Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 10 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i10.48331

PAPER

An Algorithm for the Estimation of Hemoglobin Level from Digital Images of Palpebral Conjunctiva Based in Digital Image Processing and Artificial Intelligence

Guillermo Moreno[1], Abdigal Camargo[1], Luis Ayala[2], Mirko Zimic[2], Christian del Carpio[1](✉)

[1]Laboratorio de Investigación en Inteligencia Artificial, Robótica y Procesamiento de Imágenes, Universidad Nacional de Ingeniería, Lima, Perú

[2]Laboratorio de Bioinformática y Biología Molecular, Universidad Peruana Cayetano Heredia, Lima, Perú

cdelcarpiod@uni.edu.pe

ABSTRACT

Anemia is a common problem that affects a significant part of the world's population, especially in impoverished countries. This work aims to improve the accessibility of remote diagnostic tools for underserved populations. Our proposal involves implementing algorithms to estimate hemoglobin levels using images of the eyelid conjunctiva and a calibration label captured with a mid-range cell phone. We propose three algorithms: one for calibration label segmentation, another for palpebral conjunctiva segmentation, and the last one for estimating hemoglobin levels based on the segmented images from the previous algorithms. Experiments were performed using a data set of children's eyelid images and calibration stickers. An L1 norm error of 0.72 g/dL was achieved using the SLIC-GAT model to estimate the hemoglobin level. In conclusion, the integration of these segmentation and regression methods improved the estimation accuracy compared to current approaches, considering that the source of the images was a mid-range commercial camera. The proposed method has the potential for mass screening in low-income rural populations as it is non-invasive, and its simplicity makes it feasible for community health workers with basic training to perform the test. Therefore, this tool could contribute significantly to efforts aimed at combating childhood anemia.

KEYWORDS

anemia detection, image processing, machine learning, SLIC-GAT, palpebral conjunctiva, hemoglobin level

1 INTRODUCTION

According to the World Health Organization (WHO), an estimated two billion people worldwide are affected by various levels of anemia, with a higher incidence observed in low-resource areas that are also impacted by infectious diseases such as malaria and hookworm infections [1]. This represents approximately one-third of

Moreno, G., Camargo, A., Ayala, L., Zimic, M., del Carpio, C. (2024). An Algorithm for the Estimation of Hemoglobin Level from Digital Images of Palpebral Conjunctiva Based in Digital Image Processing and Artificial Intelligence. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(10), pp. 33–46. https://doi.org/10.3991/ijoe.v20i10.48331

Article submitted 2024-02-02. Revision uploaded 2024-04-07. Final acceptance 2024-04-07.

the global population, predominantly concentrated in economically disadvantaged countries [2].

Anemia is defined as a pathological reduction in hemoglobin (Hb) levels, leading to a decrease in the oxygen-carrying capacity of the blood. It is caused by several factors, with iron deficiency in the diet being the most significant. The low concentration of Hb in the blood is the most reliable indicator for diagnosing anemia, which varies based on age, gender, and physical condition [3].

The tools developed in the state of the art for anemia detection can be categorized into two main approaches. The first approach is centered on the detection of anemia itself, while the second approach aims to implement a comprehensive service incorporating detection models.

In the field of non-invasive diagnosis for anemia, various methods have been developed. One approach involves the use of spectroscopy sensors. Kim et al. [4] utilized a reflectance model of the conjunctiva and compared the results with actual Hb measurements in adults. The method achieved an Hb level estimation MSE (mean squared error) of 1.67 g/dL and an 86% sensitivity score. McMurdy et al. [5] used visible range diffuse reflectance spectra from the palpebral conjunctiva. Hb levels were estimated using partial least-squares multivariate regression and discrete spectral region models, resulting in MSE values of 0.67 g/dL and 1.07 g/dL, respectively.

However, the methods mentioned above are expensive and not easily accessible everywhere. In contrast, there are alternative methods that do not require additional devices or manual selection of the region of interest (ROI). Chen et al. [6] selected a squared ROI and used a Kalman filter and regression method to estimate nonlinear curves for Hb levels from the mean value of the red component in digital images. They proposed a risk evaluation scheme achieving accuracy scores of 0.64, 0.68, and 0.88 for high-risk, doubtful, and low-risk classes, respectively. Muthalagu et al. [7] employed an Elman neural network for correlating and classifying anemic cases, achieving 77% sensitivity and 96% specificity. Noor et al. [8] manually cropped images and used KNN for feature extraction and LSVM and decision trees for classification. The best accuracy score of 82.61% was obtained using the decision tree model. Park et al. [9] achieved a 0.91 R2 score by reconstructing subconjunctival redness (SSR) from RGB images of the palpebral conjunctiva. Roychowdhury et al. [10] presented a method for ROI detection of the tongue and eyes. They used multiple machine learning models on images of varying sizes (from 155 × 240 to 960 × 1280 pixels) and achieved accuracy scores of 86% for eye images and 98% for tongue images.

Another set of methods utilizes smartphones for anemia detection. Mannino et al. [11] used images of the fingernail bed captured with an iPhone 5s. They manually selected the ROI and applied a linear fit to the pixel data, achieving 92% sensitivity and 76% specificity scores. Jain et al. [12] used images from the palpebral conjunctiva captured with a Google Pixel PL. They employed data augmentation on 99 images and used the mean of the red and green channels as input to their model, achieving an accuracy score of 0.97. Chen et al. [13] manually selected the ROI and used image processing and machine learning methods to propose two models: one based on digital image processing and another based on machine learning.

