Medicine in Novel Technology and Devices 20 (2023) 100269

Contents lists available at ScienceDirect

Medicine in Novel Technology and Devices

journal homepage: www.journals.elsevier.com/medicine-in-novel-technology-and-devices/

Application of ensemble models approach in anemia detection using images of the palpable palm

Peter Appiahene [a,**], Samuel Segun Dzifa Dogbe [a], Emmanuel Edem Yaw Kobina [a], Philip Sackey Dartey [a], Stephen Afrifa [a,b], Emmanuel Timmy Donkoh [c], Justice Williams Asare [a,*]

[a] Department of Computer Science and Informatics, University of Energy and Natural Resources, Ghana
[b] Department of Information and Communication Engineering, Tianjin University, China
[c] Department of Basic and Applied Biology, University of Energy and Natural Resources, Ghana

ARTICLE INFO

Keywords:
Machine learning
Anemia detection
Ensemble learning
Algorithms

ABSTRACT

Anemia is a public health issue with serious ramifications for human health globally. Anemia particularly affects pregnant women and children from 6 to 59 months old even though every individual is at risk. Anemia occurs when the Hb level is below its normal threshold or when the red blood cells are weakened or destroyed. To discover medical remedies on time, early detection or diagnosis of anemia assist patients to understand their condition.

The invasive approach for anemia detection is costive and time-consuming as compared to the non-invasive approach which is reliable and suitable for developing communities where medical resources and personnel are inadequate. This study uses palpable palm images (dataset) collected from 710 participants in selected hospitals in Ghana. The images were extracted, segmented and converted into RGB percentile to train, validate and tested the machine learning models. A hybrid model was developed with the application of ensemble learning models using the R programming language on the R Studio platform. Stacking, voting, boosting and bagging ensemble model techniques were used to build the hybrid models, the stacking ensemble model achieved an accuracy of 99.73 %. The study justifies that ensemble models are efficient for medical disease diagnosis or detection such as anemia.

1. Introduction

Technology is influencing the future of healthcare where the use of non-invasive medical techniques aided by technology is preferred in medical diagnostics. Computer-aided disease diagnosis is less expensive, saves time, is more accurate, and reduces the need for additional labor in medical decision-making [1–3]. Anemia is a worldwide public health issue with serious ramifications for human health [4]. Anemia is one of the most common medical disorders affecting people all over the world, particularly pregnant women and children. Tackling this issue with improved technology would significantly lower the occurrence.

Anemia is a condition in which the hemoglobin level is low. Hemoglobin is an iron-rich protein that gives blood its red colour and is responsible for assisting red blood cells in transporting oxygen from the lungs throughout the body, including vital organs such as the heart, kidneys, and other organs [1,5,6]. The World Health Organization discovered that 42 % of children below the age of 59 months and 40 % of pregnant females universally are anemic [7–9] and this affects 33 % of the global population due to iron deficiency [10].

Anemia is common in people with heart failure, with a frequency ranging from 9 % to 69.6 % depending on the patient's features [4]. Early detection and diagnosis of anemia would assist persons to understand their condition and discover medical remedies on time [7]. It would also be beneficial to the economy through efficient productivity. A high prevalence of anemia in the community has implications for both mortalities, and the physical and mental health of people affected [11].

Other non-invasive methods of identifying anemia include examination of the conjunctiva eyes, the colour of the fingernail, palpable palms, and tongue. The use of images of the palpable palm in the diagnosis of anemia is uncommon, although it is a non-invasive procedure that yields reliable results. The traditional method of detecting anemia with the palm was to examine how pale or yellowish the individual's palm is, and

* Corresponding author.
** Corresponding author.
E-mail addresses: peter.appiahene@uenr.edu.gh (P. Appiahene), justice.asare.stu@uenr.edu.gh (J.W. Asare).

https://doi.org/10.1016/j.medntd.2023.100269
Received 16 May 2023; Received in revised form 1 July 2023; Accepted 11 October 2023

then proceed to the laboratory for a clinical confirmation test to distinguish either one is anemic or not with the use of a blood sample.

The clinical approach for the detection of anemia involves collecting blood samples from patients, either by pricking their fingertips or with a syringe and then performing a laboratory test on these blood samples. The extraction of blood samples is an intravenous procedure that requires sophisticated surgical equipment. Nonetheless, diagnostic procedures are costive, and take enough time and effort, to obtain the findings. This also exposes biomedical scientists and medical laboratory technicians to the danger of blood-borne diseases, and patients may experience pain when blood samples are taken, while frequent blood collection causes discomfort for patients [12].

The novelty of this study is the combination of several individual or single machine learning algorithms to develop an ensemble model to detect anemia by comparing both ensemble learners and single or individual machine learning models. In addition, most studies used the conjunctiva of the eye images for the detection of anemia, however, this study utilizes medical images of the palpable palm to detect anemia in children for the reason that the human palm is one of the sensitive spots for detecting anemia in children. This is due to the reason that children may be exposed to falling objects or infected with bacteria when the conjunctiva of the eye is exposed for examination during the period of diagnosis.

The main contribution to knowledge by this study is the use of palpable palm images in the detection of anemia since most studies used the conjunctiva of the eye images. Moreover, the size of the dataset (images of the palpable palm from 710 participants) used for the study is made publicly available on the Mendeley dataset repository which is the highest number of palpable palm medical images used for the detection of anemia.

In this study, we used an ensemble model, that is, stacking, bagging, voting, and boosting to detect anemia in children aged 6–59 months using the palpable palm images. The ensemble model techniques were used to build the hybrid models which are used to diagnose and detect anemia at its preliminary stage since early detection of anemia would help prevent serious complications such as heart problems, pregnancy complications and developmental delays in children [13].

2. Related works

Over the years, machine learning has been used for the automated detection of anemia using deep convolutional models, Ensemble models or individual machine learning classifiers. Machine learning is applied to pick characteristics and trends, analyze these trends and make predictions or diagnoses based on the trends. These predictions or diagnoses have been seen to possess high accuracy levels and have aided in replacing non-invasive anemia detection progressively.

Dalvi & Vernekar [13] presented a comparative analysis of the performance of machine learning classifiers with the application of ANN, Decision Tree, k-NN, and Naïve Bayes which was used to classify the dataset. Among the individual classifiers, ANN performed better with an accuracy of 97.63 %, while k-NN had the least accuracy. In addition, the application of the Stacking ensemble learner was assigned to the k-NN and Decision tree for the classification of anemia which achieved 92.12 % accuracy better than the ANN.

Tamir et al. [14] analyzed the anterior conjunctival polar of the eye, quantifying the conjunctival colour from digital images. The digital images were taken with a smartphone camera under proper lighting and had a good resolution. The images were processed to obtain their red and green (RG) component spectra of the conjunctiva and compared against a threshold to determine if a patient is anemic or non-anemic.

A system application was developed by Dimauro et al. [4] to examine anemic conditions. In their study, a non-invasive dimension was utilized for the demonstration of the results while the conjunctiva of the eye pictures was taken with a smartphone connected to a microlens. 113 datasets (images) of the conjunctiva of the eyes were captured, which were made up of anemic and non-anemic patients. A k-NN algorithm was utilized to evaluate the results with the use of 113 datasets which identified a correspondence of 0.745 higher than the other two models.

With the utilization of the k-Means algorithm, Sevani et al. [15] detected anemia which was based on conjunctiva pallor images. The study achieved an accuracy of 90 % with a proposed clustering technique employed. Through the application of six machine learning algorithms for the detection of iron deficiency anemia, the palpable palm images were used to detect anemia by comparing the performance of CNN, Naïve Bayes, Decision Tree, k-NN and SVM algorithms by Appiahene [7]. The CNN achieved the highest accuracy of 99.92 % while the SVM had the least accuracy of 96.34 %. The study concluded that the palpable palm is one of the vital features for detecting anemia in children mostly with the use of a non-invasive approach.

