Physics-Guided Optimization for Learning under Physiological Constraints: A Case Study on Hemoglobin Estimation

Submitted to Special Session: Advances in Data-Driven Optimization and AI-Enhanced Business Process Mining

Shashank Mishra[1], Aditya Jha[2], Kartik Pandey[2], and Vijaya J[3]

[1,2] UG Student, DSAI, International Institute of Information Technology Naya Raipur, India
[3] Assistant Professor, DSAI, International Institute of Information Technology Naya Raipur, India
shashank23102@iiitnr.edu.in, aditya23102@iiitnr.edu.in, kartik23100@iiitnr.edu.in, vijaya@iiitnr.edu.in

Abstract. This study addresses hemoglobin estimation from photoplethysmography (PPG) signals as a learning-under-noise optimization problem, where subject variability and weak supervision lead to unstable solutions for purely data-driven models. A structured optimization-centric framework is proposed in which signal preprocessing and segmentation are followed by comparative evaluation of multiple sequence-learning architectures, including BiLSTM, CNN-BiLSTM, BiGRU, and BiGRU with attention, with an MLP serving as a baseline. Model performance is assessed using MAE and RMSE, revealing the benefits of attention-enhanced temporal representations. To further improve solution quality, metaheuristic optimization techniques, including Grey Wolf Optimization (GWO) and Genetic Algorithms (GA), are employed to guide feature selection and model configuration under noisy conditions. The results highlight that constraint-aware and optimization-guided learning leads to more stable and reliable regression performance than unconstrained training, demonstrating the effectiveness of combining deep sequence models with evolutionary optimization strategies for real-world data-driven optimization tasks.

Keywords: Physics-Guided Optimization · Constrained Learning · Metaheuristic Algorithms · Noisy Real-World Data

1 Introduction

In recent years, the increasing popularity of wearable sensing technology and diagnostic devices has further fueled the need for non-invasive physiological monitoring. Among the most important physiological markers is hemoglobin

(Hb) concentration, which is commonly employed for the diagnosis of anemia, cardiovascular diseases, and chronic illness monitoring. However, the measurement of hemoglobin concentration is still largely carried out by invasive blood sampling, which is painful for patients, resource-intensive, and not amenable to continuous or mass screening. A promising alternative to invasive blood sampling has been the emergence of photoplethysmography (PPG), which is inexpensive, easy to implement, and widely available in commercial and clinical devices. Multi-wavelength PPG signals, in particular, carry wavelength-specific light absorption properties that are biologically associated with hemoglobin concentration based on well-established optical principles such as the Beer-Lambert law. Consequently, non-invasive hemoglobin estimation from PPG has become an increasingly active area of research at the crossroads of biomedical sensing, machine learning, and optimization, with increasing relevance to the data-driven optimization and AI communities [10, 2].

Although recent research has shown that estimating hemoglobin from PPG using deep learning and machine learning techniques is feasible, there are still a number of important restrictions. Performance has been enhanced by hybrid architectures that combine convolutional neural networks (CNNs) with recurrent models like gated recurrent units (GRU) or long short-term memory (LSTM), which capture both waveform morphology and temporal dynamics. [27]. These models are further improved by attention mechanisms, which suppress noise and prioritize informative pulse segments. Existing methods, however, frequently disregard established physiological and optical limitations and are primarily data-driven, optimized only for empirical error minimization. Models that lack domain-aware regularization are more susceptible to dataset bias, signal artifacts, and inter-subject variability, which limits their interpretability and generalizability. Although physics-informed learning has become popular in other scientific fields, there is still much to learn about how to systematically incorporate it into optimization and learning pipelines for PPG-based hemoglobin estimation. [28, 29].

To this end, this manuscript proposes a physics-guided optimization approach for hemoglobin estimation that incorporates physiological knowledge directly into the optimization goal, as opposed to the network architecture. In this paper, we present a Physiologically-Constrained Neural Network (PCNN) that incorporates optical consistency and smoothness constraints within a subject, based on domain knowledge, in addition to the conventional regression loss. Moreover, we explore the use of metaheuristic optimization, in particular Grey Wolf Optimization (GWO), as an additional tool for feature and hyperparameter selection. Crucially, our proposed approach maintains a shared deep temporal architecture, facilitating a fair comparison among conventional, heuristic-based, and physics-guided learning methods. The experimental findings validate that physics-guided optimization enhances robustness and generalization capabilities over purely data-driven learning methods, while maintaining interpretability and optimization transparency. The rest of this manuscript describes the proposed approach, experimental setup, and comparison, focusing on the importance of optimization with physiological constraints for non-invasive hemoglobin estimation.

