Medicine in Novel Technology and Devices 18 (2023) 100244

Contents lists available at ScienceDirect

Medicine in Novel Technology and Devices

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

CP-AnemiC: A conjunctival pallor dataset and benchmark for anemia detection in children

Peter Appiahene [a], Kunal Chaturvedi [b,*], Justice Williams Asare [a], Emmanuel Timmy Donkoh [a], Mukesh Prasad [b]

[a] University of Energy and Natural Resources, Sunyani, 214, Ghana
[b] School of Computer Science, FEIT, University of Technology Sydney, Sydney, 2007, Australia

ARTICLE INFO

Keywords:
Anemia
Hemoglobin
Conjunctiva pallor
Convolutional neural network
Deep learning
Classification
Regression

ABSTRACT

Anemia is a universal public health issue, which occurs as the result of a reduction in red blood cells. This disease is common among children in Africa and other developing countries. If not treated early, children may suffer long-term consequences such as impairment in social, emotional, and cognitive functioning. Early detection of anemia in children is highly desirable for effective treatment measures. While there has been research into the development of computer-aided diagnosis (CAD) systems for anemia diagnosis, a significant proportion of these studies encountered limitations when working with limited datasets.

To overcome the existing issues, this paper proposes a large dataset, named CP-AnemiC, comprising 710 individuals (range of age, 6–59 months), gathered from several hospitals in Ghana. The conjunctiva image-based dataset is supported with Hb levels (g/dL) annotations for accurate diagnosis of anemia. A joint deep neural network is developed that simultaneously classifies anemia and estimates hemoglobin levels (g/dL) based on the conjunctival pallor images. This paper conducts a comprehensive experiment on the CP-AnemiC dataset. The experimental results demonstrate the efficacy of the joint deep neural network in both the tasks of anemia classification and Hb levels (g/dL) estimation.

1. Introduction

Anemia is an important modifiable health risk for childhood morbidity and mortality [1]. The most recent WHO estimates indicate that 42% of children less than 5 years of age are anemic owing to a myriad of factors such as malnutrition, infectious diseases, parasitic infections, and congenital disorders of hemoglobin. However, in regions with a disproportionate burden of these risk factors, the burden of anemia exceeds this average. In Ghana, for instance, the prevalence of anemia in children 6–59 months have ranged from 59.5% to 68.4% over the past decade [2]. Childhood anemia is considered a public health priority owing to the magnitude of the condition and the debilitating effects on growth and development. It is strongly associated with serious short- and long-term consequences including impaired motor development, reduced immune and cognitive function, and increased mortality and morbidity [3]. Thus, early recognition of anemia becomes important for safe and effective treatment. Early diagnosis of anemia is important for reversing the health risks posed early in life and mitigating the eventual impact [4]. Laboratory detection of anemia is usually achieved by measurement of blood hemoglobin (Hb) concentration and involves a blood draw, technical expertise, and laboratory infrastructure.

In practice, the laboratory investigation of anemia following clinical suspicion suffers several drawbacks such as inadequate medical test financing, shortages of technical expertise in remote areas, quality requirements, and client hesitancy leading to abstention [5]. Also, there is a risk of exposure to bloodborne infections for healthcare personnel performing invasive procedures [6]. To overcome these issues, AI based Computer-Aided diagnosis (CAD) has become a feasible and a popular tool in medicine [7–9]. This is due to the surge in growth and interest in computer vision-based approaches that are low-cost and transparent in decision-making [10–12]. Non-invasive machine learning approaches [13–19] have attracted a lot of attention for their usage in anemia detection as a way of overcoming the diagnostic hurdles. Such methods use attributes based on the pallor of the exposed tissues such as conjunctiva [13–17], fingernail bed [18], palms [19], and tongue, which is caused by the reduction in hemoglobin concentration and healthy red blood cells. Interestingly, digital analysis based on the conjunctiva pallor is a more accurate factor than the palm, tongue, and nail bed pallor [20].

