IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 February 2024, accepted 20 April 2024, date of publication 8 May 2024, date of current version 22 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3398353

RESEARCH ARTICLE

Fingertip Video Dataset for Non-Invasive Diagnosis of Anemia Using ResNet-18 Classifier

HUMERA SABIR[1], KIFAYAT ULLAH KHAN[2], OMER ISHAQ[3], ABDULWAHAB ALAZEB[4], HANAN ALJUAID[5], ASAAD ALGARNI[6], AND JEONGMIN PARK[7]

[1]School of Computing, National University of Computer and Emerging Sciences (FAST-NUCES), Islamabad 44000, Pakistan
[2]College of Accounting, Finance and Economics, Birmingham City Business School, Birmingham City University, B5 5JU Birmingham, U.K.
[3]Deep Learning Expert, Islamabad, Pakistan
[4]Department of Computer Science, College of Computer Science and Information System, Najran University, Najran 55461, Saudi Arabia
[5]Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), P.O. Box 84428, Riyadh 11671, Saudi Arabia
[6]Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia
[7]Department of Computer Engineering, Tech University of Korea, Siheung-si, Gyeonggi-do 15073, South Korea

Corresponding author: Jeongmin Park (jmpark@tukorea.ac.kr)

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2024-RS-2022-00156326) supervised by the IITP (Institute of Information Communications Technology Planning Evaluation). The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work under the Research Group Funding program grant code (NU/GP/SERC/13/18). Also, this research was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R54), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number "NBU-FFR-2024-231-04".

ABSTRACT Hemoglobin is the iron containing protein in red blood cells which carries oxygen from lungs to rest of the body tissues. Accurate measurement of hemoglobin is essential for diagnosing anemia, a condition characterized by a deficiency of red blood cells. This measurement is particularly vital before initiating blood transfusions for thalassemia patients. Non-invasive estimation of hemoglobin levels can be achieved through photoplethysmography (PPG)-based methods. PPG is an optical method to measure blood volume changes in successive heart beats. PPG signals can be obtained from fingertip videos using a light source and a photodetector. SmartphonePPG utilizes a smartphone's flashlight as a light source and its camera as a photodetector to acquire PPG signals, offering an affordable and portable point-of-care tool. Despite the ubiquity of smartphones, signals from their cameras often contain noise, making feature selection from PPG characteristics challenging. While PPG-based methods are invaluable, the lack of real-world datasets poses a significant challenge in maximizing the benefits of PPG technology. In this paper, we introduce a dataset comprising 1-minute fingertip video recordings from 150 anemic patients, obtained using a smartphone's camera. The dataset, publicly accessible for research purposes (https://forms.gle/LB4qn81ZMqEuy3V27), covers an age range of 6 months to 32 years, with diverse hemoglobin values (4.3 gm/dL - 12.4 gm/dL). Utilizing this dataset, we propose a deep learning-based technique employing the ResNet-18 architecture to estimate hemoglobin levels. This approach eliminates the need for manual feature extraction and selection from PPG signals, overcoming a limitation in existing smartphone PPG-based hemoglobin estimation systems. Our model achieves a hemoglobin level estimation with an RMSE of 0.81-1.39 when compared with the gold standard laboratory method, Complete Blood Count (CBC) test reports. In contrast, HemaApp, a state-of-the-art research utilizing a machine learning-based classifier (SVM), yields an RMSE of 1.7 on our dataset. The accuracy and simplicity of our model position it as a promising alternative to existing non-invasive hemoglobin level estimation methods.

INDEX TERMS PPG signals, fingertip video dataset, hemoglobin estimation, non-invasive, computer vision, image processing, deep learning.

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja.

I. INTRODUCTION

Research and development for the improvement of healthcare and assistive systems for various other areas of life has been ever been on rise, for instance in CT scans [19], medical imaging and image processing [6], [11], [14], [17], [25], [42], [48], [49], [52], [56], using machine learning algorithms [4], [21], [29], [38], [40], [47], [50], [51], data compression [28], [30], laser-assisted surgery [35], video data processing [26], robots intelligence [27] and in similar areas [20], [31], [32], [46], [53]. In this regard, the research in neural networks and its advancements [55] has also been on continuous ascent. The aforementioned recent research efforts to improve the life standards, emphasize the possibilities of further exploration. In this regard, research in estimating Hemoglobin (Hb) is also of high importance.

Hb is the oxygen carrying protein molecule present in the blood. Normal Hb level (for males $\geqslant$ 13 gm/dL and for females $\geqslant$ 12 gm/dL) is necessary to supply oxygen to body organs. Hemoglobin abnormalities like sickle cell anaemia, iron deficiency anaemia, and thalassemia major are serious health issues [5]. Iron deficiency is a major cause of anemia and is more prevalent in the developing countries. It is common among the women of reproductive age group, pregnant women, and children under five years of age. According to WHO, 1.62 billion people suffer from anemia worldwide including 300,000 babies alone. According to [2], in a developing country like Pakistan, statistics show anemia prevalence of almost 51.3% of entire population and where more than one-fifth of women suffer from it. Considering its severity, Hb level estimation is necessary not only for anemic disorder but also for assessing the response of iron supplement treatments [43], [44], treatment of dengue fever [23], blood transfusion [3], and for the diagnoses of childhood malnutrition. For this purpose, laboratory test [1] is most common procedure which accurately estimates Hb level however, at the cost of time and money for a patient and significant financial investment for laboratory equipment.

Complete Blood Count (CBC) test of Hb level is referred as invasive approach which requires skin puncturing to draw around 3mL blood [43], [44]. It is the laboratory gold standard test for Hb level estimation. Whereas the test accurately measures Hb level, it is not comfortable for pregnant women, infants, elderly, and specially when a patient needs frequent assessments. Moreover, the traditional laboratory test demands specialized equipment for Hemoglobin estimation [33], and results are not readily available [22].

As an alternative, non-invasive procedures are gaining popularity which overcome the said issues of their counterpart. In this regard, there are many medical devices available in the market that use Electrocardiography (ECG) and Photoplethysmography (PPG) as underlying technology. ECG monitors the electric activity of the heart whereas PPG is an optical method to measure the blood volume changes [45]. Out of these two techniques, the non-invasive devices to measure Hb level uses PPG technology. In this way, PPG technology is specially beneficial in critical health care units and for regular monitoring [3]. Besides their advantages, non-invasive PPG-based medical devices are often expensive and also patients are reluctant to carry them all the time [12]. Additionally, advancements in non-invasive Hb estimation methods include the utilization of conjunctiva [8], lip mucous [34] and fingernail bed images [7], [33] offering further possibilities for accessible and convenient diagnostics.

PPG technology involves throwing different light intensities on skin, and then photo detector of the corresponding non-invasive device observes the amount of reflection and absorption. Blood is pumped to the arteries during a systole and hence increases their volume. In return, the arteries allow lesser amount of light to pass through them and more light is bounced back to the photo detector and vice versa during a diastole. In this way, the device captures the difference in light intensities and convert them into a PPG signal. This phenomenon of PPG signal generation can nicely be executed using camera and flashlight of a smart phone [13]. This approach entails placing fingertip before the camera to measure PPG signal as shown in Figure 1. The camera records video of the fingertip for certain amount of time and signals are obtained by averaging the pixel intensity values of video frames.

Since the prevalence of smart phones all around, therefore many researchers are using it for Hb level estimation. HemaApp [43], [44] is the pioneering work in this regard. It uses three different hardware embodiment as an add-on to collect fingertip video data which is then processed by SVM, a machine learning model, to estimate Hb level. It estimates Hb level with the mean error of 1.56 gm/dL,1.44 gm/dL and 1.26 gm/dL for each embodiment. It is evaluated on 31 subjects. Based on these encouraging results, the authors extend HemaApp [43] using only the built-in camera and white LED through a manual color channel gain balancing technique. To extract the features from PPG, a series of signal processing steps are applied followed by linear regression to estimate Hb level on a dataset of 32 participants. The results were tested against Mosimo Pronto 7 and RMSE of 1.27 is reported with a pearson correlation value R as 0.62. In another research [22], the authors use HSV color space for Hemoglobin estimation. They used data of 60 participants. RGB values from video frames are converted to HSV color space, histogram for each frame is calculated, and averaged to make one observation which act as a feature vector. HSV space is used instead RGB color space [16] as it can separate luminance from chromaticity and is more intuitive. Partial Least Square (PLS) regression model is used and R^2 of 0.95 is reported. Reference [15] apply neural networks over the PPG signals/data of 75 participant. It receives 200 frames as input for all three channels, along with age and gender information as features. The data set is divided into three sets each containing either Red, Blue, or Green frames along with their respective demographic information. The data set consisting of Red channel values shown best results with R2= 0.941. Reference [24] used PPG characteristics features and their derivatives for Hemoglobin estimation. They collected data from 127 participants using IR Plethysmograph transducer.

They used seven features from PPG signals, along with their 1st and 2nd derivative as input to linear regression model. In this way, they obtained MAE, MSE and RMSE values of 0.737 gm/dL, 0.738 gm/dL and 0.86 gm/dL respectively. Although the mentioned studies produce promising results for non-invasive hemoglobin estimation, several of these approaches require hardware adons and manual extraction of PPG signals from the video and selection of features from PPG signals. The manual feature selection process requires a lot of human effort, domain specific knowledge, and signal processing skills.

