Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 14 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i14.50141

PAPER

Non-Invasive Detection of Iron Deficiency Anemia in Young Adults Through Finger-Tip Video Image Analysis

Miguel Angel Valles-Coral(✉), Jorge Raul Navarro-Cabrera, Lloy Pinedo, Richard Injante, Luz Karen Quintanilla-Morales, María Elena Farro-Roque

Universidad Nacional de San Martín, Tarapoto, Perú

mavalles@unsm.edu.pe

ABSTRACT

In this study, we propose the use of a convolutional neural network (CNN) (2+1) D-based model for the non-invasive detection of iron deficiency anemia using smartphones, addressing the need for accessible and efficient diagnostic methods. We have collected fingertip images from 909 young people, creating a robust dataset to train and validate the model through machine learning (ML), providing us with both spatial and temporal information. Our approach stands out for its simplicity and potential to be implemented in diverse contexts, which facilitates early and accurate diagnosis of anemia. The results show remarkable performance indices in the validation set: an accuracy of 0.9840, a precision of 0.9830, a sensitivity (recall) of 0.9840, an F1-Score of 0.9835, and an AUC ROC of 0.9878, demonstrating the model's high ability to effectively classify anemic states. These findings validate the efficacy of the model to significantly improve the detection of iron deficiency anemia, promoting the use of mobile technologies and machine learning for more accessible and efficient diagnosis globally. This breakthrough marks a significant step towards improving access and efficiency in health diagnostics through smartphone technology and machine learning, opening new avenues for the effective management of anemia and other health conditions.

KEYWORDS

artificial intelligence, convolutional neural network (CNN), iron deficiency anemia detection, non-invasive technique, supervised learning

1 INTRODUCTION

According to Stevens et al. [1], it is estimated that a quarter of the world's population is anemic, mostly due to iron deficiency. Camaschella and Girelli [2] mention that iron deficiency anemia is the final stage of a prolonged period of negative iron balance in the body. The World Health Organization [3] defines anemia as a disorder in which the number of red blood cells (and, consequently, the oxygen-carrying capacity of the blood) is insufficient to meet the body's needs.

Valles-Coral, M.A., Navarro-Cabrera, J.R., Pinedo, L., Injante, R., Quintanilla-Morales, L.K., Farro-Roque, M.E. (2024). Non-Invasive Detection of Iron Deficiency Anemia in Young Adults Through Finger-Tip Video Image Analysis. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(14), pp. 53–70. https://doi.org/10.3991/ijoe.v20i14.50141

Article submitted 2024-05-15. Revision uploaded 2024-09-06. Final acceptance 2024-09-06.

The United Nations [4] approved the 2030 Agenda for Sustainable Development, structured in 17 goals focused on improving the quality of life of the entire world population. Specifically, goals two and three seek to ensure healthy lives, promote well-being, ensure food security, promote efficient technological initiatives that completely eradicate diseases such as anemia, and address numerous and varied persistent and emerging health issues.

In Peru, the National Center for Strategic Planning [5] aims to achieve the goal of covering food needs and adequate nutrition by 2050 in order to minimize the rate of chronic malnutrition and anemia in the population. In this context, the San Martin Regional Government [6] has been developing strategies to guarantee a healthy population and is implementing the Regional Development Plan, which emphasizes actions to improve health coverage levels in order to improve the quality of life of the population.

However, monitoring these goals is a complex task since the communication channels are slow and the logistic resources assigned to the entities in charge are limited [7]. Considering that the San Martin region is a marginal urban and rural area where the highest percentages of people with this morbidity problem are concentrated [8]. The complexity of the problem is increased by the evidence collected in the field, which shows how most health centers, clinics, laboratories, and hospitals in the region rely exclusively on invasive devices for the measurement of blood hemoglobin. This invariably requires at least one puncture. On the other hand, there is a notable absence in the adoption of noninvasive devices, which eliminate the need for puncture. This situation is mainly attributed to the high costs associated with their acquisition and the demands of specialized technical care.

The field review allowed us to confirm that invasive methods provoke fear and rejection in the majority of young university students. According to studies by Cardoso-Sánchez et al. [9] and Brimson et al. [10], these groups experience significant changes in their eating habits, adopting inadequate and not very varied diets. This contributes to a prevalence of mild anemia of 68.71% and moderate anemia of 30.03%, according to Ysihuaylas Blas [11]. Such conditions can lead to symptoms such as inattention, drowsiness, and weakness, negatively affecting academic performance [12].

Ensuring that university students maintain a healthy diet to prevent anemia presents difficulties. Among other factors, many are unaware of their hemoglobin levels due to the need to go to specialized centers for the collection and analysis of blood samples to determine if they are anemic. In addition, the time required for sample processing, delivery of results, and especially the high costs associated with laboratory testing further complicate this process [13].

Therefore, considering that 89.4% of Peruvian university students own smartphones [14] and that these devices have operating systems with advanced features and functionalities [15] [16], their use as tools to detect and prevent possible diseases has the potential to improve access to and quality of health services [17] [18] [19]. In this regard, Dominguez Miranda and Rodriguez Aguilar [20]; Sembay and Jeronimo de Macedo [21]; Sotillos-González et al. [22]; Abdul-Jabbar et al. [23] emphasize the concern about the guarantees that should accompany these technologies, and in particular mobile health applications, whether in terms of security, privacy, reliability, impact, and effectiveness.

Based on the rationale, we propose the development of a non-invasive screening model for iron deficiency anemia using smartphones and machine learning to be used on an outpatient basis and have a high level of confidence in its results. This could allow the implementation of strategies in the university population to reduce the prevalence of this condition. In this context, we rely on Brimson et al. [10] who state that if young college students have the knowledge and understanding of their underlying health problems or diseases, they themselves can prevent serious health conditions and improve their college academic performance.

2 RELATED JOBS

In the last decade, the area of biomedical data analysis has seen a significant amount of research that has paved the way for current studies. Therefore, this segment addresses the noninvasive methods that have previously been used to detect anemia.

Magdalena et al. [24] developed a non-invasive method for anemia detection by imaging the palpebral conjunctiva using a convolutional neural network (CNN). The CNN, designed with five hidden layers and filters of various sizes (3×3 to 11×11), aims to identify specific features to distinguish between normal and anemic conditions. With a sample of 2000 images, divided into 1440 for training, 160 for validation, and 400 for testing, the model achieved an accuracy of 94% and averages of precision, recall, and F1 score of 0.935, 0.94, and 0.935, respectively. The results suggest that CNN can effectively classify anemic conditions with minimal errors, offering a potential system for implementation in Android mobile applications.

Appiahene et al. [25] introduced a non-invasive method to detect anemia through digital palpable palm images collected from 710 participants. Using a machine learning (ML) approach in R Studio, the images were processed and analyzed using ensemble techniques, including stacking, voting, boosting, and bagging. The evaluation of hybrid models was conducted using a number of metrics, including accuracy, sensitivity, specificity, F1 score, precision, and area under the curve (AUC). Among the algorithms employed, the stacking model, which combined Naive Bayes (NB) as base learners with Random Forest (RF), Decision Tree (DT), support vector machine (SVM), artificial neural networks (ANN), and NB as stacking learners, excelled by achieving an accuracy of 99.73%.

