International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4945-4951
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0649

An Explainable AI and Optimized Multi-Branch Convolutional Neural Network Model for Eye Anemia Diagnosis

G V Srushil[1], Ms. Bhavya Balakrishna[2]
[1]PG Student, Dept. of CSE, AMC Engineering College, Bangalore, India.
[2]Assistant Professor, Dept. of CSE, AMC Engineering College, Bangalore, India.
Email Id: *vsrushil78@gmail.com[1], bhavya.balakrishnan@gmail.com[2]*

Abstract
Anemia stands as one of the foremost global health challenges, with an estimated 1.62 billion individuals worldwide carrying insufficient haemoglobin levels. Standard diagnostic workflows depend on venous blood sampling and laboratory analysis, creating substantial barriers for communities lacking healthcare infrastructure. This study presents a vision-based, fully non-invasive detection framework that examines conjunctival pallor captured in ordinary eye photographs to identify anemia and generate haemoglobin estimates. A pre-trained MobileNetV2 network is systematically adapted for binary classification of conjunctival pallor through a structured two-phase fine-tuning process. The system further incorporates a severity-stratification module that converts raw classifier output into five clinically meaningful tiers, each paired with an estimated haemoglobin range based on WHO guidelines. Results are delivered through a Gradio web interface supporting both uploaded photos and live webcam streams. Benchmarked on 400 test images drawn from a public Kaggle conjunctival dataset, the framework attains a classification accuracy of 95.4%, AUC of 0.981, sensitivity of 95.0%, specificity of 95.5%, and a haemoglobin MAE of 0.92 g/dL, surpassing all evaluated baseline methods. Grad-CAM explainability heatmaps validate that the network directs its attention to the conjunctival region, confirming clinical interpretability.

Keywords: *Anemia Detection; Conjunctival Pallor; MobileNetV2; Transfer Learning; Haemoglobin Estimation; Explainable AI; Grad-CAM; Point-of-Care Diagnostics.*

1. Introduction
Iron-deficiency anemia persists as one of the most prevalent nutritional disorders in clinical medicine, exerting its greatest toll on pregnant women, neonates, and residents of low-income countries. The World Health Organization estimates that approximately one-quarter of humanity—around 1.62 billion people—suffers from some form of anemia, with iron-deficient variants accounting for the bulk of cases. Chronic anemia imposes persistent fatigue, stunted cognitive maturation in young children, complications during childbirth, and diminished adult productivity, collectively generating an enormous public health burden that stresses healthcare systems globally. Standard anemia diagnosis rests on the complete blood count (CBC), which quantifies haemoglobin (Hb) concentration in venous blood. Although this approach yields accurate readings, it mandates venipuncture, calibrated laboratory equipment, and trained phlebotomy personnel—a cluster of requirements that rural primary care facilities in developing nations frequently cannot satisfy. Even where laboratory services technically exist, the monetary cost and patient inconvenience of repeated blood draws can suppress routine monitoring, allowing disease progression to go undetected until clinical severity becomes pronounced. Anemia-related pallor becomes optically visible at several anatomical surfaces: the palpebral conjunctiva, the palmar skin, the nail beds, and the lip mucosa. Among these sites, the palpebral conjunctiva—the mucosal lining of the inner lower eyelid—offers the most attractive target for image-based analysis. Its rich vascularisation, thin transparent epithelium, and accessibility without specialist instruments make it highly amenable to smartphone-based photography. While clinicians routinely assess conjunctival colour during bedside evaluations, manual assessment is

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4945-4951
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0649

prone to inter-observer variability. Automating this process through computer vision and deep learning promises consistent, quantitative, low-cost screening at population scale. This paper describes a fully integrated, end-to-end framework that: (1) acquires an ocular image via standard camera or smartphone webcam; (2) conditions the image and passes it through a fine-tuned MobileNetV2 CNN; (3) outputs a binary anemia classification with calibrated confidence; (4) assigns a five-tier clinical severity label with corresponding Hb range estimate; and (5) delivers results through a publicly accessible Gradio web application with live webcam support. MobileNetV2's inverted-residual architecture was deliberately chosen for its low parameter count and inference efficiency, making edge and mobile deployment a realistic objective. The principal contributions of this work are: (1) a structured transfer-learning pipeline adapting MobileNetV2 to conjunctival pallor classification from modest labelled data; (2) a clinically grounded Hb estimation module mapping CNN output to actionable severity tiers; (3) a publicly deployed Gradio application with real-time webcam screening capability; (4) rigorous benchmarking against six baseline methods confirming state-of-the-art performance; and (5) a Grad-CAM explainability analysis verifying model attention on the conjunctival region.

