of shades from brown to black, and hemoglobin, which causes erythema or redness; minor chromophores such as DNA and certain proteins are impactful in absorbing radiation, specifically in the ultraviolet (UV) spectrum, and thus play a more dominant role in skin protection, rather than altering visible color [92]. However, upstream effects can be demonstrable in cases such as DNA damage due to excessive UV exposure, where variations in melanin and/or hemoglobin can appear on the skin surface as a result. Thus, the interplay of these chromophores not only determines skin color variability among individuals but can also provide insight on the skin's physiological state [58, 92, 94].

Melanin is a complex polymer derived from the amino acid tyrosine and plays a pivotal role in determining skin color, as well as acting as a natural sunscreen by absorbing potentially harmful UV radiation. Analysis of melanin content is crucial for the assessment of various dermatological conditions which may manifest through abnormal melanin levels [45, 55, 59, 63, 94]. An increase in melanin might suggest conditions like hyperpigmentation or lentigo, whereas a decrease could be indicative of conditions such as vitiligo or hypopigmentation. Furthermore, melanin content can directly impact susceptibility to sunburn, as well as skin cancer [55, 94]. Skin melanin can be categorized into two basic types: eumelanin, which is visually black-brown, and pheomelanin, which is visually red-yellow. The ratio of these melanin derivatives directly determines skin composition, including how the skin

may appear visually; pheomelanin, for example, is found at high-levels in those with pale skin and red or fair hair, while eumelanin is found at high-levels in those with darker skin pigmentation [55, 68].

Erythema level is another significant constituent of skin-tone that directly influences the visual appearance of skin. Erythema is a common dermatological condition characterized by redness or rash on the skin resulting from capillary dilation and increased blood flow. The characteristic redness of erythema can be attributed to an increased concentration of hemoglobin in the affected area. Erythema is particularly important in the detection of early signs of skin wounds or infections and is present in skin that is affected by conditions that may be considered cosmetic, such as acne. It can also serve as an indicator of systemic diseases or reactions to medications. However, assessing erythema can be challenging, especially in those with darker skin pigmentation, potentially leading to underdiagnosis or delayed treatment of skin conditions [1, 76, 92].

Clinical evaluation of skin must involve analysis of color and composition, among other factors; both melanin and erythema are critical in these analyses [76]. However, the influence of melanin on clinical measurements has historically complicated clinical assessment. Erythema, for example, has largely been more difficult to assess in darker skin-tones due to melanin's dominant effect on skin color [18, 22, 32]. Healthcare provider difficulty in assessing erythema in darker skin tones has been implicated in health disparities associated with conditions such as pressure injuries, sexual assault injuries, Lyme disease, and the identification of limb ischemia among racial or ethnic groups with dark skin tones [6, 29, 50, 51, 61, 65–67, 75]. Further, the influence of melanin on these assessments extends to other clinical measurements, such as the potential for inaccurate pulse oximetry readings, where melanin's absorption properties can affect device accuracy. There has been substantial research on the impact of melanin on pulse oximeter measurements, which has been directly linked to negative health outcomes in those with darker skin pigmentation [25, 26, 35, 38, 40, 42, 57, 70, 73, 78, 84]. The underdetection of hypoxemia as a result has led to differences in treatment decisions, including administration of supplemental oxygen, intravenous dexamethasone, and hospital admission, which was particularly evident during the COVID-19 pandemic [25, 26, 78]. Furthermore, these differences may also extend towards other optically-obtained biomarkers such as heart rate, where it is still unclear whether there is a racial bias in reporting of heart rate in wrist-worn wearables [5, 8, 13, 14, 44, 69]. These challenges have necessitated the need for tools and devices to quantify and classify skin-tone and skin composition.

2.2 Methods of Skin-tone Quantification and Classification

Dermatological methods for classification of skin-tone have ranged from subjective assessment to quantitative approaches of skin composition. Traditionally, visual assessment has been a critical component of these practices using scales such as the Fitzpatrick [28], Von Luschan [87], Pantone [60], and the recently developed Monk Skin Tone Scales [54], all of which utilize visual and descriptive criteria to categorize skin color and its potential response to UV exposure. These systems are particularly focused on classifying skin-tones, which is mainly dictated by melanin; Clinical Erythema Assessment (CEA) [81] is a descriptive scale that is used to categorize erythema severity. While these systems differ in the depth and scope of their classifications, they share a common limitation in providing an objective, unbiased classification of skin composition. DermaGlow offers quantitative measurement of melanin, erythema, and skin-tone, without the need for subjective classification.

Given the optical properties of melanin, reflectance spectroscopy has been a prominent method for its study. There has been substantial research dedicated to examining the reflectance spectra of chromophores, particularly melanin, in human skin. Many of these studies have reported strong absorption characteristics in the approximate spectral range of 600-800nm by measuring the skin's reflectance in various conditions [46, 64, 93, 94]. Similarly, the spectral response of erythema, primarily influenced by hemoglobin content, has been evaluated using similar methods. While traditional pulse oximetry uses red (660nm) and infrared (940nm) wavelengths to measure blood oxygenation [10, 56, 83] - wavelengths chosen for their tissue penetration depth and relative isosbestic

properties - the assessment of erythema at the skin level focuses on different spectral regions. Erythema is typically characterized in the visible spectrum, particularly between 500nm-700nm with strong absorption in the green region [24, 77]. Furthermore, the spectral response of both melanin and hemoglobin is dependent on the location of measurement [64, 93], as different body sites vary in their blood perfusion and tissue composition, thereby necessitating careful consideration of measurement sites to ensure accurate and consistent assessment.

Table 1. CIE System Axis Definitions

Axis Name	Reporting Range	Colors
L*	0 - 100	Light to dark
a*	$-\infty$ - $+\infty$	Green to red
b*	$-\infty$ - $+\infty$	Blue to yellow

Recently, colorimeters have been developed as a quantitative measure of skin-tone using optical spectroscopy by examining remittance spectra across the visible spectrum. Current commercial colorimeters are handheld devices that are used to objectively measure skin color using a three-dimensional color space following the Commission Internationale de l'Eclairage (CIE) system [90]. This system is used to represent colors in a way that is more aligned with human vision, as opposed to the traditional RGB or CMYK systems used in digital or print mediums. The CIE categorizes its axes as shown in Table 1. While a* and b* are technically unbound, they are often clamped to the range of -128 to 127 for use with integer code values.

In addition to these values, colorimeters also often report the individual typology angle (ITA). ITA is a measure used to quantify skin color using the CIE scale, specifically using the L* and b* components. The formula for calculating ITA is as follows:

$$\text{ITA} = \arctan\left(\frac{L* - 50}{b*}\right) \times \frac{180}{\pi} \tag{1}$$

The result, expressed in degrees, helps categorize skin into various typological groups. Skin is classified based on ITA measurement using the ranges specified in Table 2. ITA provides a quantitative measure of skin-tone that is classified into interpretable groups, as opposed to more subjective approaches that are typically used in this context such as the Fitzpatrick [28] skin-tone scale, thus making it the preferred approach to quantify skin-tone if available. This technique is particularly useful in cosmetic science and dermatology to analyze skin characteristics and potential cosmetic products based on skin undertones [16, 41, 48].

Table 2. Skin Classification based on ITA

Classification	**ITA Range**
Very light skin	ITA > 55°
Light skin	41° < ITA ≤ 55°
Intermediate skin	28° < ITA ≤ 41°
Tan skin	10° < ITA ≤ 28°
Brown skin	-30° < ITA ≤ 10°
Dark brown to black skin	ITA ≤ −30°

Colorimeters that utilize the CIE system can accurately measure skin color independent of the actual device that is being used for measurement. Since melanin and erythema directly impact red and green measurements, many colorimeters also report an erythema and melanin value as well. However, a confounding factor in the analysis of these two skin constituents is that their absorption spectra overlap (600-800nm range). Thus, implementation across devices for the extraction of these two biomarkers can differ, especially in the wavelengths that are utilized to infer the measurements. A summary of colorimeters and their analysis techniques is shown below in Table 3. All of these devices utilize reflectance spectroscopy, but may or may not report CIE LAB, melanin, and erythema levels altogether [41, 48].

Table 3. Commercial colorimeters and their detected spectral ranges.

Device	Reported Spectral Range	CIE LAB?	Melanin/Erythema?
Mexameter MX18[12]	568, 660, and 870nm	No	Yes
SkinColorCatch[82]	460, 540, and 620nm	Yes	Yes
Cortex DSM IV[15]	470, 510, and 660nm	Yes	Yes
Colorimeter CL400[19]	440 - 670nm	Yes	No
AMT510 Colorimeter[2]	400 - 700nm	Yes	No
Minolta CM-700D[52]	400 - 700nm	Yes	No

As previously mentioned, accurate and comprehensive skin analysis requires examining two key components: skin physiology (specifically melanin and erythema levels) and overall skin color. While melanin and erythema are fundamental physiological markers, skin color can vary independently due to various factors including diet [91], environmental conditions [21], and the use of cosmetics. This makes measuring both skin color and the underlying physiology essential for a complete assessment. Of the colorimeters available on the market today, only two, the Cortex DSM IV [15] and the Delfin SkinColorCatch [82], report melanin, erythema, and skin-color measurements. While these colorimeters offer valuable capabilities for clinical dermatology applications, they

face several practical limitations: high cost (often several thousand dollars, with prices typically restricted to institutional buyers), bulky clinical form factors, requirements for periodic calibration, and limited data collection frequency. In contrast, DermaGlow provides a cost-effective solution for continuous monitoring of melanin, erythema, and CIE LAB values, designed for integration across various wearable formats.

To date, there has been limited work on sensors and devices to monitor skin condition in a wearable or low-cost form-factor. Sundroid [23] was recently developed as a wearable prototype system that measures incident UV radiation using a body-worn sensing unit to alert users of harmful sunlight exposure. The system achieves this by directly measuring sunlight, rather than skin properties, and is thus not directly fixated directly on the skin but can be attached to the user's clothing. Several commercial devices use a similar principle, such as My UV Patch by La Roche-Posay [47], SunIndex [80], and LogicInk [49], which directly track UV exposure. Other devices have utilized imaging techniques to quantify skin lesions and potential dermatological conditions, such as the DermLite DL5 [17], MoleScope [53], and SkinVision [74]. Recently, a low-cost skin spectrum measurement device was developed which provides melanin, erythema, and skin-tone content using optical spectroscopy in a small 3D printed housing structure [39]. The device utilizes a deep neural network model, along with an accompanying mobile application, to provide insights on these parameters. However, the device is not designed for continuous use, as it needs to be held against the skin. Furthermore, discussion on the model, considering the small subject sample size for the device (16 subjects) and the training model used, is limited especially pertaining to efforts used to prevent potential overfitting. There is also no discussion regarding the diversity or skin-tone classification of the recruited subject group. Overall, none of these devices or applications provide detection of melanin, erythema, or skin-tone, and/or are limited to discrete measurements rather than continuous data collection. DermaGlow is a flexible solution that enables real-time, continuous analysis of skin composition directly and is developed on a diverse group of skin-tones, expanding upon commercial skin monitoring solutions and colorimeters by enabling these analyses in daily routine.