In this paper, we propose a deep learning-based hemoglobin estimation technique using video of fingertip obtained using smartphone camera and flashlight. We present the overview of the proposed solution in Figure 2. The proposed approach uses deep residual network (Resnet-18) [18] to estimate hemoglobin level. The advantage of deep learning based model is that it omits the feature engineering step. The video data is obtained from fingertip and video frames are stacked along z-axis to get a single image with consecutive video frames concatenated. The resultant image is passed through a Resnet-18 network with gold standard hemoglobin level as ground truth value. Mean Square Error is used as a loss function and ReLU as activation function. We use stochastic gradient decent optimizer to update the model weights. We collected data from 150 anemic patients having a diverse age range (0.5 years- 32 years) and hemoglobin level as low as 4.3 gm/dL and as high as 12.4 gm/dL. We release this dataset for research purposes only so that more innovative studies can be conducted. We also release our code[1] for development and experimental purposes. Experiments with different video lengths of 15 seconds and 1 minute are performed in this study to find the optimal length that can be used to estimate hb level. The proposed approach estimates hemoglobin level with an RMSE of 1.39 while using 1-minute unmodified video of fingertip. This result outperforms the results obtained from state-of-the-art work by [43] and [44] which receives RMSE of 1.7 on our dataset.

II. LITERATURE REVIEW

Usage of smartphone offers a lot of advantages over the traditional methods of hemoglobin estimation. Hence, smartphone based mobile health methods are gaining a lot of attention form researchers around the globe. Researchers are increasingly harnessing the power of smartphones, often coupled with specialized hardware components, to advance non-invasive methods for hemoglobin estimation [10], [12], [44].

A traditional smartphone based Hb estimation method involves a number of steps as shown in Figure 1. The very first step is to record a video, typically of one minute duration from the fingertip of the subject. Next the recorded video is converted into RGB frames. Averaging the pixel intensity values from the consecutive frames gives the signal values. These signals are referred to as photoplethysmography (PPG) signals. The process of converting video frames into PPG signals is of great importance and can affect all the subsequent steps involved in the process of estimation. A number of image processing steps are used to obtain high quality signals such as selecting a Region of Interest(RoI) in the frame and normalizing the pixel intensity values to remove the noisy frames. These signals are further processed to select PPG characteristic features from PPG signals, at this stage different signal processing techniques are used to select the PPG features. Some popular signal processing techniques used for PPG processing are Fast Fourier Transformation (FFT), High Pass filter, and 1st and 2nd derivative of the PPG signals. There exists a number of studies which are solely focused on PPG characteristics to find how each characteristic is useful and what features are reflected in them. PPG features are used as input to a prediction model, usually a machine learning algorithm or any other mathematical model, that converts the signal data into the Hb value. Some studies also utilize the use of deep learning approaches to convert the Video data into Hb level. This section describes in details the existing works which are being done on PPG signal extraction process and on the uses of different types of algorithms to convert the PPG characteristics features to estimate Hb level.

A. PPG SIGNAL EXTRACTION FROM VIDEO DATA

There are two settings in which the PPG signals are recorded namely Transmission mode and Reflection mode. In transmission mode, the photo detector and light source are placed in front of each other and the body site is placed in-between the both. Transmission mode PPG setting is most common in the medical devices such as finger probs. Taking into account the design of smartphone camera and flashlight, the reflection mode of PPG acquisition is used where the light source and photo detector are placed side by side.

Reference [13] study the signal obtained from Pulse Oximeter and using different models of smartphones side by side to study the differences in signals acquired using a smartphone setting and medical devices. The study proposed a threshold based on intensity value to classify frames into noisy and proper frames. The difference in color saturation and color ratio due to the usage of different smartphone models and different finger's pressure is studied. It is reported that the reflection mode usually result into an inverted PPG signal which needs to be further processed to obtain correct PPG signal.

Selection of Region of interest (RoI) is another important aspect during signal extraction from video frames as identified by [36]. The limited range of camera sensors result into some problem such as color saturation and cut-off distortions in the fingertip video [37]. Due to a direct contact of finger with the camera often times, Green and Blue channel variations are difficult to detect as compared to that of Red

[1] https://github.com/IntelliHb/Fingertip-Video-Dataset-and-its-Deep-Learning-based-Mining-for-Hemoglobin-Estimation.

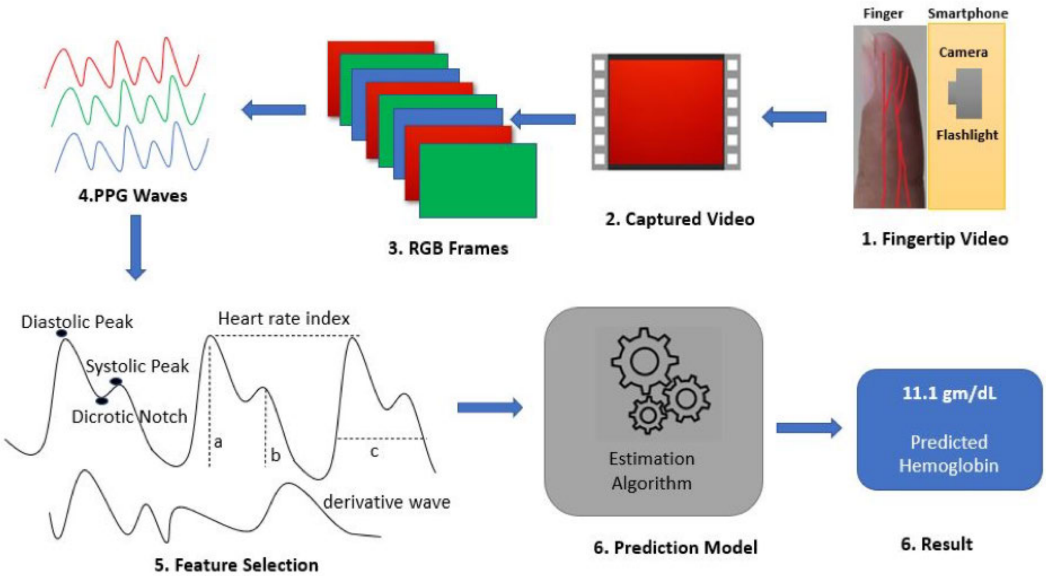

FIGURE 1. Steps involved in smartphone-PPG Based Hemoglobin estimation methods: (1) Finger is placed on smartphone covering the camera and flashlight and (2) video is recorded. (3) After the video recording the RGB frames are extracted from the video. (4) Averaging the Red, Green and Blue intensity values in consecutive frames converts the frames into PPG signal waves. (5) Features are extracted from the obtained PPG signals using various signal extraction techniques. (6) these features serve as input to a prediction model which converts the features to the Hemoglobin level.

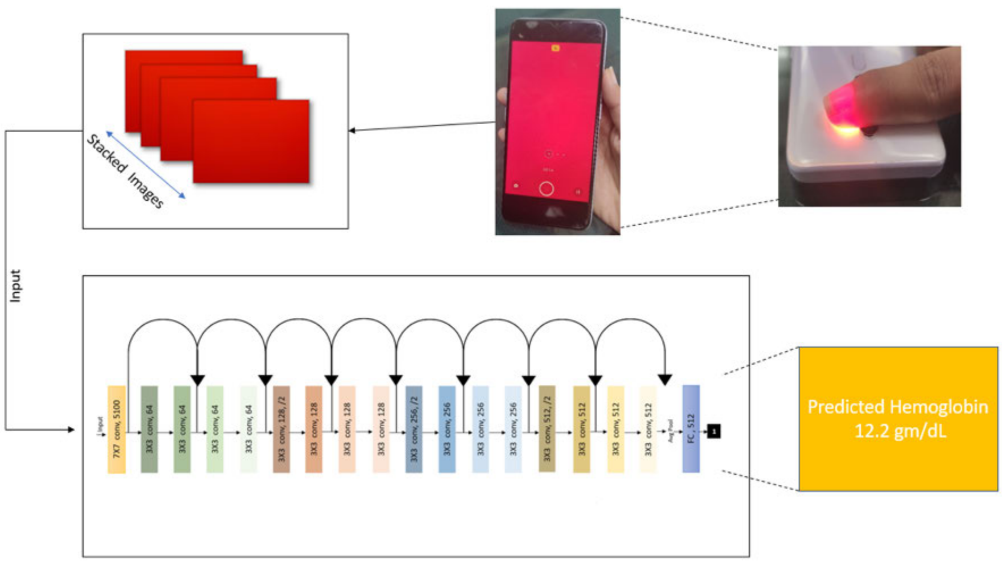

FIGURE 2. Overview of our proposed system: Fingertip video is converted to RGB frames and stacked together which serve as an input to a Resnet-18 Model that converts it to the Hb Level.

channel. To address the problem, [37] proposed an adaptive RoI selection mechanism based on the intensity values of RGB channels in the frames. The method is evaluated against an FDA approved device and it was reported that the proposed method can help minimize the noise in the PPG signals obtained from smartphone camera.

