Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 7
July 2023

Non-Invasive Measurement of Hemoglobin Using Smartphone

***R. Mala, V. Anukarani, Sivani.M, G. Vaishnavi and ** S. Devi Mahalakshmi**

Department of Biotechnology

*** Department of Computer Science and Engineering*

Mepco Schlenk Engineering College, Sivakasi-626005, Tamilnadu, India

*corresponding author : maalsindia@mepcoeng.ac.in

Abastract

Anemia is a global public health issue with serious impact on health status and the productivity of human being. Currently following haemoglobin estimation methods are invasive, expensive and requires a trained technical expert for collection and analysis of blood. Recently developed non invasive methods require either an external hardware component or additional light source for haemoglobin estimation. They are not validated with the original haemoglobin concentration through blood analysis. They are not portable. The present study aims to develop a non invasive method without an additional hardware than the existing smart phone. Digital images of eye, palm and nailbed were captured with smartphone and processed using MATLAB to estimate the concentration of hemoglobin. The images were filtered using median filter with an algorithm and segmented. The features were extracted from the images and used as an input data for ridge regression model. The model was validated with blood hemoglobin concentration. Confusion matrix was developed and the specificity and sensitivity of the proposed method was determined to be 80%. Accuracy can be further improved by sampling large number of blood samples with wide range of Hb concentration encompassing normal and anemic people of different age and gender.

Keywords: anaemia, hemoglobin, image processing, ridge regression, smartphone

1. Introduction

Anemia is global health problem with complicated consequences. According to Global Burden of Disease (GBD,2019), 40% of children of age 6 to 59 months, 37% of pregnant women, 30% of women of age 15 to 49 years are anemic. Anemia causes dizziness, weakness, tiredness and pale skin. In pregnant women, anemia increases the incidence of preterm delivery, maternal mortality and underweight in babies. In children, it decreases physical, cognitive and socio-emotional skills and increases the risk of infection and mortality.

Iron deficiency, folate deficiency and vitamin B12 deficiency account for one-third of anemia in adults. Anemic patients require frequent blood test to monitor hemoglobin (Hb) concentration (Patel, 2008). hemoglobin (Hb) concentration less than 13.5 g/dL in men and 12.0 g/dL in women is an indication of anemia. Estimation of Hb is an invasive procedure requiring technically skilled personnel for collection and analysis of blood. It is expensive and painful. It may induce blood borne infection. Diagnostic devices used in invasive methods are not portable. Blood collection on a regular basis would be challenging for patients in ICU, patients with severe anemia, pregnant women, children and infants. It is cumbersome in places without laboratory infra structure and a trained and qualified technical personnel.

Hb concentration can be determined by comparing the colour of conjunctiva with standard color scale based on anemic condition (Mannino et.al., 2018; Dimauro et al., 2020). Pallor of the skin is directly associated with the anemic status of the human being (Zucker et.al.,1997). A lack of color in the skin and mucous membranes is a sign of pallor, caused by a low level of circulating Hb concentration. The pallor is more noticeable in places where blood vessels are close to the surface, including the palm, the nail bed and mucous membranes like the tongue or conjunctiva. Point of care non-invasive method of quantifying Hb does not require a trained expert for collection and analysis of blood. Photoplethysmography and fluorescence spectroscopy of oral tissue is used for Hb determination (Dimauro et. al., 2020).

Smartphones have become a basic component and an integral part of everyone's daily life. Attractive features of mobile phones such as portability, self-handling, inbuilt camera features, storage capacity, ability to be integrated with any computer and cloud (Wightman Rojas et.al., 2019) server are exploited to be used as a point of care tool for pre-diagnosis and self-monitoring of Hb concentration. A smartphone-based point of care tool is a viable alternative to invasive clinical blood testing because of its availability, user-friendliness, and easy attachability to various biosensing devices. It is a reliable and economical point-of-care solution for remote health monitoring. Furthermore, smartphones have increased computational power, sensing capability, mobility, and widespread availability (Dimauro et al., 2018).

HemaApp, a blood Hb concentration predictor presented by Wang et al, (2017). It used fingertip video footage obtained with a smartphone utilizing various illumination sources. The system's lighting configuration made it difficult for users to illuminate the fingertip. Zhu and Ozcan, (2015) devised an optical system that illuminates a flow cytometer test strip at first. The test strip was then connected to a smartphone camera as a sensor. Karlen et. al (2013) proposed a new method for automatically detecting the best Region of Interest for an image acquired with a smartphone camera to extract a pulse waveform. As the smartphone cameras do not require any additional lighting condition, they are cost effective. Mannino et. al. (2018) developed an image analysis algorithm to measure the Hb level of an individual by capturing images of fingernail beds using smartphone. The color features from the nail bed images were converted into the CIE L*a*b* where the L value is an indicator of skin tone and Hb concentration. Hasan et. al., (2018) estimated Hb level by analyzing the HSV of a fingertip video. The RGB values of these video frames were converted into HSV color space and histogram values were developed for the HSV pixel intensities. The average histogram values were correlated with the standard Hb concentration. The frequency of H, S and V pixel intensities vary with Hb concentration.