3 DermaGlow Architecture

In this section, we will discuss the components utilized to develop the DermaGlow framework. We first describe the off-the-shelf device we utilized to gather data for algorithm development. Next, we describe the gold standard device that we utilized as reference measurement. Lastly, we discuss the algorithmic pipeline developed to report melanin, erythema, and skin-tone using the framework, and the methods used to correct pulse oximeter readings using skin-tone measures. Figure 2 demonstrates our overall framework architecture for DermaGlow. Each step is elaborated upon in subsequent sections.

3.1 Commercial Wearable

We utilized Lumos [89], an off-the-shelf wearable device, to gather optical spectroscopy data. Lumos is a multi-spectrum device that utilizes light-emitting diodes (LEDs) and photodiodes (PDs) that cover the entire visible spectrum, ranging from 400nm to 1000nm, allowing us to fully examine the spectral properties of the skin in a radiation range that is safe. The device is designed to be versatile and is available in multiple form-factors, which can measure skin content on any location in the body. Lumos utilizes off-the-shelf LEDs and PD arrays on a custom PCB, with components totaling less than $50.

Lumos contains 9 LEDs and 10 PDs across the visible spectrum to optimize spectral resolution. Each LED operates within a lumen range of 104-114 and covers wavelengths from 415nm to 910nm. The PD channel can detect wavelengths from 350nm to 1000nm, operating at a sampling rate of 0.3Hz. The operable spectral response of the LEDs and PDs for the Lumos device can be found in Figures 3a and 3b below, respectively.

While these components can be adjusted to capture wavelengths beyond this range, they are particularly suited for our current application given the specific absorption characteristics of melanin and erythema. Regardless

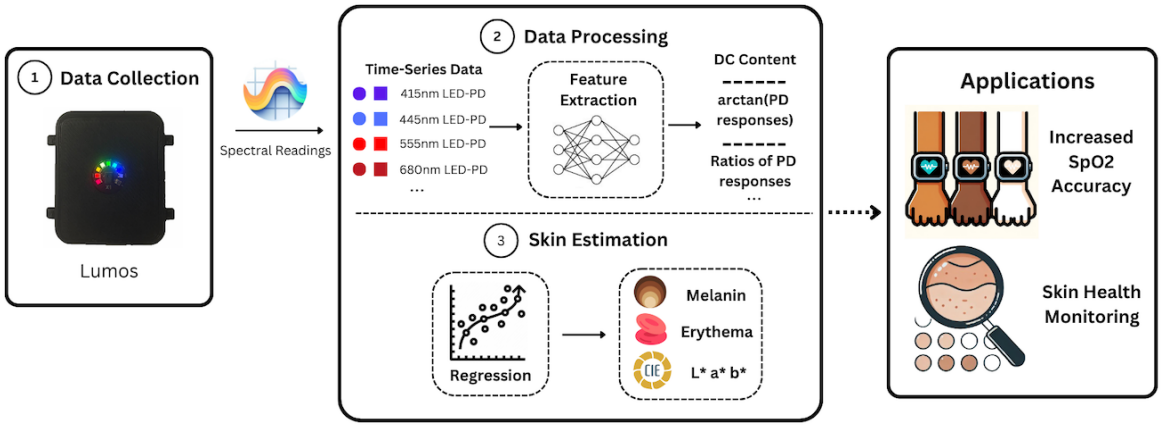
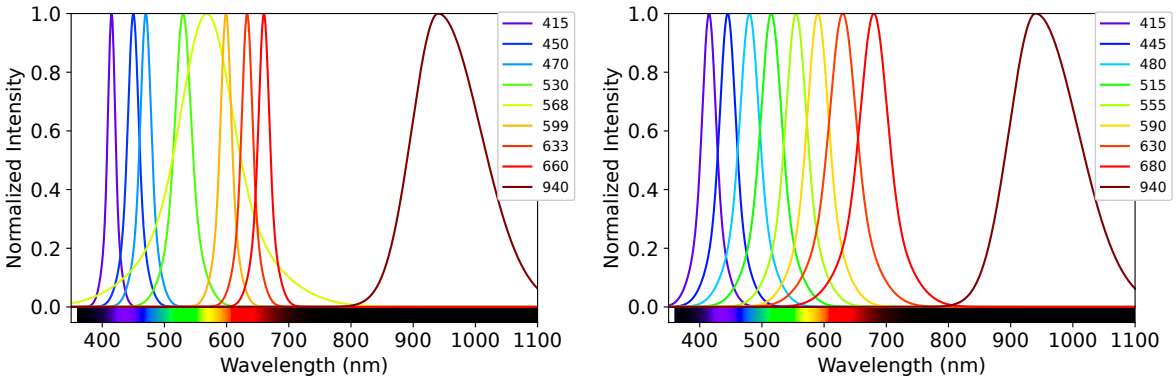


Fig. 2. DermaGlow Framework Overview: DermaGlow is a framework that utilizes off-the-shelf multi-wavelength optical sensing to predict melanin, erythema, and skin-tone values.

of configuration, Lumos captures spectral data in reflectance mode. It cycles through each LED, sequentially activating them and recording the corresponding PD responses. This cycle repeats, allowing a complete sweep of all 9 LEDs and corresponding 10 PDs in approximately 3 seconds. This systematic process ensures that we collect thorough and precise spectral data consistently.



(a) LED Wavelengths (b) PD Wavelengths

Fig. 3. Spectral Response of Lumos LEDs and PDs.

3.2 Gold Standard Comparative Device

We utilized the Cortex DSM-IV as a comparative gold-standard against Lumos output data. We did not use a visual scale such as the Fitzpatrick scale to avoid subjective bias in our model. Our goal was to ensure that DermaGlow's reporting was as comprehensive as possible by using a device capable of accurately reporting melanin, erythema, and CIE LAB values. As mentioned previously, the current colorimeters on the market that meet this criteria are

the Delfin SkinColorCatch [82] and Cortex DSM-IV [15]. Both of these devices use similar spectral ranges to report these parameters as shown in Table 3.

The comparative literature on these devices is limited but informative. Van der Wal *et al.* [85] compared the performance of these colorimeters against several visual scales and found them both to provide reliable color data on skin and scars with a single measurement. Baquie *et al.* [4] compared the SkinColorCatch with the Mexameter, and found the SkinColorCatch to be far more reliable and robust over a range of melanin and erythema values. To our knowledge, there has been no direct comparison of the SkinColorCatch against the DSM-IV. While both the DSM-IV and SkinColorCatch offer comparable technical capabilities, several factors influenced our selection of the DSM-IV colorimeter:

(1) Cost effectiveness: the DSM-IV presents a relatively more economical option compared to the SkinColorCatch without compromising measurement quality.
(2) Operational efficiency: the DSM-IV features built-in calibration protocols, whereas the SkinColorCatch requires intermittent in-lab calibration.
(3) Availability: With the discontinuation of the DSM-II and DSM-III models by Cortex Technology, the DSM-IV represents the only currently supported version of this colorimeter.
(4) Technical advancements: as an upgrade to the discontinued DSM-II and DSM-III, the DSM-IV incorporates significant improvements, including but not limited to enhanced spectral sensing, improved software algorithms, and better reproducibility.

These advantages, particularly the DSM-IV's enhanced sensor technology and updated calibration protocols, address the limitations noted in earlier DSM-II studies while matching the capabilities of the more expensive SkinColorCatch. Overall, these features make the DSM-IV an ideal choice for benchmarking Lumos performance in quantitative skin analysis.

3.3 Data Processing and Feature Engineering

Lumos provides raw spectral data in the form of analog-to-digital-converter (ADC) counts for each photodiode in each cycle of LED flashed during its operation. Thus, for each cycle of Lumos measurements consisting of 9 LEDs and 10 PDs, there are a total of 90 PD responses, where each value is the LED that is turned on at that moment in time, with its corresponding PD. Correspondingly, this data can become very dense as it is operating on multiple scales (LED, PD, ADC counts, time). To visualize and analyze this data, we utilized SpectraVue [37], an open-source platform that supports visualization of multispectrum signals. This platform is especially helpful to gain an underlying intuition for any patterns or trends that can be gathered heuristically. Utilizing both SpectraVue and techniques used in traditional photoplesmography, we developed the following features:

(1) Mean value of PD responses - calculate the mean value of each PD response over the course of the measurement.
(2) Direct current (DC) content of each PD response - this feature was calculated based on the assumption that underlying skin properties, such as melanin, erythema, or skin-tone, do not change instantaneously during a relatively short sampling period (< 1 minute). Therefore, any alternating current (AC) components detected in the spectral response are attributed to noise or biomarkers not targeted in this study.
(3) arctan of each PD response - a domain specific technique that is used to calculate ITA; the resulting angle can be interpreted as a measure of the relative contribution of different spectral components, similar to how ITA represents the relationship between color scales in skin-tone measurement.
(4) PD ratios with one another - in pulse oximetry, blood oxygenation is traditionally determined by calculating the ratio of absorption between the infrared and red light responses of the body. This method, known as the ratio-of-ratios (ROR), involves comparing the AC and DC components of both waveforms relative to each other [10, 56, 83]. We have extended this technique to a multispectral context by calculating ratios

between various responses from different LED and PD combinations. We have also taken ratios of more advanced features (such as DC content of various PD responses) as a regularization technique to maintain feature monotonicity while reducing the impact of extreme values.

To extract the baseline skin content levels for feature selection, we high-pass filter all LED and PD combinations during a measurement period using a 2nd-order Butterworth filter with a 0.1 Hz cutoff frequency. In this way, we effectively extract the DC component of the underlying signal, which we assume to be indicative of skin content. This methodology also helps remove any motion artifact that may occur during data collection. Following this filtering, as well as calculating the arctan of each LED and PD response, we then took ratios of all features with each other as an additional regularization measure.

3.4 Machine Learning

We tested several supervised learning algorithms by utilizing the aforementioned multivariate features to predict each of the desired outcome variables, namely melanin, erythema, L*, a*, and b* values. Since ITA is directly calculated using L* and b* values, we did not create a model for its prediction. We aimed to develop a generalizable and robust model by training several different linear and tree-based regression algorithms against each target variable, including linear regression, Ridge regression, RandomForest, XGBoost, LightGBM, and CatBoost models. The total feature set was reduced to 10 features for melanin and erythema, and 50 features for CIE LAB values, using recursive feature elimination (RFE) with mean absolute error (MAE) as the optimization criterion. During RFE, the native feature importance metric of each model was used to rank features: coefficient magnitude for Lasso and Ridge regression, variance reduction for RandomForest, ExtraTrees, and GradientBoosting, gain for XGBoost, and loss-function-based importance for CatBoost and LightGBM. Model hyperparameters were tuned using a grid search and selected based on achieving the optimal balance between accuracy and model complexity. Models were validated using both 5-fold cross-validation with subject stratification and Leave-One-Out Cross-Validation (LOOCV) to prevent subject data leakage between training and validation sets while ensuring robustness of our results across different validation strategies.

Model performance was evaluated using multiple metrics: for melanin and erythema predictions, we used R^2, normalized mean absolute error (NMAE), and normalized root mean squared error (NRMSE). These normalized metrics enable meaningful comparisons for melanin and erythema, whose values are abstract. For CIE LAB color predictions, we used Delta E (ΔE) the industry standard for color difference measurements in LAB space. ΔE represents the Euclidean distance between predicted and true colors, as shown in Eq 2.

$$\Delta E = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{2}$$

Unlike individual color component errors, ΔE provides a perceptually relevant measure of total color difference that aligns with human visual perception. In clinical and dermatological applications, ΔE values less than 2-3 are considered acceptable as they represent color differences that are difficult to distinguish without careful observation. Full evaluation of model performance using these metrics can be found in Section 4.

