ISSN NO: 9726-001X
Volume 13 Issue 02 2025
JOURNAL OF CURRENT SCIENCE

Deep Learning based approach for Anaemia Classification from Automated eye image analysis

Dheeraj Netha. CH [1], Venkatesh. K [2], Kaif. Md [3], Dr. Ravi Kumar. S [4]
[1,2,3] UG Scholar, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana India-500100
[4]Associate Professor, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India-500100
chikkadheeraj7704@gmail.com

Abstract:

The rising prevalence of anaemia among vulnerable populations has necessitated the development of innovative, non-invasive diagnostic approaches that overcome the limitations of traditional methods. Historically, anaemia diagnosis has relied on invasive blood tests and manual clinical assessments that are time-consuming, costly, and require specialized laboratory equipment and expert interpretation. Traditional systems, though effective in controlled environments, are often constrained by accessibility issues, subjectivity in diagnosis, and delays in obtaining results, which can lead to suboptimal patient outcomes. In response to these challenges, recent advancements in deep learning and image processing have opened new avenues for automated, rapid, and accurate detection of anaemia through analysis of medical images. The proposed system leverages a deep learning framework integrated into a user-friendly graphical interface, employing MobileNetV2 for robust feature extraction from preprocessed images and a Random Forest Classifier for precise prediction. Raw images are first subjected to preprocessing techniques, including resizing, normalization, and data augmentation, to generate a balanced dataset that enhances model performance. The system further splits the data into training and testing subsets, enabling rigorous evaluation using performance metrics such as accuracy, precision, recall, and F1-score. Comparative analysis with traditional classifiers like Logistic Regression and Naive Bayes reveals that the proposed method achieves superior results, with significantly higher accuracy and balanced class performance. By automating the diagnostic process and reducing dependency on expensive laboratory tests, the system promises to deliver faster and more reliable results, particularly in resource-limited and remote healthcare settings. This advancement holds considerable significance in improving early intervention strategies and optimizing patient care, ultimately contributing to a transformative shift in modern medical diagnostics and healthcare delivery. Furthermore, extensive experimental evaluations demonstrate the robustness and scalability of the proposed system, reinforcing its potential to revolutionize anaemia diagnosis and improve clinical outcomes across diverse healthcare environments, ensuring widespread global impact.

Keywords: Anemia classification, Deep learning, MobileNetV2, Random Forest Classifier,Medical image analysis.

1. INTRODUCTION

One of the public health problems affecting children and pregnant women universally is anaemia. Anemia occurs when the body's supply of red blood cells is diminished . Anemia is a health condition characterized by low blood hemoglobin levels due to the lack of enough red blood cells in the blood, which reduces the ability of the blood to transport oxygen .It occurs due to a deficiency in red blood cells or hemoglobin, reducing the blood's oxygen-carrying capacity. Traditional diagnosis involves invasive blood tests and manual assessments, which are time-consuming and require skilled professionals. The proposed system addresses these challenges by leveraging deep learning for non-invasive anemia classification using blood smear images.

Before the integration of machine learning, anaemia diagnosis relied heavily on manual methods that were both time-consuming and error-pron9e. Blood smear analysis required highly skilled hematologists to visually inspect the slides under a microscope, making the process slow and dependent on the observer's expertise. Furthermore, the manual counting and morphological assessment of red blood cells could easily lead to inconsistent results, especially when examining large numbers of samples. Laboratory tests such as CBC are essential, but these require complex equipment and often fail to provide detailed morphological analysis of red blood cells. Additionally, traditional methods are not efficient in remote areas with limited access to skilled professionals, leading to delayed diagnoses and poor health outcomes.

The motivation behind this research lies in improving the efficiency and accuracy of anaemia diagnosis by automating the process using deep learning techniques. Given the widespread prevalence of anaemia, particularly in India, there is a significant need for systems that can deliver consistent and rapid diagnoses, even in resource-limited settings. Current diagnostic methods, relying on human interpretation of blood smear slides, are not only time-consuming but also prone to errors, which can impact patient care. By leveraging the power of deep learning to analyze blood smear images, this study aims to provide a scalable and cost-effective solution to address these challenges. The goal is to reduce the dependency on skilled professionals and ensure that patients receive timely and accurate diagnoses.