* Corresponding author.
E-mail address: kunal.chaturvedi@uts.edu.au (K. Chaturvedi).

https://doi.org/10.1016/j.medntd.2023.100244
Received 26 March 2023; Received in revised form 24 May 2023; Accepted 27 May 2023

This is due to two reasons: (1) there is minimal connective tissue between the outer mucous membrane and blood vessels, and (2) absence of dermis, epidermis, melanin, or subcutaneous tissues [16,21]. Examining of the paleness level of the conjunctiva of the eyes can be done by analyzing the correlation between the colour of the conjunctiva of the eyes and the Hb concentration of the body. Pallor of the conjunctiva of the eyes is more sensitive than other symptoms since it has been shown to happen more frequently in people with severe anemia.

While the conjunctiva-based approach is promising, the research on anemia detection suffers due to existing datasets' limitations. First, there are no publicly available conjunctiva-based datasets. The non-availability of private datasets limits their usage in literature. Second, the private datasets consist of imbalanced classes with mostly healthy people, and the representation of anemic patients, especially those with severe anemia, is limited. As a result of this bias, machine learning algorithms tend to underperform. Third, most of the current datasets are collected from a single source. For the developed algorithm to generalize well, it is necessary to include a diverse set of examples in the training set. In this paper, we present a publicly available conjunctival pallor dataset for anemia detection in children (CP-AnemiC) to present a benchmark for research in this field. It is the largest image collection of the conjunctiva with balanced classes of anemic and non-anemic samples. The dataset contains diversified samples collected from ten different sites in Ghana. Examples of the conjunctiva pallor images from the CP-AnemiC dataset are shown in Fig. 1.

In addition, we develop a deep neural network based joint classification and regression learning framework. Unlike previous studies, which focus on either the classification task [5,21,22] or the regression task [5,15,18] separately, our framework simultaneously classifies anemia and estimates Hb levels (g/dL) using the conjunctiva pallor images. The joint deep neural network-based framework learns high-level features from conjunctiva pallor images, then concatenates these features to detect anemia and estimate Hb levels (g/dL).

This paper makes the following contributions.

- This paper proposes the CP-AnemiC dataset, which is the largest publicly available dataset based on conjunctival pallor for anemia detection in children.
- This paper develops a joint classification and regression learning framework for conjunctiva pallor-based anemia diagnosis, in which both the feature extraction and classification/regression model training processes are incorporated into a unified deep neural network .
- This paper performs an extensive set of experiments to illustrate the performance of the CP-AnemiC dataset using the joint deep CNN framework.

The remaining paper is organized as follows: Section 2 briefly reviews state-of-the-art methods, recent datasets, and challenges for anemia diagnosis. Section 3 presents the data collection, annotation, and statistical analysis of the proposed CP-AnemiC dataset. Section 4 describes the proposed methodology, experimental settings, and evaluation of the CP-Anemic dataset. At last, in Section 5, the main conclusions are outlined.

2. Related work

2.1. Anemia detection

The research field of anemia detection has been dominated by blood analysis based invasive approaches [23–25]. Constantino et al. [23] use morphological imaging and histogram analysis of RBCs to determine whether the person has anemia or not. Furthermore, Elsalamony [24,25] performs microscopic analysis of blood smears to detect anemia kinds. The analysis is done by classifying different anemia kinds by using a combination of a shape-based approach and a neural network. These methods require blood sampling from the patient, which causes pain and may increase the risks of blood disease transmission. To alleviate such issues, several image based non-invasive (non-hemanalysis) approaches have been explored (e.g., fingernails [18], conjunctiva [13–17], retina [5], etc.). The exposed tissues such as fingernails, conjunctiva, and retina are highly vascularized and are not affected by the presence of melanin, which causes changes to the skin colour among different ethnic groups [16]. These methods can detect anomalies due to reduced healthy erythrocytes (red blood cells) and hemoglobin concentrations in the exposed tissues.

Non-invasive methods based on traditional machine learning algorithms are firstly applied to this task [13–15,15,17,18]. These hand-crafted feature-based approaches may not be robust to different lighting or environmental conditions. Moreover, these approaches are generally experimented on smaller datasets, which restricts their scalability for real-world conditions. Recently, researchers have started to explore non-invasive deep learning methods [5,22] for anemia diagnosis. Mitani et al. [5] leverages deep learning to quantify Hb levels and detect anemia based on the non-invasive fundus images. Wei et al. [22] proposes AneNet, a fully-convolutional deep learning based model to diagnose anemia based on retinal vessel optical coherence tomography (OCT) images. In this study, we focus on deep learning-based anemia detection using the characteristic signs in the conjunctival pallor. The use of conjunctiva in anemia detection has additional benefits over the use of other exposed tissues. First, there is minimal distance between the blood vessels and tissue surface. Second, environmental factors such as temperature and physiological effects on blood flow have less of an impact on the diagnosis.