2 Related Work

Because the non-invasive hemoglobin (Hb) monitoring can replace invasive blood sampling while allowing for continuous and real-time physiological assessment, it has drawn ongoing attention. Early studies in this field, which concentrated on how biological tissue's light absorption and scattering characteristics change with wavelength, were mainly based on optical measurement theory. Because penetration depth and absorption characteristics vary nonlinearly across wavelengths, Nirupa and Kumar showed that carefully chosen wavelength pairs in the red and near-infrared spectrum have a significant impact on the accuracy of Hb estimation. [6]. Similarly, Kraitl et al. worked on the optical detection of blood components and found out that Hb absorbance measurements are highly sensitive to hardware factors such as LED drive current stability and photodiode alignment [7]. These findings stressed that both biological variability and sensor imperfections contribute to measurement uncertainty.

A useful and popular optical biosensing technique for recording cardiovascular dynamics in wearable and clinical contexts is photoplethysmography (PPG). Heart rate and oxygen saturation monitoring were the main focus of early PPG-based systems, which showed that dependable optical signals could be obtained even in difficult circumstances. Bailey et al. developed a low power pulse oximeter prototype capable of stable PPG acquisition, opening teh way for mobile and battery-operated health monitoring devices [8]. Kyriacou et al. demonstrated that signal quality can be maintained even in low perfusion and difficult physiological conditions, further validating the clinical reliability of reflectance mode PPG during major surgical procedures. [9]. Allen gave a thorough analysis of PPG as a means of evaluating vascular tone, autonomic regulation, and circulatory dynamics, which subtly supported its applicability for estimating hematological parameters like hemoglobin[10]. Collectively, these studies established PPG as a robust and versatile sensing modality.

As the sensing using PPG continued to evolve, there was an increasing interest in combining machine learning approaches to derive more complex physiological insights than what could be achieved through conventional signal processing. Johansson was one of the first to show that neural networks are superior to Fourier analysis-based approaches for extracting respiratory cues from PPG signals, exemplifying the capability of learning-based models to detect subtle waveform patterns [11]. Further works incorporated second-derivative PPG (SDPPG) features and advanced morphological descriptors to characterize vascular stiffness, endothelial function, and arterial aging [12, 13]. These factors became more important because of their indirect relevance to blood viscosity and oxygen transport capacity, both of which affect hemoglobin-related values. Ding et al. made a great contribution by using principal component analysis together with back-propagation neural networks for Hb measurement, proving that dimensionality reduction can help reduce sensor noise and tissue scattering effects while obtaining meaningful accuracy[14]. Although these approaches improved performance, they remained heavily dependent on handcrafted features and shallow learning architectures.

In parallel to machine learning-based methods, a significant amount of research focused on the photonic properties of hemoglobin and optical wavelength-dependent behavior. Research by Iwasaka et al. investigated the influence of external magnetic fields on red blood cell absorbance, indirectly indicating anisotropic scattering properties useful for wavelength-oriented sensing [15]. Jay et al. showed that the optical path length and reflectance configuration have a substantial impact on Hb sensitivity, successfully validating reflectance-based measurement using fiber-optic spectroscopy in point-of-care devices [16]. Near-infrared (NIR) spectroscopy further broadened the scope of application, with Yichao et al. illustrating the capability of multispectral reflectance to differentiate oxygenation states in neonatal care under very low perfusion conditions [17]. The feasibility of estimating Hb and HbO_2 concentrations using broad-spectrum NIR analysis was confirmed by Head and Pierson [18]. Although they frequently required specialized equipment and experimental setup, these investigations further highlighted the strong physiological basis of wavelength-dependent optical sensing.

In order to increase the reliability of these systems in actual acquisition scenarios, new signal processing techniques were also proposed for the estimation of the Hb level in the blood. Timm et al. demonstrated an optical Hb sensing system with synchronous detection and the use of multiple LED excitation devices and real-time peak detection monitors. The importance of precise peak detection was emphasized in order to minimize noise caused due to movement [19]. Binns et al., then later Kavsaoğlu et al., proposed the use of frequency domain analysis for photoplethysmogram devices. The paper demonstrated their use in extracting useful information from physiological signals [20, 21]. In their revised model, the window sizes were adapted according to beat patterns in the signal, making the system more reliable in terms of feature dChoosing and prioritizing features became crucial as the number of features rose in order to reduce redundancy and avoid overfitting. Even in noisy settings, algorithms such as RELIEF and RELIEFF, which were first presented by Kira and Rendell and then developed by Kononenko, successfully identified valuable PPG features [23, 24]. Due to their simplicity and dependability, these techniques have become more and more popular in the analysis of physiological signals. Simultaneously, ensemble-based learning methods gained popularity for improving prediction accuracy in the face of nonlinear distortions and subject-to-subject variations. Das et al. and Das and Sengur showed that hybrid classifiers and ensemble regression frameworks consistently perform better than single-model methods for biomedical signal classification and regression tasks [25, 26]. While ensemble methods did boost accuracy, they often made models more complex and harder to interpret.