In another study, Appiahene et al. [26] proposed a non-invasive method to detect iron deficiency anemia using palm images and compared ML algorithms: CNN, k-NN, NB, SVM, and DT. With an initial sample of 527 images, enlarged to 2635 by augmentation techniques such as rotation, flipping, and translation, the augmented data set was randomly divided into proportions of 70% for training, 10% for validation, and 20% for testing. The results showed that NB achieved the highest accuracy with 99.96%, while SVM had the lowest accuracy with 96.34%. The efficacy of these models was evaluated with metrics such as recall, precision, F1 score, and AUC, using a 10-fold cross-validation of 10% of the total data set.

Williams Asare et al. [27] detected iron deficiency anemia using a ML approach, applying NB, CNN, SVM, k-NN, and DT algorithms to images of eye conjunctiva, palpable palmar, and nail color to determine which offers higher accuracy in detecting anemia in children. Through three stages: data collection, preprocessing, and model development, the aforementioned algorithms were evaluated using metrics of accuracy, sensitivity, specificity, and F1 score. The CNN showed the highest accuracy (99.12%), standing out over the other models, while the SVM recorded the lowest accuracy (95.4%).

Mahmud et al. [28] focused on anemia detection by noninvasive analysis of labial mucosa images, using ML algorithms such as ANN, DT, KNN, LR, NB, and SVM to classify data from 138 patients, with features extracted based on RGB red color values, HSV saturation values, age, sex, and hemoglobin levels. They evaluated the efficacy of ML models using metrics such as accuracy, sensitivity, specificity, and F-score, finding that NB achieved the highest accuracy with 96%, followed by DT, KNN, and ANN with 93%, while LR and SVM presented accuracy of 79% and 75%, respectively.

Das et al. [29] proposed a system for non-invasive detection of anemia by analyzing nail pallor using a low-cost device coupled with a smartphone. The method involves inducing color changes in the nail bed by applying and releasing pressure with a customized device, followed by analysis of a video captured by a smartphone camera to measure the rate of color change. This information correlates with clinical blood hemoglobin values to quantify the hemoglobin level. With 220 subjects evaluated, the proposed model outperforms existing solutions, offering a mean RMSE and MSE of 0.63 and 0.61, respectively, with a standard deviation of 0.46 g/dL and 0.73 g/dL. This intelligent anemia care system predicts hemoglobin level with high sensitivity (0.96), specificity (0.68), and accuracy (89%), at a significantly lower cost than existing devices.

Kesarwani et al. [30] focused on the development of a non-invasive system, combining advanced computational techniques with the traditional practice of estimating blood hemoglobin levels by observing palpitation in the palm of the hand. Using a custom device to induce changes in palm color by applying and releasing pressure and capturing video of these changes with a smartphone camera, the system measures the rate of color changes and performs time domain analysis to correlate with blood hemoglobin concentration.

For video processing and analysis, deep learning models based on a tree-structured three-dimensional convolutional neural network (3D CNN) and the visual transformer (ViT) model are employed, achieving a sensitivity, specificity, accuracy, and RMSE of 96.87%, 90.90%, 94.44%, and 0.495, respectively, in a sample of 531 individuals.

In another recent study, Kesarwani et al. [31] proposed a non-invasive anemia detection method that combines time-domain analysis with the use of pre-trained deep learning models and vision transformers for feature extraction and classification or regression of hemoglobin levels based on color changes induced in the palm. This approach leverages the power of pre-trained models for feature extraction and the flexibility of transform encoders and MLP networks for efficient and effective representation learning. Two fusion approaches are proposed: decision-level and feature-level, achieving RMSE and accuracy of 0.483 and 96.296%.

Most previous approaches in anemia detection have relied on image analysis of conjunctiva and fingernails and have applied ML techniques to develop both invasive and non-invasive methods. However, there remains a need to refine the performance of these methodologies by integrating new datasets and adopting more efficient and simplified procedures. In addition, there is a notable paucity of studies focused on young populations, who also present significant susceptibility to anemia, although few investigations have directed their efforts towards this demographic group. In this context, our study proposes to expand the horizon using a (2+1) D CNN model, positioning itself as one of the first studies to apply this advanced ML technique for anemia detection in young people, aiming to fill this study gap and offer innovative and effective solutions for this vulnerable group.

3 PROPOSED METHOD

In this section, we describe the steps and processes employed in our study, organizing the proposed methodology into four main phases: (1) sample collection, (2) preprocessing, (3) processing, and (4) model validation.

For sample collection, we gathered our dataset at the University Medical Center of the National University of San Martin in Peru. We trained health professionals to use the Rad-67 device for non-invasive sample collection, which allowed us to efficiently measure hemoglobin (Hb) levels. Simultaneously, another team was involved in recording medical videos using a smartphone, which were then sent to a database for easy access and preliminary processing.

During the preprocessing stage, we extracted frames from the videos and performed cleaning and formatting procedures to obtain a structured dataset. Then, in the processing stage, we implemented a (2+1) D CNN using a ResNet18 architecture, which incorporates residual blocks to simplify training and improve the generalization capability of the model.

This system took advantage of both visual features of the videos and symptomatic data from participants to increase predictive accuracy.

Finally, in the validation phase, we conducted a continuous evaluation of the model by recording accuracy and loss metrics on the training and validation sets. We use the ModelCheckpoint technique to preserve the optimal state of the model during training. To confirm the robustness of the model, we apply additional metrics after training, thus ensuring a complete performance evaluation.

Figure 1 presents the methodology, outlining the workflow and detailing the interconnections between the different stages of the research process: data collection, preprocessing, processing, and validation. This block diagram is instrumental in understanding the sequence and technical approach applied in the study.