To address the ROI selection problem, Dimauro et al. [14] used a macro-lens with a special spacer. They performed semi-automatic segmentation using the SLIC algorithm and then applied classification tasks using SMOTE and ROSE on 102 images of size 2448 × 3246. Their approach achieved an accuracy score of 0.976. Another solution for anemia detection using phone images is presented [15], where a CNN architecture is used to classify anemia based on estimated hemoglobin levels. Photos taken with a SONY DSC-T1100 were manually segmented and processed with digital image processing and machine learning methods for feature selection and classification, respectively,

resulting in a sensitivity score of 0.78. Tamir et al. [16] also utilized phone photography and used a threshold over the R and G channels, achieving an accuracy of 78.9%.

In order to eliminate the need for additional devices to capture photos, Dimauro et al. [17] developed a method for the segmentation of the palpebral conjunctiva from images taken with a macro lens at a high-resolution close-up of 10x zoom using a Huawei P9 Lite. They employed Gaussian blurring, the Otsu algorithm, morphological opening, and contour detection. The a* channel achieved a correlation score of 0.63, and after eliminating the darkest pixels, a correlation score of 0.735 was achieved. On the other hand, Apppiahere et al. [18] performed an automatic eye conjunctiva detection model with Yolov5 and anemia detection model to run on a server and process images in real time with a sensitivity of 90%, a specificity of 95%, and an accuracy of 92.50% on average in approximately 50 seconds of processing.

The main studies mentioned above are summarized in Table 1 for comparison.

Table 1. Summary of studies

No	Methods	Advantage	Disadvantage
1	Conjunctival reflectance model [4]	Sensitivity of 86%.	Manual selection of the region of interest (ROI). MSE (mean squared error) of 1.67 g/dL.
2	Visible-range diffuse reflectance spectra of the palpebral conjunctiva [5]	It uses multivariate regression models.	Manual selection of the region of interest (ROI).
3	Feature Extraction with KNN and Tree Models [8]	It achieves an accuracy of 82.61% with the decision tree model.	Dependent on the accuracy of manual image selection.
4	Subconjunctival Redness Reconstruction [9]	Achieves an R2 score of 0.91.	Image quality may vary in real environments.
5	ROI Detection in Tongue and Eye Imaging [10]	It achieves an accuracy score of 86% for eye images and 98% for tongue images.	Reliance on multiple machine learning models.
6	Anemia Detection from Sickbed Imaging [11]	It achieves a sensitivity of 92% and a specificity of 76%.	Requires manual ROI selection.
7	Anemia Detection from Palpebral Conjunctiva Images [12]	Achieves an accuracy score of 0.97.	It does not eliminate the need for manual ROI selection.
8	Macro-lens segmentation and classification [14]	It achieves an accuracy of 0.976.	Requires semi-automatic segmentation.
9	Anemia Detection by Threshold in Color Channels [16]	It achieved an accuracy rate of 78.9%.	Depends on manual segmentation and image quality.
10	Anemia detection model using real-time machine learning techniques [18]	Achieves a sensitivity of 90%, specificity of 95% and accuracy of 92.50%.	It depends on a server and internet speed. Processing time on average 50 seconds.

According to Table 1, the common disadvantage of the state of the art is that they have a mean MSE of 1.14 g/dL, manual ROI selection, and the need for high-resolution image acquisition equipment or internet services. Therefore, the following method is proposed that achieves a mean L1 error of 0.72 g/dL with images acquired by a mid-range cell phone with a 13 MP camera and automatic ROI selection.

2 METHOD

The proposed algorithm for hemoglobin level estimation involves several steps, as illustrated in Figure 1. The first step is the segmentation of two key regions: the reference sticker and the region of interest, which is the palpebral conjunctiva. The segmentation processes for both the sticker and palpebral conjunctiva are described in separate sections: sticker segmentation and palpebral conjunctiva segmentation. After the segmentation, the information obtained from these processes is integrated to enhance the data quality extracted from the raw photo. This integration step is discussed in the section on image preprocessing. Additionally, this section provides details on the data representation used to feed the statistical models employed in the algorithm. Figure 1 provides a visual representation of the entire pipeline of the algorithm for hemoglobin level estimation.

The dataset of 500 images used for this project was provided by the Bioinformatics and Molecular Biology Laboratory at UPCH. The images of the palpebral conjunctiva of children were acquired by a camera on a Huawei Y6 II mobile device. Prior to image acquisition, the parents of the participating children were informed by the health personnel about the purpose of the study, and their consent was obtained. In addition, to obtain a relative and standardized factor of palpebral conjunctiva colors in the images, stickers were used near the eyes. To validate the hemoglobin level in children, a venous blood sample was obtained from each participant. The samples were processed using Hemocue 201 and Hemocue 301 kits, and further analysis was performed in a clinical laboratory to ensure the accuracy of the measurements.

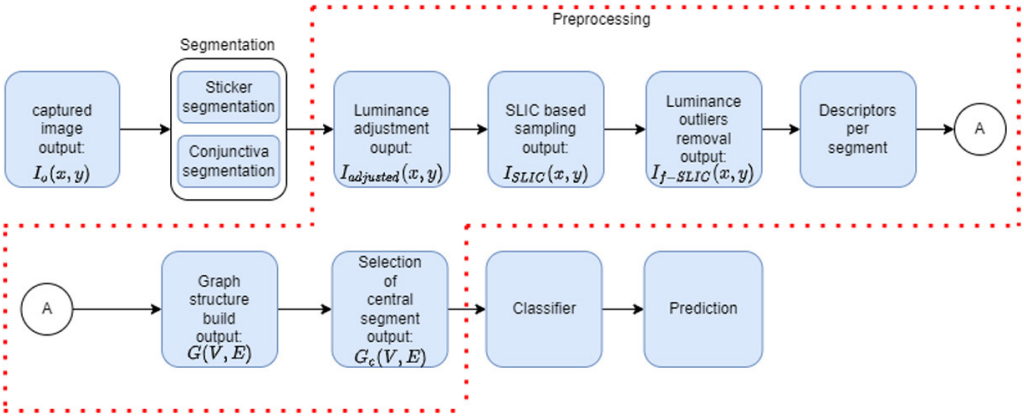

Fig. 1. Hemoglobin estimator workflow, red dotted line contains the blocks of preprocessing algorithm