The palpable palm, fingernails and conjunctiva of the eye images were used to detect iron deficiency anemia in a comparative study via the application of machine learning models by Asare et al. [16]. The study compares these three features, that is, the palpable palm, the colour of the fingernails and the conjunctiva of the eye images to justify which of them has higher efficiency in detecting anemia in children with the application of Naïve Bayes, CNN, SVM, k-NN, and decision tree algorithms. The SVM had the least accuracy of 95.4 %, while the CNN achieved a higher accuracy of 99.12 %. Based on the performance of the models, the study concludes that the non-invasive approach is an effective mechanism for anemia detection.

A “finger-mounted photo-plethysmogram (PPG) device” was created to capture PPG signals at (590, 660, 810, and 940 nm) wavelengths proposed by Chakraborty et al. [17] for the detection of anemia. Utilizing a machine-learning method, a three-way hierarchical ensemble classification system was created utilizing a set of 13 attenuation and ratio-of-ratio features that were produced from the peak and trough information retrieved from the PPG signals. In detecting anemic and non-anemic patients in women, the study achieved 92.00 % and 84.00 % sensitivity and specificity respectively. The study also demonstrated a sensitivity of 76.00 % and a specificity of 74.00 % for detecting detection based on its severity.

The “Least Absolute Shrinkage and Selection Operator (LASSO), Ridge, Elastic Net, Ada Boost and Support Vector Regression (SVR)” are two-layer stacks of regressors that were built and tested as a proposed machine learning model for anemia detection by Acharya et al. [18]. The study provided a significant Pearson's correlation coefficient (PCC) of 0.81 ($p<$; 0.01) between the Hb value determined using the suggested approach and the gold standard values of Hb, with a Root Mean Square Error (RMSE) of 1.353 0.042 g/dL. In terms of performance (low RMSE and improved CC; $p < 0.05$), the stacked regressor model performed noticeably better than the individual regressor models.

In a study by Sen et al. [19] microscopic images were processed and red blood cells were classified into three shapes in a comparative analysis. Where Microscopic images were preprocessed and results were compared among machine learning algorithms such as Random Forest, Logistic Regression Naïve Bayes and Support Vector Machine. Kasiviswanathan et al. [20] used a deep learning approach for the segmentation of the conjunctiva area for non-invasive anemia screening applications. A proposed U-Net Based Conjunctiva Segmentation Model (UNBCSM) makes use of fine-tuned U-Net architecture to effectively semantically separate conjunctiva from digital eye pictures taken by consumer-grade cameras in an uncontrolled setting. To boost the data amount and model performance, image augmentation and pre-processing were used. UNBCSM produced strong segmentation results, with an Intersection over Union (IoU) score of 96 % and 85.7 % for training and validation, respectively. Jain et al. [21] used ANN to detect anemic patients from images of eye conjunctiva. Image augmentation techniques were used to improve the number of available images. Computer vision algorithms were used for preprocessing and feature extraction. The model was able to achieve the best accuracy of 97.00 % with a sensitivity of 99.21 % and a specificity of 95.42 % on the created dataset.

Bauskar et al. [22] used conjunctiva of the eyes images and for the Region of Interest (ROI) of the images to be obtained the images were manually extracted. The red and green (RG) components of the mean intensity values of the extracted ROI images were obtained and mapped to the image pixels for corresponding. Afterwards, a patient was classified as anemic or non-anemic via the application of a tuned machine learning algorithm.

Florestiyanto & Peksi [23] used the Naïve Bayes classification method and categorized the nail and palm image data into anemia and non-anemia. This method achieved 90 % accuracy paleness classification of palm and nail images.

For optimal extraction of morphological traits and categorising the severity of anemic conditions, a "three-tier deep convolutional fused network (3-TierDCFNet)" was proposed by Shahzad et al. [24] for such a study. The study achieved a significant result of 98.95 % accuracy and 98.12 % of specificity after training, validating and testing the models.

A smart (mobile) phone application was developed by Jayakody [25] to detect anemia using images of the fingernails of different hemoglobin level patients to develop a neural network-based algorithm. The smartphone app provides the findings, which are then processed by a machine learning algorithm to produce the output.

Karagül et al. [26] proposed a system for the detection of anemia through the guidelines of conventional practice by the global public health standard. ANN, SVM, NB and Ensemble Decision Tree were used for classification. Along with accuracy values, each model was evaluated using metrics for classification error, AUC, precision, recall, and f1-score. Bagged Decision Trees, Boosted Trees, and ANN was utilized and achieved a maximum accuracy of 85.60 %, 83.00 % and 79.60 % respectively. Meena & Tayal [27] developed a decision support system using data mining methods for anemia in children. In their proposed model, they used the decision tree and association rules methods and obtained successful results.

Age, sex, symptoms, chronic illnesses, and blood parameters, along with other factors, were considered in the collaborative efforts with skilled medical specialists for the detection of anemia performed by Verma & Arjun [28]. A notable approach like an artificial neural network, support vector machine, decision tree, k-NN and Naïve Bayes algorithms were applied to the Hb values which were utilized for the study. Also, clinical symptoms which were centered on the pallor palm and fingernail images with intense colour for the detection of anemia through the utilization of a non-invasive approach were proposed by Peksi et al. [29]. The red, green and blue mean intensity colour components of 20 medical images were extracted and a Naïve Bayes technique was used for the study and resulted in an accuracy of 90 %.

Acar et al. [30] proposed the use of a colour fundus image-based automated approach to identify cataract disease. The study used a deep learning-based method, specifically the VGGNet and DenseNet architectures, to extract the ROI of the conjunctiva from the photos and discovered anomalies for the description of the region of interest in the human eyes. The proposed system achieves a diagnostic accuracy for cataract disease of 97.94 % by the VGGNet and 95.07 % by DenseNet, respectively.

A rule-based approach was presented by Belginova & Uvaliyeva [31] to detect iron deficiency anemia with the application of machine learning models. The authors outlined a decision-support system for specialized medical advisors that would incorporate patient data (identity, socioeconomic status, medical history, complaints or sensations, medical indicators, and statistical data on the condition) as well as information on the patient's complaints or sensations. The medical professional's recommendations regarding the disease might be more precisely made through the incorporation of the aforementioned data.

In a study conducted by Appiahene et al. [32], a smartphone application is developed to diagnose anemia. The study centres on pallor analysis and uses images of the conjunctiva of the eyes to detect anemia using machine learning techniques. A publicly available dataset of 710 conjunctiva of the eye images was obtained with a unique tool that eliminates any interference from ambient light. Convolutional Neural Networks, Logistic Regression, and Gaussian Blur algorithm were combined to develop a conjunctiva detection model and an anemia detection model which runs on a Fast API server connected to a frontend mobile app built with React Native.

In the study conducted by Noor et al. [33] 104 participants were enrolled for the study which consisted of 54 females and 50 males with their correlated clinical Hb value from the laboratory to indicate their anemic status and corresponding images of the conjunctiva of the eyes. For image quality to be derived, a smartphone camera was used to capture the images. The percentile of the RGB pixels of the images was extracted and processed with the application of MATLAB. 81 out of the 104 images were used to train the models while 23 of the remaining datasets were used for testing the models. The Decision tree algorithm achieved the highest accuracy of 82.61 % and was tested and plotted with the Hb values.

A conjunctival pallor dataset and benchmark for anemia detection in children were proposed by Appiahene et al. [34], The study focuses on conjunctiva image-based datasets supported with Hb levels (g/dL) annotations for accurate diagnosis of anemia. The authors deploy a joint deep neural network that concurrently classifies anemia and estimates hemoglobin levels (g/dL) based on the conjunctival pallor images. The results of their study show the efficacy of the joint deep neural network in both the tasks of anemia classification and Hb levels (g/dL) estimation.

Yang [35] proposed MATLAB-based image processing for medical engineering applications. The author highlighted the importance of image data in medical engineering as the main source of information exchange and proposed that the medical engineering application of digital image processing attracts a huge cost-effective mechanism, it also produces efficient results by minimizing noise effects and enhancing the image quality.

Sharma et al. [36] used Ensemble learning of various Transfer Learning (TL) Architecture to improve the performance of predictive models. Six TL architectures viz, InceptionV3, ResNet152V2, Xception, DenseNet201, InceptionResNetV2, and VGG19 were considered by the authors, and their various ensemble models were used to carry out the task of deficiency diagnosis. The ensemble-based architecture enhanced the highest classification accuracy of 100 % from 99.17 % in the Mendeley dataset, while for the Kaggle dataset, the accuracy was enhanced from 90 % to 90 %.

