

Ensemble Learning-Based Non-Invasive Anemia Detection Using Conjunctival Image Analysis

Sona P V¹, Dr. Anil George K

¹St. Thomas College (Autonomous), Thrissur

²St. Thomas College (Autonomous), Thrissur

DOI: <https://doi.org/10.51583/IJLTEMAS.2026.151000018>

Received: 12 October 2026; Accepted: 17 October 2026; Published: 30 October 2026

ABSTRACT

Anemia is a global health issue characterized by low hemoglobin levels, leading to reduced oxygen transport in the body. Traditional diagnosis relies on invasive blood tests, which can be costly and inaccessible. This study proposes a machine learning-based ensemble framework for non-invasive anemia detection using conjunctival images. Various classifiers, including SVM, Random Forest (RF), XGBoost, and ANN, were trained and evaluated. Ensemble learning techniques such as bagging, boosting, stacking, and voting were applied, with the stacking ensemble (RF + ANN + XGB) achieving the highest accuracy (95.77%). Extra Trees emerged as the best-performing individual model with 96.48% accuracy. The results confirm that ensemble models significantly improve classification performance over individual classifiers. Future work should integrate multi-modal approaches to further enhance diagnostic accuracy. This study highlights the potential of ensemble learning for scalable, non-invasive anemia detection in clinical settings.

Index Terms: Anemia detection, machine learning, ensemble learning, stacking, conjunctival images, Random Forest, XGBoost, Extra Trees, medical image classification.

INTRODUCTION

Anemia is a widespread public health concern that impacts populations across underdeveloped, developing, and developed nations [1]. It is characterized by a reduced concentration of hemoglobin and a decreased number of red blood cells, impairing the blood's ability to transport oxygen efficiently [2]. This condition can lead to various health complications, including chronic fatigue, weakness, dizziness, and, in severe cases, organ dysfunction [2]. Among the primary causes of anemia, iron deficiency remains the most prevalent, making early detection and diagnosis critical to mitigating long-term health consequences.[1]

Traditional diagnostic methods for anemia primarily rely on invasive blood tests, such as complete blood count (CBC) analysis, which require specialized laboratory infrastructure, trained personnel, and financial resources [2]. While effective, these methods are often costly, time-consuming, and uncomfortable for patients, particularly in resource-limited settings where access to healthcare is restricted.[2] As a result, there is an increasing demand for non-invasive, cost-effective, and accessible diagnostic alternatives [3].

Machine learning (ML) techniques have emerged as promising tools for advancing anemia detection and monitoring hemoglobin concentration. Recent studies have demonstrated the potential of ML-driven approaches to automate diagnosis through image analysis of conjunctival pallor, palm coloration, and fingernail bed characteristics [4]. However, significant challenges remain, including the lack of standardized datasets, the complexity of existing diagnostic equipment, and the need for models that generalize well across diverse populations. Addressing these challenges requires innovative solutions that leverage artificial intelligence to enhance diagnostic accuracy while maintaining practical applicability.

This study presents a non-invasive ML-based framework for anemia detection using conjunctival images. The research utilizes a dataset of segmented conjunctival images, extracted color and texture-based features,

and ensemble learning techniques to classify images as anemic or non-anemic [4-5]. The dataset was split into training, validation, and test sets, ensuring a realistic representation of real-world conditions. To enhance classification performance, various feature selection techniques, including Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE), were employed to reduce dimensionality while preserving critical information [4]. Multiple ML models—including Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Networks (ANN), and XGBoost—were trained and evaluated. Further, ensembling techniques such as bagging, boosting, stacking, and voting were implemented to improve overall performance [6].

The primary objective of this research is to develop a robust, accessible, and non-invasive anemia detection system capable of operating efficiently in resource-constrained environments. By leveraging machine learning, this study aims to address the limitations of traditional diagnostic methods and contribute to the growing body of work on AI-driven medical diagnostics. The findings of this research have the potential to aid in early anemia detection, enabling timely intervention and improving patient outcomes, particularly in underserved regions where traditional diagnostic tools remain inaccessible.

LITERATURE REVIEW

Anemia has traditionally been diagnosed using invasive laboratory tests such as complete blood count (CBC) and hemoglobin measurements, which require specialized equipment, trained personnel, and medical consumables [1–3,7]. While these methods are accurate, their accessibility is limited in resource-constrained settings, leading to delayed diagnosis and increased health risks. Visual assessment of conjunctival pallor has been explored as a simpler alternative, but its reliability is hindered by subjective interpretation [8].