Wang et. al. (2017) developed a smartphone application containing hardware components that record data which are processed by an algorithm to calculate the Hb concentration. The algorithm extracts the R, G, and B time series waveform for each video by averaging each RGB channel independently for each frame. The algorithm extracts machine learning features including peak and trough measurements for each light source, and also the interaction terms between light sources. Support Vector Machine (SVM) based regression is applied to estimate the Hb concentration. Studies by Chen et. al., (2016) Collings et. al., (2016) Chen and Miaou, (2017), Tamir et. al., (2017), Bauskar, et.al.(2019) , Kasiviswanathan et. al. (2020) and used conjunctiva images for estimating Hb concentration. Kalantri et. al., (2010) demonstrated the correlation between Hb concentration and tongue, conjunctiva, nail and palm pallor. Zucker, (1997) , Florestiyanto et. al. (2020) have stated that palms have different shades of reddishness dependent on the Hb concentration. Düntsch and Gediga (2019), performed anemia screening using cloud-based mobile photography application.

All the currently available studies estimated Hb concentration based on the images of the pallor alone. They are not validated by original Hb concentration estimated by blood analysis. They are not specific and sensitive. Hence, in the present study we aimed at quantifying the Hb concentration based on the blood test and the image processing of photographs of skin pallor of conjunctiva of the eye , palm and nailbed captured of the same individual in a smart phone using machine learning technique.

2. Materials And Methods

The general steps and the overview of the image processing stages adopted in the present study are represented in Figure 1. In short, the images of palm, nailbed and eye conjunctiva were processed. The parameter for model was fitted and the model was trained using online mathworktool [https://www.mathworks.com/products/image.html]. The trained model was tested with testing data and the performance metrics.

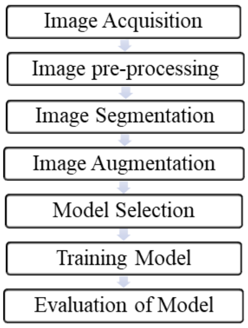


Figure 1 Steps in The Digital Image Processing

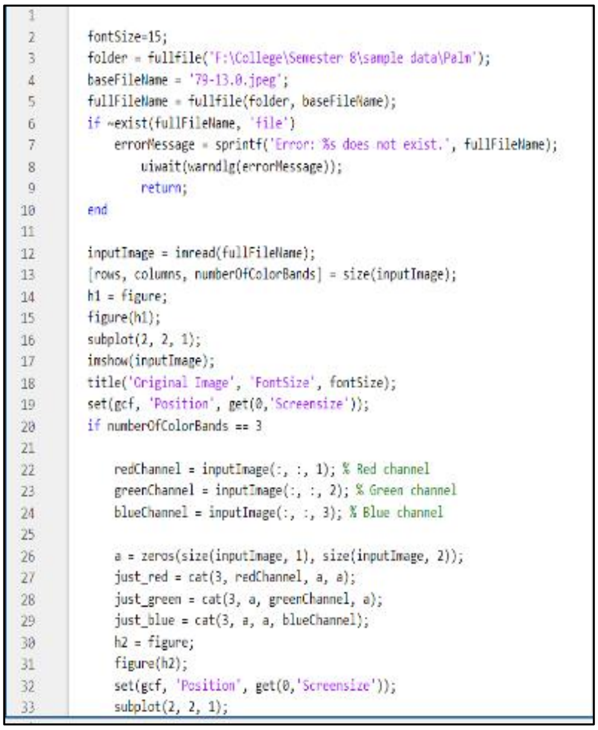
2.1 Image Acquisition

As the aim of the present work was to collect blood sample and the pallor images of palm, nail bed and eye conjunctiva from the same individual, new images were collected for image processing. Blood sample from the individuals and the respective images of palm, nail bed and eye conjunctiva were collected in association with Sri Sai Laboratory, Palladam, Tirupur, Tamilnadu, India. The images were captured using Samsung M12 phone camera. The camera specification was 48 megapixels (MP). Images of patient's eye, nailbed and images were captured. During the capturing of eye and nailbed images the flash option was disabled and the flash was enabled during the capturing of palm image. The details of the sample collected are presented in Table 1.

