

Systematic Evidence Synthesis for Non-Invasive Hemoglobin Estimation Using Multispectral Photoplethysmography and Interpretable AI

Devarsh Hitesh Kapadiya
2nd Year Student Computer Science
Nottingham Trent University
Nottingham, United Kingdom

Hitesh Kapadiya
Lecturer in Electrical Engineering
Government Polytechnic, Ahmedabad
Gujarat, India

Abstract - Over 1.6 billion people worldwide suffer from anemia, with low-resource groups, pregnant women, and small children being the most affected. Invasive venipuncture and laboratory spectrophotometry are used in standard diagnosis; these procedures are uncomfortable, involve a risk of infection, and need infrastructure that restricts repeated testing, especially in rural and isolated areas. This paper examines current developments in continuous non-invasive hemoglobin monitoring, highlighting multi-wavelength photoplethysmography that spans visible bands from 660 to 730 nm and near-infrared bands from 850 to 940 nm, together with machine learning regression models. Building on the Beer-Lambert law, we compare traditional chemometric techniques with gradient-boosted trees and deep neural networks, which capture nonlinear correlations in multichannel signals. Additionally, we evaluate the best reported predictive performance, with coefficients of determination between 0.85 and 0.997 and mean absolute errors from 0.3 to 8.5 grams per liter across the surveyed studies, and subject-level screening that reaches a sensitivity of 93.3 percent and a specificity of 93.8 percent. We look at sensor possibilities such as CMOS image sensors, solar cells, and photodiode arrays, and discuss explainable artificial intelligence techniques such as SHAP that allow clinical validation and make model outputs interpretable. Lastly, we outline strategies for creating dependable, equitable wearable screening tools appropriate for routine, community-level, and quick point-of-care anemia detection. We also highlight the main technical and clinical obstacles to adoption, such as skin tone bias, motion artifacts, data leakage, and variability across populations.

Keywords - photoplethysmography; hemoglobin estimation; multispectral optical sensing; explainable artificial intelligence; light gradient boosting machine; SHapley Additive exPlanations; machine learning regression; anemia screening

I. INTRODUCTION AND BACKGROUND

A. Clinical Burden and Diagnostic Bottlenecks

The vital iron-rich metalloprotein in erythrocytes that carries oxygen throughout the body is called hemoglobin (Hb). Anemia, which results in systemic hypoxia, cardiovascular exhaustion, cognitive impairments, and increased mortality risks in pregnant and pediatric populations, is caused by pathological reductions in circulating hemoglobin levels. Anemia affects about 30% of people worldwide, according to WHO estimates ^[1], with an acute concentration in lower-income areas where chronic illnesses and nutritional deficiencies are prevalent.

For laboratory analysis (e.g., cyanmethemoglobin test or automated cell counters), standard diagnostic procedures depend on taking venous or capillary blood. Though accurate (CV < 3%), these methods require trained staff, sterile consumables, cold-chain logistics, and painful skin punctures. Proactive point-of-care screening and ongoing patient monitoring are hampered by these considerations. The estimated prevalence across demographic groups, synthesized from the WHO Global Anemia Report ^[1] and epidemiological surveys (2015–2023), is summarized in Table I.

TABLE I. ESTIMATED GLOBAL ANEMIA PREVALENCE BY DEMOGRAPHIC GROUP

Demographic Group	Estimated Prevalence
Pregnant Women	~40%
Preschool Children	~25%
Elderly Population (65+)	~30%
School-Age Children	~18%
Non-Pregnant Women	~13%
Adult Men	~12%

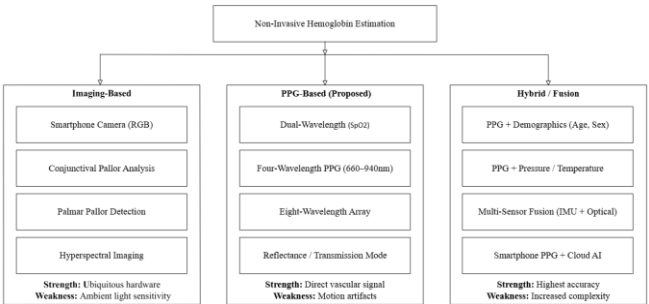


Fig. 1. System Taxonomy Architecture for Non-Invasive Hemoglobin Estimation

Existing non-invasive hemoglobin estimation techniques can be arranged into three main branches, shown as a hierarchical taxonomy in Fig. 1, which also acts as the conceptual road map of this review.