Recent advancements in machine learning (ML) and computer vision have introduced non-invasive techniques that analyze images of anatomical regions such as the conjunctiva, fingernails, palms, and sclera for anemia detection [4,5,9]. These approaches aim to automate pallor detection and reduce dependence on traditional blood-based diagnostics. Several ML algorithms have been evaluated, including Support Vector Machines (SVM), Convolutional Neural Networks (CNN), k-Nearest Neighbors (k-NN), Decision Trees, and ensemble methods [5,6].

Studies have compared SVM, CNN, and k-NN for iron deficiency anemia detection using fingernail, palm, and conjunctiva images, reporting high classification performance [5]. CNN-based approaches have generally outperformed traditional ML models, with some achieving accuracy rates above 99% [6]. Mobile-based hemoglobin estimation systems using conjunctival images have demonstrated the feasibility of portable diagnostics [12]. Methods incorporating sclera segmentation and vessel extraction have improved classification accuracy, while other works have combined MobileNetV2 with SVM for enhanced performance [5].

Dataset quality plays a critical role in ML model performance. Many available datasets are small, private, and imbalanced, which can lead to biased predictions and poor generalization [6,9]. Examples include datasets of only 100 conjunctiva images [9] and datasets from 20 pregnant patients [9]. Larger datasets, such as smartphone-based collections from 344 patients and the CP-AnemiC dataset with 710 balanced conjunctiva images, have enabled more robust model training [8,10,15].

Various techniques have been applied to improve conjunctiva-based anemia detection, including CNN for automated feature learning [4], U-Net for precise segmentation [10], and RGB thresholding for Android-based real-time screening [9]. Beyond standard imaging, alternative modalities like multiwavelength spectrophotometry [13] and smartphone-based LED illumination systems such as HemaApp [13] have been explored.

Ensemble learning has proven effective in boosting classification accuracy, with Artificial Neural Networks (ANN) achieving over 97% accuracy and stacking ensemble methods reaching up to 99.73% accuracy for

palm image classification [6,11]. Data preprocessing methods such as SMOTE have also been used to address class imbalance [13].

Despite these advancements, significant challenges remain. Model performance can be influenced by lighting conditions, camera specifications, ethnicity, and variability in image acquisition [14]. Small datasets and overfitting reduce generalization, particularly when dataset splitting occurs after augmentation, leading to data leakage and inflated accuracy metrics [16]. Best practices recommend splitting datasets before augmentation to ensure fair evaluation.

Overall, research indicates that ML-driven, non-invasive anemia detection is a promising alternative to traditional diagnostics. Future work should focus on expanding diverse datasets, improving explainability in ML models, addressing generalization issues, and optimizing real-time deployment for use in clinical and resource-limited environments.

METHODOLOGY

Anemia detection has traditionally relied on invasive blood tests, which can be costly, time-consuming, and inaccessible in resource-limited settings. This study aims to develop a non-invasive diagnostic approach using machine learning and image analysis techniques. The methodology involves processing conjunctival images, extracting relevant features, and classifying anemia using various machine learning models. By leveraging ensemble techniques, this study seeks to improve classification accuracy and demonstrate feasibility for real-world applications.

Dataset and Image Preprocessing

The dataset utilized in this study was obtained from Mendeley Data and comprises 710 segmented conjunctival images labeled as anemic or non-anemic. Labeling was based on hemoglobin (Hb) levels recorded in an accompanying Excel file, following WHO criteria, which define anemia in children aged six to 59 months as an Hb level below 11 g/dL. Among the images, 424 (60%) were labeled as anemic and 286 (40%) as non-anemic.

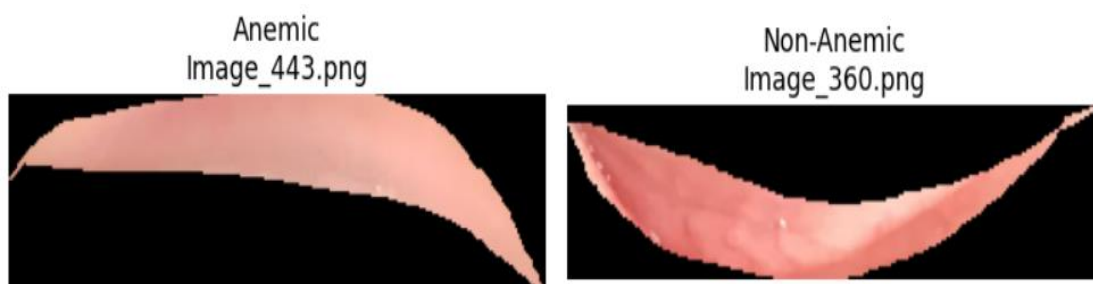


Figure 1: Representative samples from the conjunctival image dataset used for anemia detection. The left image shows an anemic case, while the right image shows a non-anemic case. These segmented eye images were used for training and evaluating the classification models.

