Deep Vision: Non-Invasive Anemia Screening with CNNs

Somya Pathak and Srimedha Bhavani Chandoo

CSCI 5922-001-003: Neural Nets and Deep Learning, University of Colorado Boulder

Abstract. Anemia is a common but often overlooked condition, specially in communities with limited access to medical testing. Blood tests is the standard method for diagnosing anemia which can be invasive, expensive, and difficult to access in remote areas. In this project, we explore an alternative solution: Using eye images and deep learning to screen for anemia. Since changes in eye pallor can be a sign of low hemoglobin levels, we trained two image classification models: a lightweight custom CNN and a MobileNetV2-based transfer learning model to detect anemia from conjunctiva images. Both models were trained and evaluated on a publicly available dataset. While MobileNetV2 achieved higher accuracy, our CNN model offered faster, more efficient performance. Together, these results show promise for building accessible, non-invasive screening tools that could run directly on everyday machines and help identify anemia early, even in low-resource settings.

Keywords: Anemia detection, Eye images, CNN, MobileNetV2, Deep learning, Non-invasive screening

1 Introduction

Anemia is one of the most common nutritional deficiencies worldwide, affecting more than 1.6 billion people according to the World Health Organization. It is especially prevalent in children and women of reproductive age in low-resource settings, where access to healthcare and laboratory testing may be limited. Traditional diagnostic methods like complete blood count (CBC) tests require invasive blood draws and laboratory infrastructure, which are often unavailable or costly in underprivileged areas. This presents a critical need for alternative, non-invasive, and accessible screening tools. With the rapid advancement of machine learning and computer vision, there is a growing opportunity to explore how AI can be used to support medical diagnostics in a cost-effective and scalable manner.

Current efforts to automate anemia detection are still in their early stages. Some research has explored image-based approaches, using photographs of eyes, nails, or palms to identify visual cues like pallor, which is a known symptom of anemia. However, many of these solutions either rely on multi-modal data (e.g., combining images with clinical tests), are limited in scope, or lack the robustness needed for real-world deployment. Additionally, existing models are often overfitted to small datasets or are too complex to be used on mobile or embedded devices, limiting their accessibility. There is still a gap in developing models that balance accuracy with simplicity and interpretability, especially for deployment in resource-constrained environments.

In this project, we address that gap by developing a lightweight Convolutional Neural Network (CNN) model to detect anemia using only eye images. Our goal is to create a system that is both accurate and practical for use on everyday devices in low-resource environments. We train our model on a publicly available dataset of conjunctiva eye images, using visual features alone, no lab data or metadata to make predictions. We also compare our custom CNN with a transfer learning approach using MobileNetV2, highlighting trade-offs between model complexity, speed, and accuracy. This work moves toward a future where AI can help make essential health screening more accessible to those who need it most.

2 Related Work

Image-Based Anemia Detection

- Appiahene et al. [1]
- Asare et al. [2]
- Sehar et al. [3]
- Tamir et al. [4]

These works employ image-based machine learning techniques to detect anemia using the conjunctiva, nails, or palms. For example, Appiahene et al. [1] develop a smartphone-based diagnostic tool, while Sehar et al. [3] propose a deep learning model for classifying conjunctiva images. Our approach differs by focusing on simplicity and accessibility: we use a lightweight CNN model trained exclusively on eye images (conjunctiva) without requiring additional sensors or patient metadata, emphasizing deployability in resource-limited settings.

Ocular Image Analysis in Healthcare

- Iqbal et al. [5]
- Tamir et al. [4]

Ocular image analysis is increasingly explored for non-invasive diagnostics. Iqbal et al. [5] review advancements in fundus image analysis for various conditions, while Tamir et al. [4] focus on threshold-based detection of anemia from eye images. Unlike fundus-based approaches that require specialized equipment, our model operates on anterior eye images, specifically targeting pallor in the conjunctiva, making it more practical for low-cost screenings.

