Journal of Information Systems Engineering and Management
2025, 10(31s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

An In-Depth Assessment and Comparison of Learning Methods for Non-Invasive Anemia Identification

Mehfooza. M[1], Haroon Basha. I[2], Deepa. D[1], Padmavathy.T[1]
[1]SCOPE, Vellore Institute of Technology, Vellore, 632014, TN, India. [2]AMET Business School, AMET University, Chennai, TN, India

ARTICLE INFO

Received: 24 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Anemia, caused by a reduced number of red blood cells or changes in cell structure, poses a significant health issue globally, particularly for children and pregnant women. Clinical diagnosis can be obstructed by patient hesitation, a lack of healthcare personnel in remote areas, and limited resources. However, due to its accessibility, affordability, and non-invasive nature, machine learning offers a compelling alternative. This paper methodically evaluates current applications of machine learning in the diagnosis of anemia. We analyze prominent algorithms, the characteristics of datasets, performance metrics, and techniques for image augmentation. Our findings illustrate the potential of non-invasive, machine learning-based methods for effective and cost-efficient screening of anemia. This paper highlights the capabilities of machine and deep learning in assessing clinical data and medical imagery to enhance anemia detection.

Keywords: Artificial intelligence; invasive; non-invasive; calculations; lack of iron.

INTRODUCTION

A serious worldwide health concern, anemia primarily affects youngsters and expecting mothers Ref. [69]. The World Health Organization estimates that anemia, which is mostly caused by an iron deficiency, affects 32.9% of the world's population, including 39.9% of pregnant women and 41.9% of children under the age of six Ref. [68]. Anaemia occurs when the number of red blood cells is reduced or when their integrity is impaired Ref. [67]. This can be caused by excessive cell death, blood loss, impaired production, or a decrease in the total number of red blood cells Ref. [69]. In order to prevent irreversible organ damage, prompt diagnosis and therapy are crucial Ref. [66].

Pregnant women and children are more likely to experience anemia-related problems, which can vary by location and raise maternal and infant mortality rates. Fatigue, weakness, light headedness, and excessive drowsiness are all symptoms of anemia. Anaemia also has a negative impact on children's development and adult productivity. Serious health issues like excessive fatigue and pregnancy troubles can result from untreated anemia. Additionally, chronic conditions can raise the risk of anemia.

Hemoglobinopathies, schistosomiasis, malaria, diabetes, renal syndrome, cancer, HIV/AIDS, inflammatory bowel disease, and cardiovascular problems are all frequently associated with anemia. Iron deficiency anemia, sickle cell disease, thalassaemia, and aplastic anemia are among the various forms of anemia, which range in severity from moderate to severe.

Traditional diagnostic techniques for anemia encounter several challenges, such as exorbitant expenses, restricted availability of qualified personnel and equipment in isolated areas, problems with quality control, and patient resistance to testing. Healthcare personnel are also at risk of infection from invasive diagnostic procedures. As a result, non-invasive methods are becoming more and more important, particularly those that use machine learning. Blood samples are usually needed for the diagnosis of anemia in current clinical procedures, which are time-consuming, costly, labor-intensive, and risky for disease transmission.

Visual assessments of the conjunctiva, palm, or fingernails are used by some medical institutions for speedier diagnosis Ref. [66,57,56], but their consistency is hampered by observer variability and limited sensitivity Ref. [60,55]. These problems make early detection difficult, especially in places where access to lab facilities is scarce Ref. [58].

Non-invasive techniques, such diagnostics using a smartphone, offer appealing substitutes for anemia monitoring. Researchers are looking toward non-invasive, efficient methods that use machine learning and medical imaging to increase the accuracy of diagnoses.

The application of machine learning for the detection of anemia is examined in this work. It compares different machine learning algorithms according to their methodology, evaluation metrics, image enhancing tactics, source and quantity of datasets, and model performance accuracy. It specifically evaluates effective non-invasive techniques using medical images. The research issues that guide this evaluation are presented in Figure 1, which will help future efforts develop accurate, non-invasive models for detecting anemia.

Because machine learning is inexpensive, easy to use, and non-invasive, it is being studied for the diagnosis of anemia Ref. [54]. Assessing pallor in tissues such as the tongue Ref. [48], palm, fingernail bed Ref. [49], and conjunctiva Ref. [64,50–53] is a common step in these methods. Given that the conjunctiva has less connective tissue and no subcutaneous layers, dermis, or epidermis, it has been acknowledged that conjunctival pallor is a more reliable signal in image analysis Ref. [51].

Since the 1960s, research on AI has advanced dramatically, especially in the last ten or so years in the healthcare industry Ref. [47]. The incorporation of AI frameworks for diagnostic purposes has been spurred by the shortage of medical experts in underdeveloped countries Ref. [46]. There are many advantages to using AI in healthcare, including lower expenses, more effectiveness, better resource management, and fewer medical errors.

Ref. [54] point out that although traditional clinical techniques for identifying anemia are useful, they can be time-consuming, error-prone, and necessitate specialist laboratory personnel. They highlight the need for reliable, automated, and affordable methods of detecting abnormalities in red blood cells.

Ref. [43] draw attention to the impact of AI on healthcare, emphasizing how it may enhance diagnostic accuracy, manage large health databases, and streamline healthcare administration. In addition to detecting anemia, machine learning algorithms can identify Ref. [70-72] those who are at risk for long-term conditions like diabetes and heart disease. Ref. [45] emphasize that by enhancing error detection, patient classification, and medication oversight, AI-powered decision support systems can improve patient safety.

Ref. [42] developed a multi-wavelength non-invasive photometry-based AI-driven anemia classification system. In order to record signals at four different wavelengths, their study used a finger-mounted photoplethysmogram (PPG) device. The collected features were examined using a machine learning-based hierarchical ensemble classification approach, which produced high sensitivity and specificity results.

Ref. [41] used a neural network-based technique to assess hemoglobin levels by collecting pictures of the fingertip. They created a non-invasive tool and a mobile app, showcasing their ability to identify anemia early.

Ref. [40] proposed a non-invasive anaemia detection device based on multi-wavelength spectrophotometry, utilizing light-emitting diodes and fiber optics. This method addresses limitations of traditional finger probes. Their system accurately detected low haemoglobin levels. They also developed an algorithm to estimate haemoglobin levels, achieving promising results.

Ref. [40] presented a non-invasive method of detecting anemia based on multi-wavelength spectrophotometry, which uses fiber optics and light-emitting diodes. This method overcomes the drawbacks of conventional finger probes. In addition to creating an algorithm for calculating hemoglobin levels, their system successfully identified reduced hemoglobin levels, yielding encouraging results.