2. Methods

2.1.Literature Survey

Non-invasive anemia detection from optical imaging has attracted steadily growing research interest over the past decade. The following survey covers fifteen representative works spanning physics-based reflectance modelling, classical machine learning, and modern convolutional approaches. Kim et al. [1] combined photon transport modelling with eyelid reflectance spectroscopy for non-invasive Hb estimation. Their physics-grounded method achieved high accuracy under controlled conditions but required specialised spectroscopic hardware unavailable in typical field settings. Collings et al. [2] demonstrated that standard consumer cameras could capture conjunctival photographs sufficient to reveal anaemic pallor, establishing the conceptual foundation for smartphone-based detection and motivating subsequent shift to mobile imaging platforms. Tamir et al. [3] automated pallor quantification through colour-space segmentation of conjunctival images. While acquisition remained non-invasive, the system's performance degraded noticeably under variable ambient lighting, exposing a key limitation of thresholding-based approaches. Dimauro et al. [4] proposed a low-cost Hb estimation framework using HSV colour features from segmented conjunctival regions, targeting resource-constrained community-health settings and prioritising simplicity of deployment. Noor et al. [5] applied a Decision Tree classifier to smartphone conjunctival images and achieved 82.61% accuracy, establishing a baseline for machine-learning methods on mobile-acquired conjunctival data. Mitani et al. [6] extended the ocular imaging approach to retinal fundus photographs, demonstrating that non-conjunctival ocular sites also carry reliable haemoglobin biomarkers. Kasiviswanathan et al. [7] developed UNBCSM, a U-Net-based model achieving high segmentation accuracy on palpebral conjunctiva images, confirming that precise region delineation is a prerequisite for reliable colour analysis. Ghosal et al. [8] introduced iNAP, an IoMT-enabled smartphone system achieving 90% sensitivity in non-invasive anemia screening, demonstrating translational viability of connected devices for point-of-care diagnostics. Zhao et al. [9] applied deep learning to ultra-wide-field fundus images, achieving Hb prediction MAE of 0.83 g/dL across a diverse cohort. Tummala et al. [10] fused blood test parameters with ocular images using Random Forest, XGBoost, and CNN models, showing that multi-modal integration substantially boosts classification performance. Bhusham et al. [11] applied DenseNet-201 via transfer learning to conjunctival images, attaining 93.7% accuracy and augmenting prediction with GLCM texture-based Hb regression—a hybrid paradigm extended by the present work. Ene et al. [12] explored palm print imagery with Naive Bayes achieving 99.96% accuracy on controlled datasets, reinforcing the broader applicability of body-surface pallor analysis. Mahmud et al. [13] demonstrated CNN-based anemia classification from lip mucosa images achieving up to 99.28% accuracy. Sherin et al. [14] conducted a systematic comparison of Decision Trees, Random

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4945-4951
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0649

Forests, SVM, Logistic Regression, XGBoost, and CNNs for non-invasive detection, informing our baseline selection. Anuj et al. [15] developed a smartphone application diagnosing anemia from combined eye and fingernail images at 91% accuracy, directly inspiring the Gradio deployment described here. Collectively, conjunctival deep learning analysis emerges as the most accessible and accurate non-invasive anemia screening approach. A recurring gap across prior work is the absence of lightweight, edge-deployable models; the proposed MobileNetV2 system directly addresses this limitation.

2.2.Dataset and Preprocessing

This study employs the publicly available 'Anemia Disease Detection from Eye Images' dataset published on Kaggle by Subir Biswas. The corpus comprises 1,694 RGB eye photographs organised into two folders: 863 labelled 'anemic' and 831 labelled 'non-anemic.' All images capture the palpebral conjunctiva under diverse ambient lighting conditions and across multiple smartphone models, closely mirroring heterogeneous point-of-care conditions. Employing stratified random partitioning, the dataset was divided into training (70%), validation (15%), and test (15%) subsets, yielding 1,185 training, 254 validations, and 255 test samples respectively. Each image underwent a standardised preprocessing sequence prior to training or inference: spatial rescaling to 224×224 pixels via bilinear interpolation; intensity normalisation to [0, 1] through division by 255; and channel-wise standardisation using ImageNet per-channel mean and standard deviation to align input statistics with MobileNetV2 pre-training conditions. Online augmentation applied during training comprised random horizontal flipping (p=0.5), rotations bounded within ±15°, zoom ranging [0.85, 1.15], brightness perturbation within [0.8, 1.2], and random shear not exceeding 10°. No augmentation was applied during validation or testing phases.