A systematic review was performed by Asare et al. [37] to examine current trends and concepts of machine learning in healthcare services and to identify viable approaches for anemia detection and diagnosis. The study compares the most successful non-invasive approaches currently in use regarding machine learning algorithms, evaluation metrics, image augmentation, and the origin and size of the dataset used for the various studies reviewed. The study concluded that non-invasive methods such as the use of machine learning algorithms to detect anemia are affordable, and provide results on time.

Regarding the related studies mentioned, it is generously clear that non-invasive approaches, such as the use of machine learning algorithms, are trustworthy in the detection of anemia due to their timely, cost-effective, and results-oriented [7] nature, as the palm is one of the priority components of the human body for anemia detection in children [7, 38] when using a non-invasive approach. Our study, however, focused on the use of the palm since it is one of the promising locations in anemia detection using the ensemble model approach. This is the reason that most of the studies employed the use of the conjunctiva of the eye images [39].

2.1. Machine learning models

- **Naïve Bayes (NB)**

It uses independent assumption and Bayes theorem, and it assumes that neither the presence nor absence of a particular attribute is related to the presence or absence of other attributes [18]. A small amount of training dataset is considered fit for the calculation of the mean and variance [7] of the variables associated [40] while splitting fields that continue into a discrete bin and target value fields. The Nave Bayes typically generalizes effectively due to the reason that there are no hyperparameters to adjust.

Mathematically, the Naïve Bayes is expressed as specified in equation (1);

$$P(c/x) = \frac{P(x/c)P(c)}{P(x)} \tag{1}$$

- **Artificial Neural Networks (ANN)**

It is computational that is similar to the structure of the human brain and gains knowledge by learning from the training dataset [21]. Knowledge learnt is applied to the classification of a dataset. It consists of the output layer, the input layer and the hidden layer [21]. The input layer receives the input, the hidden layer performs the computation and the output layer predicts the final output. The system is fed through the input layer.

The neurons from one layer are connected to another layer through channels. Also, channels are assigned values known as weights. The input values are multiplied by the corresponding weights and their sum is sent as input to the hidden layer as shown in equation (2). Its neuron is associated with a value known as a bias which is added to the input sum. The value is passed through an activation function and the results of the activation function determine which neuron would be activated. Activated functions transmit neurons to the next layer which is the output layer, forward propagation occurs when the data is propagated through the channel [41].

The AlexNet, Adam optimization and ReLu were used to train the ANN. The activation function has a regularisation of a = 0.001 and a maximum iteration of 10. The activation function's main task is to turn the nodes' signal inputs into signal outputs.

The neuron with the highest probability determines the output. The predicted output of the network is compared to the actual output. When the outputs are not the same, the system uses back-propagation to send a message to the system to reiterate changing the weights of each neuron. This process continues until accuracy from the system is achieved.

$$f\left(b + \sum_{i=1}^{n} x_i w_i\right) \tag{2}$$

where b = bias, x = input to the neuron, w = weights, n = the number of inputs from the incoming layer, and i = a counter from 1 to n.

- **Support Vector Machine (SVM)**

In n-dimensional feature space, it is relatively difficult to segregate the training data with a linear hyperplane, where “n” is the number of features in the training dataset. The fundamental algorithm of the SVM model converts the input features from the original n-dimensional space to a higher dimensional space, fragmenting datasets into multiple homogeneous fragments using linear hyperplanes [7]. The Sigmoid was used for the operation with 100 iterations as the limit, cost $(c) = 100$ and epsilon of regression (ε) *was* assigned to 1.100 while the numerical tolerance was set to 0.10000. The coefficients (a1, a2, a3, …, a and b) of the hyperplanes indicated in equation (3) are optimized throughout the training process. Optimization maximizes the boundary (δ) between the training data points and the hyperplanes.

$$a1 : x1 + a2 : x2 + \ldots\ldots + an : xn + C = 0 \tag{3}$$

Converting the training dataset from the original n-dimensional feature space to a higher dimensional space is achieved through the kernel-trick method. In the converted space, the data points are linearly divisible with the help of the hyperplanes [42]. The most commonly used Radial Basis Function (RBF) as described in equation (4) is implemented in this study, where k (xi, xj) is the kernel function amongst two points xi and xj and σ symbolize the kernel width. In this study, the value of k (xi, xj) signifies the resemblance between two almond grains. For two similar almond kernels, the value of k (xi, xj) □ 1.

$$K(x_i x_j) = e^{\left(\frac{|x_i x_j|}{2\sigma^2}\right)} \tag{4}$$

- **Random Forest (RF)**

Random forest is a scalable, simple-to-use machine learning algorithm that delivers a good result much of the time, even without hyperparameter tuning. It is also one of the most commonly used algorithms due to its simplicity and versatility which can be used as both regression and classification algorithms [43]. The random forest does both row and column sampling with a decision tree as a base. Random forest is an ensemble technique for multiple decision trees.

The maximum depth (max_depth) used was 10, 50, and 100, trees which indicate the longest path between the root and the leaf node, while the minimum sample (min_sample) split was 5, 10, and 15 indicating the number of samples required to split an internal node. The number of trees (n_estimation) for the forest is 100, 200, and 300, with the maximum feature (max_feature) as 20, 30, and 40 which is the number of features considered for splitting a node.

- **Decision Tree**

In the Decision Tree classifier, the training dataset is represented by using a tree structure, and the decision tree is used for calculating discrete-valued target functions [7]. Instances are classified by sorting them down the tree from some root to the leaf node, while each tree node represents an attribute and each branch represents a value for that attribute [33].

It splits a dataset until we are left with a real leaf root. The minimum number of instances in the leaf was set to 10 and a minimum of 100 trees, while smaller subsets of less than 5 were not divided and were used as the limit in the binary tree. It consists of a decision node that has the condition to splits the dataset and a leaf node that determines the class of data. Usually, the whole datasets are found at the root node where the dataset is split based on the provided conditions. Mostly with large datasets, a majority voting is performed and it's assigned a majority class.

The mathematical model for the Entropy of the target class is shown in equations (5) and (6).

$$\sum p(x) \log p(x) \tag{5}$$

where p(x) is

$$\text{Gini Index} = 1 - \sum_{i=1}^{n} pi^2 \tag{6}$$

2.2. Ensemble models

- **Stacking;** this technique picks combinations of machine learning algorithms to find the ones with the best compatibility [44]. The most compatible models will yield the highest accuracy.
- **Voting;** it picks combinations of the machine learning models and classifies the test data per class that obtains the maximum vote [13].
- **Bagging;** this method makes use of SVM, ANN, NB, RF and DT as individual learners and will use the training data set to draw up a sample that would be independently used to induce models when classifying data from the test data set, the induced models are

combined in a using a simple voting scheme [13,28]. The class of the new instance is generated using the most frequent among the models.

- **Boosting;** the training dataset is split into sub-datasets and the various sub dataset are used to train the individual machine learning models [17,45]. The ensemble process combines the individual models and boosts the performance of the weak models [46]. The combined results using majority voting are used in the prediction of the test set.

2.3. Proposed method

2.3.1. Conceptual framework of the study

This section discusses the various steps and processes that were used for the study. This phase is segmented into three main categories, which are; dataset acquisition, preprocessing and segmenting of the dataset and training, validation and testing of the models.

The dataset used for the study was collected from selected hospitals in Ghana. Professional biomedical scientists were given proficiency training on how to use the Kobo Collect application to collect medical images (data) with the use of an electronic tablet device and then submitted to a database for easy access and preprocessing.

The images were then segmented into their various RGB percentile components. The RGB function describes a colour giving the intensity of the 3 primary colours: red, green and blue. To use the function: RGB (red, green, blue, alpha), a quantity of red (between 0 and 1), of green and of blue, and finally transparency (alpha). To show the colours acquired for a given red value (Y axis), green value (X axis) and blue value, the transparency is set to 1. The image acquisition process, preprocessing and segmentation are detailed in section 3.2.