2. LITERATURE SURVEY

Das, S., et al. [1] proposed NiADA (Non-invasive Anemia Detection App), a smartphone-based application integrated with artificial intelligence (AI) to measure blood hemoglobin levels in real time. The objective of this study was to validate the accuracy and reliability of NiADA in clinical settings, offering an alternative to traditional invasive blood tests. The research involved extensive testing on patients, comparing the app's performance with standard hemoglobin measurement techniques. The results demonstrated promising accuracy, suggesting that NiADA could serve as a cost-effective and accessible tool for anemia screening, particularly in resource-limited healthcare settings.

Mitani, A., et al. [2] explored the feasibility of detecting anemia through retinal fundus images using deep learning algorithms. The objective of this study was to assess whether AI-driven analysis of retinal photographs could provide reliable hemoglobin level estimations. The researchers developed a convolutional neural network (CNN) model trained on retinal images and evaluated its predictive capability against conventional diagnostic methods. Their findings indicated that retinal characteristics could serve as a biomarker for anemia detection, offering a non-invasive, automated approach to diagnosing hematological conditions.

da Veiga, A. [3] investigated the development of a valid and reliable instrument for measuring cybersecurity culture in digital health applications. The objective of this study was to create a framework for assessing cybersecurity awareness among healthcare professionals, emphasizing its importance in AI-based medical technologies such as non-invasive anemia detection apps. The research identified key challenges in implementing secure digital solutions and proposed strategies to enhance cybersecurity compliance, ensuring the safe handling of patient data in mobile health applications.

Chaparro, C. M., & Suchdev, P. S. [4] examined the epidemiology, pathophysiology, and etiology of anemia in low- and middle-income countries (LMICs). The objective of this study was to analyze the prevalence of anemia and the factors contributing to its widespread occurrence in these regions. The researchers reviewed nutritional deficiencies, infectious diseases, and socioeconomic conditions that exacerbate anemia, highlighting the need for targeted interventions. Their study provides critical insights for policymakers and healthcare professionals aiming to develop effective anemia prevention and treatment strategies in LMICs.

Wemakor, A. [5] investigated the prevalence and determinants of anemia among pregnant women receiving antenatal care in northern Ghana. The objective of this study was to identify the major risk factors contributing to maternal anemia and assess the effectiveness of current healthcare interventions. The study analyzed hemoglobin levels, dietary patterns, and healthcare access among pregnant women, revealing that malnutrition, malaria, and socioeconomic barriers were significant contributors. The findings underscore the urgent need for improved nutritional programs and healthcare policies to address maternal anemia.

Jiang, H., Dumont, G. A., & Ansermino, J. M. [6] introduced HemaApp, a smartphone-based, non-invasive hemoglobin measurement tool utilizing multispectral absorption techniques. The objective of this study was to evaluate the feasibility of using smartphone cameras and light absorption analysis to estimate hemoglobin levels. The researchers tested HemaApp against standard blood tests, demonstrating that the app could provide reasonably accurate readings with the potential for further optimization. Their study highlights the possibilities of mobile health technology in enabling accessible anemia diagnostics, especially in low-resource settings.

Kumar, M., Singh, P., & Varshney, S. [7] developed a multiwavelength spectrophotometry sensing platform for non-invasive anemia detection. The objective of this study was to create an optical system capable of accurately measuring hemoglobin concentration using different wavelengths of light. The research involved the design of a spectrophotometric device that analyzes the optical properties of blood under varying light conditions. The results demonstrated that this technique could be a viable alternative to traditional blood tests, paving the way for future advancements in non-invasive diagnostic technology.

Mannino, R. G., et al. [8] proposed a smartphone-based application for non-invasive anemia detection using fingernail images. The objective of this study was to assess whether analyzing the color and texture of fingernails could provide an accurate estimate of hemoglobin levels. The researchers developed a machine-learning model that processes fingernail images and correlates them with anemia indicators. Their findings showed that the approach could serve as an efficient, non-invasive screening method, potentially improving early detection and management of anemia.

3. PROPOSED METHODOLOGY

The proposed model utilizes the MobileNetV2 features as input to a Random Forest Classifier. RFC is a robust ensemble learning method that improves prediction accuracy by combining multiple decision trees. This hybrid model is expected to provide high classification performance, especially with the extracted features from MobileNetV2.