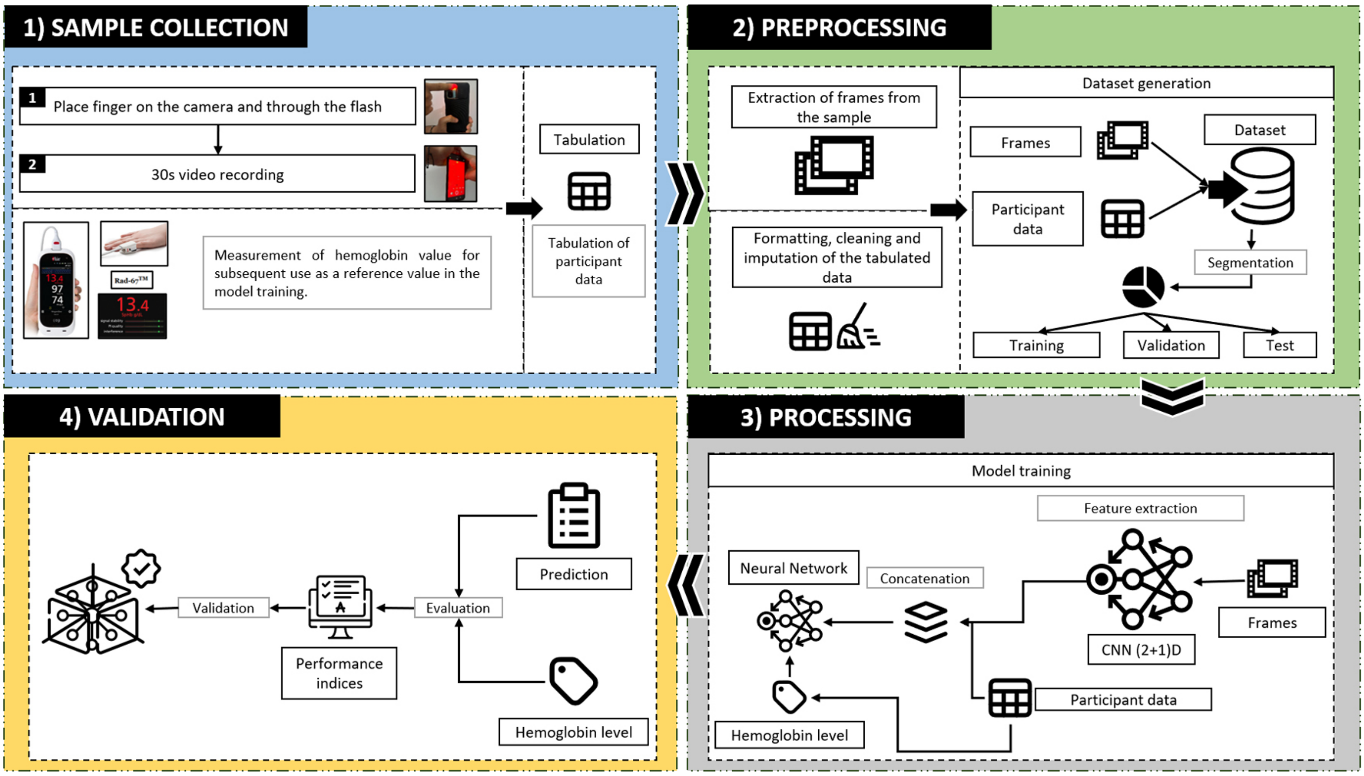

Fig. 1. Conceptual framework of the proposed methodology

3.1 Sample collection

Participants. We developed a specific protocol to collect patient information, focusing on demographic variables such as age, which for this study was limited to a range of 18 to 25 years, and sex of the participants. In addition, we included a list of symptoms associated with iron deficiency anemia that participants had experienced during the past month, including fatigue, weakness, palpitations, dyspnea (shortness of breath), dizziness or lightheadedness, angina (chest pain), cold extremities, and cephalalgia (headache).

Additionally, we recorded anthropometric data, including weight, height, body mass index (BMI), and abdominal circumference. In parallel, we performed objective measurements using the Rad-67 device that quantified key physiological parameters, including heart rate, oxygen saturation (SpO2), and Hb levels. Based on the Hb values obtained, we classified the anemia status of the participants using the values standardized by the World Health Organization [32]: in females, Hb ≥ 12 g/dL indicates no anemia (normal), 11.9–11 g/dL mild anemia, 10.9–8 g/dL moderate anemia, and <8 g/dL severe anemia; in males, Hb ≥ 13 g/dL represents no anemia (normal), 12.9–11 g/dL mild anemia, 10.9–8 g/dL moderate anemia, and <8 g/dL severe anemia. This protocol allowed us to add pertinent clinical annotations to the profile of each participant according to the established ranges. We were able to collect 909 profiles, 540 females and 369 males; Table 1 shows their distribution.

Table 1. Participant profile values

Gender	Hemoglobin Level				Anemia Prevalence (%)
	Normal	Mild	Moderate	Severe	
Female	287	179	74	0	46.85
Male	335	33	1	0	9.21

Image collection. We recorded 30-second video sequences, focusing on the left index finger of each participant (see Figure 2A), to collect the images. We kept this protocol constant throughout the study to ensure data homogeneity. We used the Samsung Galaxy A73 5G main camera, which has a resolution of 108 MP, f/1.8 aperture, 1/1.52″ sensor size, and 0.7 μm pixels, to capture the images. The camera includes optical image stabilization (OIS) technology and phase detection autofocus (PDAF). However, to optimize image quality, we manually set the parameters: ISO 250, shutter speed of 1/60, focus at 0.0, and white balance at 4400K. We instructed the participants to keep their finger still and not to exert pressure on the lens during the recording to ensure the accuracy of the measurements. In addition, we performed the sampling inside a polycarbonate box specially conditioned to nullify any possible interference from ambient light. We controlled the camera operation through external software (Figure 2B), allowing us a remote and standardized control of the smartphone during the sampling process.

(A) (B)

Fig. 2. Sample collection method

3.2 Preprocessing

We started this phase with the frame segmentation protocol, from which we created the dataset using the quantitative information obtained from the participants. We tabulated this information together with their biometric data, especially the hemoglobin value, which we used as a reference during training and validation of the model. We then subjected the dataset to a formatting, cleaning, and imputation process to correct errors or missing values derived from the previous tabulation, preparing it for statistical analysis and machine learning. With these data prepared, we generated a dataset that we divided into designated sets for training, validation, and model testing in proportions of 70%, 20%, and 10%, respectively.

Frame segmentation protocol. After collecting videos from the participants, we established a frame segmentation protocol to increase the amount of data available to train the (2+1) D convolutional model. We also maintained a moderate frames per second (FPS) rate to avoid overloading the model during training.

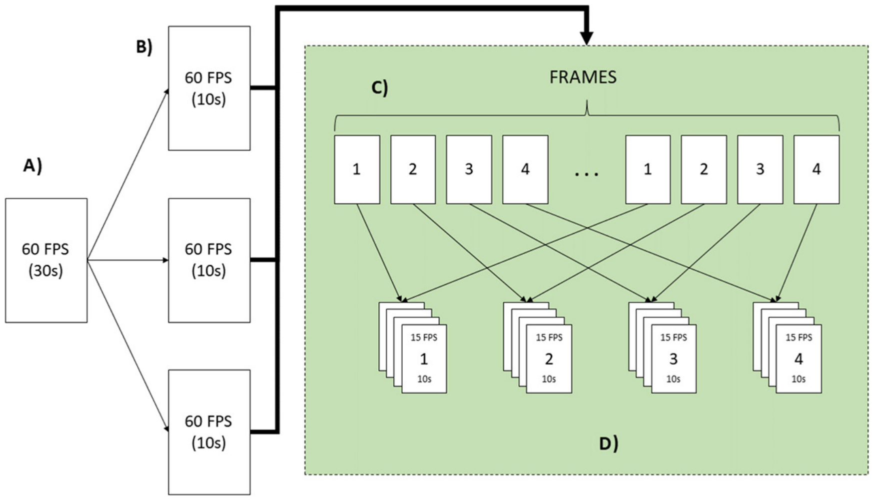

Fig. 3. Frame segmentation protocol scheme

We started the protocol with a 30-second video sample at a rate of 60 FPS (see Figure 3A). We divided this sample into three segments of 10 seconds each, maintaining the same initial FPS rate (see Figure 3B). Each video segment was then processed to dissect it into frames (see Figure 3C), assigning these frames to a specific group according to their sequence by numbering every fourth frame. This resulted in 10-second video files at a refresh rate of 15 FPS (see Figure 3D). This preprocessing method succeeded in generating 12 10-second video files at 15 FPS for each original sample.