To address class imbalance and improve model generalization, augmentation techniques were applied before splitting the dataset. These included random rotations, horizontal and vertical flipping, and brightness adjustments, which introduced variability while preserving clinically relevant features. After augmentation, the dataset was divided into three subsets: 70% for training, 10% for validation, and 20% for testing. Only the training set was balanced during this process. The validation and test sets were left imbalanced to reflect real-world clinical scenarios, ensuring that model performance was evaluated under realistic and practically relevant conditions.

Prior to feature extraction and model training, all images were resized to 224×224 pixels to standardize input dimensions. Additionally, pixel normalization was applied by scaling intensity values between 0 and 1, which facilitated stable gradient updates and improved model convergence. These preprocessing steps ensured that

the models received consistent input, enabling them to accurately capture and analyze the subtle variations in conjunctival color and texture associated with anemia.

This structured dataset preparation and preprocessing approach was essential in enhancing model robustness, maintaining class balance during training, and ensuring generalizability in real-world applications.

Feature Extraction

Feature extraction was performed to convert conjunctival images into numerical representations that capture relevant characteristics for anemia classification. This step aimed to enhance the models' ability to differentiate between anemic and non-anemic samples by leveraging both color-based and texture-based attributes.

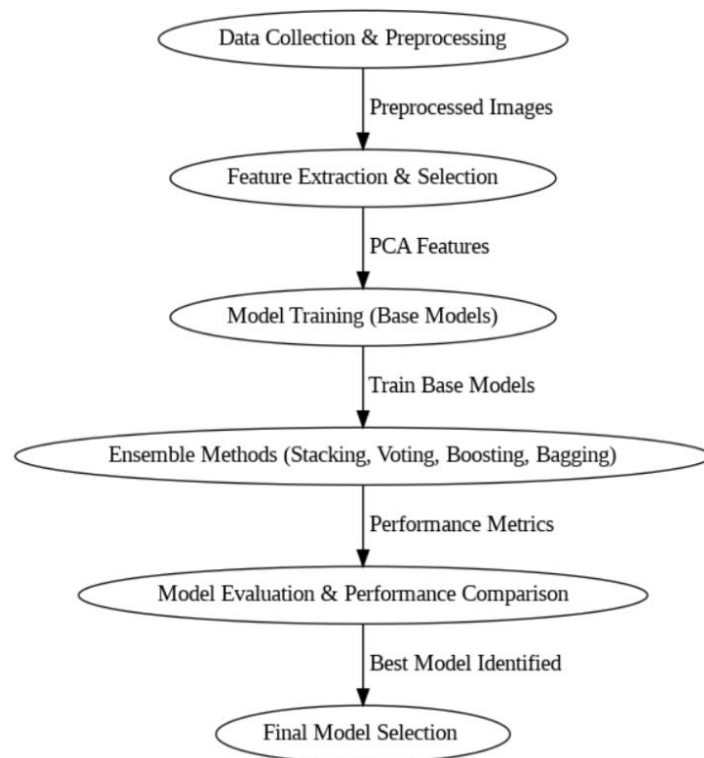


Figure 2: This flowchart illustrates the step-by-step methodology followed in this research.

Color feature extraction was conducted across multiple color spaces, including RGB, HSV, and LAB, as variations in conjunctival pigmentation serve as key indicators of anemia. From each color channel, statistical descriptors such as mean, standard deviation, skewness, and kurtosis were computed to quantify distribution patterns. These statistical measures provided insight into color intensities and variations, which are clinically relevant for detecting pallor associated with anemia.

In addition to color features, texture features were extracted using the Gray-Level Co-Occurrence Matrix (GLCM), a widely used technique for analyzing spatial relationships between pixel intensities. From the GLCM, four essential texture descriptors—contrast, homogeneity, energy, and correlation—were computed to capture textural differences between anemic and non-anemic conjunctival regions. These features helped characterize the structural patterns and smoothness of the conjunctiva, which can differ based on hemoglobin levels.

Feature selection plays a critical role in improving model efficiency by eliminating redundant or less informative features. In this study, Principal Component Analysis (PCA) was applied to perform dimensionality reduction. PCA transforms the original feature space into a set of uncorrelated principal components that capture the most significant variance within the data. By retaining only the top components, this approach reduces computational complexity while preserving essential information. This dimensionality

reduction step ensures that machine learning models receive optimized input representations, enhancing performance and generalization across different classifiers.