The taxonomy has one root node, “Non-Invasive Hemoglobin Estimation,” which identifies the general field of study, and three different branches. The left branch is imaging-based. Instead of using skin-contact sensors, these techniques make use of cameras and computer vision, such as smartphone RGB cameras and the examination of the color of the palm (palmar pallor) or the inner eyelid (conjunctival pallor) ^{[2]–[4]}, as well as hyperspectral imaging. Since almost everyone

already has a phone, their main advantage is accessibility [5]. The problem is that findings can fluctuate with room illumination and with differences in skin tone.

The center branch is based on photoplethysmography (PPG), which is the central idea of this paper. Beginning with basic two-wavelength pulse oximetry (SpO₂), the branch advances to four-wavelength PPG setups at 660, 730, 850, and 940 nm [6] before moving on to eight-wavelength arrays [7], [8]. It also distinguishes between transmission and reflectance modes. This approach is attractive because it directly scans the vascular signal and records it with a high temporal resolution.

The right branch is hybrid/fusion, the most ambitious of the three. PPG optical signals are paired with other types of data, including temperature, motion from IMU sensors, contact pressure, and patient demographics (age and sex) [9], [10]. Everything is then processed collectively by cloud-based AI.

B. Optical Biosensing and Theoretical Physics

PPG resolves these limitations by detecting optical attenuation through microvascular tissue [11]. When light is transmitted into vascularized areas (like the fingertip or earlobe), it produces a waveform with two main components: a static (DC) baseline from non-pulsatile arterial blood, venous blood, skin, and underlying tissues, and a pulsatile (AC) signal from cardiac blood volume expansion.

The Beer-Lambert law provides the mathematical basis for multi-wavelength concentration estimation [11], [12]:

$$A = \ln(I_s / I_0) = \sum \varepsilon_i(\lambda) C_i L(\lambda) \quad (1)$$

where A is the optical absorbance; I₀ is the incident light intensity; I_s is the detected light intensity; $\varepsilon_i(\lambda)$ is the wavelength-dependent molar extinction coefficient of chromophore i; C_i is the chromophore concentration (e.g., oxyhemoglobin HbO₂ and deoxyhemoglobin HHb); and L(λ) is the wavelength-dependent optical path length through scattering tissue.

In the optical window (600 nm to 1000 nm), HHb exhibits dominant extinction near 660 nm, whereas HbO₂ absorption peaks in the near-infrared region (850 nm–940 nm). Total hemoglobin (tHb) concentration can be calculated using multi-variable models by obtaining signal attenuation across these visible and near-infrared wavelengths [12].

II. CLASSIFICATION OF NON-INVASIVE DIAGNOSTIC TECHNIQUES

Non-invasive methods fall into three groups.

(1) Imaging-based modalities examine skin, palmar pallor, or palpebral conjunctiva using RGB cameras, hyperspectral imaging, or smartphone video [2]–[5].

- Strengths: hardware that is widely used and accessible.
- Weaknesses: skin surface reflections, auto-exposure artifacts, and sensitivity to changes in ambient light.

(2) Dedicated PPG optical biosensing uses contact multi-wavelength LED-photodiode systems in transmission or reflectance configurations [6]–[8].

- Strengths: strong signal-to-noise ratio (SNR), excellent temporal resolution, and direct microvascular interrogation.

- Weaknesses: prone to fluctuations in local contact pressure and to motion artifacts.

(3) Hybrid/multi-sensor fusion combines non-optical characteristics (such as skin temperature, contact pressure, and demographics like age, sex, and BMI) with optical PPG measurements [9], [10].

- Strengths: by accounting for physiological factors, prediction accuracy is increased.

III. COMPUTATIONAL FRAMEWORKS AND MACHINE LEARNING

A. Classical Models and Chemometrics

Early approaches relied on linear chemometric frameworks such as Multiple Linear Regression (MLR) and Partial Least Squares (PLS) [12]–[14]. Although PLS separates orthogonal latent structures from coupled optical channels, non-linear light scattering and skin-tissue variability are not taken into account by its linear assumptions.

B. Gradient Boosting with Decision-Tree Ensembles

Non-linear feature interactions are well captured by machine learning techniques such as Random Forest (RF), Adaptive Boosting (AdaBoost), and Extreme Gradient Boosting (XGBoost) [15], [16].

- Using an 8-wavelength PPG setup and an AdaBoost regressor, Zhu et al. [7] achieved a Mean Absolute Error (MAE) of 2.67 g/L and R² = 0.91 for a group of 56 patients.