Table 1. Total Samples collected

Hb (g/dl)	**No of Samples**
4 - 8	10
8 - 12	65
12- 20	40
Gender	
Male	17
Female	98
AGE	
1 – 20	20
21 – 30	63
31 – 51	18
51 – 60	14

The images were pre-processed for noise removal by suppressing unwanted distortions and enhancing the visual qualities. The images were restored by convolving the input with the filter function to obtain the filtered image using nonlinear filtering and blurring with a linear filter (Sonka et.al.,1993). Blurring of edges were reduced by median filtering. The images were pre-processed using following MATLAB algorithm as represented in Figure 2.



```
fontSize=15;
folder = fullfile('F:\College\Semester 8\sample data\Palm');
baseFileName = '79-13.0.jpeg';
fullFileName = fullfile(folder, baseFileName);
if ~exist(fullFileName, 'file')
    errorMessage = sprintf('Error: %s does not exist.', fullFileName);
        uiwait(warndlg(errorMessage));
        return;
end

inputImage = imread(fullFileName);
[rows, columns, numberOfColorBands] = size(inputImage);
h1 = figure;
figure(h1);
subplot(2, 2, 1);
imshow(inputImage);
title('Original Image', 'FontSize', fontSize);
set(gcf, 'Position', get(0,'Screensize'));
if numberOfColorBands == 3

    redChannel = inputImage(:, :, 1); % Red channel
    greenChannel = inputImage(:, :, 2); % Green channel
    blueChannel = inputImage(:, :, 3); % Blue channel

    a = zeros(size(inputImage, 1), size(inputImage, 2));
    just_red = cat(3, redChannel, a, a);
    just_green = cat(3, a, greenChannel, a);
    just_blue = cat(3, a, a, blueChannel);
    h2 = figure;
    figure(h2);
    set(gcf, 'Position', get(0,'Screensize'));
    subplot(2, 2, 1);
```

Figure 2 MATLAB Algorithm used for Median Filtering

2.2 Image Segmentation

The collected images were segmented in image segmenter in MATLAB. A segmentation mask was created for the threshold image using RGB colour space, HSV colour space, YCbCr color space and L*a*b* colour space. The images were cropped using the algorithm shown in Figure 3.

```
C = imread('79-13.0.jpeg');
imshow(C);
targetSize = [200 200];
r = centerCropWindow2d(size(C),targetSize);
D = imcrop(C,r);
imtool(D)
```

Figure 3. MATLAB Algorithm used for Cropping of Images

From the 115 blood samples collected and the images captured from the same individuals, only 67 images were selected based on the clarity. In some images, the nail beds were not clear due to nail polish or colored yellow shaded because of the use of turmeric. The selected 67 data were augmented by flipping or rotating the images in different angles. By using image augmentation technique, 78 datasets were created. So total of 145 data were used for training the model. MATLAB code used for image augmentation is represented in Figure 4a, 4b and 4c

```
% Read the target image file
img = imread('CRCT-39.jpg');

% Reverse the order of the element in each column
vertFlip_img = flip(img, 1);

% Display the vertically flipped image
imtool(vertFlip_img);
```

(a) MATLAB code for vertical flipping

```
I = imread('78-C.jpg');
J = imrotate(I,90,'bilinear');
figure
imshow(J)
```

(b) MATLAB code for rotation by 90°

```
I = imread('78-C.jpg');
J = imrotate(I,270,'bilinear');
figure
imshow(J)
```

(c) MATLAB code for rotation by 270°

Figure 4. MATLAB Code used for Image Augmentation

2.3 Feature Extraction

From the segmented image, features which have high correlation are extracted and it is given as input feature for the model. Mean of a*value, standard deviation of a*value and ratio between mean of red value to the mean of green value have high correlation with hemoglobin level. The Pearson correlation coefficient of these parameters with Hb value is given in Table 2.

Table 2. Correlation of input features with Hb value

Input Features	**Correlation With Hb Value**
Mean of a*value(μ_{a*})	0.424
Ratio between the mean of red and green value(μ_R/μ_G)	0.391
SD of a* value(σ_{a*})	0.344

From the RGB image, mean of red value and green value were extracted (from each pixel red value was extracted, then added and divided by total number of pixel). Then the ratio between red and green values were calculated. The ratio between the mean of red and green values (μR/μG) were extracted from the image using MATLAB code and it was given as one of the input features for the model. Then the image in RGB color space was converted to CIE L*a*b*color space. In CIE L*a*b*color space L stands for Luminance, a* stands for amount of red or green tone in an image, b* stands for amount of blue or yellow tone in an image. CIE L*a*b*colour space is useful for boosting colors and definition in images. The mean of a* value and standard deviation of a* value was extracted from CIE L*a*b*color image.