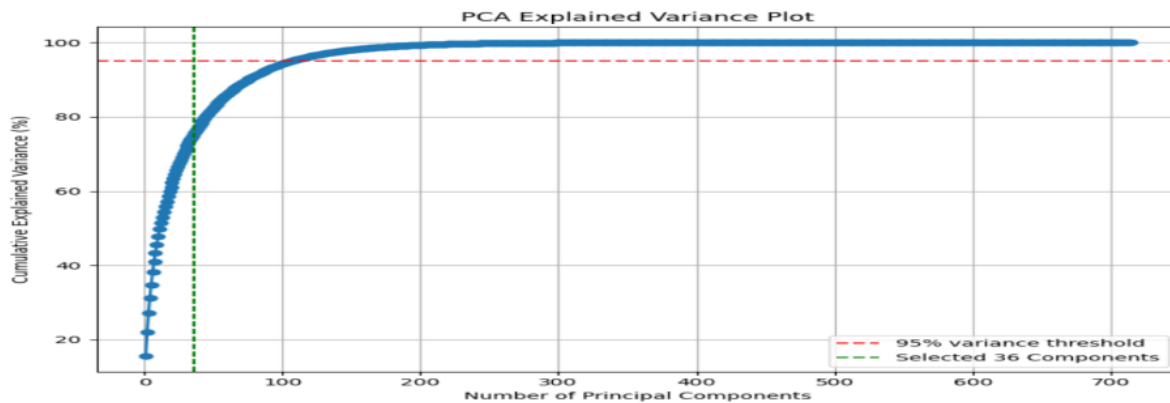


Figure 3: PCA Explained Variance Plot (Feature Selection and Dimensionality Reduction)

The PCA explained variance plot (Figure 3) illustrates the cumulative variance retained as the number of components increases. As shown, the top 36 principal components preserve approximately 95% of the total variance in the dataset. This selection represents a balance between dimensionality and information retention, making it well-suited for downstream classification tasks.

III. Model Training and Classification

Once the feature extraction process was completed, the extracted features were used as input for training various machine learning models. The goal was to classify conjunctival images as anemic or non-anemic based on their color and texture features. The study implemented individual classifiers and ensemble learning techniques to improve classification performance. The individual classifiers trained in this study included:

- Support Vector Machine (SVM): Selected for its effectiveness in handling high-dimensional data and ability to find optimal decision boundaries using kernel functions.
- Random Forest (RF): A tree-based model known for its robustness and ability to handle feature importance analysis.
- Artificial Neural Networks (ANNs): A multi-layer perceptron model designed to capture complex patterns in conjunctival features.
- XGBoost: A gradient boosting-based approach optimized for structured data classification.

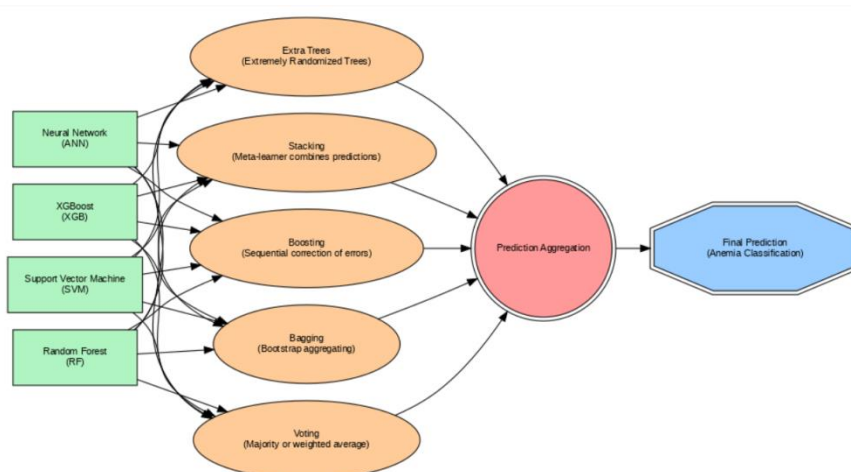


Figure 4: Visualization of Ensemble Learning Methods Used. The diagram illustrates different ensemble techniques, including Bagging (Random Forest), Extra Trees, Boosting, Voting, and Stacking.

To enhance classification performance, ensemble learning techniques were applied. These included:

- Bagging (Bootstrap Aggregation) using Random Forest to reduce variance and improve stability.
- Boosting techniques, specifically CatBoost which adjust model weights iteratively to focus on misclassified samples.
- Stacking, where multiple base classifiers were combined, and a meta-classifier (logistic regression) was used to refine final predictions.
- Voting classifiers, utilizing both soft and hard voting strategies to aggregate predictions from multiple models.

The models were trained using the training set, with hyperparameter tuning performed to optimize performance. The validation set was used to monitor performance and adjust model parameters accordingly, while the test set was used for final evaluation.