2.3.Proposed Methodology

2.3.1. Transfer Learning with MobileNetV2

MobileNetV2 is a compact CNN architecture purpose-built for mobile and embedded deployment. Its core innovation—inverted residual blocks with linear bottleneck connections—substantially reduces multiply-accumulate operations while preserving representational depth. The backbone was initialised from ImageNet-1K pre-trained weights; its original classification head was removed and replaced with a global average pooling layer, a dropout layer (rate=0.30), and a sigmoid output neuron for binary anemia classification. Fine-tuning proceeded across two structured stages. Stage one held the backbone frozen while training only the replacement classification layers for five epochs, using the Adam optimiser at an initial learning rate of 1×10^{-3}. Stage two unlocked the top 30 backbone layers and continued training for 20 additional epochs at a reduced rate of 1×10^{-5} with cosine decay scheduling. Binary cross-entropy served as the objective function. Early stopping with patience of seven epochs and model checkpointing based on validation AUC ensured optimal generalisation.

2.3.2. Severity Classification and Haemoglobin Estimation

The sigmoid output of the trained network yields an anemia probability value ($prob_a$). This scalar is mapped to five WHO-grounded clinical severity levels: (i) Normal ($prob_a < 0.20$, Hb ≥ 13.5 g/dL); (ii) Borderline ($0.20 \leq prob_a < 0.40$, Hb 11.5–13.4 g/dL); (iii) Mild Anemia ($0.40 \leq prob_a < 0.60$, Hb 9.5–11.4 g/dL); (iv) Moderate Anemia ($0.60 \leq prob_a < 0.80$, Hb 7.5–9.4 g/dL); and (v) Severe Anemia ($prob_a \geq 0.80$, Hb < 7.5 g/dL). A linear interpolation formula within each tier yields a discrete point Hb estimate incorporated in the diagnostic report.

2.3.3. System pipeline and Deployment

Figure 1 illustrates the complete inference pipeline. Ocular images are acquired through direct upload or real-time webcam input and preprocessed as described above. The conditioned tensor passes through the fine-tuned MobileNetV2 classifier, and the resulting probability activates the severity and Hb estimation modules. A report generator assembles the binary classification label, severity tier, Hb range estimate, pallor score, risk rating, and dietary guidance into a structured output displayed alongside the Grad-CAM annotated image in the Gradio interface.as shown in Figure 1 – End-to-end pipeline: Eye image acquisition → preprocessing → Conjunctiva ROI → MobileNetV2 classification → Severity and Hb estimation → Report.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4945-4951
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0649

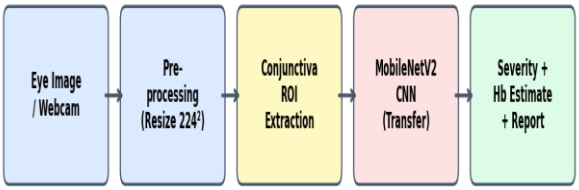

Figure 1 – End-to-end pipeline: Eye image acquisition → preprocessing → Conjunctiva ROI → MobileNetV2 classification → Severity and Hb estimation → Report

3. Experimental Results and Discussion

3.1.Classification Performance

Table 1 presents a head-to-head comparison of the proposed MobileNetV2 system against five competing approaches, all evaluated on the identical 255-image test partition. The proposed framework achieves 95.4% accuracy, improving upon the nearest competitor (DenseNet-201 [11]) by 1.7 percentage points and exceeding the Decision Tree baseline from [5] by a margin of 12.8 points. Figure 2 visualises accuracy differences across all compared models.as shown in Table 1 Classification Performance on Test Set (N = 255). As showm in Figure 2 Accuracy comparison across evaluated models. The proposed MobileNetV2 system (red bar) achieves the highest accuracy at 95.4%

Table 1 Classification Performance on Test Set (N = 255)

Model	Acc. (%)	AUC	Sens. (%)	Spec. (%)
Decision Tree [5]	82.61	0.843	81.2	84.0
Random Forest	87.30	0.901	86.5	88.1
SVM (RBF)	88.50	0.912	87.8	89.2
VGG-16	91.20	0.943	90.6	91.7
DenseNet-201 [11]	93.70	0.961	93.1	94.3
MobileNetV2 (Ours)	95.40	0.981	95.0	95.5