Deep Learning for Non-Invasive Medical Diagnostics

- Rajpurkar et al. [6]
- Appiahene et al. [1]

Deep learning has shown immense potential in non-invasive diagnostics. CheXNet [6], for example, achieves radiologist-level pneumonia detection from chest X-rays. Similarly, Appiahene et al. [1] show the feasibility of smartphone-based anemia detection. Our work builds on this foundation by refining the focus to a single visual biomarker, eye pallor, and adopting a compact architecture suitable for real-time use on mobile devices.

Anemia Eye Image Dataset

- KMIT Lab [7]

Most existing studies utilize private or imbalanced datasets, limiting reproducibility and robustness. In contrast, we leverage the Clean Augmented Anemia Dataset [7], a publicly available, balanced dataset of conjunctiva images. It includes augmented variants to simulate real-world conditions (e.g., lighting and angle changes), enhancing our model's ability to generalize during deployment.

3 Methodology

Our approach to detecting anemia from eye images is based on training a Convolutional Neural Network (CNN) to classify whether an individual is anemic or not based on visible ocular features, particularly pallor in the conjunctiva region. The overall methodology involves data preparation, preprocessing, model development, training, and evaluation.

3.1 Dataset

We used a publicly available dataset consisting of labeled eye images with corresponding hemoglobin levels. For this project, we framed it as a binary classification problem grouping images into two classes: non anemic and anemic. The dataset includes a diverse range of eye images, captured under varying lighting and imaging conditions, which makes it well-suited for building a model intended for real-world use.

3.2 Preprocessing

To ensure input consistency and improve generalization, several preprocessing steps were applied:

- All images were resized to 224×224 pixels to match the input size expected by pre-trained CNN architectures.
- Pixel values were normalized, either to the range [0, 1] or standardized according to the model's preprocessing function.
- Data augmentation techniques such as rotation, zoom, brightness adjustment, and horizontal flipping were applied to improve robustness and reduce overfitting.
- In some cases, images were cropped to emphasize the conjunctiva region and reduce background noise.

3.3 Model Architecture

We implemented two models: a custom CNN and a MobileNetV2-based transfer learning model.

The custom CNN consisted of stacked convolutional layers with ReLU activations, followed by max-pooling layers, dropout regularization, and dense layers. The model ends with a fully connected layer and a sigmoid activation function for binary classification. It was intentionally designed to be lightweight for deployment on resource-constrained devices.

In contrast, the MobileNetV2 model leveraged pre-trained ImageNet weights and was fine-tuned on our dataset. Its efficiency and strong feature extraction made it a strong candidate for mobile health applications.

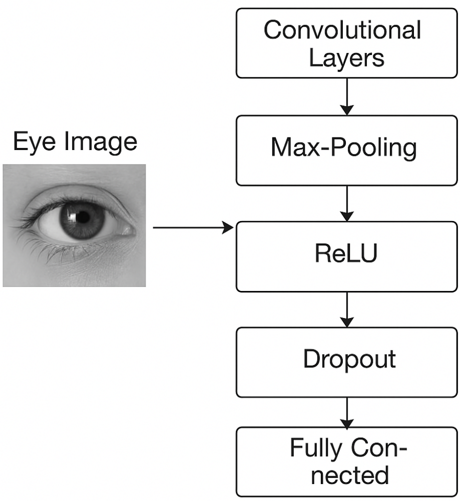

Fig. 1. Overview of the CNN architecture for anemia detection using eye images.

3.4 Training Setup

The dataset was provided with a predefined split. The validation and test sets were class-balanced to ensure fair evaluation across anemic and non-anemic categories. Both models were trained using binary cross-entropy loss and the Adam optimizer, with a learning rate of 0.0001 for initial training and 1e-5 for fine-tuning. To prevent overfitting, we employed early stopping during training by monitoring the validation loss. If the model's performance on the validation set did not improve for a fixed number of consecutive epochs, training was halted, and the weights from the best-performing epoch were restored. This helped ensure that the model generalized well without being overtrained on the training data.