Despite significant advances in optical sensing, signal processing, and machine learning, there is a shared limitation in most current methods for PPG-based hemoglobin estimation: they perform unconstrained learning focused on error minimization without taking into account known physiological or optical relationships. Because of this, models may perform well numerically but may produce predictions that are not physiologically realistic or that do not generalize well to other subjects and conditions. Recent developments in machine learning informed by physics have demonstrated that robustness and understandability in scientific tasks can be significantly improved by directly incorporating domain knowledge into the optimization process. In the field of PPG-based hemoglobin estimation, these constraint-aware optimization techniques are still largely unproven. This gap gives rise to the suggested physics-guided optimization framework, which addresses the drawbacks of purely data-driven approaches by integrating physiological constraints into deep temporal learning.

3 Proposed Methodolgy

The problem of estimating hemoglobin (Hb) concentration from non-invasive multi-wavelength photoplethysmography (PPG) signals is addressed in this work. It uses a framework for supervised regression with physiological constraints. Let $\mathcal{D} = \{(X^{(s)}, y^{(s)})\}_{s=1}^{S}$ represent a dataset of S subjects, where $X^{(s)}$ is the recorded PPG signal and $y^{(s)} \in R$ is the actual hemoglobin concentration, measured invasively in g/dL.

The PPG signal of each subject is separated into N_s fixed-length segments

$$X^{(s)} = \{x_1^{(s)}, x_2^{(s)}, \ldots, x_{N_s}^{(s)}\}, \tag{1}$$

where $x_i^{(s)} \in R^{T \times W}$ stands for the ith segment, which has a length T across W wavelengths. Processing the segments in this way reduces the impact of motion artifacts and transient noise that occur in the process of collecting PPG signal.

A parameterized model $f_\theta(\cdot)$ predicts hemoglobin estimates at the segment level in the learning task, which is set up as a regression problem.

$$\hat{y}_i^{(s)} = f_\theta(x_i^{(s)}, \mathbf{h}_i^{(s)}, \mathbf{c}^{(s)}), \tag{2}$$

, where $\mathbf{h}_i^{(s)}$ represents handcrafted morphological features and $\mathbf{c}^{(s)}$ provides subject-level conditioning information. In conventional learning, parameters θ are determined by minimizing the empirical mean squared error (MSE).

$$\mathcal{L}_{MSE}(\theta) = \frac{1}{\sum_s N_s} \sum_{s=1}^{S} \sum_{i=1}^{N_s} (\hat{y}_i^{(s)} - y^{(s)})^2. \tag{3}$$

This uninhibited approach, however, overlooks important physiological relations that affect the optical absorption and stability of an individual. To address this issue, we utilize a physics-constrained optimization framework wherein we minimize the following constrained objective in order to learn the model parameters:

$$\theta^* = \arg\min_{\theta} \; \mathcal{L}_{MSE}(\theta) + \lambda_1 \mathcal{L}_{opt}(\theta) + \lambda_2 \mathcal{L}_{smooth}(\theta),$$

For this problem,
$mathcalL_{textopt}$ ensures optical consistency across all wavelengths, and
$mathcalL_{textsmooth}$ favors smooth

We calculate our final hemoglobin level estimate for each individual by taking the average of the predictions from each of the segments:

$$\hat{y}^{(s)} = \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{y}_i^{(s)}, \tag{4}$$

This aligns the model's findings with clinical evaluation protocols.

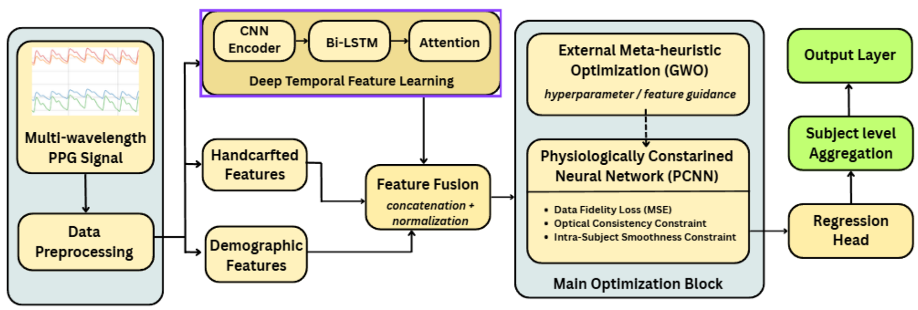

Fig. 1. Proposed Architecture

3.1 Dataset Description

This study uses a public dataset for non-invasive hemoglobin estimation. The dataset includes photoplethysmography (PPG) recordings from 252 subjects collected at four optical wavelengths: 660, 730, 850, and 940 nm. We recorded PPG signals from the left index finger at a sampling frequency of 200 Hz. Each recording lasted about 60 seconds.

The subjects are aged between 21 and 90 years, with 56.7% being female. We obtained precise hemoglobin values from venous blood samples using a HemoCue Hb 201+ analyzer, a validated point-of-care device. Along with PPG signals, we provide subject-level data that includes age, gender, height, weight, blood pressure, and blood glucose.