3.5 Pulse Oximetry Correction

We evaluated the feasibility of correcting pulse oximeter measurement inaccuracies in darker skin tones using the OpenOximetry dataset from PhysioNet [30, 33]. This centralized repository contains open-source pulse oximetry data, reference arterial line measurements, and spectrophotometer measures, including melanin, erythema, and LAB values, across a variety of skin-tones, measurement locations, pulse oximeter devices, and spectrophotometers. The repository classifies pulse oximeters based on their performance across different skin tones, enabling systematic evaluation of device accuracy among demographic groups. Within the dataset, each patient participates in multiple encounters where various pulse oximeter devices are tested against arterial blood

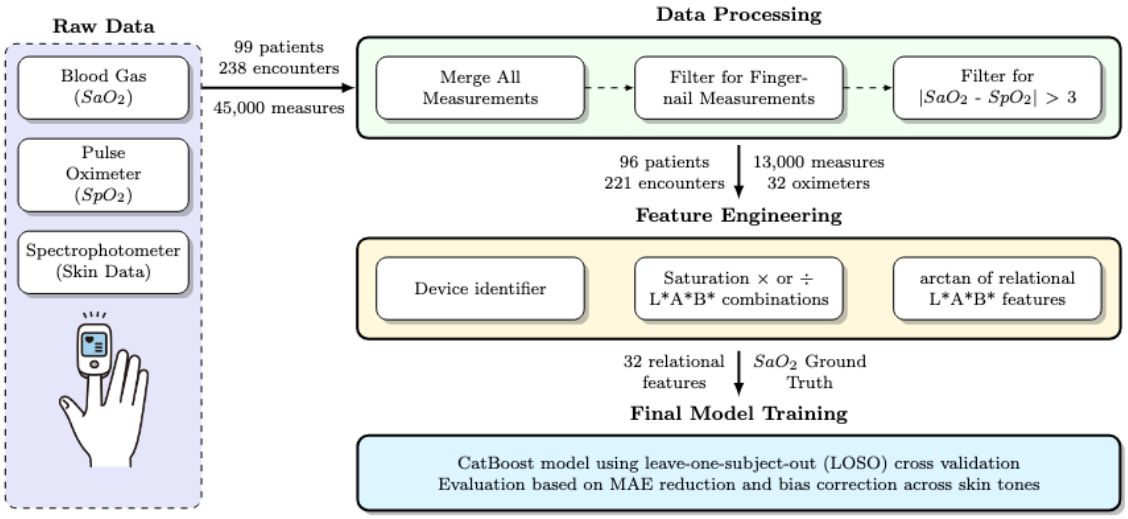


Fig. 4. Data processing, feature engineering, and model training pipeline for pulse oximeter bias correction using OpenOximetry dataset from PhysioNet.

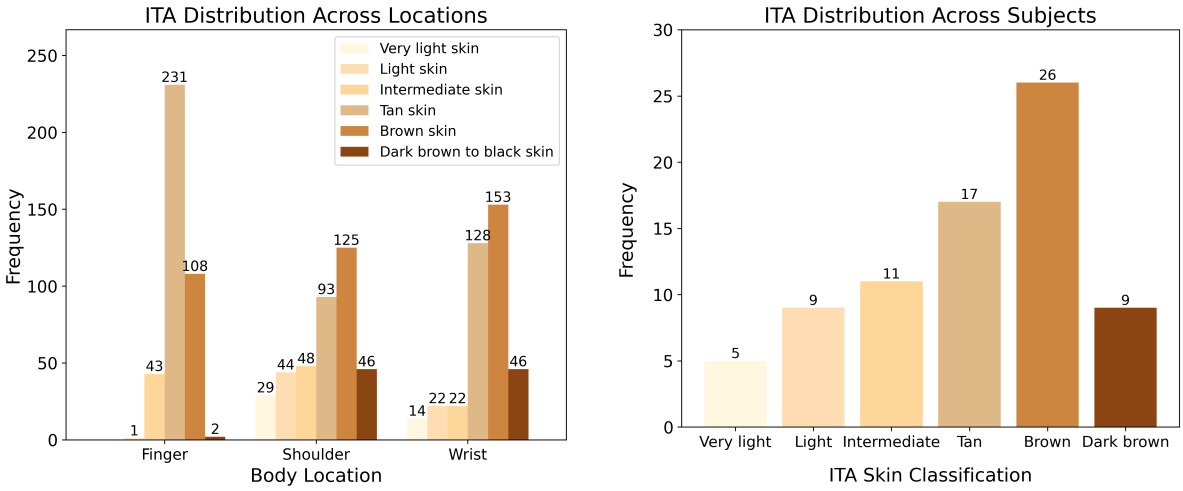
oxygen measurements, while spectrophotometer readings are simultaneously collected across different anatomical sites. We analyzed data across this repository to inspect differences in blood oxygen measurements obtained through pulse oximeters relative to arterial line measurements in various skin-tones to determine whether blood oxygen measurements can be corrected by utilizing skin-tone information. A detailed process flow diagram for the developed pulse oximetry bias correction algorithm can be found in Figure 4. The pipeline processes raw data collected from blood gas measurements (SO_2), pulse oximeter readings (SpO_2), and spectrophotometer values, comprising 99 patients across 238 encounters. Given that encounters contain spectrophotometer measurements from multiple anatomical locations, we isolated fingernail measurements to align with the exclusive use of finger pulse oximeters in the repository. We then implemented a filtering criterion to exclude cases where the difference between SO_2 and SpO_2 exceeded 3%. This threshold was selected for two primary reasons: the OpenOximetry database's pulse oximeter accuracy classification is not yet reflected in the raw relational database and thus non-compliant devices cannot be determined inherently, and FDA guidance stipulates pulse oximeter accuracy should fall within 3% of ground-truth measurements [31]. This refinement yielded a final dataset of 96 patients with 221 corresponding encounters. To minimize the influence of raw saturation values on model predictions, we engineered relational features through various polynomial and trigonometric combinations of L*, A*, and B* values. These engineered features, along with a single device identifier, serve as inputs to a CatBoost regressor model which employs leave-one-subject-out cross validation to prevent data leakage. The unique device identifier is used as different devices have different calibration curves for blood oxygen calculations, which can affect their bias for the same participant. Model performance is evaluated based on mean absolute error improvement relative to the baseline oximeter prediction as well as bias correction differences across skin tone groups; these results are presented in Section 4.

4 Evaluation

In this section, we describe the methods used to verify DermaGlow's output relative to the DSM-IV. First, we outline the experimental procedure used to gather across skin-tone data across a diverse group of subjects. Next, we evaluate our model's accuracy and generalizability across these skin-tones against the DSM-IV output. Lastly, we assess the performance of utilizing skin-tone measurements to correct for pulse oximeter measurements especially found in those with darker skin pigmentation.

We evaluated the accuracy of the DermaGlow framework in an IRB-approved (IRB-HSR #301300), 77 person user study. Participants were not directly recruited based on their skin-tone, race, gender, or ethnicity; recruitment was conducted broadly to reach a general population without specific targeting. Since skin content values can depend on location of measurement, we recorded both Lumos measurements and DSM-IV measurements on 3 sites of the body: the palmar side of the fingertip, the dorsal side of the wrist, and the shoulder. This variety in measurement locations allows the DermaGlow device to operate effectively in a location-agnostic manner. Lumos measurements were collected at each site for approximately 60 seconds, with a subsequent DSM-IV measurement at the same site. Each measurement was repeated 5 times per location, totaling approximately 1155 observations.

4.1 Experimental Procedure



(a) Observation frequency of ITA classifications across measurement locations.

(b) Observation frequency of ITA classifications per subject using a single shoulder measurement.

Fig. 5. Distribution of skin-tone classifications based on ITA across body locations (a) and across subjects (b).

A histogram of the ITA classification of study participants, as well as the ITA classifications of each body location can be found in Figures 5a and 5b. As shown in Figure 5a and 5b, which represent the frequency distribution of ITA measurements across all three body locations (fingertip, wrist, and shoulder) and the distribution of the first shoulder measurement of each subject alone, respectively, the ITA classification distributions exhibit notably different patterns when comparing measurements across all body locations versus a single shoulder measurement. The shoulder measurements (Figure 5b) were used as the primary indicator of subjects' natural skin tone, as this measurement location typically experiences minimal environmental exposure compared to the fingertip or wrist,

making it less susceptible to temporary alterations like tanning. While both distributions approximately follow a normal distribution centered around the tan and brown categories, Figure 5a demonstrates how measuring different body locations can significantly skew the apparent distribution of skin tones, with a markedly higher frequency of tan classifications compared to the shoulder-specific measurements, especially with regards to finger measurements. This finding suggests that finger measurements naturally cluster toward intermediate skin tones, regardless of the subject's overall skin-tone.

4.2 DermaGlow Accuracy

In this section we present our results for melanin, erythema, and CIE LAB prediction using both LOOCV and 5-fold cross validation across several regression models. We contextualize these results across measurement locations, as well as skin-tones, and provide important features for each predictor.

Table 4. LOOCV model results for melanin, erythema, and CIE LAB.

Model	Melanin		Erythema		CIE LAB
	NMAE (%)	NRMSE (%)	NMAE (%)	NRMSE (%)	ΔE
RandomForest	5.25(±3.56)	7.77(±5.53)	4.59(±9.55)	6.34(±11.94)	< 2.5
ExtraTrees	5.41(±3.85)	7.90(±5.32)	4.70(±9.57)	6.44(±11.73)	< 2.5
XGBoost	5.36(±3.66)	7.81(±5.42)	4.63(±9.55)	6.40(±11.84)	< 2.5
GradientBoosting	5.44(±3.57)	7.76(±5.62)	4.62(9.32±)	6.22(±11.79)	< 2.5
CatBoost	5.17(±3.45)	7.64(±5.03)	4.49(±9.27)	6.17(±11.68)	< 2.5

4.2.1 LOOCV Results. The results of the five highest performing models using leave-one-out-cross-validation (LOOCV) across all target outputs is summarized in 4 using NMAE and NRSME; R^2 was not used as an evaluation metric for this validation method due its high variance when computed on small validation sets. Overall, our models perform well on melanin predictions, with relatively low NMAEs and NRMSEs. Erythema exhibits higher variance in prediction, likely due to outliers in the dataset that the framework may struggle to generalize to. This is expected as recruiting for diverse erythema levels can be extremely challenging. For CIE LAB values, we achieve errors below 2.5 ΔE units, indicating excellent color accuracy that falls within generally accepted thresholds for perceptible color differences. We expect higher sampling of minority classes to strengthen these results in future studies. Nonetheless, the results are relatively strong across all prediction metrics, especially considering the range of values that can cause large amounts of variance, even within the same measurement site on the same subject.

Table 5. 5-fold CV Regression model results for melanin , erythema, and CIE LAB.