The field benefits from this review by: Analyzing 14 original research that used medical photos to identify anemia using machine learning. Sorting results into four main groups: (1) machine learning methods, (2) assessment methodologies, (3) dataset properties (size, augmentation), and (4) model performance. Providing information and recommendations to direct future studies in the development of efficient, non-invasive models for the identification of anemia.

TECHNICAL BACKGROUND

This narrative review provides an objective and critical assessment of the current body of evidence about the use of machine learning in anemia detection. Using the flowchart, the reporting in this study complies with the standards established. On 2021, the review protocol was registered in the Mendeley dataset repository.

From 2012 to 2025, relevant publications were found by searching the Cochrane, Google Scholar, and PubMed databases. Terms related to clinical testing, examinations, and image-based anemia detection were included in the search strategy. To locate more pertinent research, the reference lists of the papers that were found were also looked through.

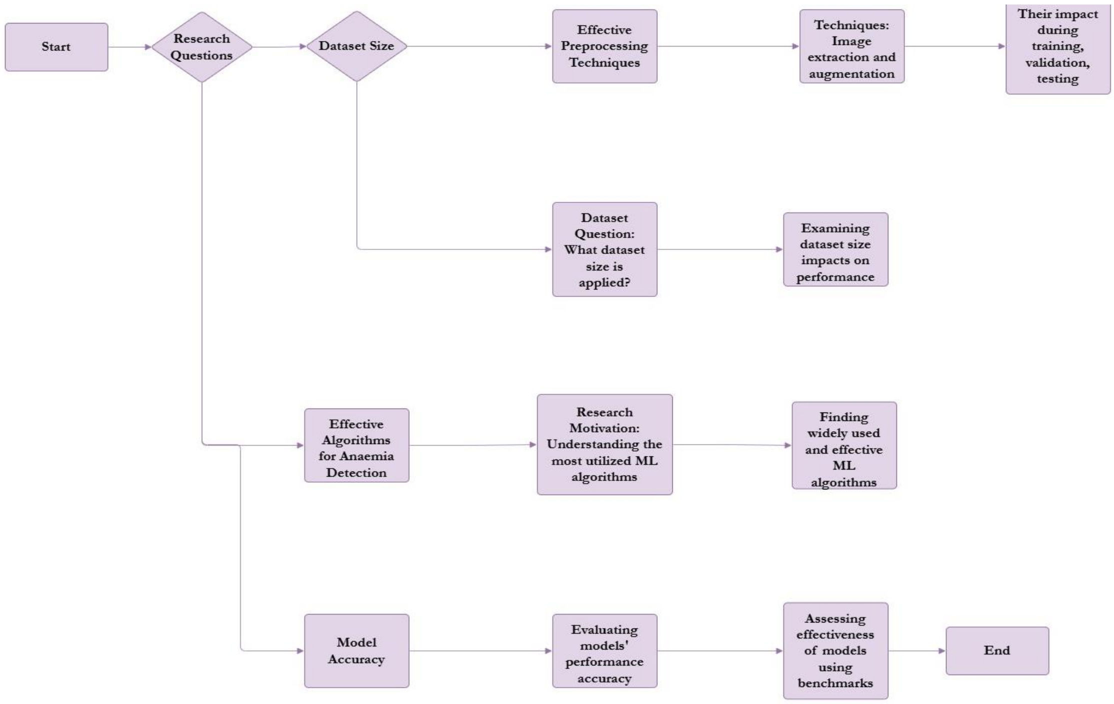

Figure 1: Guidance Of Research Questions

Following a thorough literature analysis, the search parameters were decided. Anemia detection, anemia anemia detection, anemia diagnosing, anemia diagnosing, artificial intelligence, deep learning, "detection of anemia," "diagnosing of anemia," "machine learning," and "natural language processing" were among the terms that made up the final search query. Only English-language publications were included in the search.

Excluded studies were those that did not include medical images (e.g., photos of the tongue, conjunctiva, palm, or nails), were not focused on the diagnosis or detection of anemia, or were categorized as opinion or review pieces. A conceptual framework for the study is shown in Figure 2. The search strategy's criteria are shown in Figure 3. A visual overview of the key studies mentioned in Table 1 and the machine learning models or algorithms they used is provided in Figure 4.