All experiments were conducted using Python and TensorFlow, either on a local machine or on GPU-enabled Google Colab. TensorFlow automatically detects and utilizes available GPU resources to accelerate training and inference. If a GPU is not available, the code runs seamlessly on the CPU, ensuring compatibility across different hardware setups. This flexibility allowed us to develop and test models efficiently without being limited by local hardware constraints.

3.5 Design Rationale

Our design choices were guided by a desire to create an accurate but lightweight solution. The custom CNN was selected for its simplicity and portability, making it feasible for on-device inference. The MobileNetV2 model offered a strong benchmark for accuracy and transfer learning efficiency.

By limiting the model's input to non-invasive eye photographs, we aim to develop a diagnostic tool that is not only accurate but also practical and usable in remote or under developed regions.

4 Experiments

To evaluate the effectiveness of our approach, we conducted two primary experiments. The first experiment tested a custom-built Convolutional Neural Network (CNN) trained from scratch. The second experiment used a MobileNetV2 model with transfer learning. These experiments allowed us to compare trade-offs between speed, size, and accuracy, while using the same dataset splits and evaluation metrics.

4.1 Experiment 1: Custom Convolutional Neural Network Model

Main Purpose: The goal of this experiment is to evaluate how well a custom, lightweight CNN can detect anemia using only eye images. This will serve as a performance baseline and test the feasibility of using a model that is deployable on low-resource devices.

Experimental Setup: The dataset was already split into training (80.5%), validation (9.76%), and testing (9.76%). We trained the CNN using binary cross-entropy loss and the Adam optimizer, with a batch size of 32 and early stopping to prevent overfitting. The model included convolutional layers, dropout, and dense layers leading to a sigmoid output.

Evaluation Metrics:

- **Accuracy** – to measure overall classification correctness.
- **Precision and Recall** – to understand false positives/negatives, especially important in a health screening context.
- **F1-score** – to provide a balance between precision and recall.
- **ROC-AUC** – to evaluate the model's ability to distinguish between classes across thresholds.

Experimental Results: The custom CNN achieved strong performance, particularly on non-anemic cases, with an overall test accuracy of 91%. Precision and recall scores were competitive, though slightly lower for the anemic class. The loss and accuracy curves showed steady convergence with minimal overfitting.

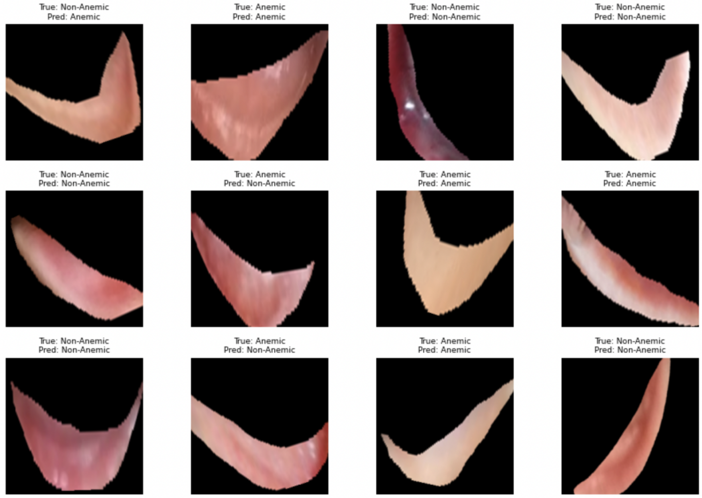

Fig. 2. Predictions on sample eye images from the test set. True and predicted labels are shown above each image.