- Chen et al. [6] used a 4-wavelength setup (660, 730, 850, and 940 nm) to extract 160 morphological characteristics. They obtained an R² of 0.997, an RMSE of 0.762 g/L, and an MAE of 0.325 g/L (n = 58) by combining Chi-square feature selection with XGBoost.

- Using 4-wavelength PPG and SHAP interpretability, Sahu et al. [17] investigated LightGBM [18] on 152 participants, obtaining an MAE of 8.50 g/L, RMSE of 8.21 g/L, and R² = 0.89 under rigorous subject-wise partitioning.

C. Deep Learning and Neural Architectures

End-to-end learning from raw waveforms is made possible by deep learning models [19]. Convolutional Neural Networks (CNNs) reduce the biases associated with human feature extraction by automatically extracting spatial-temporal characteristics from multi-wavelength pulse waveforms [20]. Multi-Layer Perceptrons (MLPs) also function well when given statistical data and demographic factors. In a 2026 study, for example, the mean relative error was lowered from 16.05% to less than 2.5% utilizing an MLP with six PPG variables in addition to age and sex [10].

D. Methodological Barriers: Preventing Information Leakage

Data leakage resulting from window-level or frame-level splitting is a crucial problem in PPG literature [17]. Metrics are artificially inflated when overlapping temporal frames from the same person are divided across training and test sets. Subject-level stratified cross-validation or held-out subject partitioning are necessary for rigorous evaluations [17].

IV. COMPARISON OF HARDWARE RECEIVERS AND FRONT-END ARCHITECTURE

A. Detector Options

Dynamic range, temporal sampling, latency, and system cost must all be balanced when choosing an optical detector. CMOS image sensors integrate the detector array and readout electronics on a single chip [21] and have also been studied as receivers outside conventional imaging, for example in camera-based visible light communication [22], [23]. The detector options are compared in Table II.

TABLE II. COMPARATIVE SENSOR MATRIX FOR OPTICAL BIOSENSING

Sensor Modality	Acquisition Mode	Active Area	Relative Cost	Latency
CMOS Image Sensor [21]	Rolling Shutter	$\sim 1.7 \mu\text{m}^2$ pixels	High	10–33 ms
CCD Array	Global Shutter	$\sim 9.0 \mu\text{m}^2$ pixels	Very High	100–500 ms
Solar Cell	Continuous	$\sim 800 \times 600 \mu\text{m}$	Very Low	1–5 ms
Single Photodiode	Direct Analog	$\sim 1000 \times 1000 \mu\text{m}$	Low	<1 ms
Multi-PD Array	Parallel Multi-Channel	Compact Array	Low-Moderate	<1 ms

Multi-channel Photodiode (PD) arrays coupled with Analog Front Ends (AFEs such as ADPD4100 or AFE4490) provide the best combination of low latency, high SNR (>50 dB), parallel wavelength sampling, and small form factors for wearables.

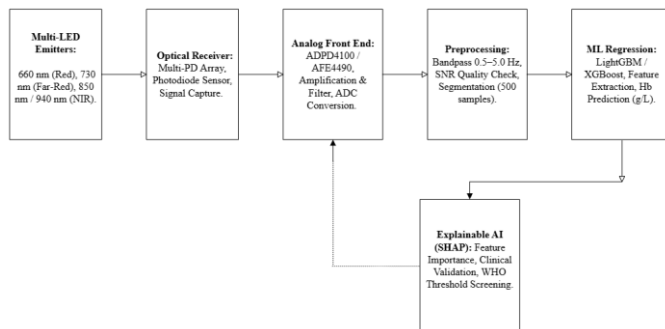


Fig. 2. End-to-End Hardware and Processing Pipeline

B. Processing Pipeline

Fig. 2 depicts the entire data journey from light emission to clinical conclusion as a linear process flow. It maps the six successive stages of the proposed non-invasive hemoglobin monitoring system, described below.