After preprocessing, we obtained the following numbers of resulting images for each class: 7464 for the "Normal" class, 2544 for "Mild anemia," and 900 for "Moderate anemia." Given the imbalance in the number of samples per class, it was necessary to limit the number of samples in the most abundant classes, "Normal" and "Mild anemia.". Thus, their numbers were randomly adjusted to 2000, 2100, and 900 for the "Normal," "Mild anemia," and "Moderate anemia" classes, respectively. Figure 4 shows the selection of random frames from the available videos within the dataset, representing different levels of intensity in the red channel values of the images.

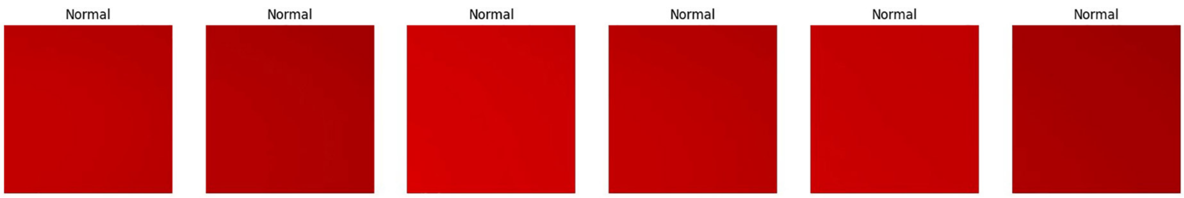

Fig. 4. *(Continued)*

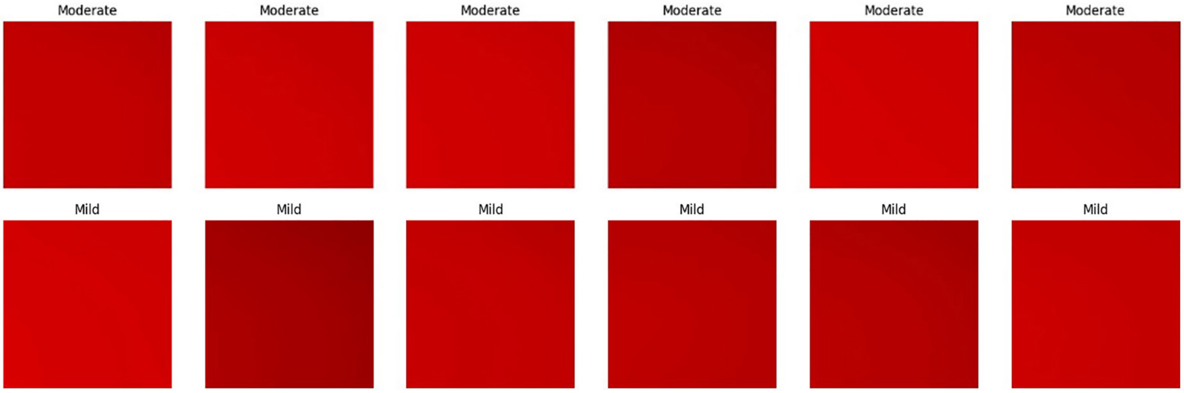

Fig. 4. Preview of the frames obtained after preprocessing

Then, the normalized values of this channel were extracted for each frame of the samples. Normalization allowed us to compare the data between different videos and ensure that they were on a uniform scale (from 0 to 1). We then calculated the mean pixel value in the red channel for each frame sequence. This mean value provided a representative measure of the brightness in that specific channel. Finally, graphs were generated to allow visualization of the variation of the values in the red channel pixels over time, as shown in Figure 5. Each row of the graph corresponds to a specific category (moderate, mild, or normal); the X axis shows the number of frames, while the Y axis shows the mean of the normalized values of the pixels of the red channel. These graphs made it possible to identify patterns, trends, and possible correlations with cardiac pulsations.

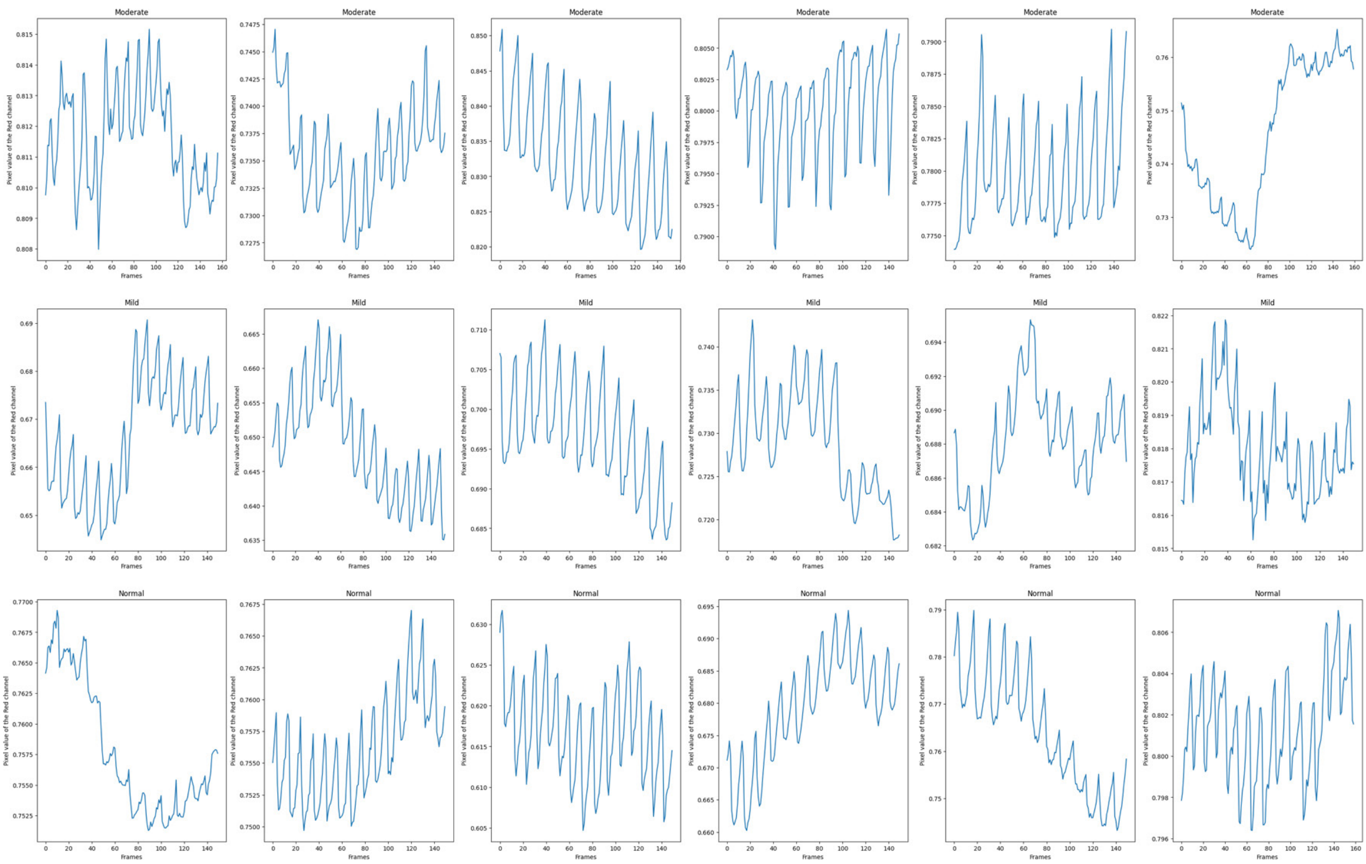

Fig. 5. Variation of the mean pixel value in the red channel as the frame sequence progresses