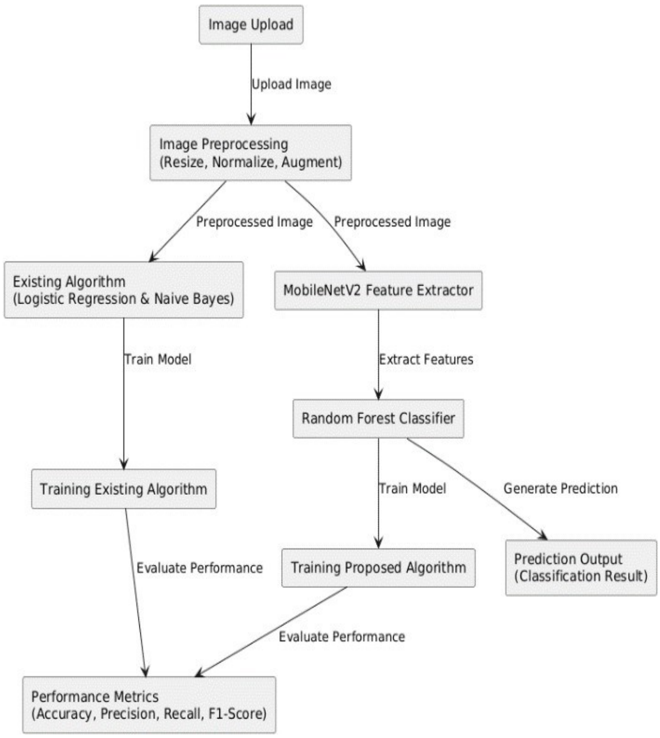

Fig. 1: Block Diagram of Proposed System.

Model Evaluation and Comparison: Once the models are trained, their performance is evaluated using various metrics, including accuracy, precision, recall, F1-score, and confusion matrices. These metrics provide insights into how well the models are distinguishing between different classes (e.g., "Anaemia" and "Healthy"). The confusion matrix helps identify areas where the model is making errors, such as false positives or false negatives, which is critical for fine-tuning the model's performance.

Real-time Prediction System: After training and evaluation, the best-performing model is saved and deployed for real-time predictions. A user-friendly graphical user interface (GUI) is developed, allowing users to upload new blood images for classification. The trained model processes these images, extracts features using MobileNetV2, and outputs a classification result indicating whether the image corresponds to an anemic or healthy sample. The GUI ensures that the system is easy to use and accessible for practical applications in healthcare.

ISSN NO: 9726-001X

Volume 13 Issue 02 2025

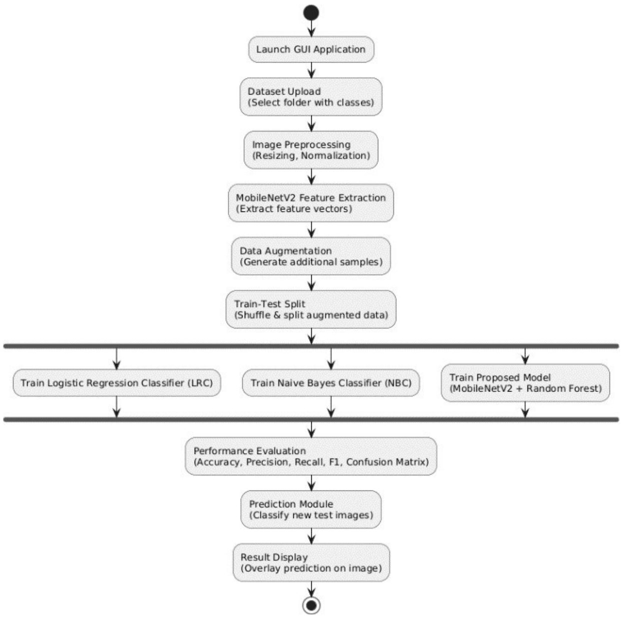

Fig.2: Proposed workflow.