The quality of PPG signals obtained from smartphone sensors is subject to a number of factors such as skin tone, skin temperature, and measurement environment [9]. The mentioned factors can result into some artifacts which can affect the signal quality and feature selection from PPG signals. The peaks and troughs in the PPG signals are representative of different factors such as heart rate, blood flow etc. Reference [9] studies these relations and shows how 1st and 2nd derivative of these PPG signals can help detect the peaks and troughs.

PPG signals are sensitive to motion artifacts and light conditioning. Hence, [54] propose a signal evaluation method

TABLE 1. Summary of the previous studies on hemoglobin estimation using fingertip video.

Reference	Input Features	Algorithm	Reported Results
[44]	PPG	SVM	MSE=1.56 gm/dL
[43]	PPG	Regression	RMSE=1.27 gm/dL
[1]	Statistical Features	Regression 4	-
[22]	HSV Histograms	PLS Regression	R2=0.95
[15]	PPG, Age, Gender	ANN	R2=0.94
[24]	PPG,1st and 2nd derivative	Regression	R2=0.76

based on frequency domain and time series parameters. The study suggested the usage of Fast Fourier Transformation (FFT) to convert the time series signal into frequency domain to eliminate the influence of amplitude differences due to motion artifacts and other environmental conditions. The study shows a method to evaluate the quality of the PPG signals.

B. PPG SIGNALS AND MACHINE LEARNING BASED SOLUTIONS

Table 1 summarizes various machine learning based solutions used in the past. HemaApp by [44] uses video data obtained using different light sources of varying intensity using a hardware ad-on and a front facing camera. The proposed algorithm performs a blood color analysis. For this purpose, the ratio of minimum and maximum intensity value is obtained that provides a measure of absorption due to different component on the blood such as plasma and hemoglobin. PPG signals are extracted from video by center cropping the images and averaging the intensity. A high pass filter with cut off frequency 0.5Hz is used to remove fluctuations due to breathing along with an FFT to extract dominant peaks. For each peak extracted, the ratio of peak and trough is calculated. Five different features are extracted from the resultant PPG signals and SVM regression model is trained and tested using leaveone-subject out validation. Video samples from 31 subjects were used in this study.

In its extension, [43] propose a technique that requires only smartphone camera and flashlight using a manual color gain setting approach to estimate the hemoglobin concentration using fingertip video data. The study performs an empirical analysis using different gain settings. Linear regression is applied on PPG signals obtained by averaging the pixel intensities from 1 minute long video. [1] study PPG signals along with some other statistical information such as median, mode, maximum, minimum, and standard deviation for each Red, Green, and Blue channels to estimate the hemoglobin level. Singular value decomposition (SVD) is used to remove the insignificant data for each frame. Multiple regression and classification methods were used such as linear regression, ADtree and J48 ADTree classification. Linear regression shows a positive correlation of red intensity and negative correlation with green intensity with the Hemoglobin values. These statistical features show weak correlation. ADTree and J48 ADTree classification methods are used to classify the Hb into low and higher classes. ADTree shows better classification results but not promising enough to be used as a diagnostic tool.

Reference [22] uses HSV color space instead of RGB for Hemoglobin estimation. The study involved data from 30 sickle cell patients and 30 other subjects having different physiological issues. RGB values from frames are converted to HSV color space and histogram for each frame is calculated and averaged to make one observation which acts as a feature vector. The feature vector is further normalized with 0 mean and standard deviation of 1. Partial Least Square (PLS) method is applied on the input feature vector. PLS component was set to 10 and R2 of 0.95 is reported.

Reference [24] proposed an approach that uses PPG signals and its derivatives for Hemoglobin estimation using machine learning. Dataset collection was done in Sree Abirami Hospitals, Coimbatore from 127 patients. The IR Plethysmograph transducer was used for signal collection and the sensor was placed in the forefinger of the subject. Wavelet transform was used for initial processing of the signal and noise removal. Seven time domain features from PPG signals, 1st and 2nd derivative of the PPG signal were extracted and used as features. These features were used as input to a linear regression model. For evaluating the performance, Mean Absolute Error (MAE), Mean Squared error (MSE) and Root Mean Square (RMSE) were used and values of 0.737 gm/dL, 0.738 gm/dL and 0.86 gm/dL were reported. R2 value of 0.768 was obtained.

C. PPG SIGNALS AND DEEP LEARNING BASED SOLUTIONS

In the realm of PPG signals and Deep Learning-based solutions, recent contributions have demonstrated the efficacy of neural network models in predicting hemoglobin levels.Among the latest work is a solution proposed by [41]. The study involves the exploration of Photoplethysmography (PPG) for hemoglobin prediction, employing a smartphone-centric methodology. The approach involves converting color channels of video frames into singular values by adding all three channel values, utilizing a dataset derived from PPG recordings of 63 patients. The study incorporates MLP, FCN, and ResNet in its experimental framework, with ResNet undergoing training for 200 epochs. Similarly [39] presents a novel non-invasive, real-time technique for hemoglobin measurement using smartphone-based Photoplethysmography (PPG) signals derived from fingertip video. The study involved 88 participants, and the video data were collected using a smartphone's primary camera and a near-infrared illumination source. The PPG signals, generated from the video, underwent a Butterworth bandpass to eliminate noise and motion artifacts, resulting in 34 time-domain characteristics. The subsequent step involved the development of an artificial neural network (ANN) model for accurate hemoglobin level estimation. The proposed approach achieved a coefficient of determination (R2) of 0.909, indicating a strong correlation between the hemoglobin level and the derived characteristics.

Reference [15] uses neural network model to predict hemoglobin level from fingertip video data. For each channel, the frame was divided into one hundred blocks. For each block, a mean value was calculated resulting into 300 values for all three channels. 200 frames for all three channels, age and gender information are used as features. The data set was divided into three sets each containing all red, blue, and green frames along with age and gender information. Each data instance was fed in to Neural Network with target value set to the ground truth obtained from gold standard hemoglobin measurement. 70% of the data set was used to train the Neural Network. Remaining 30% data was divided into two equal testing and validation sets. 75 patients, 55 women and 20 men with an age range from 20 to 56 were studied having similar skin tone with no respiratory complaints. The data set consisting of red channel shows relatively better results with R2 of 0.941.

Reference [23] used Neural network based technique to estimate hemoglobin. Data was collected from 33 subjects using a custom built data acquisition card. Hemocue Hb-201TM device was used to collect the ground truth Hb values. 40 time domain features extracted from PPG signals, 1st derivative and 2nd derivative of the signals along with gender, age, weight, and height information are used as features. Multilayer perception (MLP) was used along with many other machine learning algorithms on the input data, where MLP achieved the RMSE of 0.31.

III. MATERIALS AND METHODS

In this section, we explain our dataset, its collection process, statistics, the pre-processing applied on it, and the architecture of our deep neural network model.

A. OUR DATASET

For our study, we have collected the data of 150 subjects from a Day-care centre for thalassemia patients, in Pakistan, over a course of three weeks. A Complete Blood Count (CBC) test result was recorded as a ground truth value of hemoglobin for each patient for whom 1 minute fingertip video was recorded. The data was collected from thalassemia patients before the blood transfusion process. Patients arrive at thalassemia center and a blood sample is collected to estimate the Hb levels. After the blood sample collection video was recorded from the patient's fingertip (preferably index finger). Later after the laboratory test, results were recorded manually.

B. VIDEO RECORDING PROCESS

We recorded 1 minute long video for each subject. During video recording process, each subject was in rest position and was asked to remain as still as possible. Figure 3 illustrates the process of covering the camera and flashlight to record the video.

Meanwhile a blood sample was taken by the hospital staff for gold standard Hemoglobin estimation by drawing blood. A CBC test was performed to get the full blood picture of each patient. The test result was recorded as a ground truth value along with the age and gender information of the patient. We used frame rate of 30 fps. In most of the cases, we collected the data from index finger but in some cases other fingers were used as well. Such as for children under 1 year of age, the diameter of the finger is very small which makes it difficult to cover the camera and flashlight completely, in such cases thumb or middle finger was used.

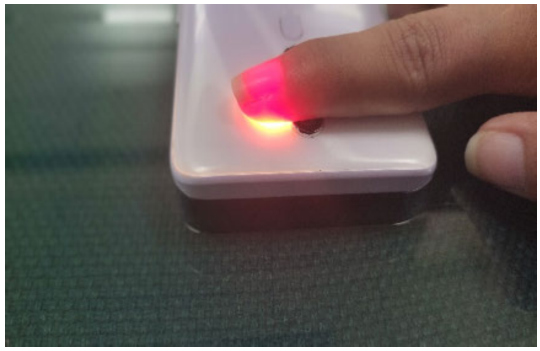

FIGURE 3. Video recording process: Finger is placed covering both camera and flashlight. A 1-minute video is recorded preferably form the index finger of the subject.

Following points were considered while collecting data from the patients.