2.4 Model Selection and Training

Ridge regression model was selected for calculating the coefficients of multiple-regression models (Spooner et.al.,2020). The mean a* of segmented eye, palm and nailbed and Standard Deviation (SD) of a* of segmented eye, palm and nailbed and ratio between mean of red and mean f green values were given as input features to the ridge regression. The model was trained using training data. Each data will have 9 parameters. The parameters are as follows:

1. Mean of a* value of segmented eye image
2. Mean of a* value of segmented palm image
3. Mean of a* value of segmented nailbed image
4. Ratio between the mean of red and green values of segmented eye image
5. Ratio between the mean of red and green values of segmented palm image
6. Ratio between the mean of red and green values of segmented nailbed image

7. SD of a* value of segmented eye image

8. SD of a* value of segmented palm image

9. SD of a* value of segmented nailbed image

The algorithm used for the extraction of RGB vales and the mean a * and SD a* are represented in Figure 5a and 5b respectively.

```
I = imread("eye.png");

Ir = I(:,:,1);
Ig = I(:,:,2);
Ib = I(:,:,3);

Igray = rgb2gray(I);
idx = Igray == 0;

Rave = uint8(mean(Ir(~idx)));
Gave = uint8(mean(Ig(~idx)));
Bave = uint8(mean(Ib(~idx)));

Ir(~idx) = Rave;
Ig(~idx) = Gave;
Ib(~idx) = Bave;
Iout = cat(3,Ir,Ig,Ib);

[Rave, Gave, Bave];
```

(a) MATLAB code for extraction of mean RGB value

```
data=imread("eye.png");
cie = rgb2lab(data);

    % Extract Hue values
    l = cie(:,:,1);

    % Extract Saturation values
    a = cie(:,:,2);

    % Extract Brightness values
    b = cie(:,:,3);

    % Find average of HSV values
    avgl = mean(l(:));
    avga= mean(a(:));
    avgb = mean(b(:));
    sdR = std(l(:));
    sdG = std(a(:));
    sdB = std(b(:));
    [avgl, avga, avgb]
    [sdR, sdG, sdB]
```

(b) MATLAB code for extraction of mean a* and SD value

Figure 5 MATLAB Code used for Feature Extraction

The training data set is displayed in Table 3. The ridge regression is the best method for analyzing multicollinearity data. When there is multicollinearity of data, there is a chance of overfitting. The ridge regression will add small amount of bias so variance of the model is reduced. Since ridge regression prevent overfitting, we have chosen ridge regression as our model (Mehta *et al.*, 2019). The general equation of ridge equation is given below:.

$$cost = \underbrace{\sum_{i=0}^{N}(y_i - \sum_{j=0}^{M} x_{ij} W_j\ )^2}_{\text{Loss function}} + \underbrace{\lambda \sum_{j=0}^{M} Wj^2}_{\text{Regularization term}}$$

The term λ is penalty or tuning parameter. Regularization term is λ times slope square. In linear regression, it tries to minimize only loss function. For linear regression, the value of λ=0. The coefficient of ridge regression is calculated by minimizing the loss function as well as regularization term. The above-mentioned parameters are given as input features to the ridge regression. The code for ridge regression is given in Figure 6.