- Multi-LED emitters (660, 730, 850, 940 nm): four LEDs emit light at wavelengths selected because oxyhemoglobin absorbs in the 850–940 nm near-infrared range, while deoxyhemoglobin absorbs substantially at 660 nm (red) [6]. The LEDs are time-multiplexed, i.e., rapidly switched, so that the photodiode can discern the signal from each wavelength.
- Optical receiver array (multi-PD): after light travels through fingertip tissue, it is captured by a multichannel photodiode array, which transforms optical intensity into an electrical current signal.
- Analog front end (ADPD4100/AFE4490): this integrated circuit carries out (a) bandpass filtering, which eliminates high-

frequency noise and low-frequency baseline drift; (b) transimpedance amplification, which transforms minute photocurrent into measurable voltage; and (c) analog-to-digital conversion (ADC). The ADPD4100 actively rejects interference from ambient light while concurrently driving up to 8 LEDs and reading 8 photodiode channels.

- Preprocessing and filtering: digital signal processing uses a third-order Butterworth bandpass filter (0.5–5.0 Hz) to separate the cardiac pulse frequency. The SNR is calculated using Welch's Power Spectral Density, and only high-quality 500-sample windows are kept; corrupted segments are discarded.
- ML regression engine (LightGBM/XGBoost): an ensemble gradient boosting model forecasts the continuous hemoglobin concentration in g/L from the cleaned feature matrix (time-domain statistics, AC/DC ratios, and logarithmic attenuation ratios) [16], [18].
- Explainable AI (SHAP) and clinical output: SHAP values reveal the wavelength ratios driving each prediction [24]. A binary anemia screening result is obtained by comparing the continuous Hb value to WHO standards (Hb < 120 g/L for females and < 130 g/L for males) [1]. SHAP insights can guide model improvement, as indicated by the dashed feedback arrow in Fig. 2.

C. Circuit-Level Implementation

A circuit-level realization of the pipeline in Fig. 2 consists of the following tangible electronic elements and their connections.

- LED array: four LEDs at 660 nm, 730 nm, 850 nm, and 940 nm, powered by a time-multiplexed LED driver circuit, so that only one LED is illuminated at any instant and the photodiode can assign the incoming light to the correct wavelength.
- Fingertip tissue: the measurement site, in which the microvasculature carries the pulsating arterial blood volume. Each wavelength penetrates the tissue to a different extent depending on its absorption and scattering characteristics. A pressure/temperature sensor beneath the digit measures contact force and skin temperature, both of which are recognized confounders of optical path length [9], [10].
- Photodiode receiver array: transforms the attenuated light, whose degree of attenuation at each wavelength carries the information on hemoglobin concentration, into a small electrical current. As a 4-channel or single multichannel device, it separates the time-multiplexed wavelength signals.
- ADPD4100 or AFE4490 integrated circuit: amplifies the weak photocurrent, mitigates noise, converts the signal to digital format, and reduces ambient light interference by synchronous modulation. It transmits clean digital data over SPI or I2C.
- nRF52840 microcontroller: processes the digital PPG data, executes preprocessing algorithms, performs machine learning inference for hemoglobin prediction, and sends the results wirelessly via Bluetooth Low Energy (BLE). It also manages LED timing and sensor coordination.
- Display/application: a smartphone or clinical tablet receives the BLE transmission and presents the predicted hemoglobin level and the anemia screening outcome.

- Power supply: a 3.7 V Li-Ion battery paired with an LDO voltage regulator ensures consistent power delivery to the LED driver, the AFE, and the microcontroller.

V. CLINICAL INTERPRETABILITY VIA SHAP METHODOLOGIES

Machine learning algorithms frequently face criticism for being “black boxes” in medical diagnosis, and healthcare professionals require clear justification before accepting AI-based evaluations.

SHapley Additive exPlanations (SHAP) address this issue by utilizing cooperative game theory to measure the contribution of each feature [24]. For a prediction $f(x)$, the Shapley value ϕ_i of feature i is determined as follows:

$$\phi_i(x) = \sum_{S \subseteq F \setminus \{i\}} w(S) [f_x(S \cup \{i\}) - f_x(S)] \tag{2}$$

$$w(S) = |S|! (|F| - |S| - 1)! / |F|! \tag{3}$$

where F denotes the complete collection of features, S signifies a subset of features that omits feature i , and $w(S)$ is the weight given by (3). Fig. 3 summarizes the SHAP explainability framework: panel (A) ranks the features by importance, panel (B) shows a waterfall explanation for an anemic subject, and panel (C) outlines the TreeSHAP computation pipeline.