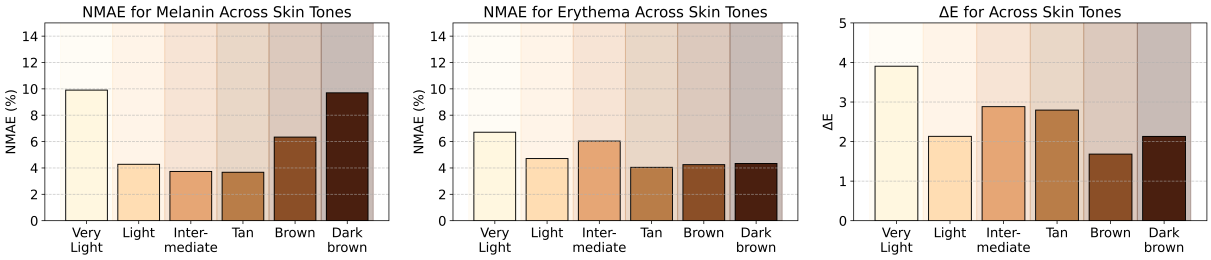
Regressor	Melanin			Erythema			LAB
	NMAE (%)	NRMSE (%)	R^2	NMAE (%)	NRMSE (%)	R^2	ΔE
RandomForest	5.37(±0.59)	8.02(±1.81)	0.81(±0.06)	4.69(±2.30)	6.22(±2.24)	0.66(±0.06)	< 2.5
ExtraTrees	5.39(±0.63)	7.99(±1.82)	0.81(±0.06)	4.18(±2.11)	6.47(±2.41)	0.66(±0.06)	< 2.5
XGBoost	5.39(±0.61)	8.01(±1.66)	0.81(±0.06)	4.46(±2.15)	6.16(±2.27)	0.67(±0.05)	< 2.5
GradientBoosting	5.50(±0.55)	8.14(±1.59)	0.80(±0.06)	4.42(±2.27)	6.06(±2.47)	0.68(±0.04)	< 2.5
CatBoost	5.33(±0.60)	7.95(±1.72)	0.81(±0.06)	4.42(±2.26)	6.07(±2.40)	0.67(±0.05)	< 2.5

4.2.2 5-fold CV Results. The results of the five highest performing models using 5-fold cross validation with subject stratification across all target outputs is summarized in Tables 5. The reported metrics - NMAE, NRMSE, and R^2 - for each regressor is a mean score ± standard deviation. The strong and consistent performance across

many regressors underscores the strength of the feature-set data in predicting skin constituents. Melanin and erythema achieved less than 6% and 5% NMAE, respectively, relative to the reference measurement across all regression models. For LAB color space evaluation, all models achieved ΔE values below 2.5, demonstrating high color prediction accuracy that meets industry standards for acceptable color differences. Overall, the general accuracy of the DermaGlow framework is relatively strong, especially considering the variance that can occur from measurement to measurement; we expect even stronger results with more diverse, targeted recruitment.

4.2.3 Contextualizing Model Accuracy. Although there are no established clinical thresholds for acceptable prediction errors in melanin or erythema, which limits formal clinical interpretation of model performance, the NMAE values achieved by our framework represent small deviations relative to the full physiological ranges of these biomarkers. Melanin index values commonly range from 20 to 120 across skin tones, while erythema indices typically vary from 0 to over 50 depending on inflammation and vascular status. A 5% NMAE corresponds to an error of approximately 1–6 units for melanin (depending on baseline skin tone) and roughly 0.5–2.5 units for erythema. In both cases, these small deviations are unlikely to alter clinical interpretation or influence treatment decisions; Vasudevan *et al.* [86] reported intra-observer coefficients of variation below 10% across various commercial melanometers and colorimeters, indicating that our measurements fall comfortably within the range of industry-accepted variability.

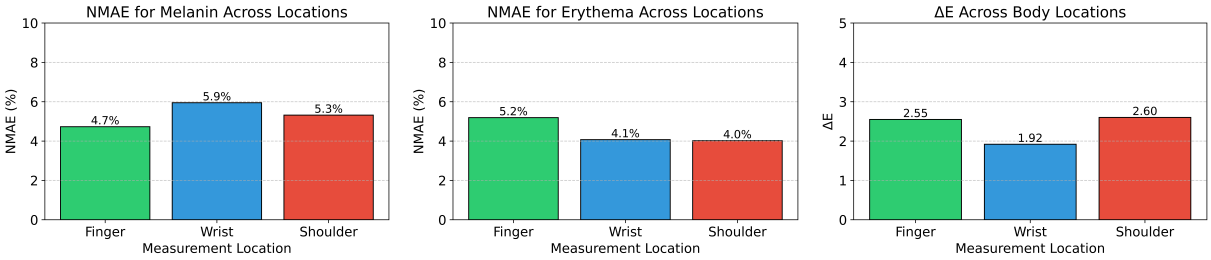
4.2.4 Accuracy across Skin-Tones and Locations. To evaluate if this accuracy is consistent across skin-tones, we grouped NMAE and ΔE results based on their ITA classifications; these results are presented in Figure 6. The analysis demonstrates that our models achieve consistent performance across most skin tones for melanin and erythema predictions, with marginally higher error rates observed in the very light and dark brown categories. These slight variations in accuracy can be attributed to the lower representation of these skin tones in our study population. Very light skin tones likely had higher ΔE values due to their low sample size and the impact of redness that can be seen visibly in these skin-tones. Future studies with more balanced representation across all skin tone categories, particularly in the edge classes, are expected to further improve these results.



(a) Melanin (b) Erythema (c) CIE LAB

Fig. 6. Normalized mean absolute error (NMAE) and ΔE values across target variables grouped by ITA skin-tone classification.

NMAE and ΔE values across each body location are presented in Figure 7 below. The errors reveal consistent performance across different body locations, demonstrating the ability of DermaGlow to operate in a location-agnostic manner despite the variations in skin-tone distributions seen across body locations. The highest ΔE values were observed in the shoulder and finger measurements. This may be due to the uneven distribution of ITA measurements seen in finger measurements, as well as the scales of these values having much higher ranges, which may cause variation especially in underrepresented groups. We expect more stability in these measurements with a larger and more diverse dataset.



(a) Melanin (b) Erythema (c) CIE LAB

Fig. 7. Normalized mean absolute error (NMAE) and ΔE values across target variables grouped by location of measurement.

4.2.5 Feature Importance. We evaluated feature importance of the resulting models using their default regression importance metrics. These native techniques revealed that the most important features for melanin and L* prediction were arctan features, specifically in the 600-700nm range, while the DC features were the most important predictors for erythema, A*, and B*, particularly in the 500-800nm range. These results coincide with the expected spectral contribution of melanin [46, 64, 93, 94] and erythema [24, 77] found in previous lab studies. A list of top feature importances for each output target and best performing model can be found in Appendix A.

4.3 Application to Pulse Oximetry

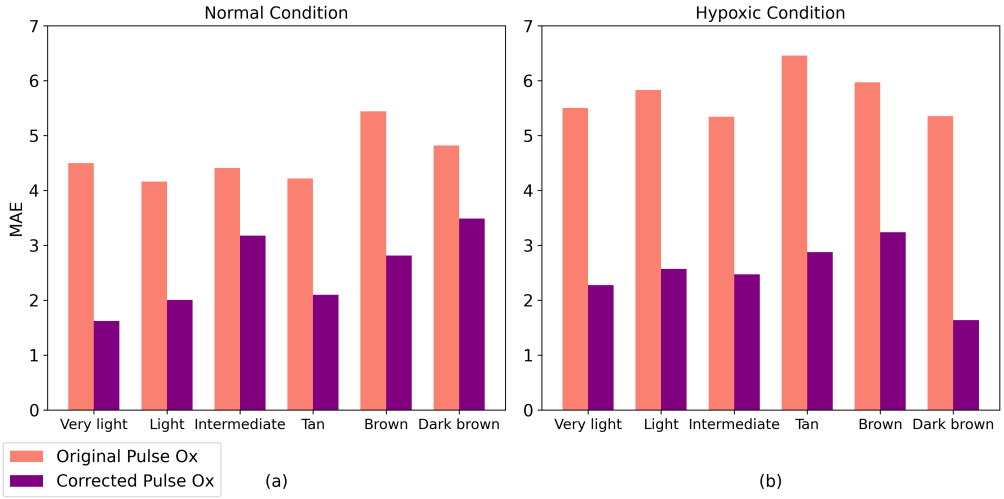
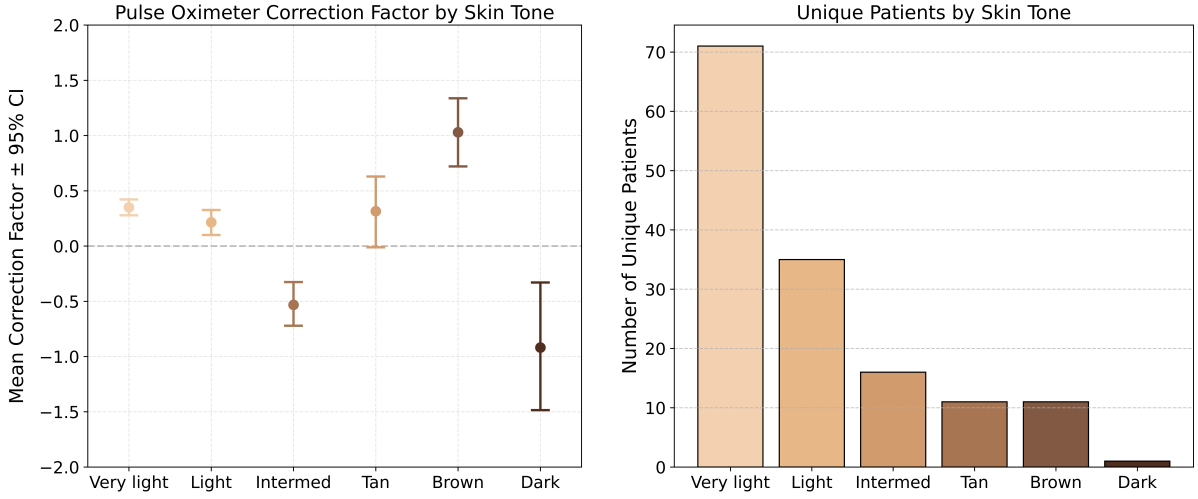


Fig. 8. Mean Absolute Error (MAE) is decreased in both normal (>=90 SpO_2) and hypoxic (<90 SpO_2) conditions for devices with high SpO_2 error (>3%) by incorporating skin-tone measurements.

Utilizing the PhysioNet Open Oximetry repository, we examined differences between measurement accuracy in pulse oximeters relative to arterial blood measurements across skin-tone groups. We clustered subjects into skin-tone groups based on ITA and associated skin-tone classification. Results indicated a significant variation in pulse

oximeter measurement accuracy across skin-tones (p-value < 0.0001) in both normal and hypoxic conditions across many commercial-grade and hospital-grade devices. MAE is also generally elevated across skin-tones in hypoxic conditions. This finding has been confirmed over a variety of previous works [25, 26, 35, 38, 40, 42, 57, 70, 73, 78, 84].

The model we developed as a corrective measure for this error resulted in a significant improvement, achieving an up to 75% reduction in Mean Absolute Error (MAE) across different skin tones, and notably reducing variability among these groups, as illustrated in Figure 8. Notably, this same MAE decrease was observed in hypoxic conditions, where many racial disparities in pulse oximeter measurements have been reported [25, 26, 35, 38, 40, 42, 57, 70, 73, 78, 84]. We have also confirmed this finding using the complete pulse oximeter dataset, demonstrating that our correction attenuates biases in poor measurements while preserving the accuracy of already reliable readings. Utilizing feature importance techniques, we discovered the most significant feature in this model was the relation of L* and b* with the calculated pulse oximeter saturation. These results highlight the potential of DermaGlow to enhance the accuracy and reduce biases in skin-tones in existing pulse oximetry devices, thereby improving their generalizability across diverse populations.