Fig. 2. Conjunctive segmentation workflow

2.1 Segmentation

Once the images have been captured, in order to estimate the hemoglobin level, it is necessary to extract sub-regions from the image $I_o(x, y)$. The association between conjunctiva pallor and hemoglobin deficiency, as indicated in [19], drives the extraction of these sub-regions for estimation of hemoglobin levels. Therefore, the

palpebral conjunctiva provides valuable information for hemoglobin level estimation, while the sticker is utilized to adjust the image color.

Palpebral conjunctiva segmentation. The entire workflow for palpebral conjunctiva segmentation consists of the following steps. Initially, as depicted in Figure 2, the input $I_o(x, y)$ is processed to obtain the final output $I_{conjunctive}(x, y)$, Figure 3 shows an example of the output. The first step involves detecting a regular region that contains the eye using a robust, tree-based machine learning model known as the Haar Cascade detector [20]. The output of the Haar Cascade detector, denoted as $I_{hcc}(x, y)$, provides the bounding rectangle within the original image I_o that encompasses the largest eye in the image, assuming vertical image orientation. The bounds of $I_{hcc}(x, y)$ are then expanded to ensure that the conjunctiva lies within this new rectangle, referred to as I_{eyeRGB}, following Equation (1):

$$I_{eyeRGB}(x, y) = I_o(y_n{:}y_n + w_n, x_n{:}x_n + h_n,:) \tag{1}$$

Where x_n, y_n, w_n and h_n are indicated in (2), (3), (4) and (5) respectively using (x_0, y_0) as left upper bound corner of $I_0(x, y)$; (w, h) as width and height of $I_{hcc}(x, y)$ and finally (W, H) as width and height of original image.

$$x_n = \max(0, x_0 - 0.75 \times w) \tag{2}$$

$$y_n = \max(0, y_0 - 0.75 \times h) \tag{3}$$

$$w_n = \min(H - x_n, 3 \times w) \tag{4}$$

$$h_n = \min(W - y_n, 3.3 \times h) \tag{5}$$

To ensure that the conjunctiva lies entirely within the rectangle, the bounds are stretched. The maximum and minimum functions are used to ensure that the bounds remain within the coordinates of the original image. This process results in the conjunctiva being contained within the rectangle, as shown in Figure 3c. Next, the image is cropped based on the adjusted rectangle, and two different color map transformations, namely YCbCr (Equation 6) and NTSC (Equation 7), are applied to the cropped image (see Figure 4a), following the approach described in [21].

$$I_{eyeYCbCr}(x, y) = \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \times \begin{bmatrix} I_{eyeR} \\ I_{eyeG} \\ I_{eyeB} \end{bmatrix} + \begin{bmatrix} 0.0 \\ 0.5 \\ 0.5 \end{bmatrix} \tag{6}$$

$$I_{eyeNTSC}(x, y) = \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \times \begin{bmatrix} I_{eyeR} \\ I_{eyeG} \\ I_{eyeB} \end{bmatrix} \tag{7}$$

Where I_{eyeR}, I_{eyeG}, I_{eyeB} are the three separated channels from I_{eyeRGB} image. From the YCbCr color space, the Cr channel is selected, while from the NTSC color space, the I and Q channels are chosen. These three channels are then equalized to cover the entire range of pixel intensities from 0 to 1. To generate the initial mask, the three equalized channels are multiplied element-wise, as described in Equation (8). The resulting image is shown in Figure 4b. Otsu Thresholding [22]

is applied using Equation (9) to obtain the final mask, which will be referred to as $I_{eyebin}(x, y)$.

$$I_{pmask}(x,y) = I_{eyeCr}(x,y) \odot I_{eyeI}(x,y) \odot I_{eyeQ}(x,y) \tag{8}$$

Where the operator $\odot$ indicates element-wise operation.

$$I_{eyebin}(x,y) = \begin{cases} 1 \; if \; I_{mask}(x,y) > \mathbb{E}\left(I_{pmask}(x,y)\right) + \sigma\left(I_{pmask}(x,y)\right) \\ 0 \qquad\qquad otherwise \end{cases} \tag{9}$$

Where $\mathbb{E}$ and σ indicate mean and standard deviation for $I_{pmask}(x, y)$ respectively. After obtaining the mask, all the contours within the mask are sorted based on their area. Only the three largest contours are retained for further processing. These masks are then sorted again based on their vertical distance from the centroid to the center of the cropped image, as shown in Figure 4c.

The obtained mask is dilated using a disk shape morphological element, K_{disk}, with radius of 100 as showed in (10) and hole filled to overwrite itself $I_{eyebin}(x, y)$, that is showed on Figure 4d.

$$I'_{eyebin} = I_{eyebin} \oplus K_{disk} \tag{10}$$

Once that the binary mask $I'_{eyebin}(x,y)$ is obtained, it's used to mask the original I_{eyeRGB} to obtain the ROI in RGB as it's indicated in (11).

$$I_{conjunctive}(x,y) = I'_{eyebin}(x,y) \,\&\, I_{eyeRGB}(x,y) \tag{11}$$

Use of the & operator to denote the operation.