3.3 Processing

The implementation of our predictive model focused on the use of a CNN for the extraction of essential features from the frames. The main challenge resided in the handling of a dataset composed of both video files and still images, which led to the need to extract relevant information from the temporal dimension of these files. To overcome this obstacle, we adopted an innovative approach by using a (2+1) D convolutional layered model as proposed by Tran et al. [33]. This model is distinguished by combining the advantages of traditional convolutional layers with an improved ability to handle temporality, thus offering greater effectiveness and efficiency compared to traditional 3D convolutional solutions.

The processing of our proposed methodology begins with the extraction of the data extracted by the convolutional model (see Figure 6A), subsequently concatenated with the biometric information and the symptoms declared by the participant to generate a complete profile (see Figure 6B). We used this profile to train the neural network (see Figure 6C) using the baseline hemoglobin value taken with the Rad-67 device, with the aim of predicting the hemoglobin level.

Fig. 6. Complete architecture of the proposed model

For model training, we opted for the Adam algorithm [34] as an optimizer with a learning rate of 0.001 across 200 epochs and a training batch size of 32 units. As evaluation metric, we chose accuracy; this added to the ModelCheckpoint callback function that allowed us to store an improved version of the model based on the selected metric (accuracy) during the training process.

Figure 7 shows the training evolution of the proposed model, both in terms of accuracy and loss. We observe that the model responded in a great way with the training set, allowing uniform and constant learning throughout the epochs. However, this pattern was not replicated with the validation set, which only managed to stabilize around epoch 130 and then exhibited erratic behavior in subsequent epochs. Such behavior suggested that the number of epochs set for training may have been excessive, which highlighted the importance of implementing the ModelCheckpoint callback function during this phase to optimize the training process and avoid overfitting.

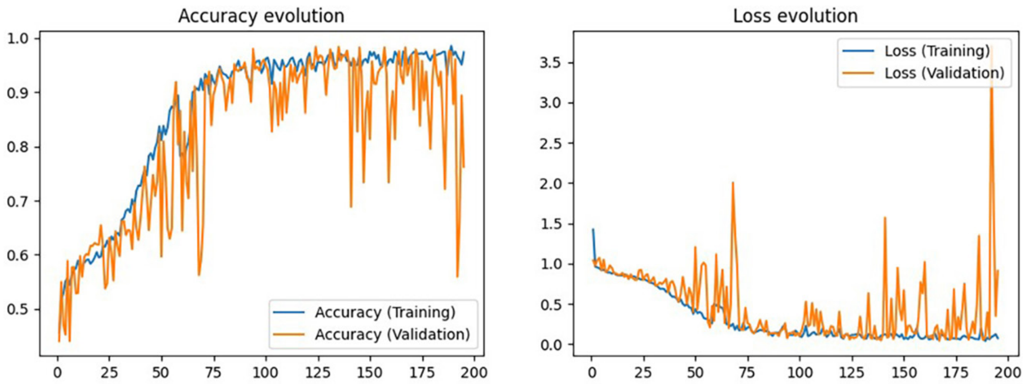

Fig. 7. Training evolution of the proposed model

3.4 Validation

With the help of the validation set we obtained at the end of the preprocessing; we subjected the model to a set of performance metrics in order to evaluate its ability to correctly classify a dataset with which it had not had any previous approach. In addition, this allowed us to evaluate the generalization capability of the proposed model, which was crucial to rule out any possibility of overfitting.

Evaluation metrics. Model evaluation plays a key role in quantifying the performance of a classifier or model in general. Its goal is to ensure that the relationships learned from the training dataset are applicable and effective on a validation or test dataset [35]. Based on the study by Appiahene et al. [25], we selected the following performance metrics:

Accuracy: measures the number of correct predictions made by the model as a percentage of the total predictions. Although it provides a quick overview of the effectiveness of the model, its reliability decreases in scenarios where there is a significant imbalance between classes.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision: calculates the percentage of positive predictions that are correct, i.e., the number of true positives divided by the sum of true positives and false positives.

It is especially valuable in contexts where the costs of false positives are high and it is important to ensure high reliability in identifying positive cases.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall: indicates the percentage of all true positive cases that the model manages to correctly identify. It is crucial in situations where it is vital to capture as many positive cases as possible, thus minimizing the incidence of false negatives. It is especially important in contexts were failing to detect a positive case has significant consequences.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1-Score: metric that harmonizes precision and recall into a single indicator, providing a measure of the effectiveness of the model in classifying with a trade-off between avoiding false positives and detecting all positive cases. It is particularly useful in scenarios with class imbalance, where a trade-off between precision and recall is needed to obtain a comprehensive assessment of performance.

$$F1 - Score = \frac{2(P * R)}{P + R} \quad (4)$$

Area Under the Curve (AUC): It represents the ability of the model to distinguish between the two classes and is calculated from the Receiver Operating Characteristic (ROC) curve. A higher AUC indicates a better discrimination capacity of the model.

$$AUC = \frac{TPR - TNR}{2} \quad (5)$$

4 RESULTS AND DISCUSSION

As part of the results, we highlight the usefulness of the confusion matrix to visually corroborate the performance of our model in the classification task, both with the validation set (see Figure 8A) and with the test set (see Figure 8B). This tool allowed us to intuitively assess the model's ability to correctly classify among the proposed classes.

Regarding the validation set, the results were: for Class 0, we achieved 180 correct predictions versus three incorrect ones; for Class 1, we achieved 415 correct predictions and seven incorrect ones; and for Class 2, the model predicted correctly 390 times, with only six failures. On the other hand, when evaluating the test set, we observed that for Class 0, there were 106 correct predictions and only one incorrect one; for Class 1, there were 206 correct predictions against four errors; and, finally, for Class 2, there were 178 correct predictions and five errors. These results, represented graphically in the confusion matrices, not only highlight the high sensitivity and specificity achieved by the proposed model but also demonstrate its accuracy in the classification task, confirming its efficacy in discriminating between the different classes evaluated.