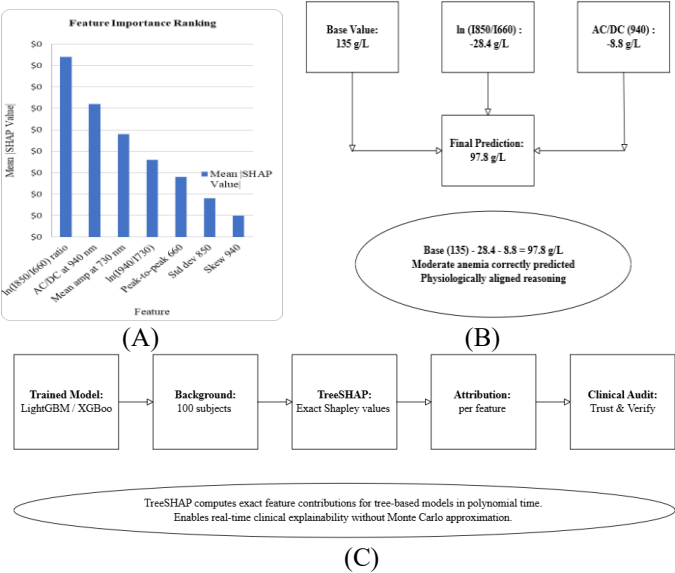


Fig. 3. SHAP Explainability Framework: (A) Feature Importance Ranking, (B) Waterfall Plot: Anemic Subject, (C) TreeSHAP Computation Pipeline

Applying SHAP to multi-wavelength PPG models verifies that the optical attenuation ratios across wavelengths, particularly near-infrared (850 nm, 940 nm) compared with visible light (660 nm, 730 nm), have the most significant impact on model predictions. This is in direct accordance with the Beer-Lambert equation, confirming that the models capture authentic physiological processes instead of dataset artifacts [17], [24].

VI. BENCHMARKS AND CLASSIFICATION EFFICACY

A. Regression Summary

TABLE III. OVERVIEW OF NOTABLE REGRESSION RESEARCH

Study/ Year	Cohort	Wavelengths	Model	R ²	MAE (g/L)
Kavsaoglu et al./2015	n = 100	Dual (660/940)	SVR + RFS	0.85	~7.50
Acharya et al./2020	n = 150	Dual + Demographics	Stacking Regression	0.88	~6.20
Kwon & Kim/2022	n = 80	4-Wavelength	Grad. Boosting	0.93	~5.50
Chen et al./2023	n = 58	4-Wavelength	XGBoost	0.997	0.325
Zhu et al./2024	n = 56	8-Wavelength	AdaBoost	0.91	2.670
Sahu et al./2026	n = 152	4-Wavelength	LightGBM + XAI	0.89	8.500
Frontiers Study/2026	n = 68	Dual + Demographics	MLP	0.96	~7.80

Table III summarizes notable regression studies on PPG-based hemoglobin estimation [6], [7], [9], [10], [15], [17], [25].

B. Efficacy of Binary Diagnostic Screening

Upon transforming continuous estimates into binary anemia screening determinations according to WHO standards (Hb < 120 g/L for females, < 130 g/L for males) [1], advanced LightGBM and ensemble frameworks [17], [26] attain the metrics listed in Table IV, measured on a separate test partition at the subject level.

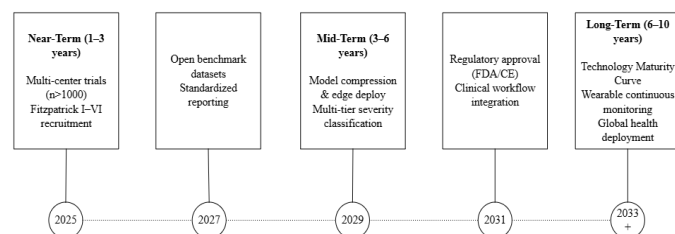
TABLE IV. PERFORMANCE METRICS FOR DIAGNOSTIC SCREENING

Diagnostic Metric	Value
Sensitivity (True Positive Rate)	93.3%
Specificity (True Negative Rate)	93.8%
Positive Predictive Value (PPV)	91.7%
Negative Predictive Value (NPV)	95.2%
Overall Classification Accuracy	93.5%
AUC-ROC	0.962

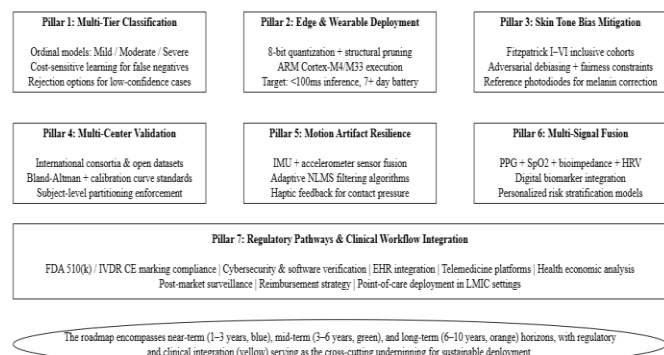
To ensure that anemic patients are rarely missed, point-of-care screening must have a high Negative Predictive Value, and the 95.2% achieved here meets this requirement.