To detect whether a patient is anemic or non-anemic, a machine learning model, that is, ANN, SVM, Naïve Bayes, Decision Tree and RF were used for the classification of the anemia status. The classifier with the highest accuracy was recorded after training, validating and testing the machine learning models with the dataset. Afterwards, ensemble learning techniques such as stacking, bagging, voting and boosting were used to create a hybrid system to check for optimal accuracy. The performance of each ensemble learning technique is measured using mathematical models [13,46] like accuracy.

The experimental setup for this study used for detecting anemia utilized image processing capabilities of MATLAB 2018 and R programming language and the R studio platform for training, validating and testing models. The methodology as illustrated in Fig. 1 using a block diagram further explains in detail the various steps such as image acquisition, preprocessing and segmentation. It also includes the feature extraction of the ROI of the images, calculating the percentage of RGB pixels in the images as indicated in Table 1, class labelling and classification and cross-validation of the models. This means, the conceptual framework gives a pictorial summary of how the study is carried out, which is detailed in sections 3.2–3.4.1.

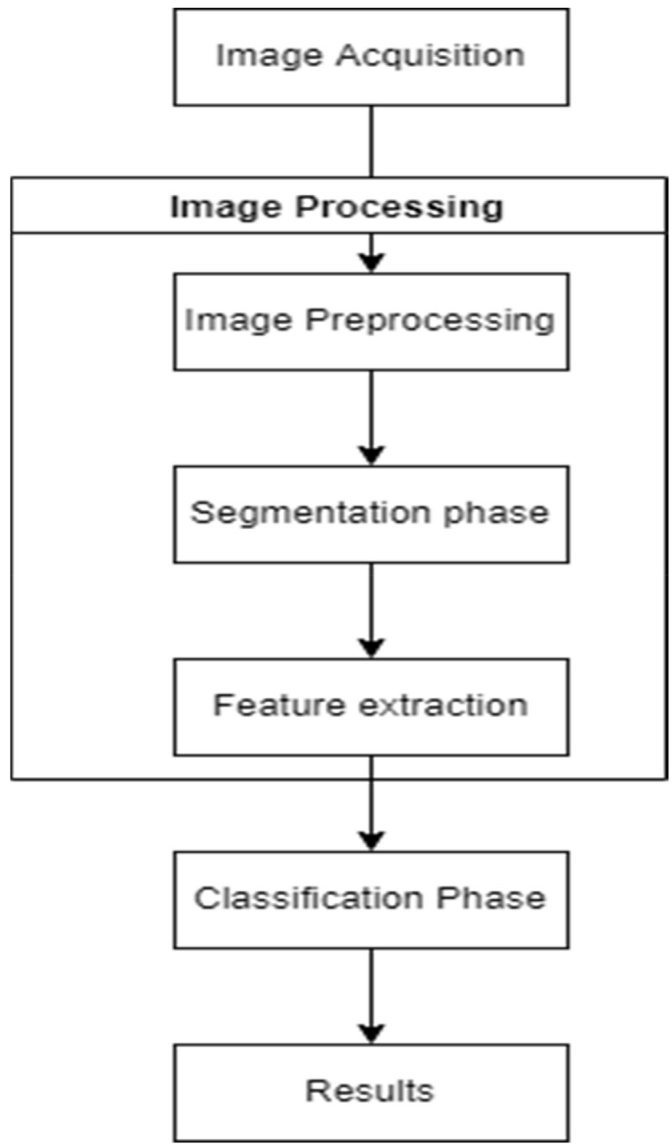

Fig. 1. Block diagram of proposed methodology.

2.4. Image acquisition, pre-processing and segmentation

A Kobo collect app was developed and deployed on an electronic mobile tablet. The participants' biodata which includes Hb values, age, and sex, and based on the Hb value result a comment (anemic or non-anemic) was added with the palpable palm images and uploaded to the cloud storage for easy access. The age range of the participants in the dataset for this study is 5–59 months. The datasets utilized the Hb value; 11 g/dL indicates anemia, whereas≥11 g/dL indicates non-anemia.

Certified biomedical scientists who underwent a weeklong competence training at the "Centre for Research in Applied Biology, University of Energy and Natural Resources, Sunyani, Ghana", on how to take pictures of infants between the ages of 6 and 59 months collected the dataset. Before the study began, it was reviewed and approved by the "Committee for Human Research and Ethics (Reference number: CHRE/CA/042/22)". Parents or guardians were given their consent for their

Table 1
Samples of the converted palm images into components of the RGB values.

Image ID	Red (R)	Green (G)	Blue (B)	Status
Image 001	39.03780	34.134139	26.82800	anemic
Image 002	37.01250	33.848009	29.13950	anemic
Image 003	43.87580	30.547063	25.57720	anemic
Image 004	40.80080	32.910173	26.28900	anemic
Image 005	45.47100	31.396006	23.13300	anemic
Image 006	41.97590	32.606409	25.41770	anemic
Image 007	48.48130	31.811414	19.70720	anemic
Image 008	46.19070	31.875195	21.93410	anemic
Image 009	35.57260	32.911054	31.51640	anemic
Image 010	50.50530	29.025078	20.46960	anemic
Image 011	38.65060	31.666920	29.68250	non-anemic
Image 012	41.66790	30.727582	27.60460	non-anemic
Image 013	41.66790	30.727582	27.60460	non-anemic
Image 014	38.56460	31.732560	29.70280	non-anemic
Image 015	38.30320	29.544891	32.15190	non-anemic
Image 016	41.66790	30.727582	27.60460	non-anemic
Image 017	41.71130	30.715082	27.57370	non-anemic
Image 018	40.43840	32.181344	27.38030	non-anemic
Image 019	42.57700	30.442059	26.98100	non-anemic
Image 020	43.23470	30.973125	25.79220	non-anemic

ward(s) before enrolling them for the study.

For the motive that participants were children, their hands (of the participants) were held and their palms were stretched through the fingers by the biomedical scientists before taking pictures of the palm. A mechanism was put in place to serve as an immense approach to eliminate the impact of ambient light on the images. When taking pictures of the images, the cameras' flashlights were turned off to prevent disproportionate gleam effects on the picture quality that could have an impact on the models' performance. Figs. 2 and 3 show a sample of the raw image of the palm indicating the ROI and extracted images of the palm of anemic and non-anemic images respectively.

The images were converted to grayscale images. A median filter was used to emit noise from the grayscale images. The final step was to enhance the image. The Region of Interest (ROI) of the palm images were then extracted from the captured image. It involved two steps which were the conversion of the image from RGB to YCbCr and thresholding. From the thresholding process, a binary image was generated that separated the object from the background. The images were converted back to RGB with the background removed.

2.5. Feature extraction

RGB value was used as a parameter of this study. From the segmented image, the system then calculated the percentage of RGB pixels in the image. The obtained results which are the percentile RGB pixels of the palms as Table 1 show samples of the converted images into RGB values.

Each image consists of a certain percentage of red, green and blue colour pigment hence, we calculated the percentages that make up each image using MATLAB. The RGB percentage value of the training data was entered into the database, while the RGB value of the data was used as input in the classification process.

2.6. Class labelling and classification

This is done if the data entered is in the form of training data. There are two types of classes in this study, namely Non-Anemic and Anemia with a number matching to each. After feature extractions, flipping, rotation and translation techniques were used to increase the dataset size through the application of image augmentation techniques. The dataset was grouped into 10 subsets of 426 instances. 7 subsets were used to train the models, 1 subset would be used to validate the model using K folds validation techniques and the remaining subset would be used to test the model. The process is repeated 10 times to generate 10 results. The final estimation is obtained by averaging the 10 results.

The classification of anemia detection was designed and implemented utilizing the application of a probability approach for anemic and non-anemic conditions indicated in equations 7–11.

$$\mathrm{P(Status = Anemic) = Anemic/Dataset} \tag{7}$$

$$\mathrm{P\ (Status = Non\ Anemic) = Non\ Anemic/Dataset} \tag{8}$$

The probabilities of individual attributes known as conditional probabilities are calculated.

$$V_{nb} = argmax\ P(vj)\pi i\ P\left(\frac{ai}{vj}\right) \tag{9}$$

where vj ϵ (Anemic, Non-Anemic) and aj are the attributes

$$V_{nb}(Anemic) = \frac{V_{nb}(Anemic)}{V_{nb}(Anemic) + V_{bn}(Non - Anemic)} \tag{10}$$

$$V_{nb}(Non - Anemic) = \frac{V_{nb}(Non - Anemic)}{V_{nb}(Anemic) + V_{bn}(Non - Anemic)} \tag{11}$$

Prior probabilities and conditional probabilities were used to classify the new instances.