- Finger should be clean and dry before capturing the video.
- The hand should be in rest position.
- Subject is either sitting or lying on the bed during the video capturing.
- The index finger should be preferred for sample collection, but other fingers may be used as well in case of an injury or other issues.
- Cover the smartphone camera and flashlight properly so that no ambient light can penetrate through.
- It was made sure that that the females participants does not have any Henna or nail polish applied on their fingers and nails.

C. DATA STATISTICS

Our collected data covers a wide ranges of Hb levels and ages. Mostly the data is collected from thalassemia major patients having different skin tones. Brown being the most common color tone and a few observation of lighter and darker color tones as well. There were 77 males and 73 females that were part of the study, as shown in Figure 4. The age ranges of all the patients are from 6 months to 32 years, displayed in Figure 6. Figure 5 shows observed Hb levels are from 4.3 gm/dL to 12.4 gm/dL. Most of the values in our dataset are between 7.0 gm/dL - 8.0 gm/dL. Figure 7 shows gender and age wise distribution of the values. It can be observed that the Hb range for female subjects in lesser than that of males in our dataset, as shows in Figure 8.

D. DATA PRE-PROCESSING AND ARCHITECTURE OF DEEP NEURAL NETWORK

The recorded videos are in MP4 format with 30 fps which gives 1800 frames per record. Multiple experiments are

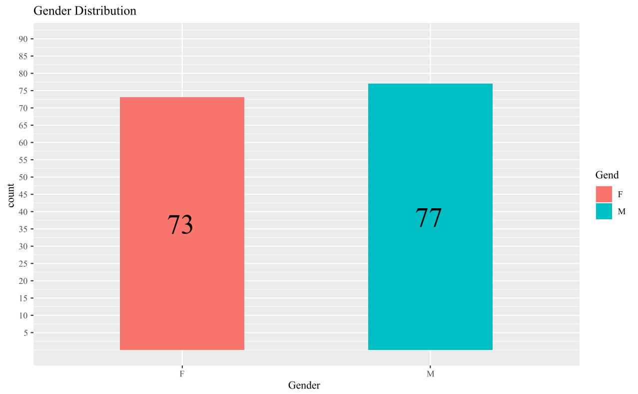

FIGURE 4. Gender distribution: Out of 150 participants, 77 were male and 73 were females.

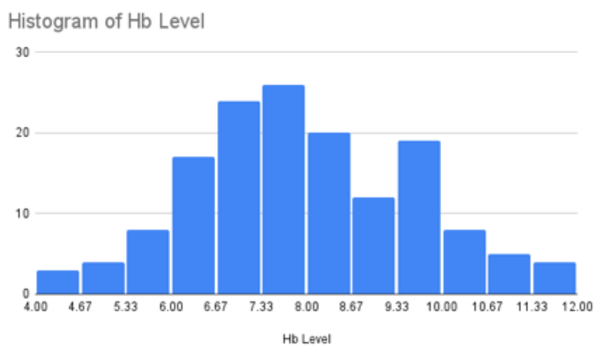

FIGURE 5. Hb distribution: The data is collected from thalassemia patients, the majority of Hb level of patients was between the range 7.0 gm/dL-8.0 gm/dL.

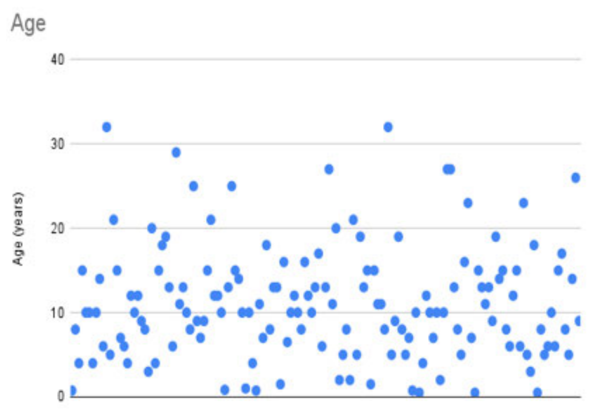

FIGURE 6. Age distribution: The minimum age was below 1 year (6 months) and maximum age was 32 years.

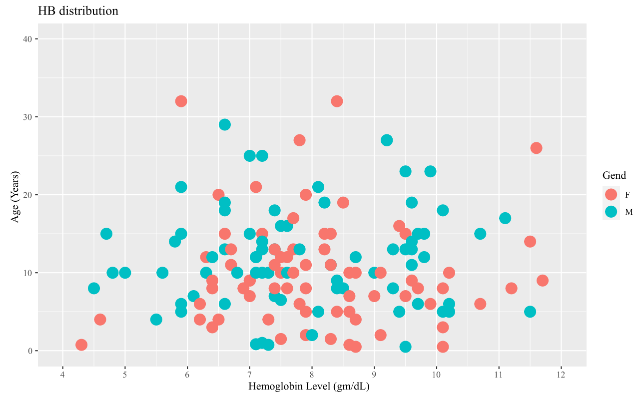

FIGURE 7. Hemoglobin distribution w.r.t Age and Gender: Hemoglobin level (x-axis) distribution of the data is shown w.r.t the age (y-axis). Gender information is represented using color.

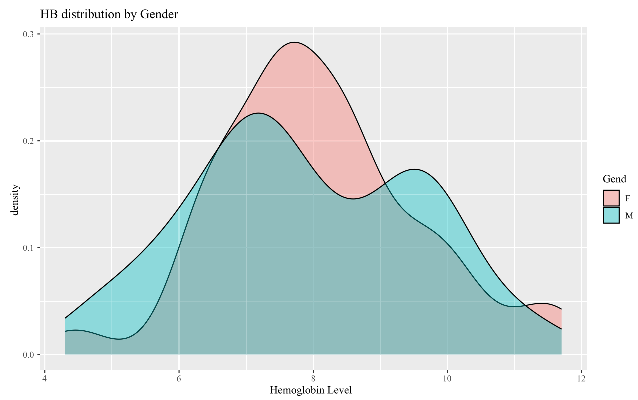

FIGURE 8. Hemoglobin distribution w.r.t Gender: The plot shows the density of Hb level w.r.t. gender. It is observed that the Hb level of male and females have a different distribution which is same for the Hb level of normal (non-thalassemia) people.

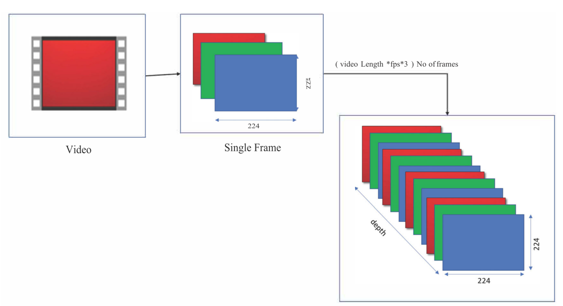

FIGURE 9. The process of conversion of Video to stacked Image: RGB Frames are extracted from the video and stacked along z-axis to obtain an image.

performed on different subsets of the dataset. For each experiment every video is converted into frames to obtain images having three channels Red, Green, and Blue. RGB channels obtained from consecutive video frames are stacked along z-axis to obtain a single image (Figure 9). The lateral size of the image after center cropping the frames is set to 224×224 pixels. The resultant image depth is frameRate × videoDuration ×3 (no. of channels).

The image serves as an input to a deep residual network Resnet-18 (Figure 10). ResNet, short for Residual Network was introduced by [18]. ResNet is a standard image classification network and we have used it for performing regression. We changed the last layer to convert the classification network for regression problem.

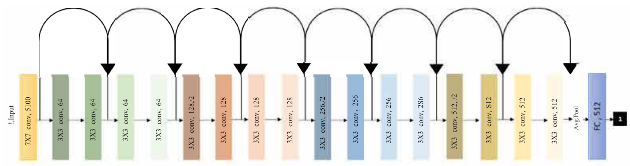

FIGURE 10. **Resnet-18 architecture: The standard Resnet-18 architecture was used by changing the last layer to contain a single neuron for regression.**

TABLE 2. **Training parameters used to train the proposed approach using Resnet-18 architecture.**

Sr. No	Parameter	Value
1	No of Epochs	50
2	Activation Function	ReLU
3	Loss Function	Mean Square Error
4	Learning Rate	0.001

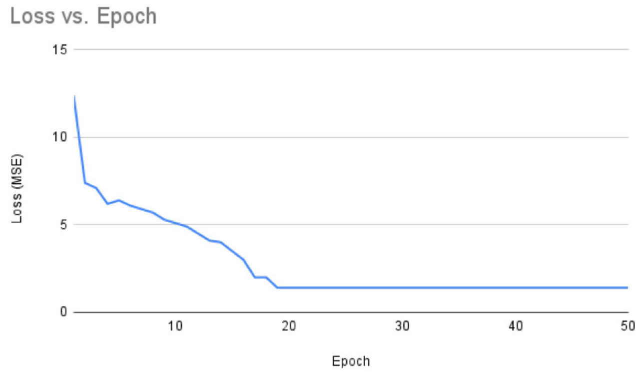

FIGURE 11. **Model training: In this experiment 450 Records of 15 seconds were used for training a Resnet-18 model. The model was trained for 50 epochs. On training data, the MSE loss of 2.1 was observed upon convergence.**

Table 2 shows summary of different parameters that were used during the experiments. In contrast to traditional methodologies for extracting Photoplethysmography (PPG) signals from video data, our proposed approach takes a unique path by directly leveraging the information encoded within the video frames. Instead of extracting PPG signals by averaging the pixel values, we stack the individual frames along the z-axis, forming a three-dimensional representation of the video data. This approach taps into the natural changes over time captured in the video frames, providing a straightforward and effective way to estimate hemoglobin levels. By skipping the usual PPG extraction steps, our method capitalizes on the inherent temporal variations present in the video frames, offering a direct and efficient means of estimating hemoglobin levels. This unique approach highlights the flexibility of our method, showing how it can make hemoglobin level estimation more straightforward by directly analyzing video data. This innovation underscores the adaptability of our methodology, illustrating its potential to streamline hemoglobin level estimation through a direct analysis of video data.