VII. RESEARCH OPPORTUNITIES AND FUTURE DIRECTIONS

A robust framework for non-invasive hemoglobin quantification has been developed by combining multi-wavelength photoplethysmography (PPG) with machine learning and explainable AI [6], [7], [17]. However, additional technological development is required to transform experimental prototypes into screening devices that are scalable, clinically feasible, and cost-effective. Important directions for further research are outlined here and summarized in Fig. 4.



(a) Development timeline and technology maturity trajectory



(b) Future PPG hemoglobin systems: seven foundational elements of strategic research

Fig. 4 Prospects for Future Non-Invasive Hemoglobin Monitoring and a Plan for Strategic Research

The two parts of Fig. 4 illustrate the directions that research into non-invasive hemoglobin monitoring is likely to take in the coming years. Part (a) is the development timeline (2025–2033), with three stages.

- **Near-term (1–3 years, 2025–2027):** establish credibility through large-scale, multi-center clinical trials with more than 1,000 participants covering a range of skin tones (Fitzpatrick types I–VI), develop publicly available benchmark datasets, and harmonize the reporting of study results.
- **Mid-term (3–6 years, 2027–2031):** make the technology more usable and deployable. Primary objectives include regulatory permissions such as FDA 510(k) and CE marking, multi-tier anemia severity categorization (mild, moderate, severe), and the compression of machine learning models for execution on small microcontrollers (edge deployment).
- **Long-term (6–10 years, 2031–2033 and beyond):** broad practical use, such as wearable hemoglobin monitoring devices and worldwide rollout in nations with limited resources.

The technology maturity curve indicates that the middle period is when development picks up speed the most. Part (b) shows the seven foundational elements of strategic research, i.e., the interrelated clinical and technical domains that must progress in tandem:

- Three-tier classification, using ordinal models to differentiate between mild, moderate, and severe anemia;
- Edge and wearable deployment, miniaturizing models for tiny chips with rapid inference and extended battery life;
- Reduction of skin-tone bias, using fairness algorithms and inclusive datasets so that the device works for all skin tones;
- Multi-center validation across various hospitals and nations using standardized evaluation methodologies;
- Motion artifact resilience, using motion sensors and filtering to ensure reliable readings regardless of patient movement;

- Multimodal fusion of PPG with other signals such as bioimpedance, heart rate variability, and oxygen saturation; and
- Regulatory and clinical integration (FDA/CE approval, cybersecurity, electronic health record integration, telemedicine, and cost-effectiveness analysis), which forms the basic pillar.

A. Gradual Classification of Anemia Severity

Existing techniques focus mainly on continuous hemoglobin regression and simple binary screening (anemic vs. non-anemic). The therapeutic efficacy of future frameworks will be greatly enhanced if they include the categorization of mild, moderate, and severe states. To represent the natural course of illness severity while significantly penalizing dangerous misdiagnoses, ordinal classification systems, cost-sensitive loss functions, and uncertainty quantification can be utilized. Moreover, algorithms should flag predictions with low certainty or borderline accuracy, and conventional laboratory follow-up should be prompted when confidence levels are not met.

B. Embedded and Edge Deployment

Wearables and point-of-care hardware have limited resources. Therefore, continuous monitoring models must run locally on these devices, without relying on high-performance remote servers or reliable network connectivity, if they are to be useful in the real world. Model compression procedures, including parameter quantization, weight pruning, and knowledge distillation, reduce memory footprints and processing demands [20]. The next step for engineers is to find a middle ground between diagnostic accuracy, latency, and power consumption by benchmarking compressed models on low-power microcontrollers.

C. Reducing Skin-Tone Bias

Systemic bias can be introduced across different skin tones because variations in epidermal melanin affect optical absorption patterns. To thoroughly assess efficacy across all Fitzpatrick skin types, future studies should aim to recruit cohorts that are representative of all relevant demographics. These optical differences can be reduced with the help of algorithmic fairness frameworks, domain adaptation techniques, additional optical channels, and reference sensors at the hardware level. Federated learning also allows cross-site collaboration on training without jeopardizing private patient data.