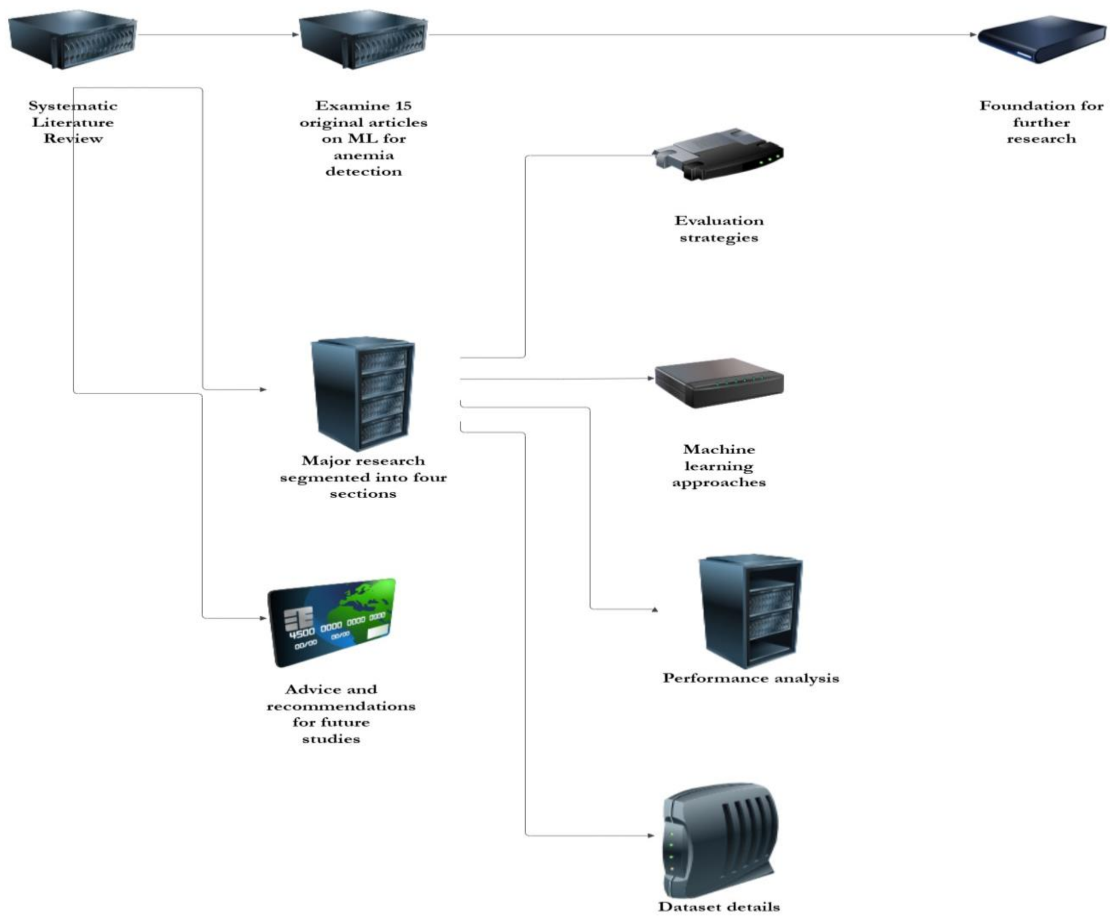

Figure 2: Conceptual Framework for the Study

Application of Machine Learning Algorithms or Models in Anaemia Detection

The science and engineering discipline devoted to developing computer systems that display intelligent behavior commonly associated with human behaviors is known as artificial intelligence Ref. [47]. One important branch of artificial intelligence is machine learning, which focuses on statistical methods and models that let computers learn from data and past experiences without explicit programming Ref. [39, 47].

For instance, Ref. [38] notes that "Plan Analyzer," a system developed for medical education research, was utilized to diagnose cardiac disease in the late 1990s as part of early attempts at computer-aided diagnosis.

Finding correlations between variables, categorizing ideas based on certain standards, forecasting, and identifying objects with similar patterns are just a few of the problems that machine learning can solve Ref. [37]. Support Vector Machines (SVM), Random Forest (RF), Gradient Boosting Decision Trees (GBDT), Artificial Neural Networks (ANN), Convolutional Neural Networks, Naïve Bayes, k-Nearest Neighbors (k-NN), and ensemble models are examples of frequently used machine learning algorithms Ref. [36,37]. Support Vector Machines (SVMs) are popular classification methods that expand upon previous approaches Ref. [36,37]. Similar to discriminant analysis, support vector machines (SVMs) operate on the assumption that data is "separable," meaning that a functional separator can separate it into different categories Ref. [35]. SVMs expand on this premise by highlighting a number of concepts based on statistical learning theory. The way that items are arranged on a decision plane is known as classification Ref. [48].

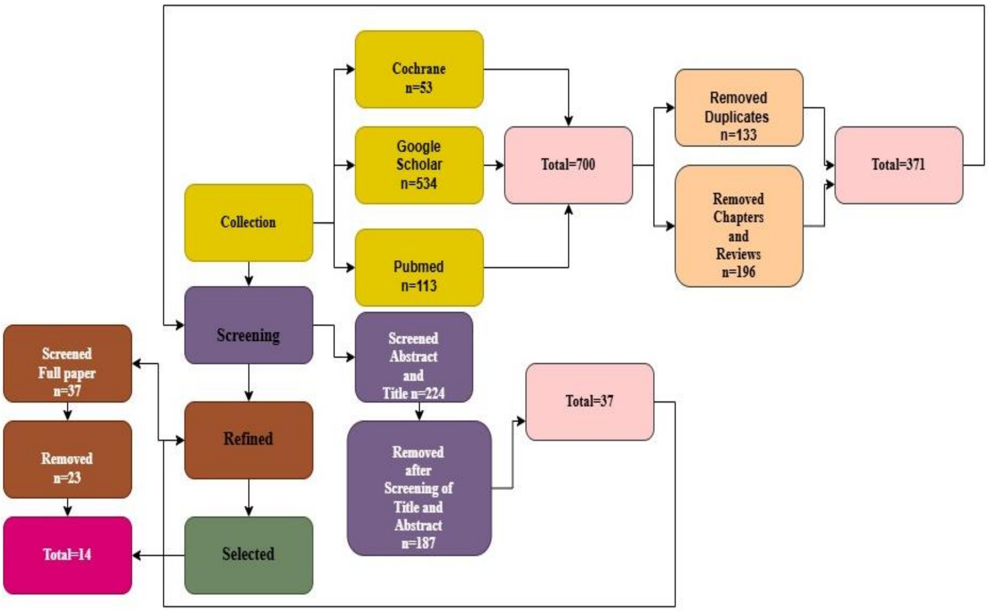

Figure 3: The Search Strategy Criteria

After being trained on a randomly chosen portion of the original training data, Random Forest uses several Classification and Regression Trees (CART), each of which casts a vote Ref. [35]. The average of the votes from each individual tree determines the final classification. Traditional methods of identifying medical conditions like anemia are sometimes complex, expensive, error-prone, time-consuming, and require the expertise of medical professionals Ref. [34].

This emphasizes the need for reliable and reasonably priced techniques that facilitate the early identification and diagnosis of anemia Ref. [34]. presents novel methods for detecting iron deficiency anemia by using Plain Convolutional Neural Networks to analyze peripheral blood imaging samples. Very Deep Convolutional Networks (VGG19), Residual Networks (RESNET-50), and PCNNs with different layers and data augmentation have all been researched.

Table 1: List of Primary Studies and its Learning Algorithms Employed

Paper Reference	Detailed Information
Ref. [14]	The CNN utilized ResNet, MobileNetV2, and DenseNet architectures. Performance was evaluated using weighted log loss, precision, and AUC. Data enhancement techniques included rotation, zooming, translation, and flipping. The dataset, comprising 402 samples, was sourced from the Sankara Nethralaya in Chennai, India. MobileNetV2 produced the most favorable outcomes, achieving 74.4% WLLM, 67% AUC, and 567% accuracy.
Ref. [13]	An SVM model assessed accuracy, specificity, and sensitivity without employing data augmentation. This dataset, consisting of 100 samples, was obtained from the Maulana Azad National Institute of Technology in Bhopal, India. The model reached an accuracy of 92.99%.
Ref. [57]	Various metrics, including accuracy, specificity, sensitivity, and confusion matrix, were applied in an ANN model. Techniques for data augmentation such as translation, rotation, and mirroring were utilized. The dataset was sourced from Italy and India, comprising a total of 100 samples. The model achieved an accuracy rate of 97.18%.
Ref. [33]	To assess precision, recall, accuracy, and F1-score, a CNN utilizing the AlexNet architecture was implemented. Data augmentation methods were applied, although the specific techniques were not detailed. The dataset consisted of 2007 samples. Employing the Adam optimizer with a learning rate of 0.000112, the accuracy of the model reached 94.116%.
Ref. [30]	The Naïve Bayes model utilized metrics such as accuracy, specificity, and sensitivity. Data augmentation was not performed. The dataset comprised 21 instances. The model achieved an accuracy rate of 90.44%.