Model Evaluation and performance Metrics

The models that underwent training were rigorously assessed utilizing a comprehensive array of performance metrics to evaluate their proficiency in the classification of anemia. The evaluation process was conducted on the test dataset, which was intentionally left imbalanced to accurately represent real-world clinical scenarios. This methodology guaranteed that the models were scrutinized under authentic conditions where class distributions are not uniform, especially characterized by a scarcity of non-anemic instances. The subsequent evaluation metrics employed were:

- Accuracy: The ratio of accurately classified samples across both categorical classes.
- Precision: The ratio of true anemic instances among all cases predicted as anemic, thereby demonstrating the reliability of the model in identifying positive instances.
- Recall (Sensitivity): The ratio of actual anemic instances accurately identified by the model, which is crucial for minimizing false negatives within a medical context.
- F1-Score: The harmonic mean of precision and recall, providing a singular metric that reconciles both dimensions in the context of class imbalance.
- Matthews Correlation Coefficient (MCC): A resilient metric that accounts for true positives, false positives, true negatives, and false negatives, rendering it particularly appropriate for imbalanced datasets. An MCC value approaching +1 signifies robust predictive capability.
- Confusion Matrix: Offers a meticulous delineation of correct and incorrect predictions across each class, thereby illuminating the model's efficacy in differentiating between anemic and non-anemic cases.
- Receiver Operating Characteristic (ROC) Curve and AUC-ROC: The ROC curve elucidates the relationship between the true positive rate and the false positive rate across varying thresholds, while the Area Under the Curve (AUC) quantifies the model's overall classification proficiency.
- Precision-Recall (PR) Curve: Particularly pertinent for imbalanced datasets, the PR curve accentuates the trade-off between precision and recall, thereby indicating the model's performance in recognizing the minority class (anemic instances).

A comprehensive comparative analysis was undertaken between distinct classifiers (including SVM, ANN, RF, and XGBoost) and ensemble techniques (such as Bagging, Boosting, Voting, and Stacking) to determine the most effective strategies for the identification of anemia.

RESULT AND DISCUSSION

The performance of the proposed anemia classification models was evaluated using multiple statistical and graphical metrics. The results provide insights into the strengths and limitations of individual classifiers and ensemble learning techniques.

Classification Performance

Table 1 summarizes the classification performance of all individual and ensemble models. Among individual classifiers, the Extra Trees Classifier outperformed all others, achieving the highest accuracy of 96.48%, with corresponding precision, recall, and F1-score values above 96%, making it the strongest standalone model in this study.

TABLE 1. Classification performance of different models.

Model	Accuracy	Precision	Recall	F1 Score	MCC
Bagging (RF)	0.9296	0.9303	0.9296	0.9298	0.8547
Extra Trees	0.9648	0.9659	0.9648	0.9649	0.9282
CatBoost	0.9296	0.9339	0.9296	0.93	0.8592
Voting (All Models)	0.9155	0.9163	0.9155	0.9157	0.8255
Voting (SVM + XGB + RF)	0.9366	0.9397	0.9366	0.937	0.8723
Voting (RF + ANN + XGB)	0.9507	0.9519	0.9507	0.9509	0.8992
Voting (RF + ANN + SVM)	0.9225	0.9224	0.9225	0.9224	0.8384
Stacking (All Models)	0.9577	0.9618	0.9577	0.958	0.917
Stacking (SVM + XGB + RF)	0.9507	0.9561	0.9507	0.9511	0.904
Stacking (RF + ANN + XGB)	0.9577	0.9618	0.9577	0.958	0.917
Stacking (RF + SVM + XGB)	0.9507	0.9561	0.9507	0.9511	0.904

Other ensemble techniques like Voting and Boosting CatBoost also showed solid performance, outperforming individual models such as SVM and ANN. SVM, in particular, exhibited the lowest performance metrics, with an accuracy of 78.17% and recall of 78.17%, indicating limited sensitivity to anemic cases.

Confusion Matrix Analysis

The confusion matrix analysis further reinforces the superior performance of Extra Trees and Stacking models. The Extra Trees classifier correctly identified 81 out of 85 anemic cases, demonstrating high recall and minimal false negatives—an essential requirement in medical diagnostics, where missed anemic cases can lead to severe health consequences. The Stacking ensemble maintained similarly high sensitivity while reducing false positives, ensuring a balanced performance across both classes. This makes it particularly suitable for real-world deployment, where reliability across varied patient cases is critical.