D. Standardized Benchmarking and Cross-Institutional Validation

Reliance on localized, small-scale clinical datasets is a common problem in current studies and limits generalizability to real-world situations [27]. Extensive clinical trials involving multiple centers and populations that differ in terms of geography, ethnicity, and socioeconomic status are necessary. To facilitate direct baseline comparisons, standardized, publicly available benchmark datasets should combine raw PPG waveforms with reference laboratory measurements, demographic variables, and sensor characteristics; a four-wavelength PPG dataset for hemoglobin assessment [28] is a step in this direction. Repeatability can be improved by consistently reporting Bland-Altman agreement analyses, calibration curves, and confidence intervals.

E. Reducing Motion Artifacts and Ensuring Signal Accuracy

For wearable PPG devices, physical movement is still the biggest cause of signal deterioration. Integrating optical and inertial measuring devices (such as gyroscopes and accelerometers) enables systems to dynamically identify and filter out movement noise. Signal fidelity can be preserved during everyday tasks by combining sensor-fusion techniques with adaptive noise cancellation. Ergonomic sensor upgrades such as contact-force monitoring can further ensure consistent skin contact and lessen motion-induced measurement drift.

F. Multimodal Health Monitoring

Integrating PPG-derived hemoglobin metrics with supplementary vital signs, including pulse rate variability, arterial oxygen saturation (SpO₂), skin temperature, and tissue hydration indices, might further improve non-invasive assessment [9], [10]. A more complete and contextualized picture for individual health monitoring and early risk detection can be achieved by incorporating broader digital biomarkers, such as self-reported fatigue, physical exertion logs, and dietary trends.

G. Clinical Workflow Integration and Regulatory Compliance

The successful implementation of these technologies in healthcare settings will require thorough testing, approval from relevant authorities, strong cybersecurity measures, and ongoing monitoring. Research needs to provide detailed reports on diagnostic reliability, subgroup dependability, repeatability, and specificity. Adoption will be facilitated by the seamless integration of these tools with electronic health records (EHRs), telemedicine software, and clinical decision support systems. Finally, health-economic analyses are crucial for proving cost-effectiveness, especially in areas with limited resources that could benefit the most from non-invasive screening.

CONCLUSIONS

From the earliest days of dual-wavelength pulse oximetry to the most recent state-of-the-art four- and eight-wavelength PPG systems integrated with ensemble machine learning and explainable AI [6], [7], [17], this review has followed the development of non-invasive hemoglobin monitoring. This area is on the cusp of clinical viability thanks to the convergence of multispectral optical hardware, stringent signal filtering, domain-informed feature engineering, and interpretable regression models. With sensitivity levels above 93% for anemia detection [17], [26] and reported mean absolute errors below 3 g/L in controlled conditions [6], [7], PPG-based hemoglobin estimation can be a useful screening and triage tool, especially in areas with limited laboratory infrastructure and resources.

However, a lot of work needs to be done before these technologies may be confidently used by varied communities all over the world. It is of the utmost importance to eliminate skin tone bias, verify results in pediatric and geriatric groups, and accomplish strong motion artifact rejection. Clinicians must ensure that model decisions are based on physiologically meaningful optical attenuation patterns instead of dataset-specific artifacts, and an explainable AI framework such as SHAP is an essential component of this process.

The future is bright: the line between home health aides and professional diagnostic tools will become increasingly blurred as smaller, less power-hungry, and smarter optical biosensors

enter the market. Continuous, non-invasive hemoglobin monitoring for billions of persons at risk of anemia becomes a more realistic goal as model compression techniques improve and open datasets grow.