Ref. [12]	Accuracy, precision, and recall were evaluated with SVM, K-NN, DT, and ANN models. No data augmentation was applied. The dataset contained 84 instances. The accuracy of the models reached 85.99%.
Ref. [18]	RMSE, LSVR, and MLR were used as evaluation metrics for the MLR, DT, and SVR models. No data augmentation was performed. The dataset (n=100) was obtained from the Chittagong Medical and Cox's Bazar College Hospitals in Bangladesh. The Decision Tree model achieved the highest accuracy at 89%.
Ref. [66]	An algorithm known as RGB Thresholding was utilized. There was no information provided on metrics or data augmentation. The dataset (n=20) was sourced from Dhaka Medical College in Bangladesh. The accuracy of the algorithm was recorded at 79%.
Ref. [19]	A confusion matrix was used as a performance statistic for the Decision Tree (DT), Support Vector Machine (SVM), and K-Nearest Neighbors (K-NN) models. Techniques for data augmentation were not used. The 100 samples in the dataset were from the Bangladeshi hospitals of Chittagong and Cox's Bazar Medical College. At 83%, the Decision Tree model's accuracy was the greatest.
Ref. [11]	Metrics including as accuracy, recall, specificity, and Mean Average Precision (MAP) were used to assess a YOLO v5 Neural Network. Techniques for augmenting data, like rotation and flipping, were used. The 450-sample dataset was acquired from Peru's Universidad Peruana Cayetano Heredia. The accuracy of the model was 82% on mobile devices and 92.89% on PC applications.
Ref. [65]	Metrics such as accuracy, F1-score, AUC, precision, and recall were employed to evaluate the performance of the CNN (AlexNet), DT, SVM, Naïve Bayes, and K-NN models. Data augmentation techniques included translations, rotations, and flipping/mirroring. The dataset comprised 530 samples from ten hospitals in Ghana. The Naïve Bayes model achieved an impressive accuracy of 99%.
Ref. [55]	For the SVM, Random Forest, K-NN, and MobileNet-V2 models, accuracy, sensitivity, and specificity were assessedThe dataset, totaling 220 samples (100 from India and 120 from Italy), was sourced from Karapakkam, Chennai, India, and Bari, Italy. The accuracy of MobileNet-V2 was recorded at 68.189% for the Indian dataset and 90.89% for the Italian dataset.
Ref. [10]	The models used included K-NN, RF, SVM, and Extreme Learning Machine, with performance evaluated using metrics such as F1-score, sensitivity, specificity, and accuracy. No data augmentation was performed. The dataset, which consisted of 201 samples, was obtained from Dr. Sardjito General Hospital located in Indonesia. The Extreme Learning Machine model achieved an accuracy of 99%.
Ref. [9]	For the Scleral Segmentation algorithm, There was no data augmentation applied. The dataset used came from Italy and India, totaling 220 samples. The highest scores recorded were 88% for precision, 83% for recall, 84% for F1-score, and 86% for F2-score.

According to their findings, non-invasive techniques can be used to diagnose anemia. Boosting algorithms are gaining popularity because of their interpretability, ease of use, and forecast accuracy. The XGBoost algorithm, gradient boosting method, and AdaBoost algorithm are among the various boosting algorithms that are available. The primary objective of boosting algorithms is to create a fundamentally weak classifier for continuous learning by building a decision tree from a sample of training data. A popular technique is an Artificial Neural Network (ANN), which is composed of processing units called nodes or neurons and may have several layers. Through the network's interconnected nodes, information is exchanged. Neurons fall into three categories: input (which receives data), hidden (whose primary function is to spot patterns), and output (which determines the network's outcome).

One type of machine learning technique designed for the collection and examination of picture data is the Convolutional Neural Network (CNN) Ref. [33]. The two main components of a CNN are a classification layer and feature extraction. The feature extraction component concentrates on gathering important information and uses distinct descriptors to increase the accuracy of the processed data. The classification layer uses fully connected neurons to transform the data into several dimensions after feature extraction Ref. [31, 32].

The Naïve Bayes method uses classifiers that don't rely on assumptions in addition to Bayes' Theorem. The Naïve Bayes approach is predicated on the idea that an attribute's presence or absence is uncorrelated. The mean and variance of the associated variables can be found using a small portion of the training dataset. When evaluating training, validation, and testing datasets, the k-Nearest Neighbor (k-NN) technique identifies related data. A small positive integer is indicated by the letter "k" in k-NN. In a testing situation, the alternative that benefits most of the neighbors is used.

The identification of anemia is one of the many medical disorders for which machine learning has become essential, successful, and efficient. The use of computer algorithms, such as machine learning, to determine hemoglobin (Hb) levels and diagnose conditions like anemia has recently gained popularity due to its excellent accuracy in assessing the color of palms, nailbeds, and eye conjunctiva from smartphone photos.

Using RGB component extraction, the study described in Ref. [30] uses a non-invasive technique to detect anemia by analyzing clinical symptoms associated with the color intensity of the palms and fingernails.

Using a sample of 20 photos and the Naïve Bayes method, this study achieved 91% accuracy. The use of patient images, including those of the conjunctiva, fingernail color, palm look, and tongue condition, for non-invasive anemia identification has been investigated in a number of studies. One study, cited as Ref. [66], used a Support Vector Machine (SVM) to analyze the conjunctiva and create a non-invasive model for detecting anemia. Using a dataset of 18 photos with known hemoglobin levels, the study's accuracy was 78.93%.

Ref. [14] examined the use of computer vision and image processing methods to detect anemia in conjunctival images by employing the Least Square Support Vector Machine (LS-SVM). The technique achieved 84% accuracy, 93% sensitivity, and 71% specificity by training a model on 78 photos (22 of non-anemic people and 57 of anemic people).

Analyzing segmented conjunctiva pictures for anemia identification using a Convolutional Neural Network, Ref. [29] achieved a sensitivity of 78% when compared to laboratory tests. Ref. [64] used a novel tool using the k-Nearest Neighbor (k-NN) method to identify anemia from conjunctiva photos of 28 anemic and 85 non-anemic individuals, achieving an accuracy of 90%.

Ref. [28] processed images of blood samples using image processing methods such as noise reduction, greyscale conversion, and image enhancement. In order to detect sickle cell anemia, their study compared k-NN, SVM, and other machine learning classification techniques using both graphical and mathematical features.