Table 3 Training dataset

Hb	rpalm	reye	rnail	apalm	aeye	anail	stdpalm	stdeye	stdnail
4.2	1.407407	1.078818	1.233333	5.7981	0.0851	0.1486	7.8153	0.7328	1.0668
5.1	1.389381	1.138365	1.254545	5.7601	0.1494	0.1312	7.5253	1.127	1.1193
5.1	1.396226	1.132075	1.265306	6.2275	0.2108	0.1524	7.6089	1.335	1.1791
5.1	1.390909	1.132075	1.266667	6.3748	0.2667	0.194	7.6442	1.4944	1.3025
5.1	1.385321	1.132075	1.26087	6.7339	0.2431	0.1856	7.6898	1.4287	1.2952
5.1	1.386792	1.132075	1.255319	3.2286	0.0841	0.0615	6.3033	0.8484	0.7517
5.1	1.396226	1.132075	1.26087	3.4483	0.0889	0.0973	6.4515	0.875	0.9318
6	1.459459	1.397163	1.527778	7.5543	0.3708	0.7846	9.8231	3.0544	3.8191
6	1.45283	1.397163	1.5	8.3444	0.5031	0.8187	9.9513	3.5484	3.7987
6	1.454545	1.397163	1.48	8.3476	0.5847	0.852	9.9696	3.8167	3.7801
6	1.466667	1.397163	1.484848	8.5292	0.5626	0.9083	10.0119	3.7457	3.973
6	1.451923	1.397163	1.466667	4.1489	0.2051	0.418	8.1573	2.2765	2.7232
6	1.45283	1.407143	1.482759	4.4724	0.2226	0.4784	8.3865	2.3746	2.8603
6.6	1.270968	1.196721	1.222222	6.2055	0.932	0.2531	7.3451	3.8659	1.6544
6.6	1.273292	1.196721	1.193548	6.931	1.036	0.2272	7.4751	4.0622	1.5546
6.6	1.272727	1.196721	1.2	6.2687	0.9609	0.2009	7.3588	3.9214	1.448
6.6	1.266234	1.196721	1.193548	3.2988	0.4504	0.0963	6.2016	2.7322	1.0078
6.6	1.272727	1.196721	1.193548	3.367	0.4548	0.1135	6.234	2.7441	1.0951
6.8	1.403101	1.315385	1.382979	6.8217	0.2326	0.1758	9.0805	2.0162	1.5308
7	1.361345	1.669643	1.42	5.6731	0.5615	0.2316	7.4983	4.0792	1.8484
7.6	1.65	1.387931	1.296296	7.9997	0.1783	0.1371	10.4211	1.7731	1.1191
7.7	1.39759	1.337748	1.094595	4.992	0.5538	0.0837	5.9421	3.3833	0.765
7.7	1.395062	1.337748	1.097087	5.4341	0.7822	0.3705	5.9415	4.0002	1.7177
7.7	1.392405	1.337748	1.09375	5.5269	0.8152	0.1012	5.9366	4.0827	0.8315
7.7	1.402597	1.337748	1.106061	5.7318	0.7676	0.1133	5.9639	3.9663	0.8759
7.7	1.379747	1.337748	1.092308	2.8456	0.3654	0.0352	5.0751	2.7606	0.4874
7.7	1.392405	1.337748	1.092308	2.9408	0.3694	0.0452	5.1438	2.7753	0.5462
8.4	1.447619	1.239264	1.431818	8.5415	0.7007	0.4317	9.7164	3.3058	2.481
8.4	1.447619	1.239264	1.422222	8.2121	0.6658	0.5741	9.6586	3.2353	2.862
8.4	1.448598	1.239264	1.428571	8.4805	0.7964	0.5834	9.7156	3.5238	2.8587
8.5	1.525773	1.65625	1.444444	7.4234	0.3792	0.3285	9.6081	3.211	1.8508
8.5	1.521277	1.673684	1.4	8.2665	0.5019	0.3751	9.7507	3.6924	1.8526
8.5	1.525773	1.65625	1.411765	8.1723	0.6276	0.3735	9.7259	4.111	1.8589
8.5	1.521739	1.65625	1.375	8.2582	0.6416	0.4156	9.7554	4.1554	1.9496
8.5	1.51087	1.65625	1.333333	4.1342	0.1993	0.1751	8.022	2.3352	1.2818
8.5	1.521739	1.65625	1.333333	4.61	0.2214	0.1982	8.3382	2.4608	1.3424
8.7	1.325758	1.464286	1.166667	5.9706	0.7368	0.4358	7.1686	4.084	2.5037
8.7	1.325397	1.464286	1.194444	6.0824	0.8126	0.2952	7.1396	4.2743	2.0594
8.7	1.325397	1.464286	1.180556	6.3244	0.7412	0.3098	7.2245	4.0888	2.1246
8.7	1.328	1.464286	1.2	2.9868	0.3579	0.1381	5.8696	2.8668	1.4229
8.7	1.328	1.464286	1.191781	3.3256	0.4109	0.1538	6.0988	3.0731	1.4978
8.7	1.543478	1.303226	1.322581	4.43	0.1218	0.0312	8.6886	1.5169	0.4966
8.7	1.561798	1.303226	1.28125	4.769	0.1201	0.0597	8.9276	1.5057	0.6809
8.9	1.578313	1.253086	1.295082	7.5191	0.2352	0.2539	9.7988	2.0395	1.7659
8.9	1.656716	1.454545	1.461538	7.3863	0.3807	0.2465	9.2424	3.0514	1.8099
8.9	1.56962	1.253086	1.290909	8.499	0.326	0.3088	9.9853	2.3893	1.9234
8.9	1.573171	1.253086	1.296296	8.6139	0.3853	0.3291	9.9482	2.5933	1.946
8.9	1.582278	1.253086	1.296296	8.4044	0.3813	0.3512	9.9054	2.58	2.0373
8.9	1.650794	1.451128	1.428571	8.1946	0.5138	0.2986	9.337	3.53	1.9555

```python
import pandas as pd
from numpy import arange
from sklearn.linear_model import Ridge
from sklearn.linear_model import RidgeCV
from sklearn.model_selection import RepeatedKFold
hbvalue=pd.read_excel("/content/drive/MyDrive/Final year project/Final-dataset/trail2excel/training2.xlsx")
data=hbvalue[["Hb","rpalm","reye","rnail","apalm","aeye","anail","stdpalm","stdeye","stdnail"]]
x=hbvalue[["rpalm","reye","rnail","apalm","aeye","anail","stdpalm","stdeye","stdnail"]]
y=hbvalue[["Hb"]]
cv = RepeatedKFold(n_splits=10, n_repeats=3, random_state=1)
model = RidgeCV(alphas=arange(0, 1, 0.01), cv=5, scoring='neg_mean_absolute_error')
model.fit(x,y)
print(model.alpha_)
```