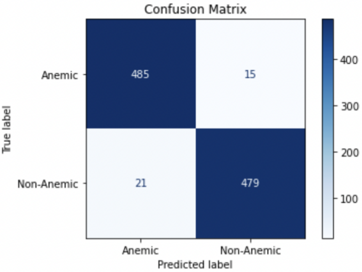

Fig. 3. Confusion matrix on test set. The model correctly classified 485 anemic and 479 non-anemic images, with minimal misclassifications.

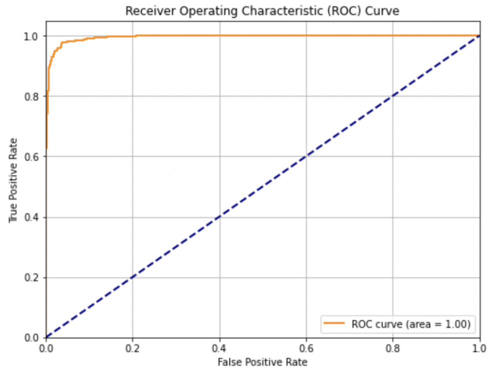

Fig. 4. ROC curve with an AUC of 1.00, showing excellent discriminative ability.

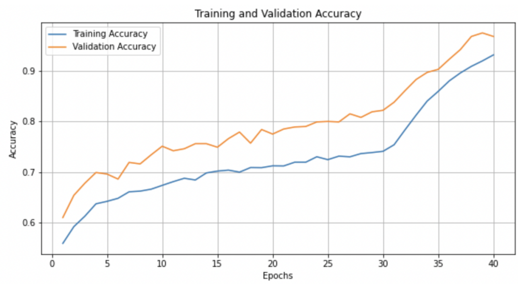

Fig. 5. Training and validation accuracy over 40 epochs. The model shows consistent improvement without overfitting.

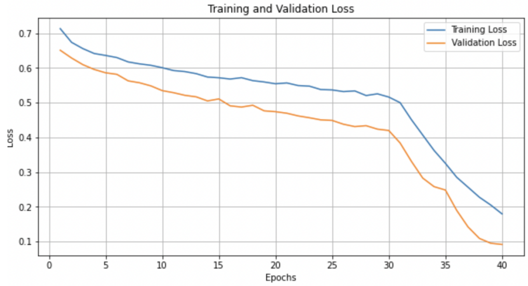

Fig. 6. Training and validation loss over 40 epochs. Both curves show steady decline, with validation loss closely tracking training loss indicating effective learning and limited overfitting.

4.2 Experiment 2: Transfer Learning with MobileNetV2

Main Purpose: To assess whether a pre-trained MobileNetV2 model, fine-tuned on conjunctiva images, could outperform our custom CNN.

Experimental Setup: We loaded MobileNetV2 with ImageNet weights, froze the base layers, and replaced the top with a global average pooling layer, dropout, and dense output. After initial training, we fine-tuned the full model using a low learning rate of 1e-5.

Experimental Results: MobileNetV2 achieved a test accuracy of 92%, slightly higher than the custom CNN. It also yielded stronger performance on anemic predictions, with improved recall and F1-score. However, the model was significantly heavier and slower than the custom CNN, which may limit real-time low-resource device deployment.

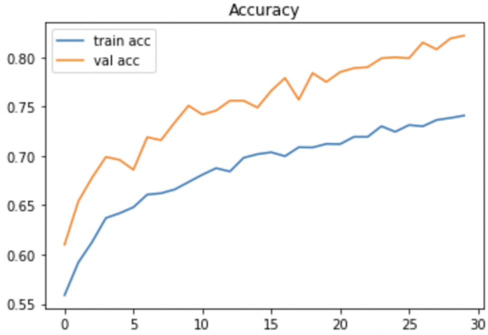

Fig. 7. Training and validation accuracy of the MobileNetV2 model. Validation accuracy steadily improves, suggesting strong generalization.