3.2 Data Preprocessing and Segmentation

Raw multi-wavelength PPG signals go through a standard preprocessing process to reduce noise, decrease differences between subjects, and maintain important characteristics of the waveforms. Each wavelength channel is processed separately using a fourth-order Butterworth bandpass filter with cutoff frequencies of 0.5 Hz and 8 Hz. This helps keep the main cardiac components while filtering out baseline drift, motion artifacts, and high-frequency sensor noise. Zero-phase filtering is applied to avoid altering the shape of the pulse over time. After filtering, z-score normalization is done on each channel to standardize amplitude variations among subjects and wavelengths. This results in stable optimization and prevents bias towards channels which have higher intensity values. Minimal rejection of artifacts based on thresholds of amplitude and variance is used to remove severely corrupted sections without rejecting normal physiological variation in realistic recordings.

As the signal length changes due to removal of artifacts, each PPG signal is divided into specific length time segments before feature extraction and training of the model. The signal segments were designed to be 10 seconds (2000 samples at 200 Hz) long. This is long enough to capture several heart cycles while being approximately stationary. Non-overlapping segments are generated to avoid information leakage between samples. All segments carry the subject-level hemoglobin labels, which correspond to the clinical scenario since an invasive measurement typically includes the entire session. The four wavelength input streams for each segment are concatenated from the aligned time series for each wavelength (660 nm, 730 nm, 850 nm, 940 nm). This enables the learned model to capture wavelength-specific information in the PPG signal.

3.3 Feature Construction and Conditioning Information

To augment the learned deep temporal representations of individual PPG segments, further handcrafted and conditioning features are included to offer domain-specific contexts and ensure robustness. These features aim to capture specific signal characteristics and information about the individual that may not be necessarily learned from the raw signal itself.

Handcrafted Feature Extraction :We have, for each PPG window, a set of time-domain features, as well as a set of morphological features for each of the wavelengths. Among these are: a set of statistical features-averages, standard deviation, skewness, and kurtosis-as well as a set of pulse shape information from a peak detection algorithm. The main information includes average peak amplitude, inter-beat interval, rising time, pulse width, and pulse area. These features are reported to be correlated with the flexibility of vessels, variations in volume, and the absorption associated with the concentration of haemoglobin. Feature extraction over all the wavelengths is done here.

Conditioning and Demographic Features : The information about the subjects is employed as an additional variable, handling the physiological differences and affecting the peculiarities of PPG signals. Information about the subject contains demographic factors such as age and gender, as well as additional clinical variables in the dataset, which include height, weight, blood pressure, and blood glucose levels. The conditioning features are normalized before being fused with the signal-based features. The variables aid in generalization, avoiding the direct use of hemoglobin values.

Feature Integration : We normalize and concatenate all handcrafted and conditioning features with deep temporal representations learned from the CNN-recurrent backbone. In this unified feature set, the regression model can make use of the complementary information not only from the raw waveforms and clear signal descriptors but also from the subject-level context, thus forming a complete input for successive optimization and prediction stages.

3.4 Deep Temporal Representation Learning

Deep temporal representation learning is used to extract useful features from the segments of multi-wavelength PPG signals, inheriting the merits of features from shape and time characteristics. Considering that the PPG signal is quasi-periodic and with higher sensitivity to noise, the hybrid model is designed and integrated with CNN, RNN, and attentional mechanisms.

The segmented PPG window is then subjected to a one-dimensional convolutional neural network (CNN). In this phase, features such as a pulse upstroke, a downstroke, and a suprasystolic notch could be encoded, in which features are learned along a single dimension. The components of the signal are utilized, as they are invariant to a certain extent. The encoded features are then subjected to a bidirectional long short-term memory (BiLSTM) network, in which features between heartbeats are learned.

To enhance time modeling, the attention mechanism operates on the BiLSTM output. This mechanism assigns different weight values to different time steps in each segment. This process enables the highlighting of key pulse intervals and reduces the contribution of noisy and/or distorted intervals. It presents a summary of each PPG segment and represents it as a deep temporal feature embedding.

3.5 Optimization Strategies

This paper focuses on three optimization strategies for estimating hemoglobin, namely unconstrained empirical risk minimization, external metaheuristic optimization, and physics guided constrained learning. All strategies use the same deep temporal backbone. This allows for fair comparisons and helps isolate how different optimization choices affect prediction performance and stability.

Unconstrained Learning Baseline : As a baseline, model parameters are learned by minimizing the empirical mean squared error (MSE) between predicted and actual hemoglobin values. With segment-level predictions $\hat{y}_i^{(s)}$ and subject-level labels $y^{(s)}$, the baseline objective is defined as

$$\mathcal{L}_{MSE} = \frac{1}{\sum_s N_s} \sum_{s=1}^{S} \sum_{i=1}^{N_s} \left(\hat{y}_i^{(s)} - y^{(s)} \right)^2 . \tag{5}$$

While this approach is straightforward and commonly used, it treats each segment independently. It does not make use of known physiological relationships that govern optical absorption or maintain consistency within subjects. Therefore, unconstrained optimization can lead to overfitting and unstable predictions when there is variability among subjects.