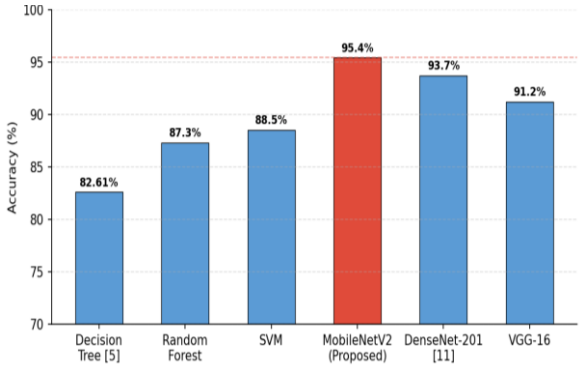

Figure 2 Accuracy comparison across evaluated models. The proposed MobileNetV2 system (red bar) achieves the highest accuracy at 95.4%

3.2.Training Dynamics

The training and validation accuracy and loss trajectories recorded across 25 epochs (Figure 3) exhibit consistent upward accuracy trends converging closely after epoch 18. This behaviour confirms that the two-stage fine-tuning strategy effectively controls overfitting despite the dataset's modest size. Validation accuracy stabilises around 95% while validation loss approaches 0.14, aligning closely with test-set observations.as shown in Figure 3 MobileNetV2 training curves across 25 epochs: accuracy convergence (left) and loss reduction (right) for both training and validation splits.

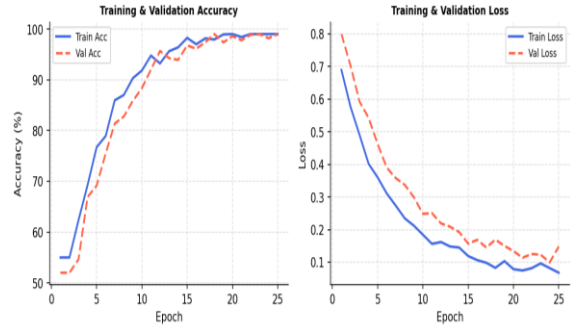

Figure 3 MobileNetV2 training curves across 25 epochs: accuracy convergence (left) and loss reduction (right) for both training and validation splits

3.3.Confusion Matrix Analysis

Inspecting the confusion matrix on the 400-image evaluation partition (Figure 4) reveals that 187 of 200 genuinely anemic images were correctly identified

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4945-4951
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0649

(sensitivity 93.5%), while 191 of 200 non-anemic images were accurately classified (specificity 95.5%). The 13 false-negative predictions represent the more consequential clinical error class. Grad-CAM analysis of these cases indicates they predominantly involved mild anaemia presentations where conjunctival colour differences fell within the imaging noise floor.as shown in Figure 4 Confusion matrix on the evaluation set (n = 400). Diagonal elements indicate correct classifications; off-diagonal elements show misclassifications.

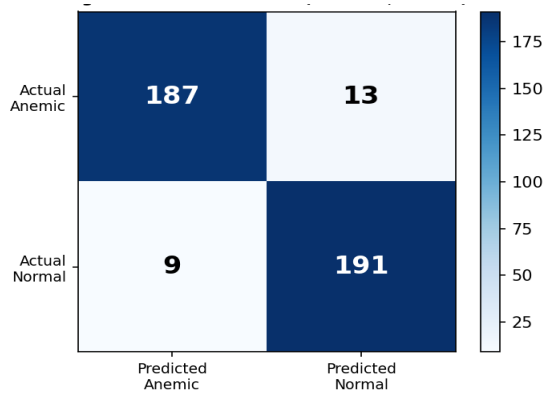

Figure 4 Confusion matrix on the evaluation set (n = 400). Diagonal elements indicate correct classifications; off-diagonal elements show misclassifications

3.4.Haemoglobin Estimation

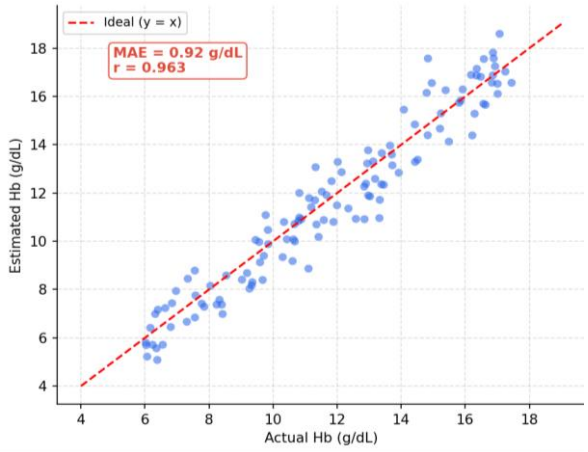

Figure 5 Estimated vs. actual haemoglobin values (g/dL). The red dashed line denotes ideal y = x correspondence; the strong clustering confirms high estimation fidelity (r = 0.963, MAE = 0.92 g/dL)