2.6.1. Cross-validation

The codes used for cross-validation for the models trained are noted here. 5 folds of cross-validation were performed.

- Cross Validation for Individual Models

```
#defining training controls for model
Cv < - trainControl (
Method = “cv",
Number = 5,
SavePredictions = ‘final',
ClassProbs = TRUE).
```

- Cross Validation for Ensemble Models.

```
#defining training controls for model
Cv < - trainControl (
Method = “cv",
Number = 5,
SavePredictions = ‘final',
ClassProbs = TRUE,
Index = createResample (training_set$Status, 25))
```

3. Results and discussions

The results of the various individual machine learning and ensemble models are deduced, analyzed and discussed in this section. The segmentation phase of the palm images has been shown to play an essential role in the detection of anemia. After training and validation of the models, each model was evaluated with the use of accuracy, specificity, recall, precision and f1 score.

3.1. Individual models

As stated in the methodology, 70 % of the dataset was used to train the model, 10 % was used to validate the model where k folds validation was used and 20 % of the dataset was used to test the model. Features

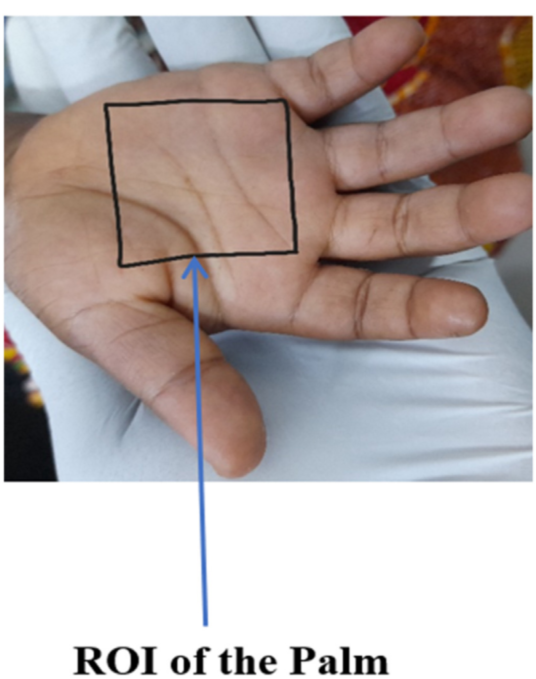

Fig. 2. A sample of the Region of Interest (ROI) of the palm [7].

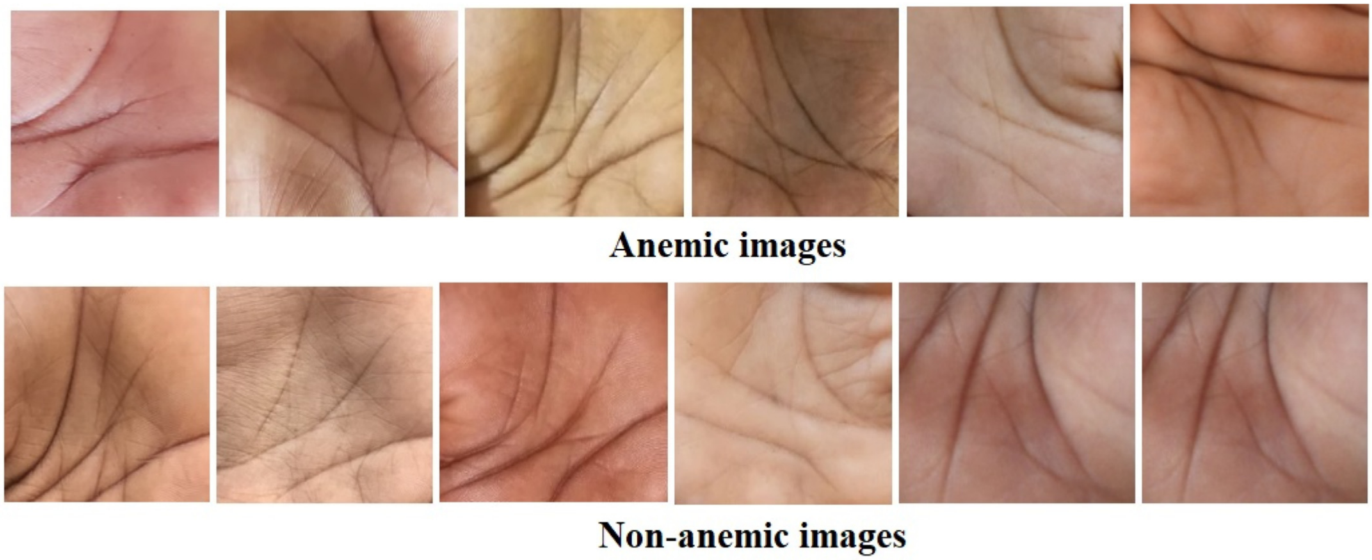

Fig. 3. Sample of extracted anemic and non-anemic images of the palm.

have been extracted from the segmented images and are fed into the individual machine learning models and the final results are presented in Table 2.

As shown in Table 2, the classification accuracies of the individual machine learning models for the detection of anemia using palm images varied in the range of 99.53 % and 64.45 %. Random forest recorded the highest accuracy which was 99.53 % while ANN recorded the lowest accuracy which was 64.45 %. Other individual models such as SVM, Naïve Bayes and Decision Tree recorded accuracies of 75.80 %,70.16 % and 99.20 % respectively. For specificity, random forest and decision tree performed better than the other individual models. Random forest and decision tree also recorded the highest value in precision, sensitivity and F1 score.

From the results achieved by the individual machine learning models, the random forest and decision tree are the best accuracy of 99.53 % and 99.20 % respectively in performance by the individual models used for the classification of anemic and non-anemic statuses of patients using the images of the palm in children. Whereas, ANN had less accuracy of 74.45 % in performance based on the accuracy of classifying images into their right status (anemic or non-anemic). So, among the individual classifiers, the random forest performs better, followed by Decision Tree, SVM classifier, and naïve Bayes with an accuracy of 99.20 %, 81.80 % and 78.16 % respectively while the model with the least performance is the ANN with an accuracy of 74.45 %.

Fig. 4 shows the accuracy, specificity, precision, sensitivity and f1 score of the five classifiers used. The RF achieved the highest accuracy of 99.53 % in terms of performance in all individual models, followed by the Decision tree, SVM, NB and ANN which achieved 99.20 %, 81.80 %, 78.16 %, and 74.45 % respectively. Fig. 5 is a graphical representation of the area under the curve for each single model. The AUC of the individual models are recorded as follows; RF (1.000000), DT (0.999939), NB (0.782753), ANN (0.763514) and SVM (0.812904), while Fig. 6 shows a fgraphical representation of the AUC of the Stacking Ensemble models.

Table 2
Evaluation metrics for individual machine learning models.

Classifiers	Evaluation Metrices%				
	Accuracy	Specificity	Precision	Sensitivity	F1 Score
Random Forest	99.53 %	100 %	100 %	99.22 %	99.61 %
SVM	81.80 %	82.62	84.24 %	87.49 %	80.72 %
Naïve Bayes	78.16 %	79.19 %	73.09 %	77.45 %	74.13 %
ANN	74.45 %	76.47 %	92.71 %	78.14 %	75.83 %
Decision Tree	99.20 %	100 %	100 %	98.71 %	99.35 %

3.2. Ensemble methods

The accuracy of each ensemble method for the detection of anemia related to the general achieved accuracy, specificity, sensitivity, precision and f1-score as indicated in Table 3. In stacking, even though naïve Bayes is among the classifiers with the lowest or weakest accuracy when combined with the random forest it recorded the highest accuracy of 99.98 %, and 100 % for specificity, sensitivity, precision and f1 score. Table 4 shows a comparative analysis between the results of other related studies and that of this (proposed) study.