Ref. [27] developed a smartphone app that measures hemoglobin non-invasively to determine a patient's anemia condition. Ref. [26] used 273 datasets from adult patients and 178 from children to create an application for measuring hemoglobin levels and identifying anemia. The HemoCue Hb 300, a point-of-care device that takes a tiny blood sample, and laboratory blood analyses—the gold standard—were used to validate the hemoglobin readings from the app. The app's accuracy and bias were 4.5 g/dl and 0.4g/dl in the clinic-based sample and 3.6 g/dl and 1.5 g/dl in the children's sample, respectively.

In their work, Ref. [65] used 526 palm photos to detect anemia using CNN, Naïve Bayes, Decision Tree, k-NN, and SVM. After augmenting the dataset to 2634 images using flipping and rotation, they were able to get an astounding accuracy of 99% with Naïve Bayes. Given the dangers associated with examining the conjunctiva, the study suggests that the palm is a useful area for identifying anemia in children.

Using smartphone photos to analyze nail color is often associated with good accuracy in hemoglobin level calculation and computational techniques for anemia detection Ref. [49]. According to research by Ref. [25], a neural network regressor produced the best results after establishing thresholds for AI-detected anemia, reaching 78% at Hb=11.23.

The extraction and processing of the Region of Interest (ROI) from conjunctiva images is the main focus of several studies Ref. [56,57,64,66]. The High Hue Ratio (HHR) was used for picture extraction in the study by Ref. [53]. The SVM with a polynomial kernel produced the best results, according to Ref. [24]. The Viola-Jones algorithm was used in Ref. [56] to automatically segment the ROI. In order to distinguish between cases of anemia, moderate anemia, and non-anemia, Ref. [23] also used color features.

Analyzed the RBC count and its fluctuations in both normal and anemic blood smears, and presented a technique for automatically counting red blood cells (RBCs) in blood samples Ref. [22]. Ref. [21] assessed the sensitivity of current models, the impact of different devices and acquisition settings, and the efficacy of sclera segmentation models using photos taken with mobile devices.

To lessen the impact of ambient lighting on the conjunctiva area of interest (ROI), digital image calibration is used. There are two algorithms used: one that uses Mahalanobis distance Ref. [53] and another that uses features based on color and texture. Using Kalman filters to reduce residual uncertainty, the authors investigated support vector machines (SVM) and artificial neural networks (ANN) as efficient algorithms Ref. [23,52,53].

Using conjunctiva images taken with a new device and the k-Nearest Neighbor (k-NN) method, Ref. [64] detected anemia with an accuracy of 90% on photos from 28 anemic and 85 non-anemic people.

Ref.[20] used hemoglobin levels, age, sex, and other clinical data from Düzce University Hospital in Turkey to develop a system for multi-class anemia detection under typical clinical conditions using machine learning techniques, while other studies concentrate on detecting anemia through the palpebral conjunctiva, nail beds, gums, tongue, and hands.

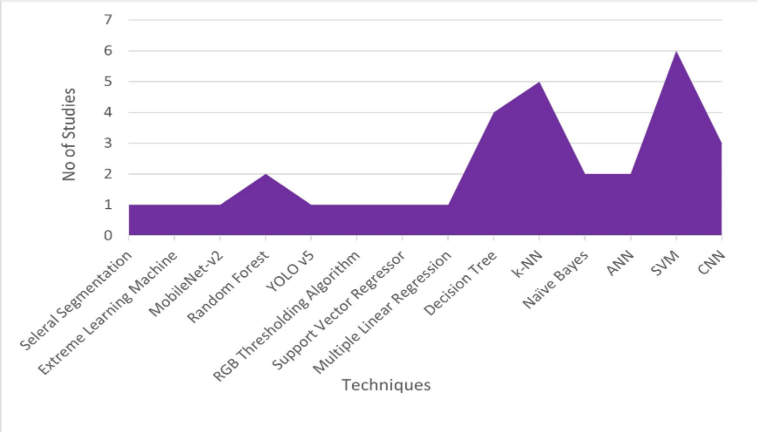

Figure 4: Representation of the Primary Studies

Ref. [19] used regression analysis with clinical information from blood samples and conjunctiva images in an attempt to calculate hemoglobin (Hb) levels. Machine learning techniques can be reliable for forecasting outcomes on Hb values, according to their evaluation of the Multivariate Linear Regression, Decision Tree, and Linear Support Vector Regression algorithms.

Physical examinations of anemic individuals include examinations of the nails, tongue, conjunctiva, and cardiorespiratory system. According to the authors, clinical evaluations, a thorough physical examination, and a careful review of medical history can all lead to an accurate analysis. Laboratory testing is common for anemic people and might be costly.

Ref. [17] presented sHEMO, a smartphone-based, autonomous, non-invasive method for detecting anemia in real time. By automatically recognizing the anterior conjunctival pallor as the area of focus and then extracting smaller sections to forecast the RGB spectrum based on spectroscopic data, sHEMO uses the smartphone's camera to take and analyze image spectroscopy.

Ref. [16] used deep learning models (UNet, UNet++, FCN, PSPNet, and LinkNet) to identify anemia in pediatric patients' palpebral pictures. The best results were obtained by LinkNet, which achieved an accuracy of 94%. Ref. [15] used deep learning techniques with retinal fundus images to propose a non-invasive way to monitor anemia during extended spaceflights.

Gaps Identified in the Reviewed Literature

Many research used small datasets Ref. [12,13,18,19,30,66]. For example, Ref. [29] examined 56 photos of the tongue and 27 photographs of the conjunctiva. Using picture augmentation is helpful for improving model performance and reducing overfitting Ref. [57,65], even though using limited datasets could be required Ref. [65].

In contrast to Ref. [30], which solely used Naïve Bayes, Ref. [57] only used one model (ANN). The scope of the analysis is limited as a result. Various research Ref. [12,30,57] do not reveal the provenance of datasets and do not comment on the data collection methods. Existing methods rely on pre-existing photos. Models' effectiveness would be better demonstrated by testing them on real-time datasets Ref. [7]. The temporal complexity of the proposed models is often overlooked, but it is essential for evaluating effectiveness Ref. [6].

Limitations of this Study

The specific kind of data taken into consideration and the reliance on previous research serve as limitations for this study. In order to identify anemia, it looks at medical photos (such as the color of the fingers, the appearance of the tongue, the color of the conjunctiva, and the color of the palm). Studies that used different kinds of data (such as complete blood counts and hemoglobin levels) were excluded. Furthermore, studies and pre-print documents that have not yet been published in peer-reviewed journals were excluded.