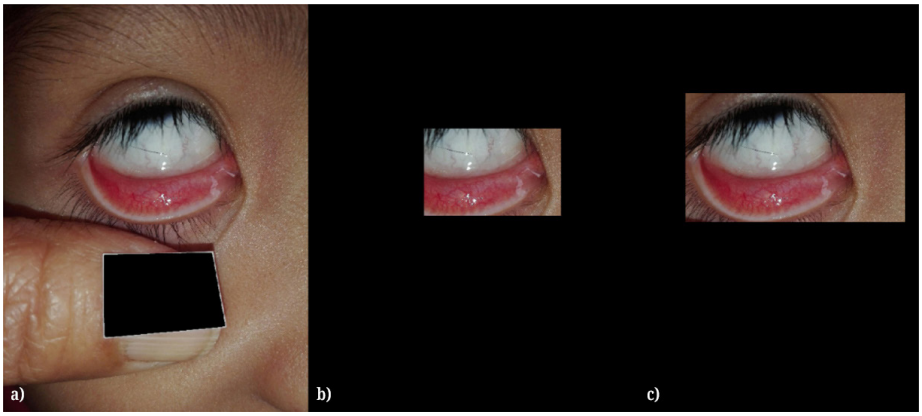

Fig. 3. Initial conjunctive detection (a) Original image without sticker, (b) The detected eye, (c) The output of rectangle expansion

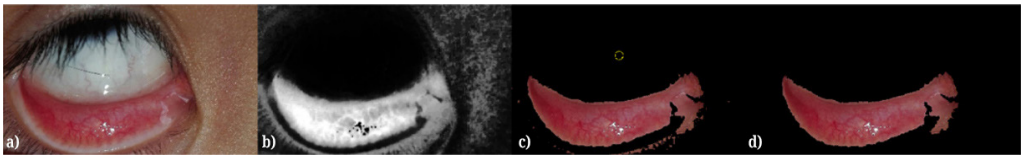

Fig. 4. Conjunctive segmentation step by step (a) Initial cropped image, (b) The product of equalized channels Cr, Y, and Q, (c) The output of Otsu filter applied to the b and the center of the cropped image, (d) The final result of conjunctive segmentation

Sticker segmentation. The sticker segmentation procedure, as illustrated in Figure 5, begins with the value channel, denoted as $I_v(x, y)$, from the HSV color space. To reduce the influence of high reflectance caused by the camera flash, an average low pass filter with a kernel size of 25 × 25, represented by $M(x, y)$, is applied to the value channel.

Next, a thresholding operation is performed on the filtered value channel. Only pixels with values greater than 0.8 are selected, as the sticker is white and thus have a high luminance value. This thresholding process generates a preliminary mask, denoted as $I_{pm}(x, y)$, where each pixel is assigned a value of 1 if it satisfies the threshold condition, and 0 otherwise, as described in (12). The resulting pre-mask serves as an initial segmentation of the sticker region.

$$I_{pm}(x,y) = \begin{cases} 1 \, if \, I_v(x,y) * M(x,y) > 0.8 \\ 0 \qquad otherwise \end{cases} \tag{12}$$

After filtering, the image may still show large areas of high reflectance. To solve this problem, contours are detected using the running squares algorithm, based on the running cubes algorithm [23]. From all detected contours, only the three largest and closed contours ($I_{3BC}(x, y)$) are selected for further processing.

Among these selected contours, those that contain other contours are considered to represent the complete sticker. This contour is denoted as $I_{pre-sticker}(x, y)$. However, to ensure a squared shape an analysis of shape is performed by using the PCA algorithm [24] over all centered n points of the i selected contour, this is apply PCA over $X_{c,i} \in R^n$ (Vector of x coordinates) and $Y_{c,i} \in R^n$ (Vector of y coordinates) projected over two components for $i = 1, ..., n$ where n is number of detected contours. The procedure is showed in (13) and (14), this process is made with all the found contours. The result is called $X_{r,i}$ and $Y_{r,i}$ referring to rectified coordinates of each contour.

$$\begin{bmatrix} X_{c,i} & Y_{c,i} \end{bmatrix} \begin{bmatrix} X_{c,i}^T \\ Y_{c,i}^T \end{bmatrix} = Q \Lambda Q^{-1} \tag{13}$$

$$\begin{bmatrix} X_{r,i}^T \\ Y_{r,i}^T \end{bmatrix} = Q \begin{bmatrix} X_{c,i}^T \\ Y_{c,i}^T \end{bmatrix} \tag{14}$$

The area of the contour, that is the number of pixels inside the contour N_{pixels} is compared with a rectangle with side lengths defined by the maximun and minimun values in $X_{r,i}$ and $Y_{r,i}$ for each contour.

After that, the contour with the highest $Similarity_i$ with a rectangle is selected as the contour of the sticker ($I_{sticker}(x, y)$), this is the mask of that detect the sticker in the image, the similarity is computed with (15). The input and output of this process is depicted in Figure 6.

$$Similarity_i = \frac{\left(\max(X_{r,i}) - min(X_{r,i})\right) \cdot \left(\max(Y_{r,i}) - \min(Y_{r,i})\right)}{N_{pixels}} \tag{15}$$

2.2 Image preprocessing

After the segmentation algorithms, a preprocessing step is performed to improve the data quality for subsequent machine learning models. This preprocessing step is illustrated by the blocks contained within the red dotted line in Figure 1.