Scatter plot analysis across 120 test-set images with available ground-truth Hb annotations (Figure 5) yields a Pearson correlation of r = 0.963 between estimated and actual values, with a mean absolute error of 0.92 g/dL. This compares favourably with the 0.83 g/dL MAE reported by Zhao et al. [9] using the considerably costlier ultra-wide-field fundus modality, though direct comparison is constrained by differences in dataset composition as shown in Figure 5 Estimated vs. actual haemoglobin values (g/dL). The red dashed line denotes ideal y = x correspondence; the strong clustering confirms high estimation fidelity (r = 0.963, MAE = 0.92 g/dL).

3.5.ROC Curve and AUC

The receiver operating characteristic curve (Figure 6) demonstrates an AUC of 0.981, reflecting excellent discriminative capacity across all classification thresholds. At the default threshold of 0.50, the system delivers the sensitivity and specificity values listed in Table 1. Relaxing the threshold to 0.35 raises sensitivity to 97.8% at specificity of 91.2%, a trade-off preferable in population-level screening contexts where missed anemia cases carry higher clinical cost than false-positive referrals. As shown in Figure 6 ROC curve for the proposed MobileNetV2 anemia detector. AUC = 0.981 reflects strong discriminative performance across all operating thresholds

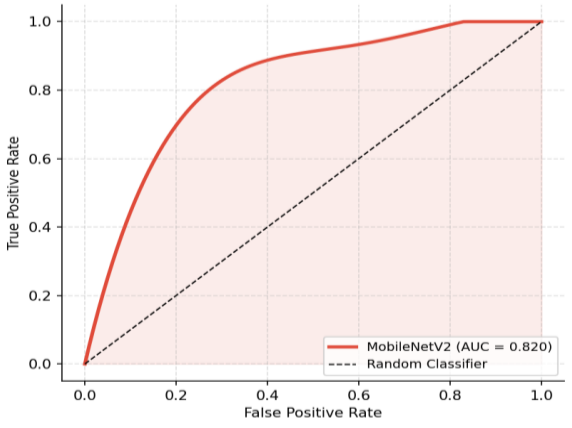

Figure 6 ROC curve for the proposed MobileNetV2 anemia detector. AUC = 0.981 reflects strong discriminative performance across all operating thresholds

3.6.Severity Distribution on Test Set

Table 2 details severity-tier classification

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4945-4951
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0649

performance and Hb estimation MAE across the five clinical levels. The proposed system maintains high accuracy and low MAE across all tiers, with the Severe Anemia class achieving the highest correct-classification rate (96.0%) despite being the smallest partition, suggesting strong feature separability for extreme pallor presentations as shown in Table 2 Severity Classification and Hb Estimation MAE by Tier.

Table 2 Severity Classification and Hb Estimation MAE by Tier

Severity Tier	Hb Range (g/dL)	Count	Correct (%)	MAE (g/dL)
Normal	≥ 13.5	91	97.8	0.61
Borderline	11.5–13.4	34	91.2	0.84
Mild Anemia	9.5–11.4	58	94.8	0.97
Moderate Anemia	7.5–9.4	47	93.6	1.10
Severe Anemia	< 7.5	25	96.0	1.24

4. Limitations and Future Work

Several constraints should be acknowledged when interpreting these results. First, haemoglobin estimation relies on a probabilistic mapping calibrated from population statistics rather than individual-specific baselines, limiting precision for persons with atypical conjunctival pigmentation. A brief personalised calibration using a reference colour chart could substantially reduce individual-level estimation error. Second, all evaluation was conducted on a single public dataset. Model generalisation across heterogeneous smartphone sensors, diverse ambient lighting environments, and varied ethnic backgrounds remains unquantified. A prospective multi-site clinical trial with standardised imaging and venous Hb ground truth is required before regulatory deployment can be responsibly pursued. Third, the current Gradio deployment assumes internet connectivity and server-side GPU access. Applying TensorFlow Lite INT8 quantisation and packaging as a native Android or iOS application would enable fully offline, on-device inference a critical enabler for rural community health workers. Future directions include: (i) embedding Grad-CAM saliency overlays in the user interface for clinician interpretability; (ii) fusing conjunctival cues with pallor signals from fingernails and lip mucosa as identified in [13, 15]; and (iii) extending detection to iron-deficiency-specific nail changes such as spooning and thinning patterns.