SUMMARY AND DISCUSSIONS

Machine learning is a vital tool for diagnosing conditions like anemia, allowing physicians to make prompt and precise judgments, especially for young patients. Reference Ref. [14] reported a precision rate of 58.9% and an accuracy of 62%, whereas Ref. [33] used CNN to reach a noteworthy 93.9% accuracy. Reference Ref. [65] achieved a remarkable accuracy of 99%, surpassing the 90.22% recorded by Ref. [30]. In another study, Ref. [13] used SVM and achieved 93.24% accuracy, which is somewhat less than the 96% accuracy that Ref. [65] reported using SVM. In the meantime, Ref. [45] used ANN to record an accuracy of 97.23%, Ref. [44] used Decision Tree to report an accuracy of 89%, and Ref. [66] used RGB Thresholding to obtain 79% accuracy. With results of 99% as opposed to 91%, the ELM approach by Ref. [10] outperformed Ref. [8]. As seen in Figure 4, SVM has become the most often used algorithm, with CNN, k-NN, and Decision Tree following closely behind.

RESULTS

According to this study, k-NN, CNN, and Decision Tree are the next most popular algorithms for identifying non-invasive anemia, after SVM. Even while machine learning is clearly beneficial, it is still unclear which model is appropriate for each age group. It is impossible to exaggerate the significance of big datasets Ref. [65]. Data augmentation is essential in the field of medical image processing Ref. [5]. Deep learning can help reduce the abuse of antibiotics, and advanced techniques may improve the detection of anemia. Further research is required for their wider applicability in haematology and general healthcare, particularly for children with recurrent anemia, even if retrospective studies about AI models indicate promise.

Lack of enough datasets is a major problem because bigger datasets improve model performance and reduce overfitting Ref. [5]. Although some authors Ref. [4,57,65] suggest that augmentation can help with smaller datasets, others Ref. [3] argue that it might not accurately reflect effectiveness in the actual world. By using techniques like data warping and oversampling, augmentation expands the size of datasets Ref. [5]. While biases from the original dataset might still exist, different approaches should be explored Ref. [1,2].

DISCUSSION

This study shows how machine learning may be able to replace traditional laboratory tests. In order to save expenses and minimize dangers (such as needle pricks), machine and deep learning approaches can evaluate medical pictures and signals for clinical evaluations. The machine learning techniques used in image analysis for the diagnosis of anemia (conjunctiva, palm, fingernail,tongue). The importance of large datasets and the need for image improvement are emphasized in this review. The most widely used algorithms are SVM, CNN, ANN, k-NN, and Decision Tree. Evaluation metrics are also covered in the review. Non-invasive machine learning offers a significant, effective, economical, and reliable way to diagnose anemia quickly. Since larger datasets improve model dependability, picture augmentation is recommended for growing datasets in medical disease detection, including anemia.

REFRENCES

[1] Appiahene P, Arthur EJ, Korankye S, Afrifa S, Asare JW, Donkoh ET. Detection of anemia using conjunctiva images: a smartphone application approach. Med Nov Technol Devices 2023;18:100237.https://doi.org/10.1016/j.medntd.2023.100237.

[2] Asare JW, Appiahene P, Donkoh ET, Dimauro G. Iron deficiency anemia detection using machine learning models: a comparative study of fingernails, palm andconjunctiva of the eye images. Eng Rep 2023. https://doi.org/10.1002/eng2.12667.

[3] Engati Simply Intelligent. Data augmentation. Published,https://www.engati.com/glossary/data-augmentation. [Accessed 13 May 2023].

[4] Alomar K, Aysel HI, Cai X. Data augmentation in classification and segmentation: a survey and new strategies. J Imaging 2023;9(2):46. https://doi.org/10.3390/jimaging9020046.

[5] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019;6(1):60.https://doi.org/10.1186/s40537-019-0197-0.

[6] Saurabh MD, Shanti RDSKG. A non-invasive medical tool for anemia disease diagnosis. Eur J of Mol & Clin Med 2021;7(8):5393–408. https://ejmcm.co m/article_8345.html.

[7] Mitra S, Rathore S, Gupta SK. Non-invasive medical tool to estimate anemia. J for all Subj 2019;8(9):1–8.

[8] Muthalagu R, Bai VT, Gracias D, John S. Developmental screening tool: accuracy and feasibility of non-invasive anaemia estimation. Technol Health Care 2018;26 (4):723–7. https://doi.org/10.3233/THC-181291.

[9] Dimauro G, Camporeale MG, Dipalma A, Guarini A, Maglietta R. Anaemia detection based on sclera and blood vessel colour estimation. Biomed Signal Process Control 2023;81:104489. https://doi.org/10.1016/j.bspc.2022.104489.

[10] Saputra DCE, Sunat K, Ratnaningsih T. A New Artificial intelligence approach using extreme learning machine as the potentially effective model to predict and analyze the diagnosis of anemia. Healthc 2023;11(5):697. https://doi.org/10.3390/ healthcare11050697.

[11] Rivero-Palacio M, Alfonso-Morales W, Caicedo-Bravo E. Anemia detection using a full embedded mobile application with YOLO algorithm. In: Int conf of the Inst of electr and electron eng (IEEE) Colomb conf on appl of Comput Intell (ColCACI);2022. p. 3–17. https://doi.org/10.1007/978-3-030-91308-3_1.

[12] Roychowdhury S, Sun D, Bihis M, Ren J, Hage P, Rahman HH. Computer aided detection of anemia-like pallor. Int Conf of the Inst of Electr and Electron Eng (IEEE) EMBS Int Conf on Biomed & Health Inform (BHI). Published online 2017: 461–4.

[13] Bauskar S, Jain P, Gyanchandani M. A noninvasive computerized technique to detect anemia using images of eye conjunctiva. Pattern Recogn Image Anal 2019; 29(3):438–46. https://doi.org/10.1134/s1054661819030027.

[14] Agrawal A. Detecting anemia from retinal images using deep learning. Indian Stat Inst; 2021.

[15] Waisberg E, Ong J, Zaman N, Kamran SA, Lee AG, Tavakkoli A. A non-invasive approach to monitor anemia during long-duration spaceflight with retinal fundusimages and deep learning. Life Sci Space Res 2022;33:69–71. https://doi.org/ 10.1016/j.lssr.2022.04.004.

[16] Dhalla S, Maqbool J, Mann TS, et al. Semantic segmentation of palpebral conjunctiva using predefined deep neural architectures for anemia detection. Procedia Comput Sci 2023;218:328–37. https://doi.org/10.1016/j.procs.2023.01.015.