One important aspect of the preprocessing is the adjustment of luminance in the cropped area obtained from the conjunctiva segmentation. This adjustment is done using Equation (16), where $I_{conjunctiva}(x, y)$ is the conjunctiva palpebral segmentation, γ represents the white mean value of $I_{sticker}(x, y)$ and $I_{adjusted}(x, y)$ is the conjunctiva illuminance adjustment to normalize before entering the SLIC algorithm. This adjustment process is based on the method described in [15] for image adjustment.

$$I_{adjusted}(x, y) = I_{conjunctiva}(x, y) \cdot \frac{\gamma}{200} \tag{16}$$

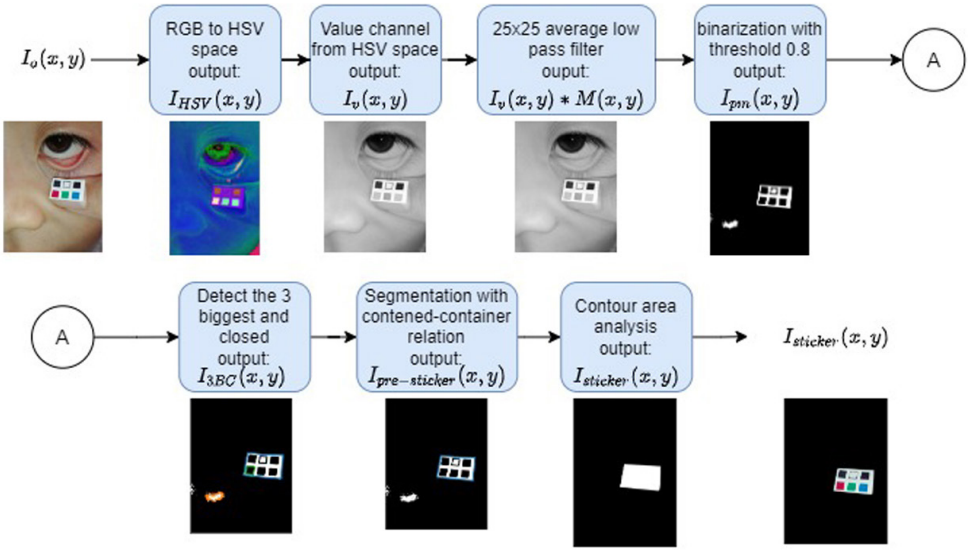

Fig. 5. Sticker segmentation workflow

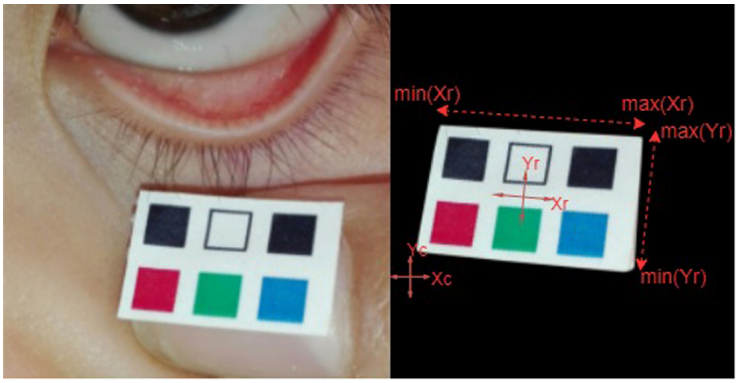

Fig. 6. Sticker segmentation result Left image: the original captured image; Right image: the segmented sticker, red marks depict some of the quantities used in the rectangle similarity computation

After the reflectance adjustment, the SLIC algorithm [25] is applied to divide the conjunctiva area into 20 segments. An additional channel is stacked to create a four-channel image $I_{SLIC}(x, y)$, where the extra channel represents the map of segments. Within each segment, the darkest and brightest pixels are filtered by selecting only those pixels for which the three channels of the RGB color space fall within the range of 0.3921 to 0.7843 in the scale of 0 to 1. This filtering process results in an image with filtered values, denoted as $I_{f\text{-}SLIC}$. Next, statistical descriptors are extracted from each segment. Through multiple experiments, specific statistical descriptors that demonstrate the best performance for hemoglobin estimation have

been identified. The subscripts L, a, x, and y refer to the L and a channels in the Lab color space, as well as the spatial coordinates x and y.

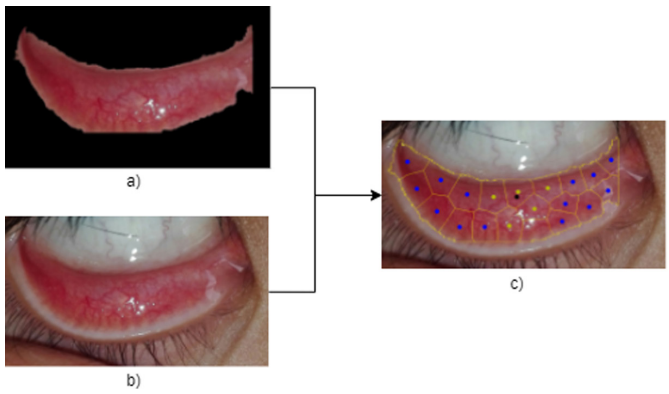

Fig. 7. (a) Conjunctiva mask as input for preprocessing. (b) Original image. (c) Referential image of result after preprocessing method (Blue points shows the centers of each sub-segment. Yellow points shows the centers of the selected sub-segments for posterior statistical processing. Black point shows the center of all the ROI.)

Using the data per segment, a graph structure $G(V, E)$ is constructed, where each node contains a vector of the statistical descriptors, and the edges are determined by the adjacency of the segments. Finally, considering the empirical experience in the medical field that indicates a higher correlation between the central zone of the conjunctiva and hemoglobin estimation, only the central nodes ($G_c(V, E)$) are selected for further processing. Figure 7 shows the result of the preprocessing steps.