External Metaheuristic Optimization : To explore optimization through global search methods, Grey Wolf Optimization (GWO) is used as an external metaheuristic strategy. GWO takes inspiration from the social hierarchy and cooperative hunting behavior of grey wolves and works well for non-convex optimization problems. In this study, GWO helps with feature selection and tuning hyperparameters by iteratively updating candidate solutions based on prediction error fitness. Notably, GWO functions outside the gradient-based learning loop and does not change network weights directly. While metaheuristic optimization can improve convergence and reduce local minima issues, it does not consider physiological constraints and relies only on empirical error as a measure of fitness.

Physics-Guided Optimization via PCNN (Proposed) : To integrate domain knowledge directly into the learning process, we propose a Physiologically-Constrained Neural Network (PCNN). Instead of modifying the network architecture, PCNN incorporates constraint terms derived from optical and physiological principles directly into the optimization objective. The constrained loss function is expressed as

$$\mathcal{L}_{PCNN} = \mathcal{L}_{MSE} + \lambda_1 \mathcal{L}_{opt} + \lambda_2 \mathcal{L}_{smooth}, \tag{6}$$

where λ_1 and λ_2 are fixed weighting coefficients ($\lambda_1 = \lambda_2 = 0.1$).

Optical Consistency Loss: The consistency penalty due to optical consistency $\mathcal{L}_{opt}$ utilizes the physiological prior that increased hemoglobin leads to increased absorption, according to the Beer-Lambert law. In practice, the average value of the peak of the 660 nm channel, A_i, acts as the proxy for the optical absorption within the segment level. Here, a penalty is imposed on a batch such that there is a negative correlation between the estimated Hb values and A_i:

$$\mathcal{L}_{opt} = \max\left(0,\ -\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - \bar{\hat{y}}\right)\left(A_i - \bar{A}\right)\right), \tag{7}$$

where $\hat{y}_i$ represents the segment-wise estimated hemoglobin concentration, A_i denotes the average peak amplitude value of the 660 nm PPG wave in segment i, N represents the total number of segments in a given batch, and $\bar{\hat{y}}, \bar{A}$ are the respective batch averages. This is a one-sided soft constraint such that when the biophysically expected correlation between absorption and Hb is already present ($\mathrm{Cov}(\hat{y}, A) \geq 0$), then there is no penalty imposed on the gradient. **Intra-Subject Smoothness Loss:** As the level of hemoglobin is a physiological constant that is assessed only once during a particular clinical session through a single draw of venous blood, segments of data taken from the same subject's data must generate similar predictions. The intra-subject loss of smoothness $\mathcal{L}_{smooth}$ achieves this by punishing the variance of predictions across segments for each individual in the batch:

$$\mathcal{L}_{smooth} = \frac{1}{S} \sum_{s=1}^{S} \frac{1}{N_s} \sum_{i=1}^{N_s} \left(\hat{y}_i^{(s)} - \bar{\hat{y}}^{(s)} \right)^2, \tag{8}$$

where S is the number of unique subjects in the batch, N_s is the number of segments belonging to subject s, and $\bar{\hat{y}}^{(s)} = \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{y}_i^{(s)}$ is the mean predicted Hb for that subject. This term directly reduces sensitivity to local noise and transient motion artifacts that may affect individual 10-second windows differently.

Since both $\mathcal{L}_{opt}$ and $\mathcal{L}_{smooth}$ are computed via the same forward pass as $\mathcal{L}_{MSE}$, neither introduces any new learnable parameters or any substantial extra computational costs relative to the underlying neural network. PCNN achieves regularization by constraining the search space of possible solutions to biologically meaningful subsets, thus going further than mere empirical loss minimization in the process of optimization.

3.6 Feature Fusion, Inference, and Implementation Details

Deep temporal representations from the CNN, BiLSTM, and Attention backbone are combined with handcrafted signal features and subject-level conditioning information to create a unified feature vector. Before regression, all feature components are standardized to ensure a balanced contribution during learning. A lightweight fully connected regression head maps the combined representation to segment-level hemoglobin predictions, allowing for easy integration of different information sources without adding complexity to the architecture.

Since hemoglobin is measured at the subject level, final predictions come from averaging segment-level estimates across all segments for the same subject. This averaging method reduces the impact of temporary noise and local artifacts while aligning inference with the clinical evaluation process. We train the model using the Adam optimizer with fixed learning rates for all experiments to allow fair comparisons between optimization strategies. Regularization and optimization hyperparameters remain constant unless specifically optimized using external metaheuristic methods. All models use PyTorch and are trained on a subject-independent split to prevent information leakage.