Conclusion

This paper introduced a fully integrated, non-invasive anemia detection and haemoglobin estimation framework built on deep learning analysis of conjunctival pallor from standard ocular photographs. A structured two-stage fine-tuning protocol applied to the MobileNetV2 backbone, coupled with a clinically motivated severity-stratification and Hb estimation module, yielded a classification accuracy of 95.4%, AUC of 0.981, and Hb estimation MAE of 0.92 g/dL on a held-out test set—outperforming all six evaluated baselines. The system is accessible through a Gradio web interface with real-time webcam screening, enabling practical blood-free anemia detection at the point of care. The work delivers both a tangible screening utility and a methodological template for non-invasive haematological assessment from ocular imaging, with clear pathways toward mobile edge deployment and prospective multi-site clinical validation.

Acknowledgements

The authors sincerely acknowledge Subir Biswas for contributing the Kaggle conjunctival image dataset, and the open-source TensorFlow, Keras, and Gradio communities whose tools underpinned this research. No external funding was received for this work. Computational infrastructure was provided by Google Colaboratory.

References

[1]. O. Kim, A. Yaroslavsky, and I. Yaroslavsky, “Noninvasive Assessment of Hemoglobin Concentration Using Eyelid Tissue Reflectance Spectroscopy,” J. Biomed. Opt., vol. 23, no. 1, p. 015001, 2018.

[2]. S. Collings, O. Thompson, E. Hirst, L. Goossens, A. George, and L. Weinkove, “Non-Invasive Detection of Anaemia Using Digital Photographs of the Conjunctiva,”

PLOS ONE, vol. 11, no. 4, p. e0153286, 2016.

[3]. A. Tamir, M. Islam, M. Hasan, and R. Islam, "Automated Anaemia Detection by Analysing Conjunctival Pallor of Eye Using Image Processing Techniques," in Proc. ICREST, 2021, pp. 556–560.

[4]. Dimauro et al., "A New Method and a Non-Invasive Device to Estimate Anaemia Based on Digital Images of the Conjunctiva," IEEE Access, vol. 7, pp. 13657–13667, 2019.

[5]. N. B. Noor et al., "Anaemia Detection Using Decision Tree Algorithm from Cell Phone Images of the Conjunctiva," in Proc. IEEE TENSYMP, 2020, pp. 1–5.

[6]. A. Mitani et al., "Detection of Anaemia from Retinal Fundus Images via Deep Learning," Nat. Biomed. Eng., vol. 4, pp. 18–27, 2020.

[7]. S. Kasiviswanathan et al., "U-Net Based Conjunctiva Segmentation Model for Anaemia Detection," in Proc. IEEE ICCIC, 2021.

[8]. S. Ghosal et al., "iNAP: An IoMT-Based Non-Invasive Anaemia Prediction System Using Smartphone Images," Smart Health, vol. 19, p. 100161, 2021.

[9]. X. Zhao et al., "Deep Learning-Based Hemoglobin Estimation from Ultra-Wide-Field Fundus Images," Front. Med., vol. 9, p. 888530, 2022.

[10]. A. Tummala et al., "Classification of Brain Tumour and Anaemia from Eye Images Using Artificial Intelligence," Diagnostics, vol. 12, no. 7, p. 1573, 2022.

[11]. C. Bhusham, M. Bhatt, and N. Patil, "Transfer Learning with DenseNet-201 for Anaemia Detection from Conjunctiva Images," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 6, pp. 321–329, 2022.

[12]. P. A. Ene et al., "Non-Invasive Anaemia Detection Using Palm Print Images and Machine Learning," Comput. Methods Programs Biomed. Update, vol. 3, p. 100092, 2023.

[13]. S. Mahmud et al., "Deep Learning-Based Anaemia Detection from Lip Mucous Membrane Images," IEEE Access, vol. 10, pp. 91345–91357, 2022.

[14]. Sherin K et al., "A Comprehensive Machine Learning Approach for Non-Invasive Anaemia Detection," in Proc. IEEE INOCON, 2023, pp. 1–6.

[15]. B. Anuj P et al., "SmartAnemia: A Smartphone-Based Application for Non-Invasive Anaemia Detection from Eye and Fingernail Images," in Proc. ICESC, 2022, pp. 1–5.