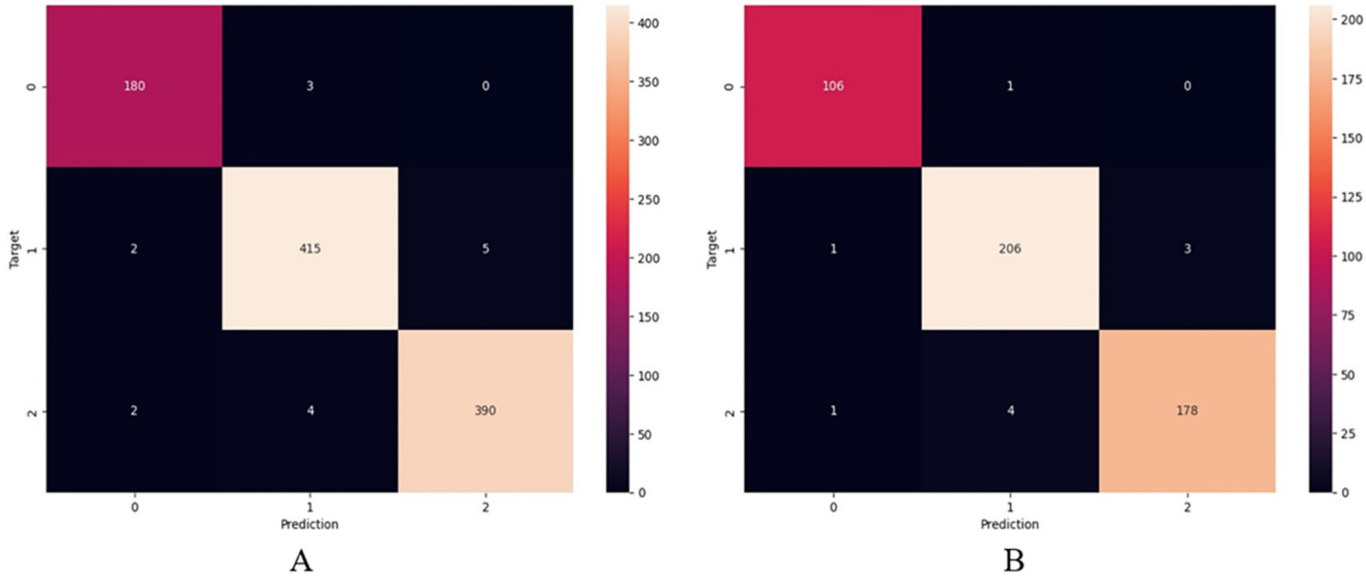

Fig. 8. Confusion matrix of the proposed model

The performance metrics presented for the validation set (refer to Table 2) and the test set (refer to Table 3) illustrate remarkable efficiency in classifying anemia across the algorithms employed, highlighting a remarkable consistency between the two sets. The accuracy and sensitivity (recall) above 98% in both sets indicate a high reliability of the model to correctly identify anemia cases, as well as its ability to minimize false negatives, which is critical in the medical context to ensure that anemic individuals are accurately identified.

The accuracy, which remains above 98% in validation and testing, suggests that the ratio of true positive identifications to all positive identifications (true and false) is also exceptionally high, implying that there are very few false positives; that is, non-anemic cases that the model misclassifies as anemic. The F1-Score, which harmonizes precision and recall, holds near 98% in both sets, reinforcing the idea of a balance between both metrics and providing a comprehensive view of the overall accuracy of the model. Furthermore, the AUC ROC, which measures the model's ability to distinguish between classes (anemic and non-anemic) under different thresholds, exceeds 98.5% in the test set, evidencing an excellent discriminative ability of the model. This is indicative of a highly effective model that can differentiate with high accuracy between anemic and non-anemic conditions, minimizing the likelihood of misclassification errors.

These results demonstrate not only the high accuracy of the model in detecting anemia but also its reliability, which is essential for its application in clinical settings where early and accurate detection can have a significant impact on the treatment and management of anemia.

Table 2. Validation set performance indices

Evaluation Metrics	Values
Accuracy	0.9840
Precision	0.9830
Recall (Sensitivity)	0.9840
F1-Score	0.9835
AUC ROC	0.9878

Table 3. Test set performance Indices

Evaluation Metrics	Values
Accuracy	0.98
Precision	0.9804
Recall (Sensitivity)	0.9814
F1-Score	0.9809
AUC ROC	0.9854

In this study, we have set a milestone in the noninvasive detection of iron deficiency anemia by developing a ML model based on fingertip images captured with smartphones. This CNN (2+1) D model, which processes both spatial and temporal information, represents a significant evolution in the diagnosis of anemia, offering an accessible and noninvasive methodology that extends previous research findings. By achieving outstanding accuracy, sensitivity, and specificity, our approach contrasts and surpasses previous studies, such as those of Magdalena et al. [24], who achieved 94% accuracy with CNN in palpebral conjunctiva analysis, and Appiahene et al. [25], who achieved 99.73% accuracy through assembly techniques. Unlike these, our method benefits from the simplicity and wide availability of smartphones, facilitating more widespread adoption.

Comparing directly with other research, we found that Appiahene et al. [26] achieved an accuracy of 99.96% using NB on palm images, while Williams Asare et al. [27] and Mahmud et al. [28] applied several ML algorithms, achieving maximum accuracy of 99.12% and 96%, respectively, in different diagnostic areas. These studies highlight the versatility and efficacy of ML techniques in the noninvasive diagnosis of anemia. However, our approach, by using the fingertip and a CNN (2+1) D model, not only simplifies data collection and analysis but also improves diagnostic accuracy without the need to induce physical changes or use additional equipment, as in the study by Das et al. [29], which presented a system for analyzing nail pallor with an accuracy of 89% and a sensitivity of 0.96.

This innovative approach is complemented by the advanced computational techniques observed in the studies of Kesarwani et al. [30] [31], which demonstrated the potential to improve anemia detection through color changes in the palm of the hand, reaching impressive sensitivities and accuracies. Our study not only aligns with these findings by leveraging advanced ML capabilities and smartphone technology, but also proposes a highly accessible and easy-to-implement model, offering a practical and efficient solution for the diagnosis of iron deficiency anemia.

The incorporation of both spatial and temporal analysis using our CNN (2+1) D model opens up new possibilities for the accurate detection of anemia, highlighting the importance of continuing to advance this intersection between technology and medicine to make diagnosis more accessible and efficient worldwide.

The importance of early detection of anemia for effective treatment and prevention of complications is underscored by our method, which offers a practical alternative to more invasive and costly traditional approaches. The choice of the fingertip as the diagnostic site highlights its accessibility and eliminates the need for specialized equipment, representing a significant advance in noninvasive disease detection and democratizing access to advanced medical diagnostic tools.

Despite significant advances, we acknowledge existing challenges in anemia detection using ML and smartphone technology, such as variability in image quality due to different device models and lighting conditions. Therefore, our study also emphasizes the need to expand the diversity of datasets in ML model training, encompassing a wider variety of demographic groups and anemia severity levels, to increase the accuracy and applicability of these methods.

As we look to the future, it is crucial to continue to explore this intersection between technology and medicine, with the goal of making the diagnosis of health conditions such as anemia more accessible and efficient worldwide. Future research could focus on the adaptability of this model across different demographics and conditions, as well as its integration into digital health systems to facilitate comprehensive tracking and management of anemic patients on a global scale.