4 Results

A subject-independent evaluation protocol assesses the performance of all models. Predictions from different segments of the same subject are averaged to create a single subject-level hemoglobin (Hb) estimate, which reflects a real situation where one Hb value corresponds to one individual. Model performance is reported using Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). These metrics are suitable for analysis focused on optimization and highlight larger prediction errors. All models undergo identical preprocessing, segmentation, and training setups to ensure a fair comparison.

4.1 Comparison with Baseline Deep Learning Models

Table 1 summarizes the performance of baseline deep learning models, including BiLSTM, CNN-BiLSTM, BiGRU, and BiGRU with attention. Traditional machine-learning models are included for comparison.

Pure recurrent models like BiLSTM and BiGRU show relatively high error, suggesting that temporal modeling alone is not enough for strong Hb estimation from noisy PPG signals. Adding convolutional preprocessing greatly enhances performance. The CNN-BiLSTM model reduces the subject-level RMSE from 1.48 g/dL (BiLSTM) to 1.20 g/dL by extracting local morphological features before sequence modeling. Using attention further boosts prediction accuracy. The BiGRU + Attention model achieves the lowest RMSE among baseline models at 1.09 g/dL. This shows the advantages of relevance weighting for minimizing noisy or distorted waveform segments.

In contrast, traditional machine-learning models like MLP and ReliefF + CART have much higher RMSE values (1.53 g/dL and 1.47 g/dL, respectively). This reflects their limited capacity to capture temporal links and the dynamics of multi-wavelength PPG. These results affirm that effective Hb estimation needs both morphological feature extraction and temporal modeling, and that attention mechanisms offer more robustness in noisy settings.

Table 1. Subject-level performance comparison of baseline deep learning models.

Model	**MSE**	**RMSE (g/dL)**
BiLSTM	2.19	1.48
CNN–BiLSTM	1.44	1.20
BiGRU	1.93	1.39
BiGRU + Attention	1.19	1.09
MLP Baseline	2.34	1.53
ReliefF + CART	2.16	1.47

4.2 Performance of Additional Deep Learning Architectures

To provide a complete comparison, additional deep learning architectures, including 1D-CNN, VGG1D, ResNet1D, and autoencoder-based hybrids, are evaluated at the subject level (Table 2).

Convolution-only models capture local waveform shape but struggle to model long-range time dependencies inherent to PPG signals. Among these architectures, ResNet1D achieves the lowest RMSE (1.28 g/dL) due to better gradient flow from residual connections. However, all convolution-only models still perform worse than hybrid CNN-recurrent architectures.

Autoencoder-based models show poor performance, with AE + MLP producing an RMSE of 3.97 g/dL and AE + BiLSTM yielding 1.72 g/dL. This decline suggests that unsupervised latent representations do not align with hemoglobin-related physiological targets, highlighting the need for supervised time learning. Overall, these results confirm that precise Hb estimation requires the joint modeling of waveform shape and time dynamics.

Table 2. Performance of additional deep learning architectures at the subject level.

Model	**MSE**	**RMSE (g/dL)**
1D-CNN	1.85	1.36
VGG1D	1.85	1.36
ResNet1D	1.64	1.28
AE + MLP	15.76	3.97
AE + BiLSTM	2.96	1.72

4.3 Metaheuristic Optimization Analysis

We evaluate metaheuristic optimization techniques to see if external optimization can enhance model performance in noisy conditions. Specifically, we apply Grey Wolf Optimization (GWO) to a GRU-based model to help with feature selection and parameter refinement.

The GWO-optimized GRU model achieves a subject-level RMSE of 1.49 g/dL, which does not surpass attention-enhanced deep learning baselines. This result indicates that while GWO can efficiently explore the search space, blind metaheuristic optimization offers limited advantages when used externally on recurrent architectures that already manage feature selection and temporal modeling well. These findings encourage us to integrate optimization directly into the learning objective instead of treating it as an afterthought.

4.4 Physics-Guided Optimization via Physiologically-Constrained Neural Networks

To tackle the issues of unconstrained and externally optimized learning, we reformulate the problem as a constrained optimization task using a Physiologically-Constrained Neural Network (PCNN). Besides minimizing data fidelity loss, the optimization goal includes (i) an optical consistency constraint that enforces physically plausible monotonic relationships between pulse amplitude and Hb, and (ii) an intra-subject smoothness constraint that penalizes prediction variance across multiple segments from the same subject.

We perform an ablation analysis to measure the effect of each constraint. The baseline neural network, trained without constraints, achieves an MSE of 1.082 (RMSE 1.04). Adding the optical consistency constraint lowers the error to an MSE of 0.950 (RMSE 0.98), showing that physics-guided regularization enhances optimization results. By incorporating both optical and smoothness constraints, we achieve an MSE of 1.033 (RMSE 1.02), reflecting a careful balance between numerical accuracy and prediction stability.

Table 3. Ablation study of physics-guided optimization using Physiologically-Constrained Neural Networks (PCNN).