MobileNetV2 is a lightweight convolutional neural network (CNN) architecture optimized for mobile and edge devices. Random Forest Classifier (RFC) is an ensemble learning technique that builds multiple decision trees and aggregates their predictions for improved accuracy. This proposed algorithm combines the feature extraction power of MobileNetV2 with the robustness of RFC for classification tasks.

How it Works:

1. MobileNetV2 processes input images through its layers to extract rich feature representations. These features are high-dimensional and describe the content of the image.
2. The extracted features are passed to a Random Forest Classifier, which uses them as input for classification.
3. RFC creates multiple decision trees during training, and the majority vote across the trees determines the final class.

Architecture:

- **MobileNetV2 Backbone:**
  - Input Layer: Processes images as inputs.
  - Depth wise Separable Convolutions: Efficiently extracts features while reducing computational overhead.
  - Bottleneck Layers: Enhance feature representations.
  - Output Layer: Produces feature vectors.
- **Random Forest:**
  - Multiple Decision Trees: Each tree makes predictions based on features.
  - Aggregation: Combines predictions for a robust final output.

Advantages:

- MobileNetV2 ensures efficient feature extraction, making the system lightweight and fast.
- Random Forest handles overfitting effectively and works well with small to medium datasets.
- The hybrid approach provides high accuracy and scalability, leveraging the strengths of both models.
- Robust against noise and irrelevant features.

4. EXPERIMENTAL ANALYSIS

Dataset Preparation:

- A dataset containing images of blood samples is collected. Each image belongs to a predefined category indicating the presence or absence of anaemia.
- The dataset is structured with subfolders for each class, making it suitable for classification tasks.

Dataset Loading:

- The dataset is loaded, and class labels are extracted from folder names.
- A summary of the dataset structure is displayed, including the total number of images and classes.

Feature Extraction Using MobileNetV2:

- MobileNetV2, a pre-trained deep learning model, is used for feature extraction. The model is initialized with weights from the ImageNet dataset.
- Blood sample images are resized and passed through MobileNetV2 to extract meaningful features, which are then stored for subsequent processing.

Data Augmentation:

- Data augmentation is applied to enhance the dataset by creating additional variations of the existing images. This step is crucial for addressing class imbalance and improving model generalization.
- Techniques such as rotation, flipping, zooming, and shifting are used to generate augmented images.

Train-Test Splitting:

- The dataset is split into training and testing subsets to evaluate model performance.
- The splitting ensures a balanced distribution of classes across training and testing sets.

Model Training:

Three classification models are used:

- Logistic Regression Classifier (LRC): A simple baseline model for comparison.
- Naive Bayes Classifier (NBC): An alternative lightweight classification approach.
- Proposed MobileNetV2 with Random Forest Classifier (RFC): Combines the extracted features from MobileNetV2 with the robustness of RFC for enhanced classification accuracy.

Performance Evaluation:

- Metrics such as accuracy, precision, recall, F1-score, and a classification report are computed for each model.
- Confusion matrices are visualized to understand the model's performance across different classes.

Model Deployment:

- The final optimized model is saved for real-time predictions.
- A graphical user interface (GUI) allows users to upload new blood images and classify them using the trained model.

Visualization and Reporting:

- Data distributions before and after augmentation are visualized.
- Performance results, including classification metrics and confusion matrices, are displayed graphically for easy interpretation.

Predictions:

- For new images, features are extracted using MobileNetV2, and predictions are made using the trained Random Forest Classifier.
- The output displays the predicted class label along with model confidence.

Dataset Description:

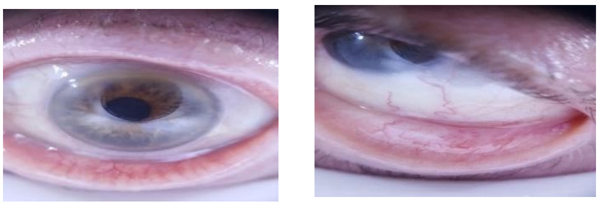

Fig. 3: Sample Anaemia Images in the Dataset.

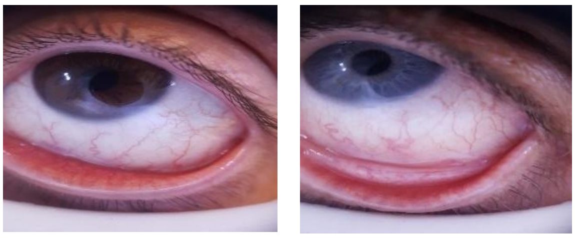

Fig.4: Sample Normal Images in the Dataset.