The varying intensity of pixel values in the consecutive frames of the video represent the PPG signal values. The consecutive frames are stacked along the Z -axis to get an image having depth equal to No-of-frames in a video. In this way, the varying intensity value of the signal can be represented inside a single image. Mean Square Error loss function is used in the experiment with Relu activation function. Gradient descent optimizer is used to update the model weights. The model was trained for 50 epochs where each epoch took approximately 100 minutes to complete on a 3070 RTX NVIDIA GPU Machine.

IV. EXPERIMENTAL EVALUATION

We performed experiments using a high-performance computer system. The system was equipped with 65 GB of DDR4 RAM and a 512 GB NVMe SSD. The system's GPU was an NVIDIA GeForce RTX 3070, which had 5888 CUDA cores, a base clock speed of 1500 MHz, and a boost clock speed of 1725 MHz. The RTX 3070 also had 8 GB of GDDR6 memory with a bandwidth of 448 GB/s. A number of experiments were performed with different sub sets of the data and with different video lengths. We used Redmi Note 10 (with a 48 MP back camera) to capture the videos. The following section describe the details of the experiments where we explain our three types of experiments.

A. EXPERIMENT TYPE 1

For the first experiment, we used all 150 video records. Since, each record is 1-minute long video with a frame rate of 30 frames per seconds, so we converted it into 3 sub clips of 15 seconds. In this way, we extracted 3 sub clips from every 1 minute video and mapped with the true values of hemoglobin against each sub-clip. Using these 3 sub-clips approach, 15 seconds from each record were intentionally dropped for the argument that it was time taken by each subject to correctly place the finger before the camera. Finally, we obtained 450 records in total. Each video consists of 450 Green, 450 Red and 450 Blue frames. All these frames were stacked together to get an image of depth 1350. 80% of the data was used for training and remaining 20% was used for testing for Resnet-18 network,which was trained for 50 epochs. The training loss of 2.3 was observed in this experiment as shown in Figure 11.

B. EXPERIMENT TYPE 2

We used 73 records of females only for this experiment. Each record was a 1-minute long video, consisting of 1800 Green, 1800 Red and 1800 Blue frames. First 100 frames from each channel were dropped for same reason as explained for Experiment Type 1. All the remaining frames were stacked together to get an image of depth 5100. Like experiment 1, we used 80% of the data training and remaining for testing for Resnet-18 network. We observed training loss of 2.9 in this experiment after training for 50 epochs, as demonstrated in Figure 12.

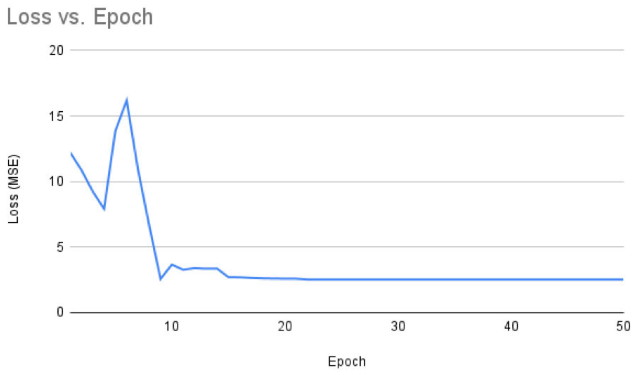

FIGURE 12. Model training: In this experiment 58 female records (80% of 73) were used for training a Resnet-18 model. The model was trained for 50 epochs. On training data the MSE loss of 2.9 was observed upon convergence.

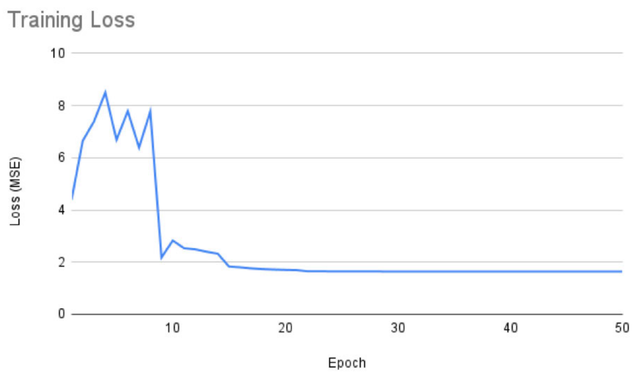

FIGURE 13. Model training: In this experiment 120 Records(80% of 150) were used for training a Resnet18 model. The model was trained for 50 epochs. On training data the MSE loss of 1.63 was observed upon convergence.

C. EXPERIMENT TYPE 3

In this experiment, we used the entire data set of 150 records, where each record is completely used of its original duration of 1 minute. We did cropping of first 100 frames for the reason stated for experiment type 1 and 2. Hence, whereas each video consists of 1800 Green, 1800 Red and 1800 Blue frames, the remaining frames after cropping were stacked together to get an image of depth 5100. By using complete dataset, we observed the most promising results. The model was trained with 80-20 and 90-10 splits and got training loss of 1.23 and 1.63 respectively. For the sake of brevity, we only present the plot for model training with 80-20 split, in Figure 13.

D. RESULTS

As stated above, we got most encouraging results using Experiment Type 3, which are presented in Figure 14 and Figure 15. The obtained RMSE lies between 0.81-1.39 against the ground truth. These promising results depict the superiority of our technique that it can effectively estimate the Hb level using fingertip video records. We also evaluated

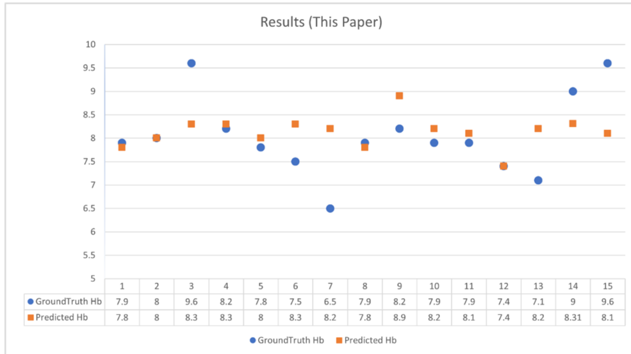

FIGURE 14. Proposed approach results on our data-set with 90-10 split: A comparison between actual Hb level and predicted Hb level is shown. Actual Hemoglobin level is the laboratory gold standard (Complete Blood Count) test result shown by circle (blue color). The predicted Hb level of proposed approach is shown by square (orange color). x-axis represents the record numbers and y-axis represents the Hb level in gm/dL. We observe the RMSE value of 0.81 in this experiment.

TABLE 3. Hb estimation - our approach vs. hemaapp.

Sr. No	Technique	RMSE
1	Proposed	0.81-1.39
2	HemaApp	1.70

the performance of our approach against HemaApp by [43] and [44], a state of the art solution. Figure 16 shows the actual and predicted Hb values of HemaApp technique on our dataset. We can see from the figure that HemaApp produced RMSE of 1.70 which is much higher compared to that produced by our algorithm which is 1.39, shown in Figure 15. The reason behind this low performance of HemaApp is its manual feature engineering process. Another detailed comparison between the proposed approach and HemaApp is shown in Figure 17. Using deep learning-based method, helps eliminate the need of manual feature extraction from the PPG signals and a fingertip video can be used to directly to estimate the Hb level. Table 3 presents the summary of HemaApp performance against proposed deep learning approach. The Root Mean Squared Error (RMSE) serves as a key metric, highlighting the superior accuracy of our method over existing techniques, especially in scenarios where manual feature extraction poses limitations.