Figure 6 Algorithm for ridge regression

Confusion matrix describes the performance of a model on a set of test data for which the true values are known (Düntsch and Gediga,2019). The matrix was prepared from the data of True Positive (TP - anemic predicted to be anemic), True Negative (TN - non-anemic is predicted as non-anemic), False Negative (FN- anemic but predicted as non-anemic condition) and False Positive (FP- Non anemic predicted as anemic). From the confusion matrix, the specificity and sensitivity of the diagnostic tool was determined. Specificity refers to its capacity to appropriately categorise an individual as disease-free.

$$Specificity = \frac{TN}{(TN + FP)}$$

The ability of a test to correctly designate an individual as 'diseased' is known as sensitivity.

$$Sensitivity = \frac{TP}{(TP + FN)}$$

2.5 Blood Collection and Hb estimation

Blood was collected from vein and the Hb concentration was estimated by Sahlis acid hematin method.

3 RESULT AND DISCUSSION

3.1 Hb Estimation from Blood and Collection of Images

Hb was estimated from the blood sample of all individuals participated in the study and the images of conjunctiva of the eye, nail bed and the palm from the respective individuals were captured. The images acquired from the patient were screened initially. Few images were neglected because of henna, turmeric in patient palm, blurred images and nail color in hand. Totally, 115 data (eye, palm, nailbed images) were collected. From that, 67 data is used for processing while the other 48 data were neglected and these neglected images can't be processed. The neglected images are shown in Table 4.

Table 4. Neglected Images

REASON FOR NEGLECTION	IMAGES	IMAGES	IMAGES
Blurred image due to inference of eyelash or shaking of photos			
Nail color in nails			
Henna in nails			
Turmeric in palm			

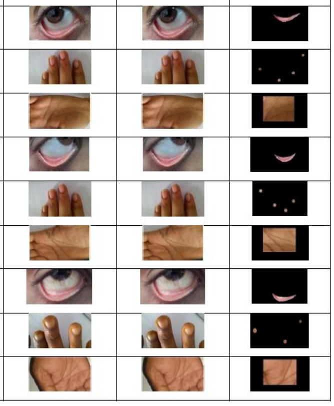
The test data were assessed and the model was tested using 25 testing data after fitting the ridge regression model. The Hb concentration was predicted and the results are depicted in Table 5.

3.2 Noise Removal and Selction of Region of Interest

The noise in the images were removed by median filter. Median filter removes noise and preserve edges. After removing the noise from the image, the region of interest (ROI) wasextracted from the image. For palm and nailbed images specific dimension of palm and nailbed images were extracted and all other regions were made black. For eye image, only conjunctiva region is extracted and all other region is made black. Making background black will have pixel value 0 and it will not affect the pixel value of ROI. These segmented images were used for further process. The raw image, filtered images and segmented images of few samples are shown in Table 5

Table 5. Filtered and segmented images

Hb	Region	Original Image	Filtered Image	Segmented Image
11.0	Eye			
	Nail			
	Palm			
10.8	Eye			
	Nail			
	Palm			
11.8	Eye			
	Nail			
	Palm			