2.3 Training

Three machine learning algorithms were developed and trained to estimate the hemoglobin level using the statistical descriptors obtained for each segment. These are the standard deviation of channel L (σ_L), standard deviation of channel a (σ_a), mode of channel L (Mo_L), mode of channel a (Mo_a), mean of spatial coordinate x (μ_x) and mean of spatial coordinate y (μ_y). To validate our results, 5-fold cross validation is employed for each algorithm.

In the training process of the aforementioned models, data augmentation was applied to the dataset of 500 images in order to prevent overfitting and improve the performance of the models. A geometric transformation was applied to rotate the image by 45° until the complete rotation was completed. Therefore, a factor of ×8 was obtained with data augmentation, and a dataset of 4000 images was obtained. Then, the dataset of images with data augmentation was divided into train and test datasets with a ratio of 70% and 30%, respectively.

Bayesian ridge regression. The machine learning model of Bayes regression is based on conditional probabilities (Bayes theorem). As this model works well on small to medium datasets, it was considered a good option. The model uses the parameters shown in Table 2.

Table 2. List of parameters for Bayesian ridge regression

Bayesian Ridge Regression Parameters		
n_iter	Tol	Verbose
1000	0.0005	true

Support vector regression. As shown in [8], support vector regression (SVR) shows good throughput on tasks such as this. Therefore, a SVR model was tested using the grid search method to determine the best-performing parameters, which were the regulation parameter C = 1 and the radial basis function kernel.

SLIC-GAT. One of the machine learning models used for estimating hemoglobin levels is SLIC-GAT [26]. This model leverages deep learning techniques and incorporates graph attention network (GAT) layers to process the graph structures derived from SLIC segmentation. The graph structure $G_c(V, E)$, as explained in the Image Preprocessing section, remains consistent throughout the model.

After conducting extensive experiments, the optimal configuration for the GAT layers is presented in Table 3, while the multi-layer perceptron (MLP) layers, employed at the end of the GAT layers, adopt the parameter settings outlined in Table 4. The input to the model comprises the five most central nodes from the graph, each containing multiple values representing image features.

During the model training process, the mean square error between the measured and estimated hemoglobin levels is minimized. This optimization is achieved using the Adam algorithm [27] with a learning rate of 0.001, $\beta_1 = 0.9$, and $\beta_2 = 0.999$.

Table 3. List of parameters for model GAT layers with the best performance

GAT Layers Components	GAT Layer Parameters						
Layer Heads	1	2	5	8	10	5	2
Activation for layer heads	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU
GAT layer size	30	32	64	64	64	32	16

Table 4. List of parameters for MLP layers with the best performance

MLP Description	MLP Layer Parameters		
MLP layers	32	32	1
Activation for MLP layer	ReLU	ReLU	ReLU

3 RESULTS

As mentioned earlier, the focus of this work is on hemoglobin estimation. The results of the hemoglobin estimation task, obtained after five-fold training, are presented in Table 5. The evaluation metric used is L1 error, which measures the absolute difference between the predicted hemoglobin levels and the ground truth values.

Table 5. Hemoglobin estimation error of each model where L1 is absolute error

Model	Mean L1 error
SLIC-GAT	0.7284
Bayesian ridge regression	0.85
Support vector regression	0.87

The regression error for the SLIC-GAT model, as shown in Table 5, is 0.7284.

4 DISCUSSIONS

The regression error achieved by the SLIC-GAT model is only surpassed by the error obtained in [5], where hemoglobin levels are estimated through spectropic examination. It is important to note that this work focuses solely on images of the eye and a calibration sticker, captured using a mid-range cell phone without any significant additional devices.

It is important to consider certain characteristics of the dataset used in this work. Firstly, the quality of the images varied as they were captured without a standardized setup. This resulted in some images being noisy. However, effective techniques such as segmentation and brightness calibration were employed to address this issue and enhance the quality of the images.

Secondly, the dataset exhibited an imbalance, which is a common challenge in disease diagnosis. Specifically, the samples representing values outside the normal range were limited in number. As depicted in Figure 8, these extreme cases accounted for less than 6.48% of the total samples. In light of this imbalance, the regression analysis was specifically performed on samples within the range of 10 to 13 g/dL, ensuring a more focused and reliable estimation of hemoglobin levels.

Despite the challenges mentioned earlier, the hemoglobin estimation error achieved in this work surpasses the results obtained in the current state-of-the-art methods, considering the specific nature of the image source used. It is important to note that the images were captured using a mid-range cell phone without any significant additional devices, making it a more accessible and cost-effective approach. Furthermore, this work proposes a standardized process and guidelines for image capture, which can serve as a valuable reference for future implementations and research.