(a) The mean bias correction factor for pulse oximeters is significantly different between ITA skin-tone classifications.

(b) The ITA distribution for unique patients in the OpenOximetry database using the fingertip measurement.

Fig. 9. (a) Mean bias correction factor for all blood oxygen measurements per skin tone group, demonstrating significant differences across groups. (b) ITA scale distribution for fingertip measurements in the OpenOximetry database, showing a skewed distribution with fewer dark-skinned individuals.

We evaluated the pulse oximeter bias corrections across all measurements stratified by skin-tone and found statistically significant differences between groups (p < 0.0001), as shown in Figure 9a. This finding suggests that the model does not apply a uniform correction across all skin tones but instead integrates skin tone information dynamically during prediction. Furthermore, the OpenOximetry dataset contains relatively few dark-skinned fingertips, leading to a distribution that is heavily skewed towards lighter-skinned participants as shown in Figure 9b. Despite this limitation, the observed trends remain consistent and statistically robust across the full range of skin tone groups, suggesting that incorporating skin tone as a predictive feature in blood oxygen estimation may be key to enabling equitable pulse oximetry.

5 Discussion and Future Work

In this section, we discuss future work planned for the DermaGlow framework. We begin by discussing the overall study limitations and plans to recruit a larger participant pool for algorithm development, with greater representation across skin-tones. Second, we explore various applications for DermaGlow, with a focus on current pulse oximeters, AI, and wearable devices. Lastly, we discuss potential commercialization avenues for DermaGlow in the dermatological and cosmetic spaces.

5.1 Dataset Limitations

While DermaGlow was evaluated across a diverse cohort, certain limitations in skin tone representation should be noted. Specifically, participants with very light and dark skin tones were underrepresented in our dataset, particularly in fingertip measurements. This underrepresentation likely contributed to marginally high error rates observed in these groups (Figure 6), and limits generalizability in edge cases of the ITA scale. While we employed objective, ITA-based classifications to avoid relying on subjective race labels, our participant pool still skewed toward intermediate skin tones. This skew was partly due to natural population demographics and practical limitations in recruiting individuals at the extremes of the ITA spectrum. We believe that future studies, with targeted recruitment efforts, will increase robustness of our framework across racial and ethnic groups.

5.2 Clinical Evaluation

One of the clear next steps for DermaGlow is large-scale validation and qualification. To improve system generalization across various skin-tones and conditions, we plan to conduct a clinical trial with recruitment of a larger and more diverse participant base. Targeted recruitment of underrepresented skin-tone groups (such as very light and darker skin-tones) should be conducted to ensure a fair representation across skin-tones. It is important to note that although ITA classifications of individuals help contexualize otherwise abstract values, a diverse set of ITA values within each skin-tone group is equally as vital when recruiting a diverse subject population. Similarly, ITA distributions within each measurement location can also vary quite significantly, and should be considered in future studies. Lastly, it would also be of critical importance to validate results using multiple colorimeters and spectrophotometers as ground-truth measurements to ensure the device is not biased to a specific colorimeter's results. We plan to conduct a validation study with these parameters in mind.

It would also be beneficial for future work to directly recruit for varying erythema levels, as erythema levels can change drastically from participant to participant, as well as over time. Future work should also consider targeted recruitment of individuals with conditions affecting the skin, such melanomas and carcinomas, cyanosis, and PVD; DermaGlow could be used in this case to analyze skin composition and monitor changes throughout the progression of these conditions. By integrating melanin and erythema levels associated with various skin and environmental conditions, we could gain valuable insights for the long-term monitoring of various health issues.

These large-scale validation studies can also enable the examination of UV exposure's effects on melanin and erythema levels. It would be particularly valuable to observe relative changes between melanin and erythema, especially in dermatological or environmental conditions where these values can change over time. While localized UV exposure effects are well-documented, [7, 72], the systemic impacts of UV radiation on overall skin health and its relationship to melanin and erythema changes across different body sites remain largely unexplored. Furthermore, the effects of UV exposure on other biomarkers, such as heart rate and oxygenation, can be examined more easily using the DermaGlow framework.

5.3 Optical Sensor Considerations

Environmental factors such as ambient light variations and participant movement can significantly impact optical measurements in wearable devices [9].While these conditions have been thoroughly explored within

the context of PPG, our current implementation builds upon these findings and incorporates several signal processing strategies that show promise in maintaining measurement integrity in real-world conditions. First, our baseline DC filtering approach effectively removes low-frequency noise from ambient light variations and gradual posture changes. The normalization features we developed, particularly the optical density ratios between different sensor pairs, further reduces the impact of motion artifacts by considering relative rather than absolute changes. Additionally, the multi-wavelength nature of our system provides inherent advantages for noise reduction. By capturing spectral data across multiple wavelengths simultaneously, our system achieves a form of measurement redundancy. This multi-modal approach potentially increases signal fidelity and robustness to environmental noise, as artifacts typically don't affect all wavelengths equally. However, we note that the specific contributions of multi-wavelength measurements to noise reduction, particularly in different lighting conditions and movement scenarios, warrants further investigation. Future work should systematically evaluate these effects through controlled studies comparing single versus multi-wavelength performance under various environmental conditions.

5.4 Pulse Oximeter Bias Correction - Considerations and Feasibility

While the OpenOximetry database is a valuable resource for evaluating pulse oximeter performance across different skin tones, it has certain limitations. One key constraint is that PPG waveform data is not available for all encounters, and even when it is present, it is not time-synchronized with arterial blood gas measurements. This lack of synchronization reduces the ability to directly analyze pulse oximetry signal dynamics in relation to skin tone. Additionally, the dataset is skewed towards lighter-skinned individuals in fingertip measurements, limiting its representativeness for darker skin tones. We observed a similar trend in our study, despite targeted recruitment of individuals with darker skin tones. Many fingertip measurements still fell within the intermediate to tan range, making it challenging to obtain sufficient data from participants with dark-skinned fingertips. This recruitment difficulty should be considered in future studies to ensure a more balanced dataset that better represents the full spectrum of skin tones.

Due to limitations in acquiring the blood oxygen algorithms for pulse oximeters as well as raw PPG waveforms in the OpenOximetry database, a regression model was built on top of the existing saturation values, a methodology we believe to be suboptimal. Ideally, the skin-tone correction would occur during the SpO_2 calculation at the PPG signal level, where a baseline measurement would be taken to assess skin-tone at the site of data collection and used directly while calculating the current saturation level. This would be analogous to current PPG LED controllers, where the LED intensity is modulated based on the surrounding light measured by the device's PDs. This method is typically used to increase light intensity when the skin-tone is dark, which can be aided by also taking the skin-tone measurements. However, recent studies have suggested that modulating LED intensity can distort pulse oximetry measurements, especially in hypoxic conditions [88]. As such, a multi-spectrum approach, such as the one used in DermaGlow, may be preferred, especially considering the reduction in observed MAE through the use of our model. To integrate this correction approach into existing pulse oximeter devices, manufacturers could extend the spectral range beyond red and infrared to include wavelengths relevant to melanin absorption and utilize a prediction algorithm for skin-tone constituents, specifically CIE LAB values.

5.5 AI and Remote Patient Monitoring

DermaGlow offers significant advantages for remote patient monitoring. Specifically, the system can facilitate the recruitment of a broader participant base, ensuring a more representative sample of a subject population. This can also enable remote recruitment of patients, while also obtaining information about participant demographics. Furthermore, DermaGlow's skin color measurement can significantly enhance representation and fairness when developing predictive models. This is particularly important in the field of medical AI, where biased datasets can

lead to skewed algorithms that perform poorly for individuals outside the majority demographic. Furthermore, enhanced dataset diversity not only improves AI system performance, but also fosters inclusivity and equity in healthcare outcomes. As the AI space continues to grow, particularly in healthcare, robustness across all demographics is essential.

5.6 DermaGlow in Wearables

As a standalone wearable device, DermaGlow can be tailored to skin analysis by incorporating sensors with LED and PD wavelengths that provide a greater spectral resolution in the 500-900nm range. This may aid in improving the accuracy of our algorithm, as melanin and erythema have high absorption bands in this area, thus allowing us to effectively differentiate between the two more effectively. This may also help in analyzing micro-trends in skin, which could be crucial for certain dermatological conditions. Despite this limitation, our algorithm has shown great results using a commercial solution, which can be improved upon with this modification.

DermaGlow's sensors and algorithm can be readily adapted to other wearable wrist-worn devices that utilize optical spectroscopy, such as the Apple Watch [3] or Google Pixel Watch [34]. Both devices have recently included multi-spectrum sensors in these devices to improve optical biomarker detection. DermaGlow can be utilized in a similar fashion by expanding the sensor resolution to include wavelengths pertinent to skin composition measurement, along with the DermaGlow model. This would aid in detection of optically-obtained biomarkers across skin-tones, while also enabling skin monitoring continuously on the wrist. We recognize that UV monitoring on the wrist may be counter-intuitive, especially considering the site of analysis would be covered by the watch, and thus potential UV overexposure may not be be identified. However, it may be helpful to analyze melanin and erythema levels in this case, as there may be compositional changes in the skin due to this exposure that are not necessarily evident visually. Melanocytes - cells that produce melanin - have been shown to release signaling molecules that influence each others functions. Upon UV exposure, melanocytes can modulate the behavior of surrounding cells, contributing to an integrated defense mechanism against UV-damage [11]. DermaGlow can allow further exploration into the effects of surrounding UV radiation in this context, which has remained largely unexplored.

5.7 Dermatological Use Cases

DermaGlow has several potential dermatological use cases that we have identified. The first of these is a wearable patch that can be prescribed by a physician for either short or long-term monitoring of a potential skin disorder. For example, during a telehealth visit, a dermatologist might examine a patient's skin; if further monitoring is required, the physician could prescribe DermaGlow to examine the affected area and monitor it remotely, thereby preventing a potentially expensive visit and/or invasive biopsy, reducing costs to both the patient and the healthcare system. These patches could also be used to monitor allergic reactions or infections at incision sites, offering significant applications in the clinic, such as for peri-operative monitoring, or during ENT appointments where potential allergies are being determined. Typically, allergies are determined via skin or blood tests, where the skin is exposed to a suspected allergen. Allergies are then determined by examining the width and overall extent of resulting erythema, which is analyzed and recorded by the attending physician. This is a subjective measure, and it may be more accurately quantified by DermaGlow for this application. Furthermore, since DermaGlow can measure the skin continuously, the progression of erythema could be tracked over time. This application could also be extended towards monitoring of cosmetic laser treatments, where colorimeters are often used to determine the type of treatment needed. The continuous nature of DermaGlow data collection can provide crucial insights to enhance these therapies to maximize therapeutic benefit and patient comfort.

5.8 Cosmetic Applications

DermaGlow has potential cosmetic applications for clinical monitoring and substantiating the efficacy of cosmetic products. Cosmetic companies often perform small to large-scale clinical trials where products are tested on a subject population to examine skin traits such as overall brightening for products like serums and creams. However, a limiting factor in these analyses is many of these trials are conducted on-site using colorimeters, which limits the ability to analyze skin reactions long-term. Often, this is solved by monitoring patients in-clinic over time, which increases subject burden and overall cost. DermaGlow could be used in a small wearable patch form-factor for this application, enabling analysis of the skin in unapplied and applied product areas to enhance research on product efficacy and overall claims substantiation. Since the data collected is continuous, scientific comparison would be more robust, and these comparisons can even be conducted remotely. This capability would greatly simplify data collection efforts, provide potentially enhanced data on skin reaction to various products, and reduce the need for prolonged in-clinic stays.