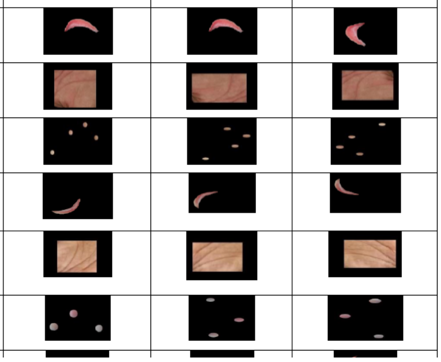


3.3 Increase of Dataset by Augmentation Technique

To train a model large number of datasets is required. Image augmentation technique like vertical flipping, rotating the image by 90° and 270° angle was done. Image augmentation was performed and 78 new dataset were created. The sample images are shown in Table 6

Table 6: Augmented images

REGION	VERTICAL FLIPPING	90° ROTATION	270° ROTATION
EYE			
PALM			
NAIL			
EYE			
PALM			
NAIL			



3.4 TRAINING THE MODEL USING INPUT FEATURES

The mean of a* value of segmented eye, palm and nailbed images, values of ratio between the mean red pixel intensities to green pixel intensities of segmented eye, palm and nailbed image, SD of a* value of segmented eye, palm, nailbed images were given as input feature for ridge regression model. Then model is trained using 67 original data and 78 augmented data. Totally, 145 data was used for training the ridge regression model. The model will fit a ridge regression equation. This equation was used to predict the Hb concentration

3.2 Model Testing and Evaluation of Model

The model was tested using 25 testing data and hemoglobin concentration is predicted. Based on predicted hemoglobin concentration, the performance of model is evaluated. The test data is shown in Table 7

Table 5. Measurement of Hb Concentration

SN	Actual	Gender	Predicted	Actual	Predicted
1	9.5	F	10.8	A	A
2	13.4	F	13.4	NA	NA
3	11.8	F	13.5	A	NA
4	9.9	F	10.2	A	A
5	10.2	F	8.1	A	A
6	11.3	F	11.1	A	A
7	7.6	F	9.2	A	A
8	8.2	F	8.8	A	A
9	11.7	F	9	A	A
10	10.7	F	9.3	A	A
11	12.3	F	11	NA	A
12	14.1	M	14.7	NA	NA
13	8.7	F	9.1	A	A
14	10.8	F	11.8	A	A
15	15	M	13.5	NA	NA
16	11.5	F	11.5	A	A
17	13.3	F	10.3	NA	A
18	12	F	12	NA	NA
19	8.4	F	9.2	A	A
20	6.6	F	8.1	A	A
21	12.4	F	10.8	NA	A
22	8.7	F	11.3	A	A
23	13.1	M	11.9	NA	A
24	11	F	9.3	A	A
25	9.6	F	9.7	A	A

Based on predicted Hb concentration, the performance of the model was evaluated using the evaluation metrices such as accuracy, sensitivity, error rate and specificity. The definitions used to evaluate are error rate (an error rate is defined as the percentage value of the difference of the observed and the actual value divided by the actual value), accuracy (capacity to appropriately distinguish between patient and healthy cases=Accuracy = 1 – Error rate). The results are represented in Table 5.

Table 5 Confusion Matrix

		ACTUAL	
		ANEMIC	NON-ANEMIC
PREDICTED	ANEMIC	16	1
	NON-ANEMIC	4	4

From this confusion matrix sensitivity and specificity were determined to be 80%. Accuracy of the model will increase by increasing the dataset for training. This model would serve as a better tool to predict Hb concentration without any assistance of external light source or any other additional hardware devices.

Conclusion

The current study demonstrated that the non invasive method of Hb determination was accurate with less deviation. It corroborates well with the original Hb concentration estimated by invasive method. But the accuracy can be increased with minimum error rate if the data is trained with higher number of samples. This model is useful for large scale screening of anemia. Further, if a web application was developed, people can predict the Hb concentration from their residence.

ACKNOWLEDGEMENT

The authors acknowledge the support rendered by the Management Board, Principal and the Department of Biotechnology for execution of the work.