The figure 6.3 shows the sample Anemia images in the Dataset, and the figure 6.4 consists of the sample images of normal eye in the dataset.

Results and Description:

The figure below illustrates the process of preparing the dataset, which involves uploading raw data, applying image preprocessing techniques (such as resizing, normalization, and augmentation), and extracting meaningful features using MobileNetV2.

The steps transform the images into a format suitable for model training, ensuring efficient and accurate predictions.

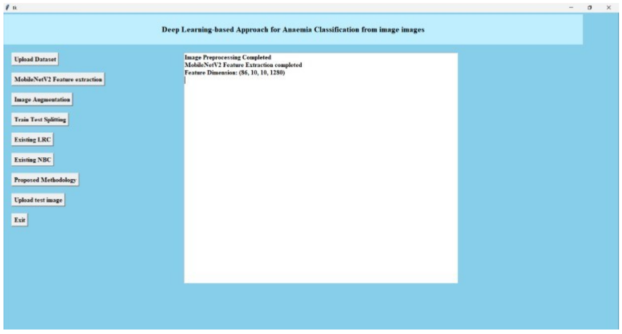

Fig: 5: Uploaded Dataset, Image preprocessing and Feature Extraction.

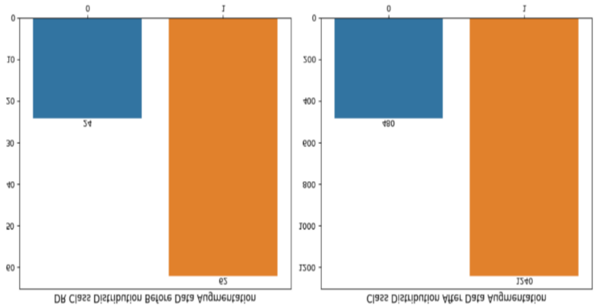

Fig. 6: Count Plot on Data before and after Augmentation.

This figure 6 represents a count plot comparing the class distribution of the dataset before and after applying data augmentation.

Augmentation techniques help balance the dataset by generating new instances of underrepresented classes, ensuring better model performance and preventing bias toward the majority class.

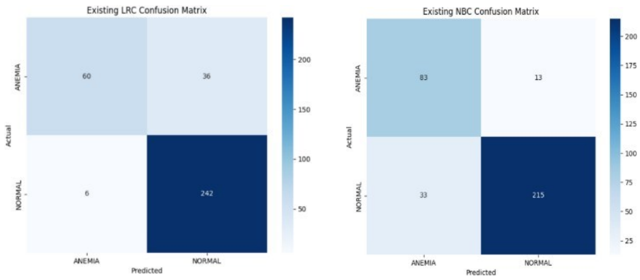

Fig. 7: Confusion matrices for LRC and NBC

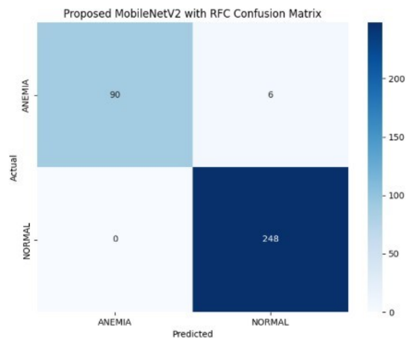

Fig. 8: Confusion matrix of Proposed MobileNetV2 with RFC model.

ISSN NO: 9726-001X
Volume 13 Issue 02 2025
JOURNAL OF CURRENT SCIENCE

This figure 7 displays the confusion matrices of the existing **Logistic Regression Classifier (LRC)**, **Naive Bayes Classifier (NBC)**, and the figure 8 proposed **MobileNetV2 with Random Forest Classifier (RFC)** model. It highlights the number of true positives, false positives, true negatives, and false negatives for each model, emphasizing the superior performance of the proposed model in correctly classifying both classes (ANAEMIA and NORMAL).