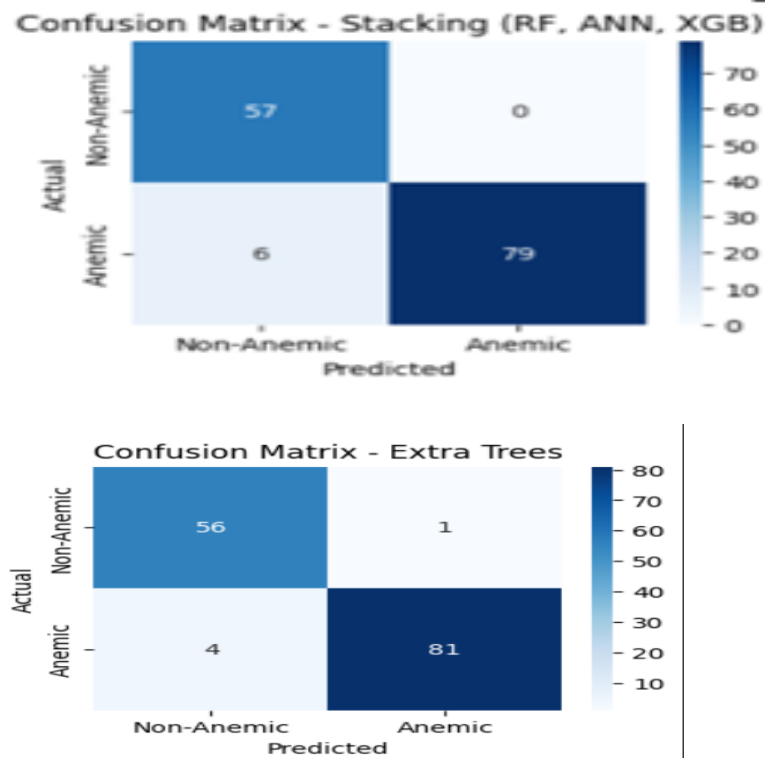


Figure 5. Confusion Matrix of Best Models

III. Receiver Operating Characteristic (ROC) and Precision-Recall (PR) Analysis

The ROC curves (Figure 6) illustrate the trade-off between true positive rate (recall) and false positive rate. The AUC-ROC values confirmed the superior performance of RF (0.98), XGB (0.97), ANN (0.96), and stacking (≈ 1.00), indicating a high discriminatory ability. In contrast, SVM had a lower AUC of 0.88, suggesting weaker generalization to the test data.

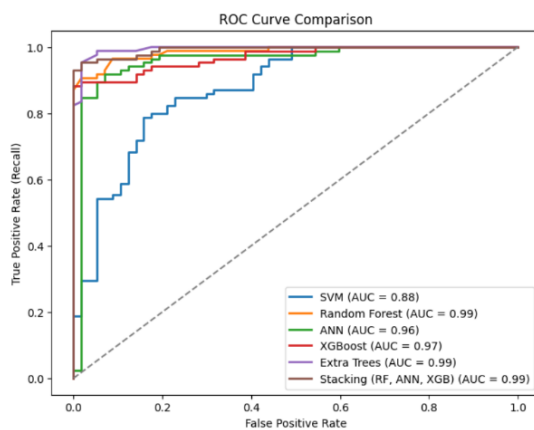


Figure 6. ROC Curve Comparison

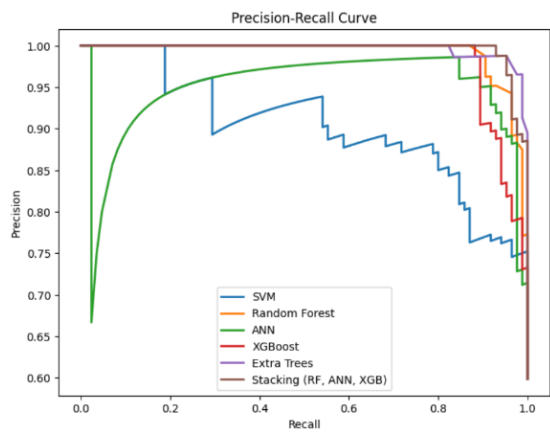


Figure 7. Precision-Recall curve

The Precision-Recall curves (Figure 7) highlight the model's ability to handle the class imbalance present in the test dataset. The stacking model consistently maintained high precision and recall across different thresholds, reinforcing its robustness. Notably, the ANN model exhibited a sharp decline in precision at lower recall values, suggesting it may generate false positives under certain conditions. SVM showed an unstable PR curve, further affirming its limited suitability for this task.

Comparative Analysis and Practical Implications

The results demonstrate that ensemble learning, particularly stacking, significantly enhances classification performance by leveraging the strengths of multiple models. Traditional classifiers like SVM struggled with imbalanced data, showing limited generalization capability, while deep learning approaches such as ANN achieved competitive recall but required careful optimization to avoid overfitting. Tree-based methods (RF and XGB) provided a strong and consistent baseline due to their robustness in handling structured, feature-rich datasets.