REFERENCES

- [1] World Health Organization, "Global anaemia report 2023," WHO, Geneva, Switzerland, 2023.
- [2] A. Kesarwani *et al.*, "Non-invasive anemia detection via palmar pallor analysis and computational methods," *J. Med. Syst.*, vol. 47, no. 1, p. 45, 2023.
- [3] X. Fan *et al.*, "Regression-based smartphone biosensor for hemoglobin quantification from RGB features," *Biosens. Bioelectron.*, vol. 198, Art. no. 113805, 2022.
- [4] G. Dimauro *et al.*, "Eyes-Defy-Anemia: A dataset and RUSBoost classification for anemia detection from conjunctival images," *Sci. Rep.*, vol. 13, no. 1, p. 7824, 2023.
- [5] M. A. Hasan *et al.*, "Smartphone-based non-invasive hemoglobin estimation using artificial neural networks," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–20, 2018.
- [6] Z. Chen, H. Qin, W. Ge, S. Li, and Y. Liang, "Research on a non-invasive hemoglobin measurement system based on four-wavelength photoplethysmography," *Electronics*, vol. 12, no. 6, p. 1346, 2023.
- [7] J. Zhu *et al.*, "A non-invasive hemoglobin detection device based on multispectral photoplethysmography," *Biosensors*, vol. 14, no. 1, p. 22, 2024.
- [8] Y. Liu *et al.*, "Portable eight-channel PPG system for non-invasive hemoglobin estimation," *IEEE Sensors J.*, vol. 20, no. 18, pp. 10842–10851, 2020.
- [9] P. Acharya *et al.*, "A multi-model stacking regressor for non-invasive hemoglobin prediction using PPG signals," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 8, pp. 2345–2354, 2020.
- [10] "An approach to machine learning-based non-invasive hemoglobin estimation using multi-wavelength PPG signal features," *Front. Physiol.*, Apr. 21, 2026.
- [11] K. Li and J. Sun, "Understanding the physiological transmission mechanisms of photoplethysmography signals: A comprehensive review," *Physiol. Meas.*, vol. 45, no. 9, Art. no. 09TR01, 2024.
- [12] B. Yuan *et al.*, "Near-infrared spectrophotometric hemoglobin measurement with DOSC-PLS," *J. Biomed. Opt.*, vol. 20, no. 3, Art. no. 037001, 2015.
- [13] Y. Kumar, "Multi-wavelength pulse wave acquisition and linear regression for hemoglobin estimation," *Biomed. Signal Process. Control*, vol. 78, Art. no. 103978, 2022.
- [14] F. Tian *et al.*, "Comparison of BP-ANN and PLS for non-invasive hemoglobin measurement using multi-wavelength PPG," *Opt. Express*, vol. 27, no. 15, pp. 21089–21102, 2019.
- [15] S. Kwon and J. Kim, "Gradient boosting and random forest models for PPG-based hemoglobin prediction," *Comput. Biol. Med.*, vol. 145, Art. no. 105438, 2022.
- [16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794.
- [17] G. Sahu, P. Verma, and N. Tapas, "Non-invasive anemia detection: A multichannel PPG-based hemoglobin estimation with explainable artificial intelligence," 2026, *arXiv:2602.17290*.
- [18] G. Ke *et al.*, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 30, 2017, pp. 3146–3154.
- [19] G. Nie *et al.*, "A scoping review of deep learning methods for photoplethysmography data," 2024, *arXiv:2401.12783*.
- [20] F. Smarandache *et al.*, "Noninvasive hemoglobin estimation with adaptive lightweight convolutional neural network using wearable PPG," *Comput. Model. Eng. Sci.*, vol. 144, no. 3, pp. 3715–3735, 2025.
- [21] E. R. Fossum, "CMOS image sensors: Electronic camera-on-a-chip," *IEEE Trans. Electron Devices*, vol. 44, no. 10, pp. 1689–1698, Oct. 1997.
- [22] C. Danakis *et al.*, "Using a CMOS camera sensor for visible light communication," in *Proc. IEEE Globecom Workshops*, 2012, pp. 1244–1248.

- [23] W. Huang and Z. Xu, "Characteristics and performance of image sensor communication," *IEEE Photon. J.*, vol. 9, no. 2, pp. 1–19, Apr. 2017.
- [24] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 30, 2017, pp. 4765–4774.
- [25] A. R. Kavsaoglu, K. Polat, and M. Hariharan, "Non-invasive prediction of hemoglobin using machine learning regression with PPG signal features," *Biocybern. Biomed. Eng.*, vol. 35, no. 3, pp. 208–218, 2015.
- [26] F. Peng, N. Zhang, C. Chen, and F. Wu, "Ensemble extreme learning machine method for hemoglobin estimation based on photoplethysmographic signals," *Sensors*, vol. 24, no. 5, p. 1547, 2024.
- [27] D. F. Silva *et al.*, "Hemoglobin estimation from smartphone-based photoplethysmography with small data," in *Proc. IEEE 36th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, 2023, pp. 345–350.
- [28] T. Abuzairi *et al.*, "A four-wavelength photoplethysmography dataset for non-invasive hemoglobin assessment," *Sci. Data*, vol. 11, Art. no. 245, 2024.