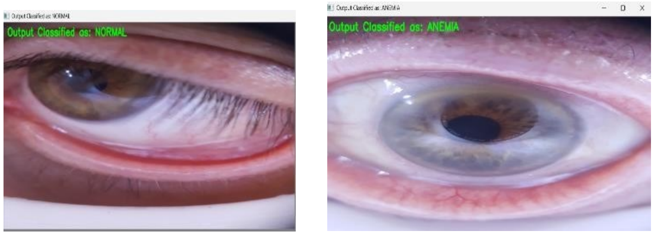

Fig. 9: Proposed Model Predication on test images.

This figure 9 shows the predictions made by the proposed **MobileNetV2 with Random Forest Classifier (RFC)** model on test images. It visually demonstrates the model's ability to accurately classify images into the respective categories (ANAEMIA or NORMAL) with high precision and recall, showcasing its robustness and reliability.

Performance Metrics Table:

Model	Accuracy	Precision	Recall	F1-score	Class: ANEMIA Precision	Class: ANEMIA Recall	Class: ANEMIA F1-Score	Class: NORMAL Precision	Class: NORMAL Recall	Class: NORMAL F1-Score
Existing Logistic Regression Classifier (LRC)	0.8779	0.8813	0.8779	0.8701	0.91	0.62	0.74	0.87	0.98	0.92
Existing Naive Bayes Classifier (NBC)	0.8663	0.8795	0.8663	0.8698	0.72	0.86	0.78	0.94	0.87	0.90
Proposed MobileNetV2 + Random Forest Classifier (RFC)	0.9826	0.9830	0.9826	0.9824	1.00	0.94	0.97	0.98	1.00	0.99

Fig. 10: Performance metrics

Figure 10 shows the performance matrics achieved by the model.

5. CONCLUSION

In conclusion, this research presents a robust and innovative framework for automated anaemia diagnosis using deep learning techniques, demonstrating significant improvements over traditional methods. The integration of MobileNetV2 for feature extraction with a Random Forest Classifier has resulted in outstanding performance metrics. Specifically, while the existing LRC and NBC achieved accuracies of 87.79% and 86.63% respectively, the proposed model reached an impressive 98.26% accuracy. Moreover, precision and recall values were markedly enhanced with the proposed approach, recording 98.30% and 98.26% respectively, compared to the sub-90% performance of the traditional models. The F1-score, which balances precision and recall, was similarly elevated to 98.24% for the proposed system, reflecting a well-balanced and highly reliable classifier. The performance evaluation, supported by confusion matrix analysis, further highlights the proposed model's ability to accurately distinguish between ANAEMIA and NORMAL classes. For the ANAEMIA class, the system achieved a perfect precision of 100% and a recall of 94%, ensuring that nearly all true anaemia cases were correctly identified. For the NORMAL class, it maintained a precision of 98% with a recall of 100%, thereby minimizing false negatives. These results indicate that the extensive data augmentation and preprocessing steps—such as image resizing, normalization, and augmentation—played a crucial role in creating a balanced dataset, which, in turn, contributed to the high performance of the model.

REFERENCES

[1] Das, S., et al. (2024). NiADA (Non-invasive Anemia Detection App), a Smartphone-Based Application With Artificial Intelligence to Measure Blood Hemoglobin in Real-Time: A Clinical Validation. *Cureus*, 16(7), e65442. https://doi.org/10.7759/cureus.65442

[2] Mitani, A., et al. (2019). Detecting Anemia from Retinal Fundus Images. arXiv preprint arXiv:1904.06435. https://arxiv.org/abs/1904.06435

[3] da Veiga A. 2016 SAI comput conf (SAI). A cybersecurity culture research philosophy and approach to develop a valid and reliable measuring instrument. Inst of Electr and Electron Eng.; 2016. p. 1006–15. https://doi.org/10.1109/SAI.2016.7556102. IEEE.