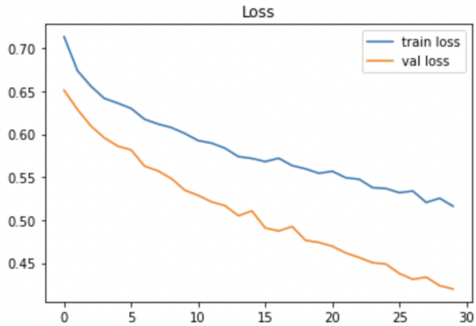

Fig. 8. Training and validation loss of the MobileNetV2 model. Both curves show consistent downward trends, indicating effective learning and minimal overfitting.

4.3 Summary and Next Steps

Both models performed well, with MobileNetV2 offering better accuracy at the cost of speed and model size. The custom CNN, while slightly less accurate, is more efficient and better suited for edge devices.

However, performance on the anemic class lagged slightly behind, suggesting a need for targeted data augmentation. Future work could involve improving localization (e.g., focusing specifically on conjunctiva regions), collecting more diverse samples, or testing on images captured with mobile cameras to evaluate real-world usability.

5 Conclusion

In this project, we developed and evaluated a lightweight CNN-based system for detecting anemia from non-invasive eye images. By focusing on pallor in the conjunctiva, a visible symptom of low hemoglobin, we explored whether deep learning could offer a scalable and accessible alternative to traditional blood tests, particularly in low-resource settings. Our approach involved training both a custom CNN model and a MobileNetV2-based model on a curated dataset of conjunctiva images.

Our experiments demonstrated that while MobileNetV2 achieved slightly higher accuracy, our custom CNN performed competitively with significantly lower complexity and faster inference time. Both models showed strong generalization and were evaluated using accuracy, precision, recall, F1-score, and ROC-AUC metrics. The results support the potential for real-time deployment of AI-powered screening tools that require only a camera and minimal computational resources.

Looking ahead, this work can be extended in several directions. One possibility is to explore severity classification instead of binary prediction, allowing for more informative outputs. Another path is integrating the model into telemedicine platforms or mobile applications for community-level use. Notably, the current model is already portable and can run on systems without dedicated GPUs or high computing power. With further optimization, it can be adapted for real-time use directly on smartphones making anemia screening even more accessible in rural or under-resourced areas. Improvements in conjunctiva localization and dataset balance could further boost reliability, especially for minority classes. Ultimately, this work contributes to building AI systems that are not only technically sound but also practical, inclusive, and impactful in real-world healthcare.

References

1. Appiahene, P., Arthur, E.J., Korankye, S., Afrifa, S., Asare, J.W., Donkoh, E.T.: Detection of anemia using conjunctiva images: A smartphone application approach. Artificial Intelligence in Medicine **139** (2023) 102521

2. Asare, J.W., Appiahene, P., Donkoh, E.T., Dimauro, G.: Iron deficiency anemia detection using machine learning models: A comparative study of fingernails, palm and conjunctiva of the eye images. Engineering Reports **5**(5) (2023) e12667
3. Sehar, N., Krishnamoorthi, N., Kumar, C.V.: Deep learning model-based detection of anemia from conjunctiva images. Journal of Healthcare Engineering **2023** (2023) 1–10
4. Tamir, A., Jahan, C.S., Saif, M.S., Zaman, S.U., Islam, M.M., Khan, A.I.: Detection of anemia from image of the anterior conjunctiva of the eye by image processing and thresholding. In: 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE), IEEE (2017) 8289053
5. Iqbal, S., Khan, T.M., Naveed, K., Naqvi, S.S., Nawaz, S.J.: Recent trends and advances in fundus image analysis: A review. Computers in Biology and Medicine **151** (2022) 106277
6. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P., Ng, A.Y.: Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017)
7. Lab, K.: Clean augmented anemia dataset. https://www.kaggle.com/datasets/t2obd1a1253kmit/clean-augmented-anemia-dataset (2023) Accessed: 2024-04-09.