E. DISCUSSION

Convolutional neural networks (CNNs) have proven to be highly effective in a range of applications, including computer vision tasks like object recognition, segmentation, and tracking. One particular type of CNN that has gained widespread attention in recent years is the Residual Network (ResNet), which uses residual connections to enable training of deeper architectures. One of the key strengths of CNNs is their ability to automatically learn features and patterns from input data through the use of convolutional layers. Another advantage of CNNs is that they require minimal manual engineering compared to traditional machine learning

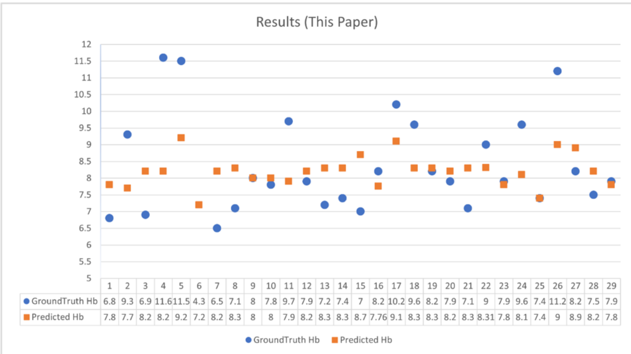

FIGURE 15. Proposed approach results on our dataset with 80-20 split: A comparison between actual Hb level and predicted Hb level is shown. Actual Hemoglobin level is the laboratory gold standard (Complete Blood Count) test result shown by circle (blue color). The predicted Hb level of proposed approach is shown by square (orange color). x-axis represents the record numbers and y-axis represents the Hb level in gm/dL. We observe the RMSE value of 1.39 in this experiment.

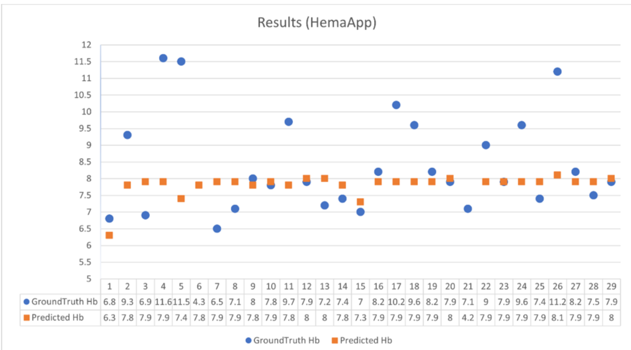

FIGURE 16. Comparison against HemaApp by [43] and [44] on our dataset: A comparison between actual Hb level and predicted Hb level is shown. Actual Hemoglobin level is the laboratory gold standard (Complete Blood Count) test result shown by circle (blue color). The predicted Hb level of HemaApp algorithm is shown by square (orange color). x-axis represents the record numbers and y-axis represents the Hb level in gm/dL. HemaApp approach achieves an RMSE of 1.70 on our dataset.

algorithms. This is because CNNs can learn the necessary parameters for feature extraction automatically from the input data, without the need for hand-crafted features or feature engineering.

In our research, we employed a ResNet-18 architecture to estimate Hemoglobin levels from fingertip video data. The network was trained using dataset of video recordings and corresponding Hemoglobin measurements. In this way, the network was able to learn complex patterns and features from the input data to make accurate predictions. One of the key advantages of using a ResNet-18 architecture is its ability to capture deep representations of the input data, enabling it to model complex relationships between features and learn more meaningful representations of the input. Overall, the use of a ResNet-18 architecture in our research allowed us to effectively leverage the power of CNNs to estimate Hemoglobin levels from fingertip video data. By automatically learning features and patterns from the

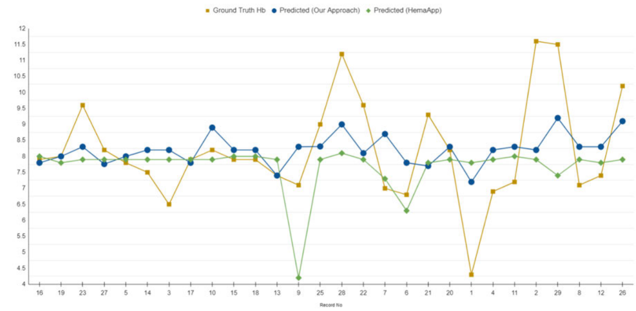

FIGURE 17. A comparison of HemaApp by [43] and [44] algorithm results and our proposed approach results on 80-20 data split is shown. Actual Hemoglobin level is the laboratory gold standard (Complete Blood Count) test result shown by circle (blue line). The predicted Hb level of proposed approach is shown by square (orange line) and HemaApp results are shown by diamond (green line). x-axis represents the record no and y-axis represents the Hb level (gm/dL).

input, the ResNet-18 was able to provide accurate predictions with minimal manual engineering. Table 1 shows that most of the comparative studies require manual feature extraction from the PPG signals. While some of these approaches have shown promising results, they have significant limitations. For example, the need of manual feature extraction from the signal may require expert knowledge. Overall these weaknesses highlight the need for more automated solution which can overcome the limitation of manual feature extraction.

One of the techniques proposed by [43] and [44] reports an RMSE value of 1.27. We have implemented and evaluated this technique on our dataset. A comparison of proposed approach with HemaApp approach by [43] and [44] on our dataset is shown in Figure 17. Clearly our approach outperforms the existing one. Our approach achieved significantly higher accuracy compared to the existing methods.

Our analysis revealed that the majority of records in our dataset fall within the range of 7.0-8.0 gm/dL, and our approach demonstrated high performance in this range, as shown in Figure 14. However, we also noticed that there were only a few records falling in the ranges of 4.0-5.0 gm/dL and 10.0-12.0 gm/dL. Given that our approach was specifically designed to estimate Hemoglobin levels from fingertip video data, it is possible that the performance of our approach can be further improved by applying the same technique on a larger dataset with more records in these ranges. This would allow the model to better learn and generalize to the full range of Hemoglobin levels, including those that are less common in our current dataset. Overall, our findings suggest that our approach has the potential to perform well across a wide range of Hemoglobin levels, but further evaluation on larger and more diverse datasets is needed to fully assess its generalizability and clinical utility.

In addition to the discussion about Hb estimation, few points about data collection are also worthwhile to mention. During the data collection process, some limitations were encountered which may have impacted the quality or quantity of data collected. Specifically, the CBC equipment used to estimate the Hb level was not working properly for 1-2 days, which may have resulted in incomplete or inaccurate data for those patients. In addition, some patients were below 1 year of age and their fingers were too small to adequately cover the flashlight and camera during video recording. While efforts were made to address these limitations, such as rescheduling appointments and adjusting the equipment, it's important to acknowledge that they may have impacted the overall data collection process. Despite these limitations, we believe that our dataset provides valuable insights into fingertip morphology among thalassemia patients, and we have taken steps to ensure that our findings are appropriately interpreted in light of these limitations.

To address the limitations encountered during data collection, we took several steps to ensure the quality and validity of our data. For instance, we excluded any data collected on days when the CBC equipment was not working properly, to ensure that our results were not impacted by incomplete or inaccurate data. Additionally, for children whose fingers were too small to adequately cover the flashlight and camera during video recording, we used the middle finger or thumb to collect data. This allowed us to collect high-quality data for all patients, regardless of age or finger size. By taking these steps, we were able to ensure that our dataset provides valuable insights into fingertip morphology among thalassemia patients, while also addressing any limitations encountered during the data collection process.

V. CONCLUSION AND FUTURE WORK

In this study, we present a dataset for hemoglobin estimation publicly available for research purposes. Our dataset has a diverse age range and a varying range of hemoglobin levels. To the best of our literature review, all the existing published research studies in this domain, have not made their datasets publicly available. Hence, this dataset is going to open doors for ground breaking research studies in future not only for Hb estimation but also for many other health avenues.

To estimate the Hb value, we propose to use deep learning based technique that uses RestNet-18 architecture. By employing ResNet-18, we eliminate the need for manual feature extraction from PPG signals, a step traditionally performed in non-invasive hemoglobin estimation techniques based on machine learning algorithms. Our proposed algorithm yields promising results, achieving a remarkably low error rate of 0.81. Additionally, experiments with varying video lengths indicate that a 1-minute fingertip video accurately estimates hemoglobin levels.

We have tested our trained model on secondary records taken from same patients whose values are used in training. We have trained and tested our model against post blood transfusion test to report the correctness of locally adopted method of testing post transfusion hemoglobin. This information can be used to further improve the current adopted methods to better analyze the patient's hemoglobin level.

Our proposed ResNet-18-based algorithm for hemoglobin estimation demonstrates notable strengths. By eliminating manual feature extraction from PPG signals, it streamlines the process, achieving high accuracy with an impressive error rate of 0.81 in a 1-minute fingertip video. On the positive side, the ResNet-18 architecture proves efficient, accurate, and versatile. The open dataset availability fosters adaptability and exploration in health-related applications.Rigorous testing against locally adopted post-transfusion hemoglobin testing methods validates the model's accuracy. Future efforts will focus on expanding the dataset, optimizing the approach for resource-constrained devices, and developing a cross-platform mobile application for widespread accessibility. Despite current limitations, our study lays the foundation for impactful advancements in non-invasive hemoglobin estimation and broader health research.

In future, we plan to collect more data from healthy patients to fill the gap of few missing ranges of hemoglobin level. We also plan to see how the proposed approach can be used to make it run on resource contained devises like mobile phones. We also plan to develop a cross platform mobile application and release to the public as a portable hemoglobin estimation tool.

ACKNOWLEDGMENT

The authors thank Dr. Zohra Jabeen Wazir from PathWell who helped them during data collection from Pakistan Thalassemia Welfare Society, Rawalpindi, Pakistan (PathWell). They also thank Dr. Ahmad Jalal for his support against analytical and experimental observations.