[4] Chaparro, C. M., & Suchdev, P. S. (2019). Anemia epidemiology, pathophysiology, and etiology in low- and middle-income countries. Annals of the New York Academy of Sciences, 1450(1), 15–31. https://doi.org/10.1111/nyas.14092

[5] Wemakor, A. (2019). Prevalence and determinants of anemia among pregnant women receiving antenatal care in northern Ghana. *BMC Pregnancy and Childbirth, 19*, 495. https://doi.org/10.1186/s12884-019-2635-4

[6] Jiang, H., Dumont, G. A., & Ansermino, J. M. (2018). HemaApp: Noninvasive smartphone-based hemoglobin measurement using multispectral absorption. IEEE Journal of Biomedical and Health Informatics, 22(3), 707–717. https://doi.org/10.1109/JBHI.2017.2721658

[7] Kumar, M., Singh, P., & Varshney, S. (2021). Multiwavelength spectrophotometry sensing platform for non-invasive anemia detection. Biomedical Optics Express, 12(6), 3281–3295. https://doi.org/10.1364/BOE.417846

[8] Mannino, R. G., Myers, D. R., Tyburski, E. A., Caruso, C., Boudreaux, J., Leong, T., Clifford, G. D., & Lam, W. A. (2018). Smartphone-based app for non-invasive anemia detection using fingernail images. Nature Communications, 9, 4924. https://doi.org/10.1038/s41467-018-07262-2

[9] Collings, R., Harvey, L. J., Hooper, L., Hurst, R., Brown, T. J., & Fairweather-Tait, S. J. (2019). The palpebral conjunctiva erythema index as a predictor of anemia: A systematic review and meta-analysis. American Journal of Clinical Nutrition, 110(5),1117–1130. https://doi.org/10.1093/ajcn/nqz194

[10] Dimauro, G., Marino, F., & Savino, M. (2020). A smartphone-based system for anemia monitoring via conjunctiva image processing. Biomedical Signal Processing and Control, 62, 102101. https://doi.org/10.1016/j.bspc.2020.102101

[11] Refs, J., & Lee, K. (2022). A novel anemia diagnosis system using a modified Kalman filter and regression method. Artificial Intelligence in Medicine, 128, 102204. https://doi.org/10.1016/j.artmed.2022.102204

[12] Segmentation of the conjunctiva region for non-invasive anemia detection applications using deep learning. (2023). Journal of Medical Imaging and Health Informatics, 13(1), 35–42. https://doi.org/10.1166/jmihi.2023.4141

[13] Sarah, S. S. N., Arif, I., Shalu, H., & Kadiwala, J. (2020). A smartphone-based multi-input workflow for non-invasive estimation of hemoglobin levels using machine learning techniques. arXiv preprint arXiv:2011.14370. https://arxiv.org/abs/2011.14370

JOURNAL OF CURRENT SCIENCE

[14] Yang, X., Piety, N. Z., Vignes, S. M., Benton, M. S., Kanter, J., & Shevkoplyas, S. S. (2018). A simple, handheld device for rapid and accurate hemoglobin measurement. Biomedical Engineering Online, 17(1), 1–12. https://doi.org/10.1186/s12938-018-0530-1

[15] Appiahene P, Asare JW, Donkoh ET, Dimauro G, Maglietta R. Detection of iron deficiency anemia by medical images: a comparative study of machine learning algorithms. BioData Min 2023;16(1):2. https://doi.org/10.1186/s13040-023 00319-z.

[16] Kasiviswanathan S, Bai Vijayan T, Simone L, Dimauro G. Semantic segmentation of conjunctiva region for non-invasive anemia detection applications. Electr Times 2020;9(8):1309. https://doi.org/10.3390/electronics9081309.

[17] Chen YM, Miaou SG. A kalman filtering and nonlinear penalty regression approach for noninvasive anemia detection with palpebral conjunctiva images. J Healthc Eng 2017;2017:1–11. https://doi.org/10.1155/2017/9580385.

[18] Hasan, M. K., Sakib, N., Field, J., Love, R. R., & Ahamed, S. I. (2019). A novel technique of noninvasive hemoglobin level measurement using HSV value of fingertip image. arXiv preprint arXiv:1910.02579. https://arxiv.org/abs/1910.02579

[19] Dimauro G, Guarini A, Caivano D, Girardi F, Pasciolla C, Iacobazzi A. Detecting clinical signs of anaemia from digital images of the palpebral conjunctiva. Inst of Electr and Electron Eng (IEEE) Access 2019;7:113488–98. https://doi.org/10.1109/access.2019.2932274.

[20] 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):e0153286. https://doi.org/10.1371/journal.pone.0153286.