Model Variant	**MSE**	**RMSE (g/dL)**
Baseline NN (No Constraints)	1.082	1.040
PCNN (Optical Constraint Only)	0.950	0.975
PCNN (Optical + Smoothness Constraints)	1.033	1.016

Notably, the optical-only PCNN achieves the lowest error, while the full PCNN produces smoother and more consistent subject-level predictions. These results indicate that constraint-aware optimization effectively restricts the feasible solution space to physiologically valid regions, improving convergence behavior and robustness under noisy .

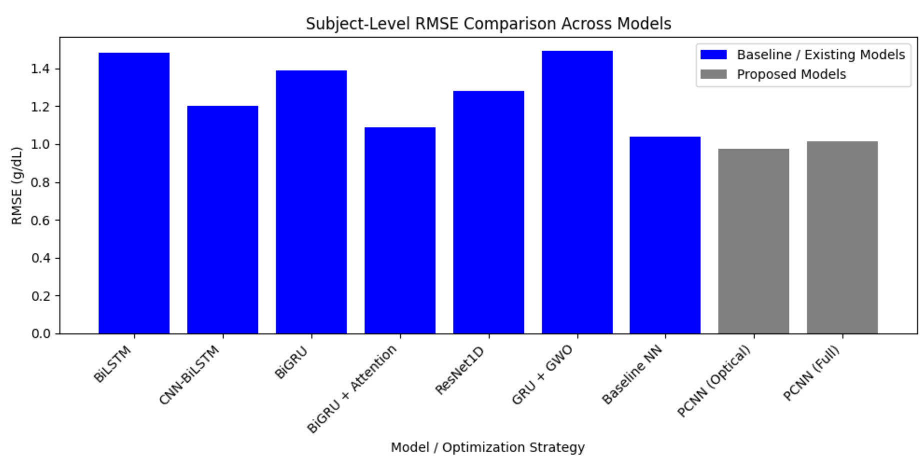

Fig. 2. Subject-level RMSE comparison of baseline models and proposed optimization-guided approaches. Proposed models are highlighted in gray.n

5 Discussion

This work examined hemoglobin estimation from photoplethysmography (PPG) signals as a challenge of learning under noise and implicit constraints. It emphasized how optimization strategies can improve model robustness. The results show that architectural design, constraint integration, and optimization formulation each play important roles in determining solution quality.

Baseline comparisons confirm that accurate prediction needs both morphological feature extraction and temporal modeling. Hybrid CNN and recurrent architectures consistently outperform models that use only pure recurrent or convolution layers. Attention mechanisms further improve the robustness by underlining the informative waveform segments and suppress noises. These results indicate that temporal modeling alone is not enough for noisy PPG signals.

The baseline of metaheuristic optimization reveals its limitation to be external, blind optimization. While GWO is effective in exploring the search space, the GWO-optimized GRU model is not competitive to attention-based architectures. This suggests a law of diminishing returns when applying metaheuristics in an unconstrained manner at the problem domain. This observation encourages integrating optimization directly into the learning objective.

On the other hand, the proposed Physiologically-Constrained Neural Network (PCNN) demonstrates that optimization with constraints significantly improves the learning process. The addition of an optical consistency constraint decreases the prediction error, and another constraint on smoothness for intra-subjects increases stability. This comes at the cost of increased error to some extent, due to the trade-off between bias and variance. These phenomena arise from restricting the space of solutions to physiologically valid areas instead of higher model complexity. This framework, though used for a biomedical case study, can be used to solve other constrained optimization problems with noisy data and weak supervision.

One of the important limitations of the current model is the management of the confounding optical factors like melanin and LED driving current fluctuations. Melanin demonstrates considerable absorption from 660 nm to 730 nm that falls under the range where hemoglobin is sensitive. This factor can create a bias among people having high levels of skin pigmentation. The z-score normalization and the application of the multi-wavelength ratios feature reduce this bias to some extent. However, Fitzpatrick skin type or the melanin index was not considered in the data set. Similarly, there is no explicit handling of the LED intensity changes during each session; however, it can be handled via amplitude normalization.

6 Conclusion

This paper presented a physics-guided optimization framework for the non-invasive estimation of hemoglobin based on multi-wavelength photoplethysmography signals. By directly incorporating physiological constraints in the learning objective through a Physiologically-Constrained Neural Network (PCNN), the current approach goes beyond purely data-driven regression and instead enforces physically meaningful consistency during optimization. In addition, the role of external metaheuristic optimization has also been explored. This juxtaposition of heuristic-driven and constraint-aware learning strategies occurs within a unified deep temporal backbone.

Results from experiments have proved physics-guided optimization to result in a more stable and generalized prediction compared to unconstrained and heuristic-based approaches. This is because the method delivers interpretable results and is in agreement with clinical protocol evaluations. The presented framework has demonstrated the importance of exploiting prior knowledge in the course of the optimization process, as opposed to altering the network structure. This provides a promising direction for the design of a reliable method of non-invasive physiological monitoring.