REFERENCES

[1] G. M. T. Ahsan, Md. O. Gani, M. K. Hasan, S. I. Ahamed, W. Chu, M. Adibuzzaman, and J. Field, "A novel real-time non-invasive hemoglobin level detection using video images from smartphone camera," in *Proc. IEEE 41st Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, vol. 1, Jul. 2017, pp. 967–972.

[2] A. Al-Jawaldeh, M. Taktouk, R. Doggui, Z. Abdollahi, B. Achakzai, H. Aguenaou, M. Al-Halaika, S. Almamary, R. Barham, F. Coulibaly-Zerbo, L. E. Ammari, J. Elati, N. A. Nishtar, N. Omidvar, M. Q. Shams, A. B. Qureshi, and L. Nasreddine, "Are countries of the eastern mediterranean region on track towards meeting the world health assembly target for anemia? A review of evidence," *Int. J. Environ. Res. Public Health*, vol. 18, no. 5, p. 2449, Mar. 2021.

[3] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Meas.*, vol. 28, no. 3, pp. R1–R39, Mar. 2007.

[4] P. Bing, Y. Liu, W. Liu, J. Zhou, and L. Zhu, "Electrocardiogram classification using TSST-based spectrogram and ConViT," *Frontiers Cardiovascular Med.*, vol. 9, Oct. 2022, Art. no. 983543.

[5] Y. Chen, T. Zhong, X. Song, S. Zhang, M. Sun, X. Liu, J. Wei, J. Shu, Y. Liu, and J. Qin, "Maternal anaemia during early pregnancy and the risk of neonatal outcomes: A prospective cohort study in central China," *BMJ Paediatrics Open*, vol. 8, no. 1, Jan. 2024, Art. no. e001931.

[6] D. Cheng, L. Chen, C. Lv, L. Guo, and Q. Kou, "Light-guided and cross-fusion U-Net for anti-illumination image super-resolution," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 12, pp. 8436–8449, Dec. 2022.

[7] S. Das, A. Kesarwani, M. Dalui, D. R. Kisku, B. Sen, S. Roy, and A. Basu, "Smartphone-based non-invasive haemoglobin level estimation by analyzing nail pallor," *Biomed. Signal Process. Control*, vol. 85, Aug. 2023, Art. no. 104959.

[8] G. Dimauro, M. E. Griseta, M. G. Camporeale, F. Clemente, A. Guarini, and R. Maglietta, "An intelligent non-invasive system for automated diagnosis of anemia exploiting a novel dataset," *Artif. Intell. Med.*, vol. 136, Feb. 2023, Art. no. 102477.

[9] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current Cardiology Rev.*, vol. 8, no. 1, pp. 14–25, Jun. 2012.

[10] Z. Fan, Y. Zhou, H. Zhai, Q. Wang, and H. He, "A smartphone-based biosensor for non-invasive monitoring of total hemoglobin concentration in humans with high accuracy," *Biosensors*, vol. 12, no. 10, p. 781, Sep. 2022.

[11] C. Fu, H. Yuan, H. Xu, H. Zhang, and L. Shen, "TMSO-net: Texture adaptive multi-scale observation for light field image depth estimation," *J. Vis. Commun. Image Represent.*, vol. 90, Feb. 2023, Art. no. 103731.

[12] M. A.-U. Golap, S. M. T. U. Raju, M. R. Haque, and M. M. A. Hashem, "Hemoglobin and glucose level estimation from PPG characteristics features of fingertip video using MGGP-based model," *Biomed. Signal Process. Control*, vol. 67, May 2021, Art. no. 102478.

[13] D. Grimaldi, Y. Kurylyak, F. Lamonaca, and A. Nastro, "Photoplethysmography detection by smartphone's videocamera," in *Proc. 6th IEEE Int. Conf. Intell. Data Acquisition Adv. Comput. Syst.*, vol. 1, Sep. 2011, pp. 488–491.

[14] X. Han, C. Zhao, S. Wang, Z. Pan, Z. Jiang, and X. Tang, "Multifunctional TiO_2/C nanosheets derived from 3D metal–organic frameworks for mild-temperature-photothermal-sonodynamic-chemodynamic therapy under photoacoustic image guidance," *J. Colloid Interface Sci.*, vol. 621, pp. 360–373, Sep. 2022.

[15] M. K. Hasan, M. M. Haque, R. Adib, J. F. Tumpa, A. Begum, R. R. Love, Y. L. Kim, and I. A. Sheikh, "Smarthelp: Smartphone-based hemoglobin level prediction using an artificial neural network," in *Proc. AMIA Annu. Symp.*, vol. 2018, p. 535. Bethesda, MD, USA: American Medical Informatics Association, 2018.

[16] F. S. Hassan and A. Gutub, "Improving data hiding within colour images using hue component of HSV colour space," *CAAI Trans. Intell. Technol.*, vol. 7, no. 1, pp. 56–68, Mar. 2022.

[17] B. He, Q. Lu, J. Lang, H. Yu, C. Peng, P. Bing, S. Li, Q. Zhou, Y. Liang, and G. Tian, "A new method for CTC images recognition based on machine learning," *Frontiers Bioengineering Biotechnol.*, vol. 8, p. 897, Aug. 2020.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[19] C. Hu, T. Xia, Y. Cui, Q. Zou, Y. Wang, W. Xiao, S. Ju, and X. Li, "Trustworthy multi-phase liver tumor segmentation via evidence-based uncertainty," *Eng. Appl. Artif. Intell.*, vol. 133, Jul. 2024, Art. no. 108289.

[20] X. Hu, Q. Kuang, Q. Cai, Y. Xue, W. Zhou, and Y. Li, "A coherent pattern mining algorithm based on all contiguous column bicluster," *J. Artif. Intell. Technol.*, vol. 2, no. 3, pp. 80–92, 2022.

[21] H. Huang, N. Wu, Y. Liang, X. Peng, and J. Shu, "SLNL: A novel method for gene selection and phenotype classification," *Int. J. Intell. Syst.*, vol. 37, no. 9, pp. 6283–6304, Sep. 2022.

[22] M. K. Hasan, N. Sakib, J. Field, R. R. Love, and S. I. Ahamed, "A novel technique of noninvasive hemoglobin level measurement using HSV value of fingertip image," 2019, *arXiv:1910.02579*.

[23] A. R. Kavsaoğlu, K. Polat, and M. Hariharan, "Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features," *Appl. Soft Comput.*, vol. 37, pp. 983–991, Dec. 2015.

[24] M. Lakshmi, S. Bhavani, and P. Manimegalai, "Investigation of non-invasive hemoglobin estimation using photoplethysmograph signal and machine learning," in *Proc. Int. Conf. Comput. Vis. Bio Inspired Comput.* Coimbatore, India: Springer, 2019, pp. 1273–1282.

[25] J. Li, L. Han, C. Zhang, Q. Li, and Z. Liu, "Spherical convolution empowered viewport prediction in 360 video multicast with limited FoV feedback," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 19, no. 1, pp. 1–23, Jan. 2023.

[26] J. Li, C. Zhang, Z. Liu, R. Hong, and H. Hu, "Optimal volumetric video streaming with hybrid saliency based tiling," *IEEE Trans. Multimedia*, vol. 25, no. 5, pp. 2939–2953, Sep. 2022.

[27] D. Liu, X. Liu, Z. Chen, Z. Zuo, X. Tang, Q. Huang, and T. Arai, "Magnetically driven soft continuum microrobot for intravascular operations in microscale," *Cyborg Bionic Syst.*, vol. 2022, pp. 1–6, Jan. 2022.

[28] H. Liu, H. Yuan, Q. Liu, J. Hou, H. Zeng, and S. Kwong, "A hybrid compression framework for color attributes of static 3D point clouds," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 3, pp. 1564–1577, Mar. 2022.

[29] H. Liu, S. Zhang, H. Gamboa, T. Xue, C. Zhou, and T. Schultz, "Taxonomy and real-time classification of artifacts during biosignal acquisition: A starter study and dataset of ECG," *IEEE Sensors J.*, vol. 24, no. 6, pp. 9162–9171, Mar. 2024.

[30] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, and H. Yang, "Reduced reference perceptual quality model with application to rate control for video-based point cloud compression," *IEEE Trans. Image Process.*, vol. 30, pp. 6623–6636, 2021.

[31] S. Lu, J. Yang, B. Yang, X. Li, Z. Yin, L. Yin, and W. Zheng, "Surgical instrument posture estimation and tracking based on LSTM," *ICT Exp.*, vol. 42, pp. 2–8, Jan. 2024.

[32] S. Ma, Y. Chen, S. Yang, S. Liu, L. Tang, B. Li, and Y. Li, "The autonomous pipeline navigation of a cockroach bio-robot with enhanced walking stimuli," *Cyborg Bionic Syst.*, vol. 4, p. 67, Jan. 2023.

[33] R. G. Mannino, D. R. Myers, E. A. Tyburski, C. Caruso, J. Boudreaux, T. Leong, G. D. Clifford, and W. A. Lam, "Smartphone app for non-invasive detection of anemia using only patient-sourced photos," *Nature Commun.*, vol. 9, no. 1, p. 4924, Dec. 2018.