Although Extra Trees achieved the highest accuracy (96.48%) among individual models, it was not selected as the most clinically suitable model due to potential overfitting and its tendency to slightly prioritize specificity over sensitivity. In medical diagnostics, minimizing false negatives (i.e., ensuring anemic cases are correctly identified) is critical. The Stacking ensemble (RF + ANN + XGB) achieved a more balanced trade-off between sensitivity and specificity, making it more reliable for detecting anemia across diverse patient cases.

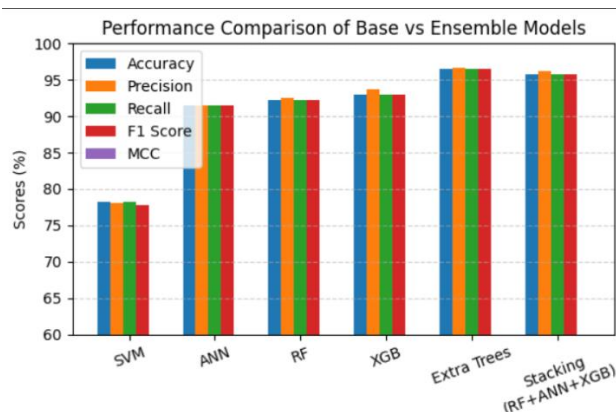


Figure 8: Model Performance Comparison

The inclusion of Matthews Correlation Coefficient (MCC) further confirmed the robustness of stacking ensembles, as they maintained high performance even under imbalanced test conditions. This balance is essential for real-world deployment, where data distribution may vary.

From a practical perspective, the proposed stacking-based approach offers multiple advantages:

- **Early and Non-Invasive Screening:** Enables rapid anemia detection from conjunctival images without requiring invasive blood tests, particularly beneficial in low-resource settings.
- **Reduced Clinical Risk:** High sensitivity minimizes missed diagnoses, while high precision reduces unnecessary follow-up tests for non-anemic individuals.
- **Scalable Deployment:** PCA-driven feature reduction ensures low computational cost, enabling integration into portable devices or mobile health applications.
- **Support for Medical Decision-Making:** Acts as a pre-screening tool, flagging high-risk individuals for confirmatory blood testing, thereby optimizing healthcare resource allocation.

LIMITATIONS AND FUTURE DIRECTIONS

This study is limited by the small, population-specific dataset and variations in image acquisition, which may affect generalizability. PCA-based dimensionality reduction, while improving efficiency, may have removed subtle features. Although severity classification was included, its performance requires further optimization. Future work should use larger, more diverse datasets and refine multi-task models for robust anemia detection, severity classification, and hemoglobin prediction in real-world settings.

CONCLUSION

This study investigated the application of machine learning models and ensemble learning techniques for non-invasive anemia detection using segmented conjunctival images. The results demonstrated that ensemble methods, particularly stacking, can substantially improve classification performance by combining the strengths of multiple models. Among individual classifiers, tree-based approaches such as Random Forest and XGBoost delivered strong predictive results, while ANN showed competitive performance but required careful parameter optimization for better generalization. Ultimately, the stacking ensemble (RF + ANN + XGB) proved to be the most effective, achieving a balanced combination of high accuracy, recall, precision, F1-score, and MCC, making it well-suited for practical clinical use.

While Extra Trees achieved the highest accuracy (96.48%) among standalone models, it was not chosen as the best due to its tendency toward overfitting and prioritizing specificity over sensitivity. In contrast, the stacking ensemble provided a more balanced trade-off, reducing false negatives—a critical factor in medical diagnostics where missing an anemic case could have serious health implications.

Despite these strong results, certain limitations remain. The dataset, though imbalanced to reflect real-world conditions, was relatively small in size and may not fully capture variability in imaging conditions. Furthermore, the study focused solely on binary classification without incorporating severity classification or additional clinical parameters such as age, gender, and hemoglobin levels, which could enhance predictive performance.

In conclusion, this research demonstrates that carefully designed ensemble learning pipelines, combined with robust feature extraction and dimensionality reduction techniques, can offer a reliable, non-invasive, and scalable approach for anemia detection. Expanding the dataset, incorporating multi-class severity prediction, and integrating clinical parameters represent promising directions for future work, ultimately moving closer to real-world clinical deployment.