[17] Ghosal S, Das D, Udutalapally V, Talukder AK, Misra S. sHEMO: smartphone spectroscopy for blood hemoglobin level monitoring in smart anemia-care. Int Conf of the Inst of Electr and Electron Eng (IEEE) Sens J. 2021;21(6):8520–9. https:// doi.org/10.1109/JSEN.2020.3044386.

[18] Noor N Bin, Anwar MdS, Dey M. An efficient technique of hemoglobin level screening using machine learning algorithms. In: 2019 4th IntConf on electr Inf and Commun technol (EICT); 2019. https://doi.org/10.1109/ eict48899.2019.9068812.

[19] Noor N Bin, Anwar MdS, Dey M. Comparative study between Decision Tree, SVM and KNN to predict anaemic condition. In: 2019 Int conf of the Inst of electr and electron eng (IEEE) Int conf on biomed eng, Comput and Inf technol for health (BECITHCON). IEEE; 2019. p. 24–8. https://doi.org/10.1109/ BECITHCON48839.2019.9063188.

[20] Karagül YT, Yurtay N, ¨Oneç B. Classifying anemia types using artificial learning methods. Eng Sci and Technol, an IntJ. 2021;24(1):50–70. https://doi.org/ 10.1016/j.jestch.2020.12.003.

[21] Vitek M, Das A, Pourcenoux Y, et al. SSBC 2020: sclera segmentation benchmarking competition in the mobile environment. In: 2020 Int conf of the Inst of electr and electron eng (IEEE) Int Jt conf on biom (IJCB). IEEE; 2020. p. 1–10. https://doi.org/10.1109/IJCB48548.2020.9304881.

[22] Nithya R, Nirmala K. Detection of anaemia using image processing techniques from microscopy blood smear images. J Phys Conf Ser 2022;2318(1):012043. https:// doi.org/10.1088/1742-6596/2318/1/012043.

[23] 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.

[24] Dimauro G, Camporeale MG, Dipalma A, Guarini A, Maglietta R. Anaemia detection based on the sclera and blood vessel colour estimation. Biomed Signal Process Control 2023;81:104489. https://doi.org/10.1016/j.bspc.2022.104489.

[25] Espinoza BS, Nú˜nez-Fern´andez D, Porras FPB, Zimic M. Portable system for the prediction of anemia based on the ocular conjunctiva using Artificial Intelligence. Electr Eng and Syst Sci 2019;4. https://doi.org/10.48550/arXiv.1910.12399.

[26] Haggenmüller V, Bogler L, Weber AC, et al. Smartphone-based point-of-care anemia screening in rural Bihar in India. Commun Med 2023;3(1):38. https://doi. org/10.1038/s43856-023-00267-z.

[27] Dewantoro P, Gandana CE, Zakaria RORH, Irawan YS. Development of smartphone-based non-invasive hemoglobin measurement. In: 2018 Int symp on electr and smart devices (ISESD); 2018. https://doi.org/10.1109/ isesd.2018.8605489. Published online.

[28] Yeruva S, Varalakshmi MS, Gowtham BP, Chandana YH, Prasad PESNK. Identification of sickle cell anemia using deep neural networks. Emerg Sci J 2021;5 (2):200–10. https://doi.org/10.28991/esj-2021-01270.

[29] Delgado-Rivera G, Roman-Gonzalez A, Alva-Mantari A, et al. Method for the automatic segmentation of the palpebral conjunctiva using image processing. In: 2018 Inst of electr and electron eng (IEEE) Int conf on autom/XXIII Congr of the Chil assoc of autom Control. ICA-ACCA); 2018. p. 1–4. https://doi.org/10.1109/ICA-ACCA.2018.8609744.

[30] Peksi NJ, Yuwono B, Florestiyanto MY. Classification of anemia with digital images of nails and palms using the naive bayes method. Telemat 2021;18(1):118. https://doi.org/10.31315/telematika.v18i1.4587.

[31] Joker PT. Binary image classifier CNN using TensorFlow. 2020. Published, https://medium.com/techiepedia/binaryimage-classifier-cnn-using-tensorflowa3f 5d6746697.

[32] Putra AT, Usman K, Saidah S. Webinar student presence system based on regional convolutional neural network using face recognition. J Tek Inform (Jutif). 2021;2 (2):109–18.

[33] Magdalena R, Saidah S, Ubaidah IDS, Fuadah YN, Herman N, Ibrahim N. Convolutional neural network for anemia detection based on conjunctiva palpebral images. J Tek Inform (Jutif). 2022;3(2):349–54. https://doi.org/10.20884/1. jutif.2022.3.2.197.

[34] Dada GE, Oyewola OD, Joseph BS. Deep convolutional neural network model for detection of sickle cell anemia in peripheral blood images. Commun Phys Sci 2022; 8(1):9–22.

[35] Djuric N, Grbovic M, Vucetic S. Distributed confidence-weighted classification on big data platforms. In: Hand book of statistics; 2015. p. 145–68. https://doi.org/10.1016/B978-0-444-63492-4.00007-1.

[36] Meena K, Tayal DK, Gupta V, Fatima A. Using classification techniques for statistical analysis of Anemia. Artif Intell Med 2019;94:138–52. https://doi.org/10.1016/j.artmed.2019.02.005.

[37] Handbook of statistical analysis and data mining applications. Elsevier; 2018. https://doi.org/10.1016/C2012-0-06451-4.

[38] Belle A, Kon MA, Najarian K. Biomedical informatics for computer-aided decision support systems: a survey. Sci World J 2013;2013:1–8. https://doi.org/10.1155/ 2013/769639.

[39] Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 2020;22(6):e15154. https://doi.org/10.2196/15154.

[40] Kumar RD, Guruprasad S, Kansara K, et al. A novel noninvasive hemoglobin sensing device for anemia screening. Inst of Electr and Electron Eng (IEEE) Sens J.2021;21(13):15318–29. https://doi.org/10.1109/JSEN.2021.3070971.

[41] Jayakody JADCA, Edirisinghe EAGA. HemoSmart: a non-invasive, machine learning based device and mobile app for anemia detection. In: 2020 Inst of Electr and electron eng (IEEE) reg10 conf (TENCON). IEEE; 2020. p. 1401–6. https://doi. org/10.1109/TENCON50793.2020.9293903.

[42] Chakraborty S, Kansara K, Dinesh Kumar R, Swaminathan D, Aatre K, Acharya S. Non-invasive estimation of clinical severity of anemia using hierarchical ensemble classifiers. J Med Biol Eng 2022;42(6):828–38. https://doi.org/10.1007/s40846-022-00750-3.