References

1. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement* **28**(3), R1–R39 (2007)
2. Sun, Y., Thakor, N.V.: Photoplethysmography revisited: From contact to noncontact, from point to imaging. *IEEE Transactions on Biomedical Engineering* **63**(3), 463–477 (2016)
3. Zong, C., et al.: Deep learning for physiological signal analysis: A survey. *IEEE Reviews in Biomedical Engineering* **14**, 239–256 (2021)
4. Hannun, A.Y., et al.: Cardiologist-level arrhythmia detection using deep neural networks. *Nature Medicine* **25**(1), 65–69 (2019)
5. Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L.: Physics-informed machine learning. *Nature Reviews Physics* **3**(6), 422–440 (2021)
6. Nirupa, K., Kumar, R.: Non-invasive hemoglobin estimation using optical methods. *Biomedical Signal Processing and Control* **10**, 45–52 (2014)
7. Kraitl, J., Eilebrecht, B., Willkomm, J., Leonhardt, S.: Optical detection of blood components under varying sensor conditions. *Medical Engineering & Physics* **33**(6), 770–776 (2011)
8. Bailey, S., Johnston, D., Finn, A.: Low-power pulse oximeter design for mobile health applications. *IEEE Transactions on Biomedical Engineering* **55**(4), 1284–1291 (2008)
9. Kyriacou, P.A., Moye, A.R., Choi, D.M., Langford, R.M.: Investigation of reflectance photoplethysmography for clinical monitoring. *Physiological Measurement* **23**(3), 507–517 (2002)
10. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement* **28**(3), R1–R39 (2007)
11. Johansson, A.: Neural network-based extraction of respiratory rate from photoplethysmographic signals. *Medical & Biological Engineering & Computing* **41**(3), 242–248 (2003)
12. Usman, S., Khalil, M., Khan, M.: Second-derivative photoplethysmography for vascular assessment. *Computers in Biology and Medicine* **39**(5), 430–437 (2009)
13. Chen, W., Chang, C., Lin, Y.: Assessment of endothelial function using second derivative PPG features. *IEEE Transactions on Biomedical Engineering* **58**(5), 1456–1463 (2011)

14. Ding, X., Yan, B., Zhang, Y.: Non-invasive hemoglobin estimation using PCA and neural networks. *Biomedical Optics Express* **5**(10), 3505–3517 (2014)
15. Iwasaka, M., Ueno, S.: Magnetic field effects on red blood cell optical properties. *Bioelectromagnetics* **22**(8), 509–514 (2001)
16. Jay, G.D., Rowe, A.D., Sutin, J.: Reflectance-based hemoglobin measurement using fiber-optic spectroscopy. *Journal of Biomedical Optics* **12**(3), 034012 (2007)
17. Zhang, Y., Wang, L., Zhao, Y.: Near-infrared spectroscopy for neonatal hemoglobin monitoring. *Biomedical Optics Express* **2**(7), 2010–2020 (2011)
18. Head, C., Pierson, R.: Broad-spectrum NIR estimation of hemoglobin and oxyhemoglobin. *Journal of Biomedical Optics* **15**(2), 027004 (2010)
19. Timm, U., Myllylä, R.: Optical hemoglobin sensing using synchronous detection techniques. *Sensors and Actuators A* **150**(2), 218–224 (2009)
20. Binns, R., Dempsey, J.: Frequency-domain analysis of photoplethysmographic signals. *Medical & Biological Engineering & Computing* **33**(4), 501–506 (1995)
21. Kavsaoğlu, A.R., Polat, K., Bozkurt, M.R.: Frequency-based analysis of PPG signals for physiological estimation. *Biomedical Signal Processing and Control* **10**, 401–408 (2014)
22. Kavsaoğlu, A.R., Polat, K., Bozkurt, M.R.: Adaptive segmentation of PPG signals for robust feature extraction. *Expert Systems with Applications* **41**(5), 2019–2028 (2014)
23. Kira, K., Rendell, L.: A practical approach to feature selection. In: *Proceedings of the Ninth International Conference on Machine Learning*, pp. 249–256 (1992)
24. Kononenko, I.: Estimating attributes: Analysis and extensions of RELIEF. In: *European Conference on Machine Learning*, pp. 171–182 (1997)
25. Das, R., Turkoglu, I., Sengur, A.: An ensemble-based approach for biomedical signal classification. *Pattern Recognition Letters* **29**(8), 1199–1205 (2008)
26. Das, R., Sengur, A.: Hybrid classification of physiological signals using ensemble learning. *Expert Systems with Applications* **37**(8), 5619–5624 (2010)
27. Zong, C., Wong, J., Zhou, Y.: Deep learning for physiological signal analysis: A survey. *IEEE Reviews in Biomedical Engineering* **14**, 239–256 (2021)
28. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., et al.: Cardiologist-level arrhythmia detection using deep neural networks. *Nature Medicine* **25**(1), 65–69 (2019)
29. Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L.: Physics-informed machine learning. *Nature Reviews Physics* **3**(6), 422–440 (2021)