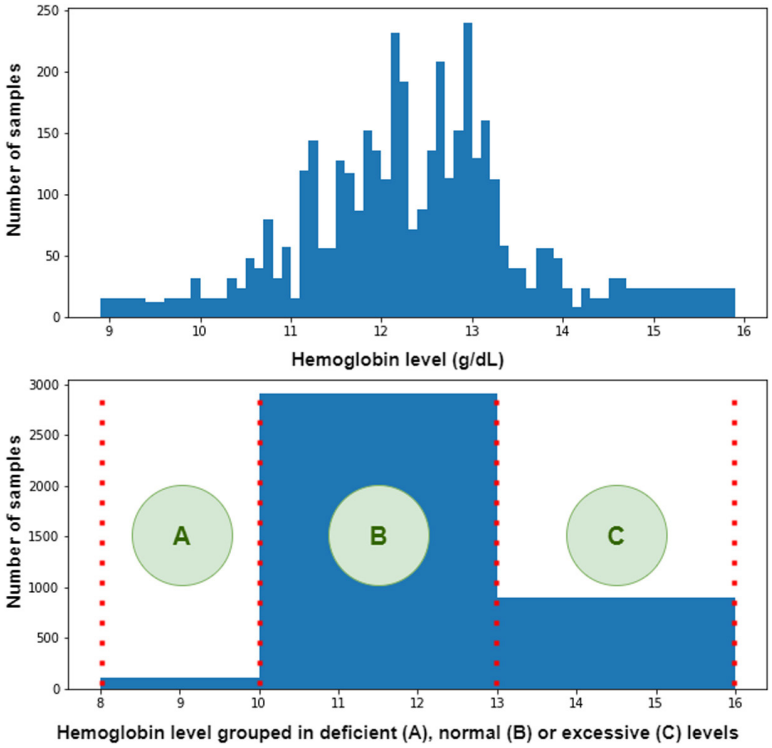

Fig. 8. (Upper) Distribution of samples with the hemoglobin level grouped by hemoglobin measurement resolution (Inferior) Distribution of samples grouped by hemoglobin ranges to diagnosis of anemia

5 CONCLUSIONS

The contributions of this work are methods for ROI segmentation of the palpebral conjunctiva and hemoglobin estimation for a dataset of images from the palpebral conjunctiva and a calibration sticker captured with a mid-range cell phone. The results of the segmentation task were evaluated using the entire dataset. To ensure the effectiveness of the segmentations in future works, the following image-capturing standard is proposed:

- Reduction of the reflections in the image
- Image must contain only the eye and the sticker
- Images should be taken with the child's eye aligned frontally toward the camera

The results of the segmentation are important since they enhance the quality of the images for the next task.

The hemoglobin estimation task was formulated as a machine learning regression problem, and multiple models were implemented using the enhanced image as input in graph form. Among these models, the best result was an L1-norm error of 0.72 g/dL, which was achieved with SLIC-GAT.

6 FUNDINGS

This study was funded by the NATIONAL INNOVATION PROGRAM FOR COMPETITIVENESS AND PRODUCTIVITY of the Peruvian Ministry of Production (CONVENIO No 556-INNOVATE PERU-ISDA-2017), and the research support program of the National University of Engineering of Peru (FIM-PF-35-2021).

7 REFERENCES

[1] World Health Organization, "Anaemia," *Who.int*. [Online]. Available: https://www.who.int/health-topics/anaemia. [Consulted: 01-Apr-2024].

[2] M. M. Sirdah, A. Yaghi, and A. R. Yaghi, "Iron deficiency anemia among kindergarten children living in the marginalized areas of Gaza Strip Palestine," *Rev. Bras. Hematol. Hemoter.*, vol. 36, no. 2, pp. 132–138, 2014. https://doi.org/10.5581/1516-8484.20140030

[3] B. De Benoist, M. Cogswell, I. Egli, and E. McLean, "Worldwide prevalence of anemia 1993–2005; WHO Global Database of anemia," *Centers for Disease Control and Prevention*, 2008. [Online]. Available: https://stacks.cdc.gov/view/cdc/5351. [Consulted: 01-Apr-2024].

[4] O. Kim, J. McMurdy, G. Jay, C. Lines, G. Crawford, and M. Alber, "Combined reflectance spectroscopy and stochastic modeling approach for noninvasive hemoglobin determination via palpebral conjunctiva," *Physiol. Rep.*, vol. 2, no. 1, p. e00192, 2014. https://doi.org/10.1002/phy2.192

[5] J. W. McMurdy, G. D. Jay, S. Suner, F. M. Trespalacios, and G. P. Crawford, "Diffuse reflectance spectra of the palpebral conjunctiva and its utility as a noninvasive indicator of total hemoglobin," *J. Biomed. Opt.*, vol. 11, no. 1, p. 014019, 2006. https://doi.org/10.1117/1.2167967

[6] Y. M. Chen and S. G. Miaou, "A Kalman filtering and nonlinear penalty regression approach for noninvasive anemia detection with palpebral conjunctiva images," *J. Healthc. Eng.*, vol. 2017, pp. 1–11, 2017. https://doi.org/10.1155/2017/9580385

[7] R. Muthalagu, V. T. Bai, and S. John, "A smart (phone) solution: An effective tool for Screening Anaemia-Correlation with conjunctiva pallor and haemoglobin levels," *TAGA J.*, vol. 14, pp. 2611–2621, 2018.

[8] N. B. Noor, M. S. Anwar, and M. Dey, "Comparative study between decision tree, SVM and KNN to predict anaemic condition," in *IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON)*, Dhaka, Bangladesh, 2019, pp. 24–28. https://doi.org/10.1109/BECITHCON48839.2019.9063188

[9] S. M. Park *et al.*, "mHealth spectroscopy of blood hemoglobin with spectral super-resolution," *Optica*, vol. 7, no. 6, pp. 563–573, 2020. https://doi.org/10.1364/OPTICA.390409

[10] S. Roychowdhury, D. Sun, M. Bihis, J. Ren, P. Hage, and H. H. Rahman, "Computer aided detection of anemia-like pallor," in *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Orlando, FL, USA, 2017, pp. 461–464. https://doi.org/10.1109/BHI.2017.7897305

[11] R. G. Mannino *et al.*, "Smartphone app for non-invasive detection of anemia using only patient-sourced photos," *Nat. Commun.*, vol. 9, no. 1, pp. 1–10, 2018. https://doi.org/10.1038/s41467-018-07262-2