REFERENCES

1. Kassebaum, N. J., Jasrasaria, R., Naghavi, M., et al. (2014). A systematic analysis of global anemia burden from 1990 to 2010. *Blood*, 123(5), 615-624. <https://doi.org/10.1182/blood-2013-06-508325>
2. McLean, E., Cogswell, M., Egli, I., Wojdyla, D., & de Benoist, B. (2009). Worldwide prevalence of anemia, WHO Vitamin and Mineral Nutrition Information System, 1993–2005. *Public Health Nutrition*, 12(4), 444-454. <https://doi.org/10.1017/S1368980008002401>

3. von Schenck, H., Falkensson, M., & Lundberg, B. (1986). Evaluation of “HemoCue,” a new device for determining hemoglobin. *Clinical Chemistry*, 32(3), 526-529. <https://doi.org/10.1093/clinchem/32.3.526>
4. Asare, J. W., Appiahene, P., Donkoh, E. T., & Dimauro, G. (2023). Detection of anemia using conjunctiva images: A smartphone application approach. *Informatics in Medicine Unlocked*, 40, 101283. <https://doi.org/10.1016/j.imu.2023.101283>
5. Asare, J., Appiahene, P., Donkoh, E., & Dimauro, G. (2023). Iron deficiency anemia detection using machine learning models: A comparative study of fingernails, palm, and conjunctiva of the eye images. *Engineering Reports*, 5(11), e12667. <https://doi.org/10.1002/eng2.12667>
6. Appiahene, P., Dogbe, S. S. D., Kobina, E. E. Y., Dartey, P. S., Afrifa, S., Donkoh, E. T., & Asare, J. W. (2023). Application of ensemble models approach in anemia detection using images of the palpable palm. *Medicine in Novel Technology and Devices*, 20, 100269. <https://doi.org/10.1016/j.medntd.2023.100269>
7. GBD 2021 Anaemia Collaborators. (2023). Prevalence, years lived with disability, and trends in anaemia burden by severity and cause, 1990–2021: findings from the Global Burden of Disease Study 2021. *The Lancet Haematology*, 10(8), e574-e586. [https://doi.org/10.1016/S2352-3026\(23\)00160-6](https://doi.org/10.1016/S2352-3026(23)00160-6)
8. Asare, Justice Williams; Appiahene, Peter; Donkoh, Emmanuel. (2023). CP-AnemiC (A Conjunctival Pallor) Dataset from Ghana. Mendeley Data, V1. <https://doi.org/10.17632/m53vz6b7fx.1>
9. Collings, S., Ellingson, K. D., Keeler, C., et al. (2016). Non-invasive detection of anaemia using digital photographs of the conjunctiva. *PLoS One*, 11(4), e0153286. <https://doi.org/10.1371/journal.pone.0153286>
10. Appiahene, P., Chaturvedi, K., Asare, J. W., Donkoh, E. T., & Prasad, M. (2023). CP-AnemiC: A conjunctival pallor dataset and benchmark for anemia detection in children. *Medicine in Novel Technology and Devices*, 18, 100244. <https://doi.org/10.1016/j.medntd.2023.100244>
11. Dimauro, G., Bevilacqua, V., Colizzi, L., et al. (2021). A machine learning approach for anemia detection based on conjunctiva images. *Computer Methods and Programs in Biomedicine*, 200, 105911. <https://doi.org/10.1016/j.cmpb.2020.105911>
12. Mannino, R. G., Myers, D. R., Tyburski, E. A., et al. (2018). Smartphone app for non-invasive detection of anemia using only patient-sourced photos. *Nature Communications*, 9(1), 4924. <https://doi.org/10.1038/s41467-018-07262-2>
13. Shao, J., Wong, J. Y., Poljak, M., et al. (2017). HemaApp: Noninvasive blood screening of hemoglobin using smartphone cameras. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3), 97. <https://doi.org/10.1145/3130952>
14. Muljono, S. A., Wulandari, H. A., Azies, M., Naufal, W. A., Prasetyanto, F. A., & Zahra, F. A. (2024). Breaking boundaries in diagnosis: Non-invasive anemia detection empowered by AI. *IEEE Access*, 12, 9292-9307. <https://doi.org/10.1109/ACCESS.2024.3353788>
15. Asare, Justice Williams; Appiahene, Peter; Donkoh, Emmanuel. (2023). CP-AnemiC (A Conjunctival Pallor) Dataset from Ghana. Mendeley Data, V1. <https://doi.org/10.17632/m53vz6b7fx.1>
16. Ali, M. A., Dornaika, F., Cherry, A., et al. (2024). Advancing medical image processing: A comparative analysis of ‘Naturalize’ and traditional augmentation techniques for enhanced diagnostic accuracy. *IC2SPM*, 138–143. <https://doi.org/10.1109/ic2spm62723.2024.10841335>