5 LIMITATIONS OF THE STUDY

Due to the non-invasive nature of this study and the implementation of the non-invasive Rad-67 device to obtain the reference value for training the model, working on the basis of the margin of error declared by the device manufacturer (±1 to ±2 g/dL in hemoglobin ranges between 8 and 17 g/dL) is a limitation.

During the collection of samples from the participants, we obtained a varied set of hemoglobin value results; however, no subject in the sample presented hemoglobin values corresponding to the classification: severe anemia. Therefore, we had to work on the remaining three classifications: normal, and moderate anemia.

Another important aspect to consider is the limited age range of the participants, which was between 18 and 25 years. This narrow age range could introduce bias into the model, limiting the generalizability of the findings to other demographic populations. The youth of the participants may reflect certain physiological, lifestyle, or health characteristics that are not necessarily found in older or younger age groups, which could affect the applicability of the model in a broader context.

Extending the age range in future studies could improve the robustness and generalization of the model, allowing its application in different demographic groups with different characteristics and health conditions.

6 CONCLUSIONS

The study presented marks a significant advance in the noninvasive detection of iron deficiency anemia through the use of smartphones and ML, using fingertip images of young people. This study demonstrates that it is possible to achieve high accuracy, sensitivity, and specificity in the classification of anemic states through a CNN (2+1) D model, overcoming the challenges presented by traditional invasive diagnostic methods. Comparing our model with previous research, we highlight its simplicity, accessibility, and the ability to be implemented in diverse contexts, offering a valuable tool for the early and accurate detection of anemia.

This innovative approach not only aligns with previous research findings that applied ML techniques to improve anemia detection but also highlights the advantages of using mobile technologies and advanced ML algorithms to develop accessible and efficient solutions. The inclusion of both spatial and temporal analysis using our CNN (2+1) D model opens up new possibilities for accurate anemia detection, emphasizing the importance of technology and medicine in making diagnostics more accessible and efficient worldwide.

7 ACKNOWLEDGEMENTS

We extend our gratitude to the Universidad Nacional de San Martín for funding the project "HemoTupunaApp: Detección clínica no invasiva de anemia ferropénica utilizando smartphones," funded by University Council Resolution No. 1063-2022UNSM/CU-R.

8 REFERENCES

[1] G. A. Stevens *et al.*, "National, regional, and global estimates of anaemia by severity in women and children for 2000–19: A pooled analysis of population-representative data," *The Lancet Glob. Heal.*, vol. 10, no. 5, pp. e627–e639, 2022. https://doi.org/10.1016/S2214-109X(22)00084-5

[2] C. Camaschella and D. Girelli, "The changing landscape of iron deficiency," *Mol. Aspects Med.*, vol. 75, p. 100861, 2020. https://doi.org/10.1016/j.mam.2020.100861

[3] World Health Organization, "Anaemia as a public health problem," in *Nutritional anaemias: Tools for effective prevention and control*, 2017. [Online]. Available at: https://iris.who.int/bitstream/handle/10665/259425/9789241513067-eng.pdf

[4] United Nations, "Sustainable Development Goals: 17 Goals to Transform our World," 2015. [Online]. Available at: https://www.un.org/en/exhibits/page/sdgs-17-goals-transform-world

[5] Centro Nacional de Planeamiento Estratégico, "Perú: Plan Estratégico de Desarrollo Nacional al 2050," 2023. [Online]. Available at: https://www.gob.pe/institucion/ceplan/campañas/11228-peru-plan-estrategico-de-desarrollo-nacional-al-2050

[6] Gobierno Regional San Martín, "Plan de Desarrollo Regional Concertado San Martín al 2030," 2018. [Online]. Available at: http://drtcsanmartin.gob.pe/documentos/PDRC_2021.pdf

[7] E. Espinoza-Portilla, W. Gil-Quevedo, and E. Agurto-Távara, "Principales problemas en la gestión de establecimientos de salud en el Perú," *Rev. Cuba. Salud Pública*, vol. 46, no. 4, 2020. [Online]. Available at: https://revsaludpublica.sld.cu/index.php/spu/article/view/2146

[8] Ministerio de Desarrollo e Inclusión Social, "Reporte regional de indicadores sociales del departamento de San Martín," 2023. [Online]. Available at: https://sdv.midis.gob.pe/redinforma/Upload/regional/San Martin.pdf

[9] L. I. Cardoso-Sánchez, A. Vaquero-Vera, N. V. Gutiérrez-Moguel, and Z. A. Acosta-Chí, "Sobrepeso y Obesidad, Anemia e Inseguridad Alimentaria en Estudiantes de la Universidad de la Cañada: Un Estudio Descriptivo," *Salud y Adm.*, vol. 5, no. 15, 2018. [Online]. Available at: https://revista.unsis.edu.mx/index.php/saludyadmon/article/view/113

[10] S. Brimson, Y. Suwanwong, and J. M. Brimson, "Nutritional anemia predominant form of anemia in educated young Thai women," *Ethn. Heal.*, vol. 24, no. 4, pp. 405–414, 2019. https://doi.org/10.1080/13557858.2017.1346188

[11] K. S. Ysihuaylas Blas, "Prevalencia, grado de anemia y clasificación según índices eritrocitarios en estudiantes de la Universidad Nacional Mayor de San Marcos. Lima, 2016," 2017. [Online]. Available at: https://hdl.handle.net/20.500.12672/7294

[12] A. A. Al-Alimi, S. Bashanfer, and M. A. Morish, "Prevalence of iron deficiency anemia among university students in Hodeida province, Yemen," *Anemia*, vol. 2018, no. 1, 2018. https://doi.org/10.1155/2018/4157876

[13] M. J. Cecilia, N. M. Atucha, and J. García-Estañ, "Estilos de salud y hábitos saludables en estudiantes del Grado en Farmacia," *Educ. Médica*, vol. 19, no. 3s, pp. 294–305, 2018. https://doi.org/10.1016/j.edumed.2017.07.008

[14] J. R. Orosco Fabian, R. Pomasunco Huaytalla, and E. E. Torres Cortez, "Uso del smartphone en estudiantes universitarios de la región central del Perú," *IE Rev. Investig. Educ. la REDIECH*, vol. 11, p. e769, 2020. https://doi.org/10.33010/ie_rie_rediech.v11i0.769

[15] F. D. Guillén-Gámez and M. J. Mayorga-Fernández, "Empirical study based on the perceptions of patients and relatives about the acceptance of wearable devices to improve their health and prevent possible diseases," *Mob. Inf. Syst.*, vol. 2019, no. 1, 2019. https://doi.org/10.1155/2019/4731048

[16] G. Moreno, A. Camargo, L. Ayala, M. Zimic, and C. Del Carpio, "An algorithm for the estimation of Hemoglobin level from digital images of palpebral conjunctiva based in digital image processing and artificial intelligence," *Int. J. Online Biomed. Eng. (iJOE)*, vol. 20, no. 10, pp. 33–46, 2024. https://doi.org/10.3991/ijoe.v20i10.48331