[12] P. Jain, S. Bauskar, and M. Gyanchandani, "Neural network based non-invasive method to detect anemia from images of eye conjunctiva," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 1, pp. 112–125, 2019. https://doi.org/10.1002/ima.22359

[13] Y. M. Chen, S. G. Miaou, and H. Bian, "Examining palpebral conjunctiva for anemia assessment with image processing methods," *Comput. Methods Programs Biomed.*, vol. 137, pp. 125–135, 2016. https://doi.org/10.1016/j.cmpb.2016.08.025

[14] G. Dimauro, A. Guarini, D. Caivano, F. Girardi, C. Pasciolla, and A. Iacobazzi, "Detecting clinical signs of anaemia from digital images of the palpebral conjunctiva," *IEEE Access*, vol. 7, pp. 113488–113498, 2019. https://doi.org/10.1109/ACCESS.2019.2932274

[15] B. Saldivar, D. Núñez, F. Porras, A. Alva, L. S. Leslie, and M. Zimic, "Portable system for the prediction of anemia based on the ocular conjunctiva using Artificial Intelligence," *arXiv preprint arXiv:1910.12399*, 2019. https://doi.org/10.48550/arXiv.1910.12399

[16] A. Tamir *et al.*, "Detection of anemia from image of the anterior conjunctiva of the eye by image processing and thresholding," in *IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, Dhaka, Bangladesh, 2017, pp. 697–701. https://doi.org/10.1109/R10-HTC.2017.8289053

[17] G. Dimauro, L. Baldari, D. Caivano, G. Colucci, and F. Girardi, "Automatic segmentation of relevant sections of the conjunctiva for non-invasive anemia detection," in *3rd International Conference on Smart and Sustainable Technologies (SpliTech)*, Split, Croatia, 2018, pp. 1–5.

[18] P. Appiahene *et al.*, "Detection of anemia using conjunctiva images: A smartphone application approach," *Med. Nov. Technol. Devices*, vol. 18, no. 100237, p. 100237, 2023. https://doi.org/10.1016/j.medntd.2023.100237

[19] T. N. Sheth, N. K. Choudhry, M. Bowes, and A. S. Detsky, "The relation of conjunctival pallor to the presence of anemia," *J. Gen. Intern. Med.*, vol. 12, no. 2, pp. 102–106, 1997. https://doi.org/10.1046/j.1525-1497.1997.00014.x

[20] J. Suryaprasad, D. Sandesh, V. Saraswathi, D. Swathi, and S. Manjunath, "Real time drowsy driver detection using haarcascade samples," in *Computer Science & Information Technology (CS & IT)*, vol. 45, 2013, p. 54. https://doi.org/10.5121/csit.2013.3805

[21] A. K. Jain, *Fundamentals of Digital Image Processing*. Upper Saddle River, NJ, Estados Unidos de América: Pearson, 1989.

[22] A. Nicolaou, R. Ingold, and M. Liwicki, "Binarization with the local Otsu filter: Integral histograms for document image analysis," in *Graphics Recognition. Current Trends and Challenges*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 176–190. https://doi.org/10.1007/978-3-662-44854-0_14

[23] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *Comput. Graph. (ACM)*, vol. 21, no. 4, pp. 163–169, 1987. https://doi.org/10.1145/37402.37422

[24] T. Minka, "Automatic choice of dimensionality for PCA," in *Proceedings of the 13th International Conference on Neural Information Processing Systems*, 2000, pp. 577–583.

[25] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels," School of Computer and Communications Sciences, EPFL Technical Report 149300, 2010.

[26] P. H. C. Avelar, A. R. Tavares, T. L. T. da Silveira, C. R. Jung, and L. C. Lamb, "Superpixel image classification with graph attention networks," in *33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, Porto de Galinhas, Brazil, 2020, pp. 203–209. https://doi.org/10.1109/SIBGRAPI51738.2020.00035

[27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization", *arXiv preprint arXiv:1412.6980 [cs.LG]*, 2014. https://doi.org/10.48550/arXiv.1412.6980

8 AUTHORS

Guillermo Moreno is a bachelor's graduate in Mechatronics Engineering from the National University of Engineering (UNI), Lima, Peru.

Abdigal Camargo is a bachelor's graduate in Mechatronics Engineering from the National University of Engineering (UNI) in Lima, Peru, and a master's candidate at the Pontifical Catholic University of Rio de Janeiro (PUC-RIO), Río de Janeiro, Brazil.

Luis Ayala is a bachelor's graduate in Electronic Engineering from the Peruvian University of Applied Sciences (UPC), Lima, Peru. His research interests include image processing and artificial intelligence, mainly in machine learning, computer vision and natural language processing.

Mirko Zimic, PhD degree in Control and Prevention of Diseases from Johns Hopkins University, MSc. Biochemistry from the Peruvian University Cayetano Heredia (UPCH) and also obtained his BSc. Physics at the National University of Engineering. Currently, he is the head of the Laboratory of Bioinformatics and Molecular Biology of the UPCH and leads a multidisciplinary working group, with whom he researches in different areas of science and technology.

Christian del Carpio received B.S. degree in Electrical Engineering in 2005 from the San Martin de Porres Private University (USMP), Lima, Peru. In the year 2015, he obtained a master's degree in science, from National University of Engineering (UNI), Peru. He is working as a research professor for the undergraduate and postgraduate programs of the schools of Electrical Engineering and the school of Mechatronics Engineering from the National University of Engineering (UNI), Lima, Peru. His research interests include image and signal processing (E-mail: cdelcarpiod@uni.edu.pe).