The next best ensemble classifier was the combination of all 5 classifiers which recorded the same values as the combination of Naïve Bayes and random forest. It was also observed that the addition of Random Forest to other classifiers produced higher percentages of accuracy, specificity, precision, sensitivity and f1 score. The combination of the three individual models that recorded the lowest accuracy (ANN + NB + SVM) recorded an accuracy of 72.12 %, specificity of 63.99 %, precision of 74.35 %, sensitivity of 78.22 % and f1 score of 76.23 % which are the lowest values in the stacking as shown in Table 3.

In evaluating the models based on their accuracies, the combination of all five models produced the highest accuracy, followed by the combination of naïve Bayes and random forest then, the combination of random forest and naïve Bayes and SVM followed by the combination of RF + NB + SVM + ANN and lastly SVM + ANN + NB.

In voting, since the method majority voting technique was used to get the right results only an odd number of models was combined. A combination of three different models was used and later the combination of all 5 models. The combination of all 5 individual models produced an accuracy, specificity, precision, sensitivity and f1 score of 87.30 %, 90.09 %, 94.40 %, 86 %, and 90.01 % respectively whereas the combination of decision tree, random forest and Naïve Bayes produced the highest results with accuracy as 99.22 %, specificity as 100 %. Precision is 100 %, sensitivity is 98.71 % and f1 score is 99.35 %. Also, the combination of ANN + NB + SVM produced an accuracy of 82.00 %, specificity of 87.60 %, precision of 81.90 %, sensitivity of 78.22 % and an f1 score of 87.89 %.

The combination of NB + RF + DT and the combination of all 5 individual models produced the highest accuracy and specificity. Precision, sensitivity and f1 score.

Bagging the individual models such as NB, SVM, ANN, and DT

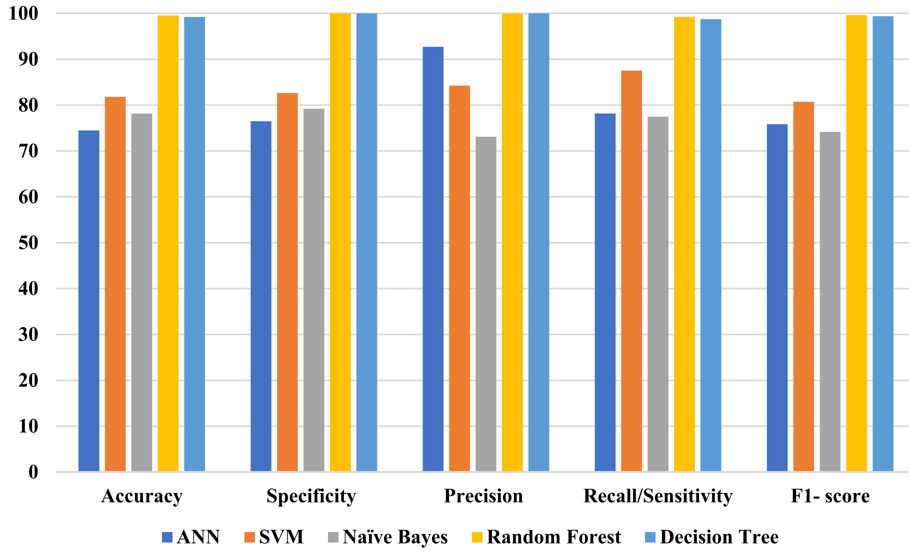

Fig. 4. Graphical representation for evaluation metric for individual models.

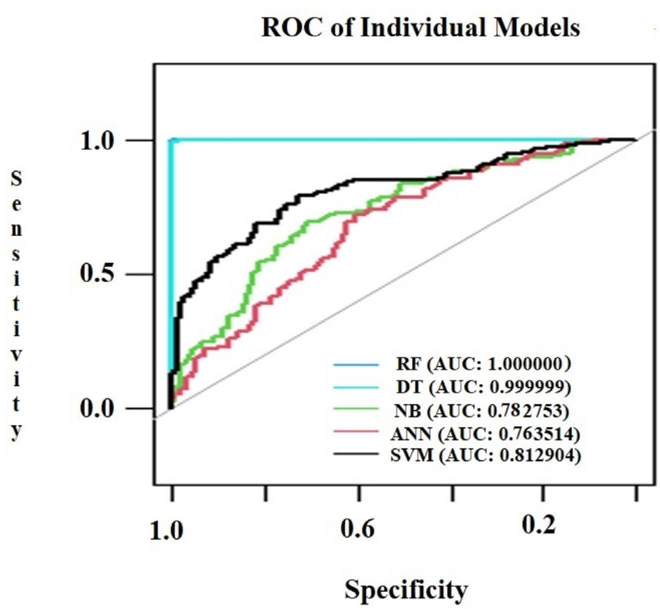

Fig. 5. ROC for the individual models.

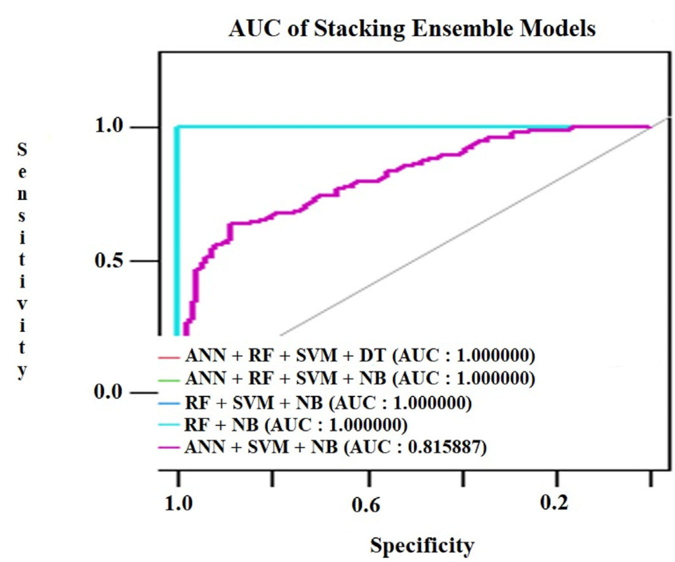

Fig. 6. ROC of the stacking ensemble techniques.

recorded a decrease in the evaluation metrics as compared to the single models. In the bagging process, the accuracy ranged between 80.48 % and 78.56 % where SVM recorded the highest accuracy and ANN recorded the lowest accuracy. Boosting the model produced an accuracy of 74.67 % with a specificity of 71.39 %, precision of 86.07 % sensitivity of 73.20 % and f1 score of 79.11 %. Based on the results in Table 3 the results achieved on the Ensemble techniques, stacking performed the best followed by voting, boosting and bagging.

Dalvi & Vernekar [13] used 4 ensemble techniques and performed 10 folds of cross-validation which were repeated, the boosting and the bagging achieved the highest accuracy of 91.806 % each by the ANN, while the k-NN achieved the least accuracy of 80.784 % and 82.332 % by the bagging and boosting respectively. However, among the voting and stacking, the Voting achieved a higher accuracy of 91.49 % with the application of (ANN + K-NN + Naïve + Decision) and lower accuracy of 90.234 % with the application of (Decision Tree + ANN) and Stacking with higher accuracy of 92.122 % by the (Decision + KNN) and lower accuracy of 90.862 % by the (Decision + ANN). As compared to the proposed study, Voting achieved a higher accuracy = 99.22 % and a lower accuracy of 72 % while Stacking also achieved a higher accuracy of 99.98 % as applied to the ensemble models.

In addition, the study by Dhakal et al. [46] achieved a lower accuracy of 92.20 % (NB + LR) and a higher accuracy of 98.40 % (RF + DT + ANN) were achieved when the voting ensemble model was used, while the Stacking ensemble model also achieved the least accuracy of 92.70 % (NB + LR), and a higher accuracy of 98.70 % by the (RF + ANN) and the (RF + LR)with the Bagging and Boosting achieving an accuracy of 98.80 % and 98.40 % all by the Random Forest respectively.