[43] Secinaro S, Calandra D, Secinaro A, Muthurangu V, Biancone P. The role of artificial intelligence in healthcare: a structured literature review. BMC Med Inf Decis Making 2021;21(1):125. https://doi.org/10.1186/s12911-021-01488-9.

[44] Aggarwal AK, Tripathy JP, Sharma D, Prabhu A. Validity of palmar pallor for diagnosis of anemia among children aged 6–59 months in North India. Anemia 2014;2014:1–4. https://doi.org/10.1155/2014/543860.

[45] Anggraeni MD, Fatoni A. Non-invasive Self-Care Anemia detection during pregnancy using a smartphone camera. IOP Conf Ser Mater Sci Eng 2017;172: 012030. https://doi.org/10.1088/1757-899X/172/1/012030.

[46] Acar E, Türk ¨O, ¨Of Ertu˘grul, Aldemir E. Employing deep learning architectures for image-based automatic cataract diagnosis. Turk J Electr Eng Comput Sci 2021;29 (SI-1):2649–62. https://doi.org/10.3906/elk-2103-77.

[47] Go´zdzikiewicz N, Zwoli´nska D, Polak-Jonkisz D. The use of artificial intelligence algorithms in the diagnosis of urinary tract infections—a literature review. J Clin Med 2022;11(10):2734. https://doi.org/10.3390/jcm11102734.

[48] Dalvi PT, Vernekar N. Anemia detection using ensemble learning techniques and statistical models. In: 2016 Inst of electr and electron eng (IEEE) Int conf on recent trends in electr, Inf & Commun technol (RTEICT). IEEE; 2016. p. 1747–51.

[49] Mannino RG, Myers DR, Tyburski EA, et al. Smartphone app for non-invasive detection of anemia using only patient-sourced photos. Nat Commun 2018;9(1). https://doi.org/10.1038/s41467-018-07262-2.

[50] 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.

[51] Suner S, Rayner J, Ozturan IU, et al. Prediction of anemia and estimation of hemoglobin concentration using a smartphone camera. PLoS One 2021;16(7):e0253495. https://doi.org/10.1371/journal.pone.0253495.

[52] 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.

[53] Chen YM, Miaou SG, Bian H. Examining palpebral conjunctiva for anemia assessment with image processing methods. Comput Methods Progr Biomed 2016;137:125–35. https://doi.org/10.1016/j.cmpb.2016.08.025.

[54] N KT, Prasad K, Singh BMK. Analysis of red blood cells from peripheral blood smear images for anemia detection: a methodological review. Med Biol Eng Comput 2022;60(9):2445–62. https://doi.org/10.1007/s11517-022-02614-z.

[55] Dimauro G, Griseta ME, Camporeale MG, Clemente F, Guarini A, Maglietta R. An intelligent non-invasive system for automated diagnosis of anemia exploiting a novel dataset. Artif Intell Med 2023;136:102477. https://doi.org/10.1016/j. artmed.2022.102477.

[56] Irum A, Akram M, Ayub SM, Waseem S, Khan MJ. Anemia detection using image processing. In: The Int conf on digit Inf process, electr, and wirel Commun.; 2016. p. 31–6.

[57] Jain P, Bauskar S, Gyanchandani M. Neural network based non-invasive method to detect anemia from images of eye conjunctiva. Int J Imag Syst Technol 2020;30(1): 112–25. https://doi.org/10.1002/ima.22359.

[58] World Health Organization (WHO). Anemia treatment, prevalence and data status. Published October 12, https://www.who.int/health-topics/anaemia#tab=tab_3.

[59] Rice BD, Tomkins SE, Ncube FM. Sharp truth: health care workers remain at risk of bloodborne infection. Occup Med (Chic Ill) 2015;65(3):210–4. https://doi.org/ 10.1093/occmed/kqu206.

[60] Mitani A, Huang A, Venugopalan S, et al. Detection of anaemia from retinal fundus images via deep learning. Nat Biomed Eng 2020;4(1):18–27. https://doi.org/ 10.1038/s41551-019-0487-z.

[61] Al-alimi AA, Bashanfer S, Morish MA. Prevalence of iron deficiency anemia among University students in Hodeida Province, Yemen. Anemia 2018;2018:1–7. https://doi.org/10.1155/2018/4157876.

[62] Diagnostic for anemia. Published, https://www.ekfdiagnostics.com/who-has-the -highest-risk-of-developing-anemia.html.

[63] Mohammed MS, Ahmad AA, Sari M. Analysis of anemia using data mining techniques with risk factors specification. In: 2020 Int conf for emerg technol, INCET 2020. Inst of Electr and Electron Eng Inc.; 2020. https://doi.org/10.1109/ INCET49848.2020.9154063.

[64] Dimauro G, Caivano D, Girardi F. A new method and a non-invasive device to estimate anemia based on digital images of the conjunctiva. Inst of Electr and Electron Eng (IEEE) Access 2018;6:46968–75. https://doi.org/10.1109/ACCESS.2018.2867110.

[65] 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.

[66] Tamir A, Jahan CS, Saif MS, et al. Detection of anemia from image of the anterior conjunctiva of the eye by image processing and thresholding. In: 2017 Inst of Electr and Electron Eng (IEEE) reg10 humanit technol conf (R10-HTC); 2017. p. 697–701. https://doi.org/10.1109/R10-HTC.2017.8289053.

[67] Tartan EO, Berkol A, Ekici Y. Anemia diagnosis by using artificial neural networks. Int J of Multidiscip Stud and Inno Technol 2020;4(1):14–7.

[68] Pasricha SR, Tye-Din J, Muckenthaler MU, Swinkels DW. Iron deficiency. Lancet 2021;397(10270):233–48. https://doi.org/10.1016/s0140-6736(20)32594-0.

[69] Kavsaoglu AR, Polat K, Hariharan M. Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features. Appl Soft Comput 2015;37:983–91. https://doi.org/10.1016/j.asoc.2015.04.008.

[70] Mehfooza M, I HB., (2021). An automated prescriptive domain data preprocessing algorithm to support multilabel-multicriteria classification for Indian coastal dataset, crop dataset, and breast cancer dataset. Int J Commun Syst. 2021;e4796. https://doi.org/10.1002/dac.4796.

[71] Mehfooza, M. and Pattabiraman, V. (2021) 'A new efficient learning approach E-PDLA in assessing the knowledge of breast cancer dataset', Int. J. Services and Operations Management, Vol. 38, No. 2, pp.153–160.

[72] M. Mehfooza & V. Pattabiraman (2018): SP-DDPT: a simple prescriptive based domain data preprocessing technique to support multilabel-multicriteria learning with expert information, International Journal of Computers and Applications, DOI: 10.1080/1206212X.2018.1547475.