[17] S. G. Saavedra Grandez, "Intervención de las TICs en redefinición de atención externa en Hospital II-2 Tarapoto en épocas de pandemia Covid 19," *Rev. Científica Sist. e Informática*, vol. 1, no. 1, pp. 58–68, 2021. https://doi.org/10.51252/rcsi.v1i1.120

[18] F. Alòs and A. Puig-Ribera, "Uso de wearables y aplicaciones móviles (mHealth) para cambiar los estilos de vida desde la práctica clínica en atención primaria: una revisión narrativa," *Atención Primaria Práctica*, vol. 3, no. 1s, p. 100122, 2021. https://doi.org/10.1016/j.appr.2021.100122

[19] A. Alam, A. Das, M. S. Tasjid, and A. Al Marouf, "Leveraging sensor fusion and sensor-body position for activity recognition for wearable mobile technologies," *Int. J. Interact. Mob. Technol. (iJIM)*, vol. 15, no. 17, pp. 141–155, 2021. https://doi.org/10.3991/ijim.v15i17.25197

[20] S. A. Dominguez Miranda and R. Rodriguez Aguilar, "Machine learning models in health prevention and promotion and labor productivity: A co-word analysis," *Iberoam. J. Sci. Meas. Commun.*, vol. 4, no. 1, pp. 1–16, 2024. https://doi.org/10.47909/ijsmc.85

[21] M. J. Sembay and D. D. Jeronimo de Macedo, "Health information systems: Proposal of a provenance data management method in the instantiation of the W3C PROV-DM model," *Adv. Notes Inf. Sci.*, vol. 2, pp. 192–201, 2022. https://doi.org/10.47909/anis.978-9916-9760-3-6.101

[22] B. Sotillos-González *et al.*, "Citizen perspectives on doctor-prescribed mobile health apps and information and communication technology usage within the Andalusian healthcare system," *J. Healthc. Qual. Res.*, vol. 33, no. 4, pp. 225–233, 2018. https://doi.org/10.1016/j.jhqr.2018.04.004

[23] S. S. Abdul-Jabbar, A. K. Farhan, and A. S. Luchinin, "Comparative study of anemia classification algorithms for international and newly CBC datasets," *Int. J. Online Biomed. Eng. (iJOE)*, vol. 19, no. 6, pp. 141–157, 2023. https://doi.org/10.3991/ijoe.v19i06.38157

[24] R. Magdalena *et al.*, "Convolutional neural network for anemia detection based on conjunctiva palpebral images," *J. Tek. Inform.*, vol. 3, no. 2, pp. 349–354, 2022. https://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/197

[25] P. Appiahene *et al.*, "Application of ensemble models approach in anemia detection using images of the palpable palm," *Med. Nov. Technol. Devices*, vol. 20, p. 100269, 2023. https://doi.org/10.1016/j.medntd.2023.100269

[26] P. Appiahene, J. W. Asare, E. T. Donkoh, G. Dimauro, and R. Maglietta, "Detection of iron deficiency anemia by medical images: A comparative study of machine learning algorithms," *BioData Min.*, vol. 16, 2023. https://doi.org/10.1186/s13040-023-00319-z

[27] J. Williams Asare, P. Appiahene, E. Timmy Donkoh, and G. Dimauro, "Iron deficiency anemia detection using machine learning models: A comparative study of fingernails, palm and conjunctiva of the eye images," *Eng. Reports*, vol. 5, no. 11, 2023. https://doi.org/10.1002/eng2.12667
[28] S. Mahmud, T. B. Donmez, M. Mansour, M. Kutlu, and C. Freeman, "Anemia detection through non-invasive analysis of lip mucosa images," *Front. Big Data*, vol. 6, 2023. https://doi.org/10.3389/fdata.2023.1241899
[29] S. Das *et al.*, "Smartphone-based non-invasive haemoglobin level estimation by analyzing nail pallor," *Biomed. Signal Process. and Control*, vol. 85, p. 104959, 2023. https://doi.org/10.1016/j.bspc.2023.104959
[30] A. Kesarwani, S. Das, D. R. Kisku, and M. Dalui, "Non-invasive anaemia detection based on palm pallor video using tree-structured 3D CNN and vision transformer models," *J. Exp. Theor. Artif. Intell.*, pp. 1–29, 2024. https://doi.org/10.1080/0952813X.2023.2301401
[31] A. Kesarwani, S. Das, D. R. Kisku, and M. Dalui, "Dual mode information fusion with pre-trained CNN models and transformer for video-based non-invasive anaemia detection," *Biomed. Signal Process. and Control*, vol. 88, p. 105592, 2024. https://doi.org/10.1016/j.bspc.2023.105592
[32] "Haemoglobin concentrations for the diagnosis of anemia and assessment of severity," *World Health Organization*, 2011. https://www.who.int/es/publications/i/item/WHO-NMH-NHD-MNM-11.1
[33] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, and M. Paluri, "A closer look at spatiotemporal convolutions for action recognition," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6450–6459. https://doi.org/10.1109/CVPR.2018.00675
[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014. https://doi.org/10.48550/arXiv.1412.6980
[35] J. Rahman Khan, S. Chowdhury, H. Islam, and E. Raheem, "Machine learning algorithms to predict the childhood Anemia in Bangladesh," *J. Data Sci.*, vol. 17, no. 1, pp. 195–218, 2022. https://doi.org/10.6339/JDS.201901_17(1).0009

9 AUTHORS

Miguel Angel Valles-Coral is a Systems Engineer with a Master's Degree in Information Technology Management. He is an Associate University Professor and PhD candidate in Systems Engineering and Computer Science and a leader of projects linked to artificial intelligence. He is noted for his scientific production and holder of Level III of the RENACYT (E-mail: mavalles@unsm.edu.pe).

Jorge Raul Navarro-Cabrera is a Systems and Computer Engineer; a Candidate in Master of Science with a Major in Information Technology; has experience of being a consultant and formulator of management and information technology projects; known for participation in the development of research with artificial intelligence and as a RENACYT Level VI Researcher.

Lloy Pinedo is a Systems and Computer Engineer; a Candidate for Master of Science with a mention in Information Technology; Qualified as RENACYT Researcher level V, with expertise in the formulation and execution of disciplinary and interdisciplinary research projects in the fields of information systems, computer science, and artificial intelligence.

Richard Injante is a Systems Engineer with a Master's Degree in Systems and Information Engineering; an Assistant Professor at the Universidad Nacional de San Martín, with experience in web technologies, data and computer sciences, highlighting pattern detection in images and classification; and a RENACYT Level V Researcher.

Luz Karen Quintanilla-Morales is a Graduate in Nursing, Master in Public Health with a mention in Health Management Planning; and graduated from the Doctorate in Nursing Sciences; with professional experience in the healthcare area and as an Associate University Professor at the Undergraduate level and Director of the Professional School of Nursing.

María Elena Farro-Roque is a Graduate in Bromatology and Nutrition, Doctor in University Management, Master in Public Health and Specialist in Clinical Nutrition. Principal Professor of the Academic Department of Obstetrics and Nursing of the Faculty of Health Sciences Universidad Nacional de San Martín.