In the study conducted by Sevani et al. [15] the k-means algorithm achieved an accuracy of 90 % when a clustering technique was used, while the Naïve Bayes also achieved an accuracy of 90 % in a study conducted by Peksi et al. [29], and 85.60 %, 83.00 % and 79.60 % accuracy achieved by Bagged Decision Trees, Boosted Trees, and ANN respectively in a study conducted by Karagül et al. [26]. However, the ensemble learners' results achieved in this study performed better than the results of the single and other models reviewed in the literature.

Moreover, the study by Dalvi & Vernekar [13] stated that the performance of a voting technique improves when additional classifiers are

Table 3
Evaluation metrics for ensemble models.

S/N	Ensemble method		Evaluation metrics				
			Accuracy	Specificity	Precision	Recall/Sensitivity	F1 Score
1	Bagging	NB	79.75 %	59.68 %	70.18 %	76.02 %	72.99 %
		SVM	80.48 %	86.75 %	82.94 %	22.14 %	77.17 %
		ANN	78.56 %	75.79 %	79.56 %	81.36 %	83.36 %
		DT	79.24 %	100 %	100 %	60.81 %	75.63 %
2	Boosting		74.67 %	71.39 %	86.07 %	73.20 %	79.11 %
3	Voting	RF + SVM + ANN	82.46 %	88.00 %	94.14 %	80.16 %	86.59 %
		RF + DT + NB	99.22 %	100 %	100 %	98.71 %	99.35 %
		SVM + NB + RF	83.16 %	80.50 %	87.76 %	84.78 %	86.24 %
		ANN + SVM + NB + DT + RF	87.30 %	90.09 %	94.40 %	86 %	90.01 %
		SVM + ANN + NB	82 %	87.67 %	81.90 %	78.22 %	87.89 %
4	Stacking	NB + RF	99.98 %	100 %	100 %	100 %	100 %
		RF + NB + SVM	99.53 %	100 %	100 %	99.22 %	99.61 %
		NB + RF + SVM + ANN	99.30 %	100 %	100 %	98.84 %	99.42 %
		ANN + DT + RF + SVM + NB	99.98 %	100 %	100 %	100 %	100 %
		SVM + NB + ANN	82.12 %	83.99 %	89.35 %	88.22 %	89.23 %

Table 4
Comparative Analysis between other studies.

Reference	Ensemble Technique	Accuracy
Dalvi & Vernekar [13]	Boosting, bagging, Bayesian Boosting, voting and Stacking Ensemble.	Boosting: Min = 82.332 % (KNN) Max = 91.806 % (ANN). Bagging: Min = 80.784 % (KNN) Max = 90.86 % (ANN) Voting: Max = 91.49 % (ANN + K-NN + Naïve + Decision) Min = 90.234 % (Decision Tree + ANN) Stacking: Max = 92.122 % (Decision + KNN) Min = 90.862 % (Decision + ANN)
Dhakal et al. [46]	Voting, stacking, bagging and boosting	Voting: Min = 92.20 % (NB + LR), Max = 98.40 % (RF + DT + ANN) Stacking: Min = 92.70 (NB + LR), Max = 98.70 % (RF + ANN) and (RF + LR) Bagging: RF = 98.80 % Boosting: RF = 98.40 %
Proposed Study	ANN, SVM, NB, DT, RF, Stacking, Boosting, Bagging, Voting.	Boosting:72.69 % Bagging: Min = 53.56 % (ANN) Max = 70.48 % (SVM) Voting: Max = 99.22 % Min = 72 % Stacking: Max = 99.98 % Min = 72.12. RF = 99.53 % SVM = 75.80 % ANN = 64.45 % NB = 70.16 % DT = 99.20 %

added to the ensemble of voting methods [13].

However, with the Stacking approach, it is not the case. The highest accuracy was attained by the Decision Tree and K-NN classifier combination alone, and extra classifiers were not able to improve the accuracy of this ensemble. This indicates that with the Stacking approach, the best classifiers are not always required in order to attain the highest accuracy. In contrast, the accuracy of diagnosis increases proportionally as the number of classifiers increases in voting.

In this study, developing the stacking models we used naïve Bayes as a meta-learning because it does not require much training and can handle both continuous and discrete data. The two models with the highest accuracy were chosen and paired with a naïve Bayes and based on the results obtained we realized that combining a strong model with a naïve Bayes increases its accuracy. We then decided to pair a strong model with a weak model and realized an increase in accuracy.

With this, we realized that the combination of any model with a strong classification model increases its accuracy. Three of the weakest models were added and they performed better as compared to the individual models. Majority voting is a technique that only allows you to add odd numbers of individual models. This is a technique used for creating voting ensemble. A combination of randomly selected strong models and weak models was chosen for building the ensemble model. The combination of the two best models and a naïve Bayes for maximum accuracy.

An essential observation was the performance of random forest and decision tree, these two models are the best individual models because the data was split into binary because our study had just two classes (anemic and non-anemic) and the number of trees used was 500 which is also a factor when using random forest and decision tree. This study used a larger number of datasets and was able to provide higher accuracy for stacking and voting. The individual models, random forest and decision tree produced higher accuracies.

3.3. Limitations to the study

This study is limited to the use of Hemoglobin and other related components such as complete blood count, mean corpuscular hemoglobin concentration and mean corpuscular volume values as the integral dataset for the training, validating and testing the models, even though they were utilized by most studies. In addition, since the study's main focus is based on the use of the palpable to detect anemia in children using an ensemble approach, other medical images such as the palpable tongue, colour of the fingernails and the conjunctiva of the eye images were exempted. The outcome of this study would compare us to conduct a comparative in anemia detection using medical images such as the palpable palm, colour of the fingernails and conjunctiva of the eye images with the application of ensemble model techniques to know which of them has higher accuracy in anemia detection in children.

3.4. Conclusion

A new and non-invasive method to detect anemia based on digital images of the palm has been presented. Realizing that the detection and diagnosis of anemia are mostly based on the use of conjunctiva of the eye images, this study utilized digital images of the palm for the detection of anemia. Images acquired were augmented, trained, validated and tested using machine learning models, that is, SVM, ANN, DT, NB and RF and evaluation metrics for every single model.

Based on various research made, it was noted that to build an efficient and effective model, then a combination of one or more models must be made. Hence, we implored the process of creating a hybrid model using various Ensemble techniques such as bagging, voting, boosting and stacking. The effectiveness of the various ensemble machine learning classifiers to predict anemia classification was evaluated and compared.

For each datum, a training set and a test set were constructed for the extensive examination of a trained model with five-fold cross-validation. Additionally, the model on the test set was evaluated using several measures, such as accuracy, recall, F-1 score, precision, specificity, and AUC metrics.

The stacking approach, which was one of the four ensemble methods evaluated, outperformed the other three ensemble methods by attaining the highest levels of accuracy, specificity, precision, sensitivity, and f1 score. Naive Bayes are used as the base learners and the other models are used as stacking learners in the stacking approach [13] to accomplish this performance. Voting, boosting, and then bagging come after the stacking procedure. The top-performing effective individual classifier is the RF, followed by the DT, SVM, NB, and ANN. Individually, the NB is part of the model that performs the worst.

The best accuracy in the Stacking ensemble, however, was achieved by a combination of NB and RF. According to the results of the study, an ensemble of several algorithms or models performs significantly better than a single classifier or model. As a result, an ensemble of learners should be utilized to produce effective, swift, and cost-effective results while developing a model based on machine learning for the detection of anemia. Future works would be based on deploying the developed proposed models on a smartphone application to assist in anemia diagnosis where medical personnel, facilities and resources and scarce, mostly in developing communities.

4. The Study's consent and ethics

Before the study began, the collection of the dataset had been authorized by the ethical committees of all hospitals where it was obtained for this research. Since the study's participants were minors, their guardians/parents provided ethical consent on their behalf and informed them of the study's purpose. Before participants in the data collection were enlisted, their parents'/guardians' consent was authorized. In addition, this work was given authorization by the University of Energy and Natural Resources, Committee for Human Research and Ethics (CHRE) (Reference number: CHRE/CA/042/22). Additionally, during the image-capturing process, participants' names and faces were hidden, making it impossible to determine their identity.