RFERENCES

1. Bauskar S, Jain P, Gyanchandani M. A noninvasive computerized technique to detect anemia using images of eye conjunctiva. Pattern Recognition and Image Analysis. 2019;29(3):438–46. doi:10.1134/s1054661819030027.
2. Chen Y-M, Miaou S-G, Bian H. Examining palpebral conjunctiva for anemia assessment with image processing methods. Computer Methods and Programs in Biomedicine. 2016;137:125–35. doi:10.1016/j.cmpb.2016.08.025
3. Chen Y-M, Miaou S-G. A Kalman filtering and nonlinear penalty regression approach for noninvasive anemia detection with palpebral conjunctiva images. Journal of Healthcare Engineering. 2017;2017:1–11. doi:10.1155/2017/9580385
4. Collings S, Thompson O, Hirst E, Goossens L, George A, Weinkove R. Non-invasive detection of anaemia using digital photographs of the conjunctiva. PLOS ONE. 2016;11(4). doi:10.1371/journal.pone.0153286
5. Dimauro G, Caivano D, Girardi F. A new method and a non-invasive device to estimate anemia based on digital images of the conjunctiva. IEEE Access. 2018;6:46968–75. doi:10.1109/access.2018.2867110
6. Dimauro G, De Ruvo S, Di Terlizzi F, Ruggieri A, Volpe V, Colizzi L, et al. Estimate of anemia with new non-invasive systems—a moment of reflection. Electronics. 2020;9(5):780. doi:10.3390/electronics9050780
7. Düntsch I, Gediga G. Confusion matrices and rough set data analysis. Journal of Physics: Conference Series. 2019;1229(1):012055. doi:10.1088/1742-6596/1229/1/012055
8. Florestiyanto, M. Y., & Peksi, N. J. (2020, October). Non-Invasive Anemia Screening Using Nails and Palms Photos. In Proceeding of LPPM UPN "Veteran" Yogyakarta Conference Series 2020–Engineering and Science Series (Vol. 1, No. 1, pp. 311-318).
9. GBD, 2019. Global burden of 369 diseases and injuries in 204 countries and territories, 1990– 2019: a systemati analysis for the Global Burden of Disease Study 2019 . Lancet 2020; 396: 1204–22
10. Hasan MK, Haque M, Sakib N, Love R, Ahamed SI. Smartphone-based human hemoglobin level measurement analyzing pixel intensity of a fingertip video on Different Color Spaces. Smart Health. 2018;5–6:26–39. doi:10.1016/j.smhl.2017.11.003
11. Kalantri A, Karambelkar M, Joshi R, Kalantri S, Jajoo U. Accuracy and reliability of pallor for detecting anaemia: A hospital-based Diagnostic Accuracy Study. PLoS ONE. 2010;5(1). doi:10.1371/journal.pone.0008545
12. Karlen W, Ansermino JM, Dumont GA, Scheffer C. Detection of the optimal region of interest for camera oximetry. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2013; doi:10.1109/embc.2013.6609988
13. Kasiviswanathan S, Vijayan TB, John S. Ridge regression algorithm based non-invasive anaemia screening using conjunctiva images.

Journal of Ambient Intelligence and Humanized Computing. 2020; doi:10.1007/s12652-020-02618-3

14. Mannino RG, Myers DR, Tyburski EA, Caruso C, Boudreaux J, Leong T, et al. Smartphone app for non-invasive detection of anemia using only patient-sourced photos. Nature Communications. 2018;9(1). doi:10.1038/s41467-018-07262-2
15. Wightman Rojas PM, Mass Noriega LA, Salazar Silva A. Hemoglobin screening using cloud based mobile photography applications [online first]. Ingenieria y Universidad. 2019;23(2). doi:10.11144/javeriana.iyu23-2.hsuc
16. Patel KV, Guralnik JM. Epidemiology of anemia in older adults. Blood Disorders in the Elderly. 2007;11–20. doi:10.1017/cbo9780511545238.003
17. Sonka, M., Hlavac, V., & Boyle, R. 1993. Chapter : Introduction In : Image processing, analysis, and machine vision. (4th edition) Cengage Learning 200 First Stamford Place, 4th Floor Stamford, CT 06902 USA
18. Spooner A, Chen E, Sowmya A, Sachdev P, Kochan NA, Trollor J, et al. A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. Scientific Reports. 2020;10(1). doi:10.1038/s41598-020-77220-w
19. Tamir A, Jahan CS, Saif MS, Zaman SU, Islam MdM, Khan AI, & ShahnazC. Detection of anemia from image of the anterior conjunctiva of the eye by Image Processing and Thresholding. 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). 2017; doi:10.1109/r10-htc.2017.8289053
20. Wang EJ, Li W, Zhu J, Rana R, Patel SN. Noninvasive hemoglobin measurement using unmodified smartphone camera and White Flash. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2017; doi:10.1109/embc.2017.8037323
21. Zhu, H., & Ozcan, A. (2015). Opto-fluidics based microscopy and flow cytometry on a cell phone for blood analysis. In Mobile Health Technologies (pp. 171-190). Humana Press, New York, NY.
22. Zucker, J. R., Perkins, B. A., Jafari, H., Otieno, J., Obonyo, C., & Campbell, C. C. (1997). Clinical signs for the recognition of children with moderate or severe anaemia in western Kenya. Bulletin of the World Health Organization, 75(Suppl 1), 97.