6 Conclusion

Current methods that quantify melanin, erythema, and/or skin-tone often suffer from being highly subjective, expensive, non-continuous or impractical for daily use. The potential for a non-invasive, wearable, and continuous quantitative skin measurement device has implications for skin monitoring, as well as correcting for racial biases in pulse oximeter measurements. In this paper, we introduced DermaGlow, a novel wearable framework that utilizes optical spectroscopy to quantify melanin, erythema, and skin-tone. Our results demonstrate that DermaGlow can accurately quantify melanin, erythema, and CIE L*, a*, and b* values, achieving a normalized mean absolute error (NMAE) of 5.33% for melanin and 4.18% for erythema, with ΔE values below 2.5 when compared to a state-of-the-art commercial colorimeter. Furthermore, we developed an algorithm leveraging DermaGlow's outputs to correct pulse oximeter inaccuracies associated with skin tone, leading to up to a 75% reduction in MAE for hypoxic readings. These findings underscore DermaGlow's potential as both a continuous skin monitoring solution and a corrective system for improving the accuracy of wearable medical devices that rely on skin perfusion measurements, particularly in addressing disparities in pulse oximetry across diverse skin tones.

Acknowledgments

We gratefully acknowledge the support and contributions of our collaborators and advisors throughout the development of this work. Special thanks to the participants and staff involved in the DermaGlow study for their time and cooperation. This research was partially supported by the University of Virginia's Wallace H. Coulter Center for Translational Research.

References

[1] R. Abdlaty, J. Hayward, T. Farrell, and Q. Fang. 2021. Skin erythema and pigmentation: a review of optical assessment techniques. *Photodiagnosis Photodyn Ther* 33 (2021), 102127. doi:10.1016/j.pdpdt.2020.102127 Epub 2020 Dec 1.

[2] Amtast. [n. d.]. AMT 510 Colorimeter. Colorimeter used for scientific and industrial color measurement. https://amtast.com/content-93.html

[3] Apple. 2024. Apple Watch. https://www.apple.com/watch/

[4] M Baquié and B Kasraee. 2014. Discrimination between cutaneous pigmentation and erythema: comparison of the skin colorimeters Dermacatch and Mexameter. *Skin Res. Technol.* 20, 2 (May 2014), 218–227.

[5] B Bent, BA Goldstein, WA Kibbe, and JP Dunn. 2020. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digital Medicine* 3 (Feb 2020), 18. doi:10.1038/s41746-020-0226-6

[6] DZ Bliss, O Gurvich, K Savik, et al. 2017. Racial and ethnic disparities in the healing of pressure ulcers present at nursing home admission. *Archives of Gerontology and Geriatrics* 72 (2017), 187–194.

[7] Michaela Brenner and Vincent J Hearing. 2008. The protective role of melanin against UV damage in human skin. *Photochemistry and Photobiology* 84, 3 (May-Jun 2008), 539–549. doi:10.1111/j.1751-1097.2007.00226.x

[8] AM Cabanas, M Fuentes-Guajardo, K Latorre, D León, and P Martín-Escudero. 2022. Skin Pigmentation Influence on Pulse Oximetry Accuracy: A Systematic Review and Bibliometric Analysis. *Sensors* 22, 9 (Apr 2022), 3402. doi:10.3390/s22093402

[9] D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran. 2018. A review on wearable photoplethysmography sensors and their potential future applications in health care. *International Journal of Biosensors & Bioelectronics* 4, 4 (2018), 195–202. doi:10.15406/ijbsbe.2018.04.00125 Epub 2018 Aug 6.

[10] ED Chan, MM Chan, and MM Chan. 2013. Pulse oximetry: understanding its basic principles facilitates appreciation of its limitations. *Respiratory Medicine* 107, 6 (Jun 2013), 789–799. doi:10.1016/j.rmed.2013.02.004 Epub 2013 Mar 13.

[11] Mirosława Cichorek, Małgorzata Wachulska, Agnieszka Stasiewicz, and Agnieszka Tyminska. 2013. Skin melanocytes: biology and development. *Postepy Dermatologii i Alergologii* 30, 1 (2013), 30–41. doi:10.5114/pdia.2013.33376

[12] CKElectronic. [n. d.]. Mexameter MX18. Device used for measuring skin melanin and erythema. https://www.courage-khazaka.de/en/scientific-products/occupational-health/occupational-health/169-mexameter-e

[13] PJ Colvonen. 2021. Response To: Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digital Medicine* 4, 1 (Feb 2021), 38. doi:10.1038/s41746-021-00408-5

[14] PJ Colvonen, PN DeYoung, NA Bosompra, and RL Owens. 2020. Limiting racial disparities and bias for wearable devices in health science research. *Sleep* 43, 10 (Oct 2020), zsaa159. doi:10.1093/sleep/zsaa159

[15] Cortex. 2024. Cortex DSM 4. Digital spectrophotometer for dermatological assessments. https://www.cortex.dk/skin-analysis/colorimeter-dsm-4/

[16] S Del Bino and F Bernerd. 2013. Variations in skin colour and the biological consequences of ultraviolet radiation exposure. *The British Journal of Dermatology* 169, Suppl 3 (2013), 33–40. doi:10.1111/bjd.12529

[17] DermLite 2024. DermLite DL5. DermLite. https://dermlite.com/products/dermlite-dl5

[18] L. Dhoonmoon and J. Fletcher. 2022. Assessing skin tones in practice: results of an international survey. *Wounds International* 13, 2 (2022), 6–9.

[19] CK Electronic. 2024. COLORIMETER® CL 400. Precision colorimeter for color consistency and quality control. https://www.enviroderm.co.uk/products/colorimeter-cl-400

[20] A. Elyas and M. Bradley. 2021. Viktigt att beakta att hudåkommor kan variera beroende på hudtyp [An introduction to skin disease in melanin rich skin]. *Lakartidningen* 118 (2021), 20208. Swedish.

[21] J S C English, R S Dawe, and J Ferguson. 2003. Environmental effects and skin disease. *British Medical Bulletin* 68 (2003), 129–142. doi:10.1093/bmb/ldg026

[22] European Pressure Ulcer Advisory Panel and National Pressure Ulcer Advisory Panel and Pan-Pacific Pressure Injury Alliance Year of publication. *Prevention and Treatment of Pressure Ulcers/Injuries: Quick Reference Guide.* European Pressure Ulcer Advisory Panel and National Pressure Ulcer Advisory Panel and Pan-Pacific Pressure Injury Alliance. URLifavailable Accessed: insert access date here.

[23] Thomas Fahrni, Michael Kuhn, Philipp Sommer, Roger Wattenhofer, and Samuel Welten. 2011. Sundroid: solar radiation awareness with smartphones. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (Beijing, China) *(UbiComp '11)*. Association for Computing Machinery, New York, NY, USA, 365–374. doi:10.1145/2030112.2030162

[24] D Fajuyigbe, A Coleman, RPE Sarkany, AR Young, and AW Schmalwieser. 2018. Diffuse Reflectance Spectroscopy as a Reliable Means of Comparing Ultraviolet Radiation-induced Erythema in Extreme Skin Colors. *Photochemistry and Photobiology* 94, 5 (Sep 2018), 1066–1070. doi:10.1111/php.12947 Epub 2018 Jul 4.

[25] A Fawzy, TD Wu, K Wang, ML Robinson, J Farha, A Bradke, SH Golden, Y Xu, and BT Garibaldi. 2022. Racial and Ethnic Discrepancy in Pulse Oximetry and Delayed Identification of Treatment Eligibility Among Patients With COVID-19. *JAMA Internal Medicine* 182, 7 (2022), 730–738. doi:10.1001/jamainternmed.2022.1906 Erratum in: JAMA Intern Med. 2022 Oct 1;182(10):1108.

[26] A Fawzy, TD Wu, K Wang, KE Sands, AM Fisher, SA Arnold Egloff, JD DellaVolpe, TJ Iwashyna, Y Xu, and BT Garibaldi. 2023. Clinical Outcomes Associated With Overestimation of Oxygen Saturation by Pulse Oximetry in Patients Hospitalized With COVID-19. *JAMA Network Open* 6, 8 (2023), e2330856. doi:10.1001/jamanetworkopen.2023.30856

[27] J.R. Feiner, J.W. Severinghaus, and P.E. Bickler. 2007. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. *Anesth Analg* 105, 6 Suppl (2007), S18–S23. doi:10.1213/01.ane.0000285988.35174.d9

[28] T. B. Fitzpatrick. 1975. Soleil et peau. *Journal of Medicine Esthetique and Dermatological Surgery* 2 (1975), 33–34. Sun and skin.

[29] A D et al. Fix. 2000. Racial differences in reported Lyme disease incidence. *American journal of epidemiology* 152, 8 (2000), 756–9. doi:10.1093/aje/152.8.756

[30] Nicholas Fong et al. 2024. OpenOximetry Repository. PhysioNet. doi:10.13026/cc78-ad74

[31] Food and Drug Administration. 2025. *Pulse Oximeters for Medical Purposes - Non-Clinical and Clinical Performance Testing, Labeling, and Premarket Submission Recommendations: Draft Guidance for Industry and Food and Drug Administration Staff.* Draft Guidance. U.S. Department of Health and Human Services. Document issued on January 7, 2025.

[32] KF Francis. 2023. Assessment and Identification of skin disorders in skin of color: an integrative review. *Journal of Wound Ostomy and Continence Nursing* 50, 2 (2023), 107–114.

[33] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220.

[34] Google. 2024. Google Pixel Watch. https://store.google.com/us/product/pixel_watch_2?hl=en-US

[35] ER Gottlieb, J Ziegler, K Morley, B Rush, and LA Celi. 2022. Assessment of Racial and Ethnic Differences in Oxygen Supplementation Among Patients in the Intensive Care Unit. *JAMA Internal Medicine* 182, 8 (Aug 2022), 849–858. doi:10.1001/jamainternmed.2022.2587

[36] M.K. Gudelunas, M. Lipnick, C. Hendrickson, S. Vanderburg, B. Okunlola, I. Auchus, J.R. Feiner, and P.E. Bickler. 2024. Low Perfusion and Missed Diagnosis of Hypoxemia by Pulse Oximetry in Darkly Pigmented Skin: A Prospective Study. *Anesth Analg* 138, 3 (2024), 552–561. doi:10.1213/ANE.0000000000006755 Epub 2023 Dec 18.