Authors' contributions

Conceptualization, PA, JWA, SA and ETD; methodology, JWA, SSDD, EDYK, and PSD; software, SSDD, EDYK, and PSD; validation, PA, JWA, SSDD, EDYK, PSD, SA, and E.T.D; data curation, PA, JWA and ETD; writing—original draft preparation, SSDD, EDYK, PSD; writing—review and editing, JWA, and SA; visualization, PA, JWA, SSDD, EDYK, PSD, SA, and E.T.D; supervision, PA. and ETD; project administration, PA, JWA, SA and ETD.

Funding source

This study received no external funding. All funding was made solely by the authors.

Published datasets on a repository

This study's datasets have been made publicly available in a repository as;

“Asare, Justice Williams; Appiahene Peter; Donkoh, Emmanuel (2022), “Anemia Detection using Palpable Palm Image Datasets from Ghana”, Mendeley Data, V1, doi: 10.17632/ccr8cm22vz.1”

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Wiafe MA, Ayenu J, Eli-Cophie D. A review of the risk factors for iron deficiency anaemia among adolescents in developing countries. Anemia 2023:1–11. https://doi.org/10.1155/2023/6406286.
[2] 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. Health 2023;11(5):697. https://doi.org/10.3390/healthcare11050697.
[3] Dhalla S, Maqbool J, Mann TS, Gupta A, Mittal A, Aggarwal P, 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.
[4] Dimauro G, Caivano D, Girardi F. A new method and a non-invasive device to estimate anemia based on digital images of the conjunctiva. IEEE Access 2018;6:46968–75. https://doi.org/10.1109/ACCESS.2018.2867110.
[5] 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.
[6] 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.
[7] 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.
[8] World Health Organ (Who). Anemia treatment, prevalence and data status. https://www.who.int/health-topics/anaemia#tab=tab_3. [Accessed 16 July 2022].
[9] Al-alimi AA, Bashanfer S, Morish MA. Prevalence of iron deficiency anemia among university students in hodeida province, Yemen. Anemia 2018:1–7. https://doi.org/10.1155/2018/4157876.
[10] 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.
[11] Dimauro G, Baldari L, Caivano D, Colucci G. Automatic segmentation of relevant sections of the conjunctiva for non-invasive anemia detection. Int Conf on Smart and Sustain Technol (SpliTech) 2019:1–5.
[12] 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.
[13] Dalvi PT, Vernekar N. Anemia detection using ensemble learning techniques and statistical models. In: 2016 IEEE int conf on recent trends in electron, inf & commun technol (RTEICT). IEEE; 2016. https://doi.org/10.1109/RTEICT.2016.7808133. 1747–1751.
[14] Tamir A, Jahan CS, Saif MS, Zaman SU, Islam M, Intisar A. Detection of anemia from image of the anterior conjunctiva of the eye by image processing and thresholding. IEEE reg10 humanit technol conf (R10-HTC) 2017:21–3. https://doi.org/10.1109/R10-HTC.2017.8289053.
[15] Sevani N, Fredicia F, Persulessy GBV. Detection anemia based on conjunctiva pallor level using k-means a lgorithm. IOP Conf Ser Mater Sci Eng 2018;420:012101. https://doi.org/10.1088/1757-899X/420/1/012101.
[16] Asare JW, Appiahene P, Donkoh ET, Dimauro G. Iron deficiency anemia detection using machine learning models: a comparative study of fingernails, palm and

conjunctiva of the eye images. Eng Rep 2023:1–28. https://doi.org/10.1002/eng2.12667.

[17] 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.

[18] Acharya S, Swaminathan D, Das S, Kansara K, Chakraborty S, Kumar RD, et al. Non-invasive estimation of hemoglobin using a multi-model stacking regressor. IEEE J Biomed Health Inform 2020;24(6):1717–26. https://doi.org/10.1109/JBHI.2019.2954553.

[19] Sen B, Ganesh A, Bhan A, Dixit S, Goyal A. Machine learning based diagnosis and classification of sickle cell anemia in human RBC. In: 2021 int conf on intell commun technol and virtual mob netw (ICICV). IEEE; 2021. p. 753–8. https://doi.org/10.1109/ICICV50876.2021.9388610.

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

[21] 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.

[22] 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.

[23] Florestiyanto MY, Peksi NJ. Non-invasive anemia screening using nails and palms photos. Proc on Eng and Sci Ser 2020;1(1):311–8. https://doi.org/10.31098/ess.v1i1.124.

[24] Shahzad M, Umar AI, Shirazi SH, Khan Z, Khan A. Identification of anemia and its severity level in a peripheral blood smear using 3-tier deep neural network. Appl Sci 2022;12(10):5030. https://doi.org/10.3390/app12105030.

[25] Jayakody J A D C A HemoSmart. A non-invasive, machine learning based device and mobile app for anemia detection. IEEE reg 10 conf (TENCON) 2020:1401–6. https://doi.org/10.1109/TENCON50793.2020.9293903.

[26] Karagül T, Yurtay N, Öneç 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.

[27] Meena K, Tayal DK. 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.

[28] Verma Dr J, Arjun P. Diabetes mellitus prediction using ensemble machine learning techniques. Int J Recent Technol Eng 2020;9(2):312–6. https://doi.org/10.35940/ijrte.b3480.079220.

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

[30] Acar E, Türk Ö, Ertuğrul ÖF, 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.

[31] Belginova S, Uvaliyeva I. Decision support system for diagnosing anemia. Int Conf on Comput and Technol Appl (ICCTA) 2018:211–5. https://doi.org/10.1109/CATA.2018.8398684.

[32] 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.

[33] Noor NB, Anwar MdS, Dey M. Comparative study between decision tree, svm and knn to predict anaemic condition. In: 2019 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.

[34] Appiahene P, Chaturvedi K, Asare JW, Donkoh ET, Prasad M. Cp-anemic: a conjunctival pallor dataset and benchmark for anemia detection in children. Med Nov Technol Devices 2023;18:100244. https://doi.org/10.1016/j.medntd.2023.100244.

[35] Yang L. Medical application of digital image processing based on matlab. Int J of Innov Sci, Eng & Technol 2016;3(3):148–55.

[36] Sharma M, Nath K, Sharma RK, Kumar CJ, Chaudhary A. Ensemble averaging of transfer learning models for identification of nutritional deficiency in rice plant. Electron 2022;11(1):148. https://doi.org/10.3390/electronics11010148.

[37] Asare JW, Appiahene P, Donkoh ET. Detection of anaemia using medical images: a comparative study of machine learning algorithms – a systematic literature review. Inform Med Unlocked 2023;40:101283. https://doi.org/10.1016/j.imu.2023.101283.

[38] Chalco JP, Huicho L, Alamo C, Carreazo NY, Bada CA. Accuracy of clinical pallor in the diagnosis of anaemia in children: a meta-analysis. BMC Pediatr 2005;5(1):46. https://doi.org/10.1186/1471-2431-5-46.

[39] Getaneh T, Girma T, Belachew T, Teklemariam S. The utility of pallor detecting anemia in under five years old children. Ethiop Med J 2000;38(2):77–84.

[40] Joker PT. Binary image classifier cnn using tensorflow. https://medium.com/techiepedia/binaryimage-classifier-cnn-using-tensorflowa3f5d6746697 [accessed 26 June 2022].

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

[42] 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.

[43] Appiahene P, Missah YM, Najim U. Predicting bank operational efficiency using machine learning algorithm: comparative study of decision tree, random forest, and neural networks. Adv in Fuzzy Syst 2020;2020:1–12. https://doi.org/10.1155/2020/8581202.

[44] Jader R, Aminifar S. Predictive model for diagnosis of gestational diabetes in the kurdistan region by a combination of clustering and classification algorithms: an ensemble approach. Appl Comput Intell and Soft Comput 2022;2022:1–11. https://doi.org/10.1155/2022/9749579.

[45] Dejene BE, Abuhay TM, Bogale DS. Predicting the level of anemia among ethiopian pregnant women using homogeneous ensemble machine learning algorithm. BMC Med Inf Decis Making 2022;22(1):247. https://doi.org/10.1186/s12911-022-01992-6.

[46] Dhakal P. Prediction of anemia using machine learning algorithms. Int J Comput Sci Inf Technol 2023;15(1):15–30. https://doi.org/10.5121/ijcsit.2023.15102.