[34] S. Mahmud, M. Mansour, T. B. Donmez, M. Kutlu, and C. Freeman, "Non-invasive detection of anemia using lip mucosa images transfer learning convolutional neural networks," *Frontiers Big Data*, vol. 6, Nov. 2023, Art. no. 1291329.

[35] H. Mo, X. Li, B. Ouyang, G. Fang, and Y. Jia, "Task autonomy of a flexible endoscopic system for laser-assisted surgery," *Cyborg Bionic Syst.*, vol. 2022, pp. 52–59, Jan. 2022.

[36] R.-C. Peng, W.-R. Yan, N.-L. Zhang, W.-H. Lin, X.-L. Zhou, and Y.-T. Zhang, "Investigation of five algorithms for selection of the optimal region of interest in smartphone photoplethysmography," *J. Sensors*, vol. 2016, pp. 1–7, Jul. 2016.

[37] L.-M. Po, X. Xu, L. Feng, Y. Li, K.-W. Cheung, and C.-H. Cheung, "Frame adaptive ROI for photoplethysmography signal extraction from fingertip video captured by smartphone," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2015, pp. 1634–1637.

[38] H. Qi, Z. Zhou, J. Irizarry, D. Lin, H. Zhang, N. Li, and J. Cui, "Automatic identification of causal factors from fall-related accident investigation reports using machine learning and ensemble learning approaches," *J. Manage. Eng.*, vol. 40, no. 1, Jan. 2024, Art. no. 04023050.

[39] S. M. Taslim Uddin Raju and M. M. A. Hashem, "Real-time hemoglobin measurement using smartphone video and artificial neural network," in *Proc. 4th Int. Conf. Electr., Comput. Telecommun. Eng. (ICECTE)*, Dec. 2022, pp. 139–142.

[40] H. Sheng, S. Wang, D. Yang, R. Cong, Z. Cui, and R. Chen, "Cross-view recurrence-based self-supervised super-resolution of light field," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 12, pp. 7252–7266, Dec. 2023.

[41] D. F. Silva, J. G. De M, L. V. Domingues, and T. Mazzu-Nascimento, "Hemoglobin estimation from smartphone-based photoplethysmography with small data," in *Proc. IEEE 36th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2023, pp. 75–78.

[42] K. Uesugi, H. Mayama, and K. Morishima, "Analysis of rowing force of the water strider middle leg by direct measurement using a bio-appropriating probe and by indirect measurement using image analysis," *Cyborg Bionic Syst.*, vol. 4, p. 61, Jan. 2023.

[43] E. J. Wang, W. Li, J. Zhu, R. Rana, and S. N. Patel, "Noninvasive hemoglobin measurement using unmodified smartphone camera and white flash," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2333–2336.

[44] E. J. Wang, W. Li, D. Hawkins, T. Gernsheimer, C. Norby-Slycord, and S. N. Patel, "HemaApp: Noninvasive blood screening of hemoglobin using smartphone cameras," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Sep. 2016, pp. 593–604.

[45] C. Wei, L. Sheng, G. Lihua, C. Yuquan, and P. Min, "Study on conditioning and feature extraction algorithm of photoplethysmography signal for physiological parameters detection," in *Proc. 4th Int. Congr. Image Signal Process.*, vol. 4, Oct. 2011, pp. 2194–2197.

[46] W. Wu, H. Zhu, S. Yu, and J. Shi, "Stereo matching with fusing adaptive support weights," *IEEE Access*, vol. 7, pp. 61960–61974, 2019.

[47] Z. Wu, H. Zhu, L. He, Q. Zhao, J. Shi, and W. Wu, "Real-time stereo matching with high accuracy via spatial attention-guided upsampling," *Appl. Intell.*, vol. 53, no. 20, pp. 24253–24274, 2023.

[48] X. Xie, X. Wang, Y. Liang, J. Yang, Y. Wu, L. Li, X. Sun, P. Bing, B. He, G. Tian, and X. Shi, "Evaluating cancer-related biomarkers based on pathological images: A systematic review," *Frontiers Oncol.*, vol. 11, Nov. 2021, Art. no. 763527.

[49] Q.-Y. Yao, M.-L. Fu, Q. Zhao, X.-M. Zheng, K. Tang, and L.-M. Cao, "Image-based visualization of stents in mechanical thrombectomy for acute ischemic stroke: Preliminary findings from a series of cases," *World J. Clin. Cases*, vol. 11, no. 21, pp. 5047–5055, Jul. 2023.

[50] X. Yi, X. Guan, C. Chen, Y. Zhang, Z. Zhang, M. Li, P. Liu, A. Yu, X. Long, L. Liu, B. T. Chen, and C. Zee, "Adrenal incidentaloma: Machine learning-based quantitative texture analysis of unenhanced CT can effectively differentiate sPHEO from lipid-poor adrenal adenoma," *J. Cancer*, vol. 9, no. 19, pp. 3577–3582, 2018.

[51] G. Zhan, W. Wang, H. Sun, Y. Hou, and L. Feng, "Auto-CSC: A transfer learning based automatic cell segmentation and count framework," *Cyborg Bionic Syst.*, vol. 2022, pp. 1–8, Jan. 2022.

[52] C. Zhang, H. Ge, S. Zhang, D. Liu, Z. Jiang, C. Lan, L. Li, H. Feng, and R. Hu, "Hematoma evacuation via image-guided para-corticospinal tract approach in patients with spontaneous intracerebral hemorrhage," *Neurol. Therapy*, vol. 10, no. 2, pp. 1001–1013, Dec. 2021.

[53] J. Zhang, G. Ye, Z. Tu, Y. Qin, Q. Qin, J. Zhang, and J. Liu, "A spatial attentive and temporal dilated (SATD) GCN for skeleton-based action recognition," *CAAI Trans. Intell. Technol.*, vol. 7, no. 1, pp. 46–55, Mar. 2022.

[54] Y. Zhang and J. Pan, "Assessment of photoplethysmogram signal quality based on frequency domain and time series parameters," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2017, pp. 1–5.

[55] M. Zheng, K. Zhi, J. Zeng, C. Tian, and L. You, "A hybrid CNN for image denoising," *J. Artif. Intell. Technol.*, pp. 93–99, Apr. 2022.

[56] W. Zheng, S. Lu, Y. Yang, Z. Yin, and L. Yin, "Lightweight transformer image feature extraction network," *PeerJ Comput. Sci.*, vol. 10, p. e1755, Jan. 2024.

HUMERA SABIR received the M.S. degree from the National University of Computer and Emerging Science (FAST-NUCES), Islamabad, Pakistan. She is currently pursuing the Ph.D. degree with The University of Texas at Arlington, USA. Her research interests include deep learning, image processing, and databases.

KIFAYAT ULLAH KHAN received the M.S. degree from the University of Greenwich, U.K., and the Ph.D. degree from Kyung Hee University, South Korea. He is currently a Senior Lecturer with the College of Accounting, Finance and Economics, Birmingham City Business School, Birmingham City University, Birmingham, U.K. His research interests include artificial intelligence, FinTech, databases, big data, and social network analysis.

OMER ISHAQ received the M.S. degree from Simon Fraser University, Canada, and the Ph.D. degree from Uppsala University, Sweden. He is currently an Expert in deep learning. His research interests include deep learning, computer vision, medical image analysis, NLP, OCR, and neural machine translation.

ABDULWAHAB ALAZEB received the B.S. degree in computer science from King Khalid University, Abha, Saudi Arabia, in 2007, the M.S. degree in computer science from the Department of Computer Science, University of Colorado Denver, USA, in 2014, and the Ph.D. degree in cybersecurity from the University of Arkansas, USA, in 2021. He is currently an Assistant Professor with the Department of Computer Science and Information System, Najran University. His research interests include cybersecurity, cloud and edge computing security, machine learning, and the Internet of Things. He received the Graduate Certificate in cybersecurity from the University of Arkansas, in 2021.

HANAN ALJUAID received the B.S. degree from KAU University and the M.S. and Ph.D. degrees in computer science from UTM University, in 2014. She is currently with the Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah Bint Abdul Rahman University (PNU), Saudi Arabia. She has given numerous invited talks and tutorials. She has published numerous articles on pattern recognition, the IoT, and data science. Her research interests include computer vision and NLP. Much of her work has been on improving the understanding, design, and performance of pattern recognition, mainly through the application of data mining and machine learning.

ASAAD ALGARNI received the Ph.D. degree in software engineering from North Dakota State University, USA. He is currently an Assistant Professor with the Department of Computer Sciences, College of Computing and Information Technology, Northern Border University, Saudi Arabia. His research interests include software engineering, computer vision applications, and machine learning.

JEONGMIN PARK received the Ph.D. degree from the College of Information and Communication Engineering, Sungkyunkwan University, South Korea, in 2009. He is currently an Associate Professor with the Department of Computer Engineering, Tech University of Korea, South Korea. Before joining the Tech University of Korea, in 2014, he was a Senior Researcher with the Electronics and Telecommunications Research Institute (ETRI) and a Research Professor with Sungkyunkwan University. His research interests include high-reliable autonomic computing mechanisms and human-oriented interaction systems.