[37] Tarek Hamid, Insup Lee, and Amanda Watson. 2023. SpectraVue - An Interactive Web Application Enabling Rapid Data Visualization and Analysis for Wearable Spectroscopy Research *(UbiComp/ISWC '23 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 146–150. doi:10.1145/3594739.3610709

[38] NR Henry, AC Hanson, PJ Schulte, NS Warner, MN Manento, TJ Weister, and MA Warner. 2022. Disparities in Hypoxemia Detection by Pulse Oximetry Across Self-Identified Racial Groups and Associations With Clinical Outcomes. *Critical Care Medicine* 50, 2 (Feb 2022), 204–211. doi:10.1097/CCM.0000000000005394

[39] LC Hsu, S Hsu, TH Tan, CH Cheng, and CC Chang. 2022. Developing Low-Cost Mobile Device and Apps for Accurate Skin Spectrum Measurement via Low-Cost Spectrum Sensors and Deep Neural Network Technology. *Sensors* 22, 22 (Nov 2022), 8844. doi:10.3390/s22228844

[40] H Jamali, LT Castillo, CC Morgan, J Coult, JL Muhammad, OO Osobamiro, EC Parsons, and R Adamson. 2022. Racial Disparity in Oxygen Saturation Measurements by Pulse Oximetry: Evidence and Implications. *Annals of the American Thoracic Society* 19, 12 (Dec 2022), 1951–1964. doi:10.1513/AnnalsATS.202203-270CME

[41] VG Kanellis. 2019. A review of melanin sensor devices. *Biophysical Reviews* 11, 6 (Dec 2019), 843–849. doi:10.1007/s12551-019-00581-8

[42] SW Ketcham, MC Konerman, VD Marshall, and SK Adie. 2024. Racial Bias in Pulse Oximetry Measurement: Considerations in Patients With Heart Failure. *Circulation: Cardiovascular Quality and Outcomes* 17, 3 (Mar 2024), e010390. doi:10.1161/CIRCOUTCOMES.123.010390

[43] A.K. Khanna, J. Beard, S. Lamminmäki, J. Närväinen, N. Antaki, and H.O. Yapici. 2024. Assessment of skin pigmentation-related bias in pulse oximetry readings among adults. *J Clin Monit Comput* 38, 1 (2024), 113–120. doi:10.1007/s10877-023-01095-1 Epub 2023 Oct 26.

[44] D Koerber, S Khan, T Shamsheri, A Kirubarajan, and S Mehta. 2023. Accuracy of Heart Rate Measurement with Wrist-Worn Wearable Devices in Various Skin Tones: a Systematic Review. *Journal of Racial and Ethnic Health Disparities* 10, 6 (Dec 2023), 2676–2684. doi:10.1007/s40615-022-01446-9 Epub 2022 Nov 14.

[45] N. Kollias and A. Baqer. 1985. Spectroscopic characteristics of human melanin in vivo. *Journal of Investigative Dermatology* 85, 1 (1985), 38–42. doi:10.1111/1523-1747.ep12275011

[46] N Kollias and AH Baqer. 1987. Absorption mechanisms of human melanin in the visible, 400-720 nm. *Journal of Investigative Dermatology* 89, 4 (Oct 1987), 384–388. doi:10.1111/1523-1747.ep12471764

[47] La Roche-Posay 2024. My UV Patch. La Roche-Posay by L'Oréal. https://www.lorealtechincubator.com/myuvpatch A stretchable skin sensor that monitors UV radiation exposure and works with a mobile app..

[48] M Langeveld, LS van de Lande, E O'Sullivan, B van der Lei, and JA van Dongen. 2022. Skin measurement devices to assess skin quality: A systematic review on reliability and validity. *Skin Research and Technology* 28, 2 (Mar 2022), 212–224. doi:10.1111/srt.13113

[49] LogicInk 2024. LogicInk. LogicInk. https://logicink.com/ Creates wearable sensors that provide visual indications of health and environmental exposures, including UV exposure..

[50] S Maya-Enero, J Candel-Pau, J Garcia-Garcia, AM Gimenez-Arnau, and MA Lopez-Vilchez. 2020. Validation of a neonatal skin color scale. *European Journal of Pediatrics* 179, 9 (2020), 1403–1411.

[51] HE McCreath, BM Bates-Jensen, G Nakagami, et al. 2016. Use of Munsell Color Charts to measure skin tone objectively in nursing home residents at risk for pressure ulcer development. *Journal of Advanced Nursing* 72, 9 (2016), 2077–2085.

[52] Konica Minolta. [n. d.]. Konica Minolta CM-700d Spectrophotometer. Spectrophotometer for accurate color measurement in a wide range of industries. https://sensing.konicaminolta.us/us/products/cm-700d-spectrophotometer/

[53] MoleScope 2024. MoleScope. MetaOptima Technology Inc. https://www.molescope.com/

[54] Ellis Monk. 2023. The Monk Skin Tone Scale. https://osf.io/preprints/socarxiv/ Web.

[55] Sarah Mosca and Aldo Morrone. 2023. Human Skin Pigmentation: From a Biological Feature to a Social Determinant. *Healthcare* 11, 14 (2023). doi:10.3390/healthcare11142091

[56] M Nitzan, A Romem, and R Koppel. 2014. Pulse oximetry: fundamentals and technology update. *Medical Devices (Auckland, NZ)* 7 (8 Jul 2014), 231–239. doi:10.2147/MDER.S47319

[57] OE Okunlola, MS Lipnick, PB Batchelder, M Bernstein, JR Feiner, and PE Bickler. 2022. Pulse Oximeter Performance, Racial Inequity, and the Work Ahead. *Respiratory Care* 67, 2 (Feb 2022), 252–257. doi:10.4187/respcare.09795

[58] R.L. Olson, J. Gaylor, and M.A. Everett. 1973. Skin color, melanin, and erythema. *Arch Dermatol* 108, 4 (1973), 541–544.

[59] J.P. Ortonne. 2002. Photoprotective properties of skin melanin. *Br. J. Dermatol.* 146 (2002), 7–10. doi:10.1046/j.1365-2133.146.s61.3.x

[60] Pantone 2024. Pantone SkinTone Guide. Pantone. https://www.pantone.com/products/fashion-home-interiors/pantone-skintone-guide
[61] EM Polfer, RM Zimmerman, E Tefera, RD Katz, JP Higgins, and Jr. Means, KR. 2018. The effect of skin pigmentation on determination of limb ischemia. *Journal of Hand Surgery (American Volume)* 43, 1 (2018), 24–32.
[62] M.L. Ramsey and S. Rostami. 2023. Skin Biopsy. *StatPearls [Internet]* (2023). https://www.ncbi.nlm.nih.gov/books/NBK470464/ Treasure Island (FL): StatPearls Publishing; 2024 Jan–.
[63] PA Riley. 1997. Melanin. *International Journal of Biochemistry & Cell Biology* 29, 11 (Nov 1997), 1235–1239. doi:10.1016/s1357-2725(97)00013-7
[64] I Sadiq, N Kollias, and A Baqer. 2019. Spectroscopic observations on human pigmentation. *Photodermatology, Photoimmunology & Photomedicine* 35, 6 (Nov 2019), 415–419. doi:10.1111/phpp.12474 Epub 2019 May 24.
[65] LK Saladin and JS Krause. 2009. Pressure ulcer prevalence and barriers to treatment after spinal cord injury: comparisons of four groups based on race-ethnicity. *Neurorehabilitation* 24, 1 (2009), 57–66.
[66] CJ Santiago, VW Weedn, and FJ Diaz. 2022. The impact of skin color on the recognition of blunt force injuries. *American Journal of Forensic Medicine and Pathology* 43, 3 (2022), 220–224.
[67] KN Scafide, MC Narayan, and L Arundel. 2020. Bedside technologies to enhance the early detection of pressure injuries: a systematic review. *Journal of Wound Ostomy and Continence Nursing* 47, 2 (2020), 128–136.
[68] D.I. Schlessinger, M.D. Anoruo, and J. Schlessinger. 2023. Biochemistry, Melanin. *StatPearls [Internet]* (2023). Updated 2023 May 1; Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459156.
[69] K Setchfield, A Gorman, AHRW Simpson, MG Somekh, and AJ Wright. 2024. Effect of skin color on optical properties and the implications for medical optical technologies: a review. *Journal of Biomedical Optics* 29, 1 (Jan 2024), 010901. doi:10.1117/1.JBO.29.1.010901 Epub 2024 Jan 24.
[70] M Sharma, AW Brown, NM Powell, N Rajaram, L Tong, PM Mourani, and M Schootman. 2024. Racial and skin color mediated disparities in pulse oximetry in infants and young children. *Paediatric Respiratory Reviews* (5 Jan 2024). doi:10.1016/j.prrv.2023.12.006 Epub ahead of print.
[71] A. Shcherbina, C.M. Mattsson, D. Waggott, H. Salisbury, J.W. Christle, T. Hastie, M.T. Wheeler, and E.A. Ashley. 2017. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *J Pers Med* 7, 2 (2017), 3. doi:10.3390/jpm7020003
[72] Seiichiro Shono, Genji Imokawa, Yoshihiro Yada, and Mitsuru Kimura. 1985. The relationship of skin color, UVB-induced erythema, and melanogenesis. *The Journal of Investigative Dermatology* 84, 4 (1985), 265–267. doi:10.1111/1523-1747.ep12265342
[73] MW Sjoding, RP Dickson, TJ Iwashyna, SE Gay, and TS Valley. 2020. Racial Bias in Pulse Oximetry Measurement. *New England Journal of Medicine* 383, 25 (Dec 2020), 2477–2478. doi:10.1056/NEJMc2029240 Erratum in: N Engl J Med. 2021 Dec 23;385(26):2496.
[74] SkinVision 2024. SkinVision. SkinVision. https://www.skinvision.com/
[75] MS Sommers, Y Regueira, DA Tiller, et al. 2019. Understanding rates of genital-anal injury: role of skin color and skin biomechanics. *Journal of Forensic and Legal Medicine* 66 (2019), 120–128.
[76] SE Sonenblum, R Patel, S Phrasavath, S Xu, and BM Bates-Jensen. 2023. Using Technology to Detect Erythema Across Skin Tones. *Advances in Skin & Wound Care* 36, 10 (Oct 2023), 524–533. doi:10.1097/ASW.0000000000000043
[77] GN Stamatas, BZ Zmudzka, N Kollias, and JZ Beer. 2008. In vivo measurement of skin erythema and pigmentation: new means of implementation of diffuse reflectance spectroscopy with a commercial instrument. *British Journal of Dermatology* 159, 3 (Sep 2008), 683–690. doi:10.1111/j.1365-2133.2008.08642.x Epub 2008 May 28.
[78] SEK Sudat, P Wesson, KF Rhoads, et al. 2023. Racial disparities in pulse oximeter device inaccuracy and estimated clinical impact on COVID-19 treatment course. *American Journal of Epidemiology* 192, 5 (2023), 703–713.
[79] S.E.K. Sudat, P. Wesson, K.F. Rhoads, S. Brown, N. Aboelata, A.R. Pressman, A. Mani, and K.M.J. Azar. 2023. Racial Disparities in Pulse Oximeter Device Inaccuracy and Estimated Clinical Impact on COVID-19 Treatment Course. *Am J Epidemiol* 192, 5 (2023), 703–713. doi:10.1093/aje/kwac164
[80] Sun Index 2024. Sun Index Wearable Watch. Sun Index. https://sunindex.co/ A wearable device that provides real-time UV index monitoring to help manage sun exposure..
[81] Jerry Tan, Hong Liu, James J. Leyden, and Matthew J. Leoni. 2014. Reliability of Clinician Erythema Assessment grading scale. *Journal of the American Academy of Dermatology* 71, 4 (2014), 760–763. doi:10.1016/j.jaad.2014.05.033
[82] Delfin Technologies. 2024. Delfin SkinColorCatch. Portable device for skin color measurement. https://delfintech.com/products/skincolorcatch/
[83] K Tekin, M Karadogan, S Gunaydin, and K Kismet. 2023. Everything About Pulse Oximetry-Part 1: History, Principles, Advantages, Limitations, Inaccuracies, Cost Analysis, the Level of Knowledge About Pulse Oximeter Among Clinicians, and Pulse Oximetry Versus Tissue Oximetry. *Journal of Intensive Care Medicine* 38, 9 (Sep 2023), 775–784. doi:10.1177/08850666231185752 Epub 2023 Jul 12.
[84] VSM Valbuena, RP Barbaro, D Claar, TS Valley, RP Dickson, SE Gay, MW Sjoding, and TJ Iwashyna. 2022. Racial Bias in Pulse Oximetry Measurement Among Patients About to Undergo Extracorporeal Membrane Oxygenation in 2019-2020: A Retrospective Cohort Study. *Chest* 161, 4 (Apr 2022), 971–978. doi:10.1016/j.chest.2021.09.025

[85] Martijn van der Wal, Monica Bloemen, Pauline Verhaegen, Wim Tuinebreijer, Henrica de Vet, Paul van Zuijlen, and Esther Middelkoop. 2013. Objective color measurements: clinimetric performance of three devices on normal skin and scar tissue. *J. Burn Care Res.* 34, 3 (May 2013), e187–94.
[86] Sandhya Vasudevan, William C. Vogt, Sandy Weininger, and T. Joshua Pfefer. 2024. Melanometry for objective evaluation of skin pigmentation in pulse oximetry studies. *Communications Medicine* 4, 1 (2024), 138. doi:10.1038/s43856-024-00550-7
[87] F. von Luschan. 1897. *Beiträge zur Völkerkunde der Deutschen Schutzgebieten.* Deutsche Buchgemeinschaft, Berlin.
[88] M. Wald, P. Erwin, N. Annon-Eberharter, and T. Werther. 2023. LED light can falsify pulse oximetry readings via the stroboscopic effect. *Phys Eng Sci Med* 46, 4 (Dec 2023), 1667–1675. doi:10.1007/s13246-023-01328-2 Epub 2023 Sep 19.
[89] Amanda Watson, Claire Kendell, Anush Lingamoorthy, Insup Lee, and James Weimer. 2023. Lumos: An Open-Source Device for Wearable Spectroscopy Research. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 187 (jan 2023), 24 pages. doi:10.1145/3569502
[90] S. Westland. 2016. The CIE System. In *Handbook of Visual Display Technology*, J. Chen, W. Cranton, and M. Fihn (Eds.). Springer, Cham. doi:10.1007/978-3-319-14346-0_11
[91] Ross D Whitehead, Daniel Re, Dengke Xiao, Gozde Ozakinci, and David I Perrett. 2012. You are what you eat: within-subject increases in fruit and vegetable consumption confer beneficial skin-color changes. *PLoS One* 7, 3 (2012), e32988. doi:10.1371/journal.pone.0032988
[92] Antony R Young. 1997. Chromophores in human skin. *Physics in Medicine & Biology* 42, 5 (1997), 789. doi:10.1088/0031-9155/42/5/004
[93] George Zonios, Julie Bykowski, and Nikiforos Kollias. 2001. Skin Melanin, Hemoglobin, and Light Scattering Properties can be Quantitatively Assessed In Vivo Using Diffuse Reflectance Spectroscopy. *Journal of Investigative Dermatology* 117, 6 (2001), 1452–1457.
[94] George Zonios, Aikaterini Dimou, Ioannis Bassukas, Dimitrios Galaris, Argyrios Tsolakidis, and Efthimios Kaxiras. 2008. Melanin absorption spectroscopy: new method for noninvasive skin investigation and melanoma detection. *Journal of Biomedical Optics* 13, 1 (Jan 2008), 014017. doi:10.1117/1.2844710

Appendix A: Feature Importance Scores

The following tables summarize the top 10 features selected for each output and their relative importance scores from the final CatBoost regression model. Feature names are abbreviated for clarity (e.g., LED680-PD415nm_arctan refers to the arctan of the 415nm PD response when the 680nm LED is illuminated.).

Table 6. Top 10 Features for Melanin Prediction (CatBoost Model)

Feature	**Importance Score**
LED680-PD415nm_arctan/LED680-PD415nm_log	17.26
LED680-PD415nm_arctan/LED680-PD590nm_log	16.02
LED680-PD515nm_arctan/LED680-PD680nm_log	12.66
LED680-PD590nm_arctan/LED680-PD590nm_log	9.39
LED680-PD590nm_arctan/LED680-PD680nm_log	9.16
LED680-PD630nm_arctan/LED680-PD630nm_log	8.14
LED630-PD445nm_log/LED820-PD680nm_log	8.06
LED630-PD480nm_log/LED820-PD680nm_log	6.84
LED630-PD555nm_arctan/LED555-PD680nm_log	6.27
LED555-PD630nm_arctan/LED555-PD680nm_log	6.18

Table 7. Top 10 Features for Erythema Prediction (CatBoost Model)

Feature	Importance Score
LED630-PD415nm_DC/LED630-PD680nm_DC	26.19
LED630-PD415nm_DC/LED630-PD680nm_mean	11.33
LED630-PD680nm_DC/LED630-PD415nm_mean	11.15
LED590-PD445nm_DC/LED590-PD680nm_DC	8.77
LED590-PD445nm_DC/LED590-PD680nm_mean	8.27
LED555-PD480nm_DC/LED555-PD590nm_DC	8.00
LED555-PD480nm_DC/LED555-PD590nm_mean	7.43
LED555-PD590nm_DC/LED555-PD480nm_mean	7.08
LED590-PD445nm_mean/LED590-PD680nm_mean	5.94
LED555-PD590nm_log/LED820-PD680nm_log	5.84

Table 8. Top 25 Features for CIE L* Prediction (CatBoost Model)

Feature	Importance Score
LED630-PD590nm_DC/LED820-PD630nm_DC	7.07
LED630-PD590nm_DC/LED820-PD680nm_DC	6.72
LED630-PD590nm_DC/LED820-PD630nm_mean	5.00
LED630-PD590nm_DC/LED820-PD680nm_mean	4.55
LED590-PD590nm_DC/LED820-PD680nm_DC	4.48
LED590-PD590nm_DC/LED820-PD680nm_mean	4.09
LED555-PD445nm_DC/LED820-PD680nm_DC	3.35
LED555-PD445nm_DC/LED820-PD680nm_mean	3.11
LED555-PD630nm_DC/LED820-PD680nm_DC	3.06
LED555-PD630nm_DC/LED820-PD680nm_mean	3.01
LED555-PD680nm_DC/LED555-PD630nm_mean	2.92
LED820-PD680nm_DC/LED630-PD680nm_mean	2.91
LED630-PD590nm_mean/LED820-PD630nm_mean	2.79
LED630-PD590nm_mean/LED820-PD680nm_mean	2.38
LED590-PD590nm_mean/LED820-PD680nm_mean	2.17
LED555-PD445nm_mean/LED820-PD680nm_mean	1.99
LED555-PD630nm_mean/LED820-PD680nm_mean	1.94
LED630-PD415nm_log/LED820-PD630nm_log	1.93
LED630-PD415nm_log/LED820-PD680nm_log	1.90
LED630-PD445nm_log/LED820-PD630nm_log	1.78
LED630-PD445nm_log/LED820-PD680nm_log	1.74
LED630-PD480nm_log/LED820-PD630nm_log	1.68
LED630-PD480nm_log/LED820-PD680nm_log	1.64
LED630-PD515nm_log/LED820-PD630nm_log	1.60
LED630-PD515nm_log/LED820-PD680nm_log	1.60

Table 9. Top 25 Features for CIE a* Prediction (CatBoost Model)

Feature	Importance Score
LED590-PD515nm_DC/LED590-PD680nm_DC	11.95
LED590-PD515nm_DC/LED590-PD680nm_mean	5.28
LED590-PD680nm_DC/LED555-PD590nm_DC	4.77
LED590-PD680nm_DC/LED590-PD630nm_mean	4.11
LED590-PD680nm_DC/LED555-PD590nm_mean	3.36
LED555-PD445nm_DC/LED415-PD680nm_DC	3.30
LED555-PD515nm_DC/LED590-PD680nm_mean	3.21
LED555-PD555nm_DC/LED415-PD630nm_DC	3.12
LED555-PD590nm_DC/LED415-PD630nm_DC	3.06
LED555-PD590nm_DC/LED415-PD680nm_DC	2.88
LED555-PD590nm_DC/LED415-PD630nm_mean	2.80
LED555-PD590nm_DC/LED415-PD680nm_mean	2.68
LED415-PD630nm_DC/LED555-PD590nm_mean	2.42
LED415-PD680nm_DC/LED555-PD445nm_mean	2.20
LED415-PD680nm_DC/LED555-PD590nm_mean	2.03
LED820-PD415nm_DC/LED820-PD630nm_mean	2.03
LED820-PD415nm_DC/LED820-PD680nm_mean	1.97
LED820-PD445nm_DC/LED820-PD630nm_DC	1.81
LED820-PD445nm_DC/LED820-PD630nm_mean	1.77
LED820-PD480nm_DC/LED820-PD630nm_DC	1.72
LED820-PD480nm_DC/LED820-PD630nm_mean	1.69
LED820-PD515nm_DC/LED820-PD680nm_DC	1.48
LED820-PD515nm_DC/LED820-PD680nm_mean	1.38
LED820-PD590nm_DC/LED590-PD555nm_log	1.33
LED820-PD590nm_DC/LED555-PD515nm_log	1.32

Table 10. Top 25 Features for CIE b* Prediction (CatBoost Model)

Feature	Importance Score
LED470-PD480nm_mean/LED820-PD680nm_mean	19.28
LED470-PD480nm_mean/LED820-PD590nm_mean	12.35
LED470-PD480nm_mean/LED680-PD555nm_log	10.17
LED470-PD480nm_mean/LED630-PD680nm_log	7.59
LED470-PD480nm_DC/LED680-PD480nm_log	3.04
LED470-PD480nm_DC/LED445-PD680nm_log	3.03
LED470-PD480nm_DC/LED590-PD445nm_DC	2.71
LED470-PD480nm_DC/LED470-PD555nm_log	1.53
LED470-PD480nm_DC/LED590-PD515nm_log	1.47
LED470-PD480nm_DC/LED415-PD630nm_DC	1.44
LED470-PD480nm_DC/LED590-PD680nm_DC	1.43
LED470-PD480nm_DC/LED555-PD590nm_DC	1.43
LED470-PD480nm_DC/LED470-PD680nm_mean	1.41
LED470-PD480nm_DC/LED470-PD680nm_DC	1.41
LED470-PD480nm_DC/LED470-PD555nm_mean	1.39
LED470-PD480nm_DC/LED630-PD415nm_DC	1.26
LED470-PD480nm_DC/LED555-PD480nm_DC	1.21
LED470-PD480nm_mean/LED555-PD590nm_mean	1.20
LED470-PD480nm_DC/LED555-PD590nm_log	1.18
LED470-PD480nm_DC/LED590-PD515nm_mean	1.16
LED470-PD480nm_DC/LED555-PD555nm_mean	1.15
LED470-PD480nm_DC/LED555-PD480nm_mean	1.13
LED470-PD480nm_DC/LED590-PD515nm_DC	1.12
LED470-PD480nm_DC/LED555-PD590nm_DC	1.04
LED470-PD480nm_DC/LED470-PD555nm_DC	1.01