2.2. Related dataset

Although numerous researchers have proposed anemia diagnosis based on palpebral conjunctiva [13,15–17,26], majority of the results are

Fig. 1. Exemplar conjunctival pallor images in our CP-AnemiC dataset. The first row represents images from the non-anemic patients and the second represents images from the anemic patients.

Table 1
Summary of representative conjunctival pallor based datasets.

Dataset	Participants	Age Group
Chen et al. [13]	100	–
Anggraeni et al. [17]	20	22–36 years
Dimauro et al. [15]	113	–
Bauskar et al. [26]	99	–
Suner et al. [16]	344	19–96 years
Noor et al. [27]	104	–
Tamir et al. [21]	19	–

reported on their private datasets, making it difficult to compare existing methods and suggest the best direction for future research. In Table 1, we summarize the conjunctiva image datasets with related information. Chen et al. [13] use Kalman filter (KF) and a regression based approach on a small dataset of 100 images. Dimauro et al. [15] use a macro-lens attached to a smartphone to collect data from 113 participants and diagnose them using a K-Nearest Neighbor (KNN) algorithm. Anggraeni et al. [17] build a conjunctiva image dataset from 20 pregnant patients to predict hemoglobin levels. The recently published dataset by Bauskar et al. [26] consists of 99 conjunctiva images that are collected from the internet and other sources. The collected dataset is further annotated by a pathologist. Even though these datasets are a valuable contribution to computer vision-aided anemia detection, they do have some limitations. One common issue is that they contain very few samples of the conjunctiva pallor images. Furthermore, a majority of the existing datasets are class imbalanced, which inevitably cause the trained model to exhibit majority class bias. Most recently, Suner et al. [16] proposes a smartphone-camera based dataset collected from 344 patients. A special care has been taken to exclude any conjunctiva image that is out of focus or has poor lighting. However, the dataset is collected from only one site by a single operator which reduces the data variability.

To overcome these limitations, we propose the largest publicly available dataset to assist the development of deep learning-based algorithms in this field. Our dataset is class balanced and contains samples from 10 different sites in Ghana.

3. Dataset

3.1. Participant recruitment and data collection

From January to June 2022, pairs of conjunctival images and hemoglobin concentrations were opportunistically retrieved by trained biomedical scientists from 10 health facilities across the country including the Komfo Anokye Teaching Hospital at Kumasi, Bolgatanga Regional Hospital at Bolgatanga, Kintampo Municipal Hospital at Kintampo, Ahmadiyya Muslim Hospital at Techiman, Sunyani Municipal Hospital at Sunyani, Manhyia District Hospital at Kumasi, Ejusu Government Hospital at Ejusu, SDA Hospital at Sunyani, Nkawie-Toase Government Hospital at Nkawie-Toase and Holy Family Hospital at Berekum. The study was reviewed and approved by the Committee for Human Research and Ethics (Reference number: CHRE/CA/042/22). Parents of eligible children provided written informed consent prior to enrolment in the study and data collection.

Fig. 2 shows the pipeline of the data collection. Licensed biomedical scientists were given a week-long proficiency training at the Centre for Research in Applied Biology, University of Energy and Natural Resources, Ghana on how to capture conjunctival images from children aged 6–59 months presenting to assigned healthcare facilities within the stated period using electronic instruments (Kobo Collect v2021.2.4, Massachusetts, USA) deployed on mobile tablets (Samsung Galaxy Tab 7 A, Samsung, Vietnam). The system consisted of a form to capture information from patients such as Hb levels, age, gender, and a remark based on the Hb Value obtained during laboratory assessment, and thereafter took a picture of the image of the conjunctiva of the eyes to upload to the database which allows for easy availability and accessibility. To capture

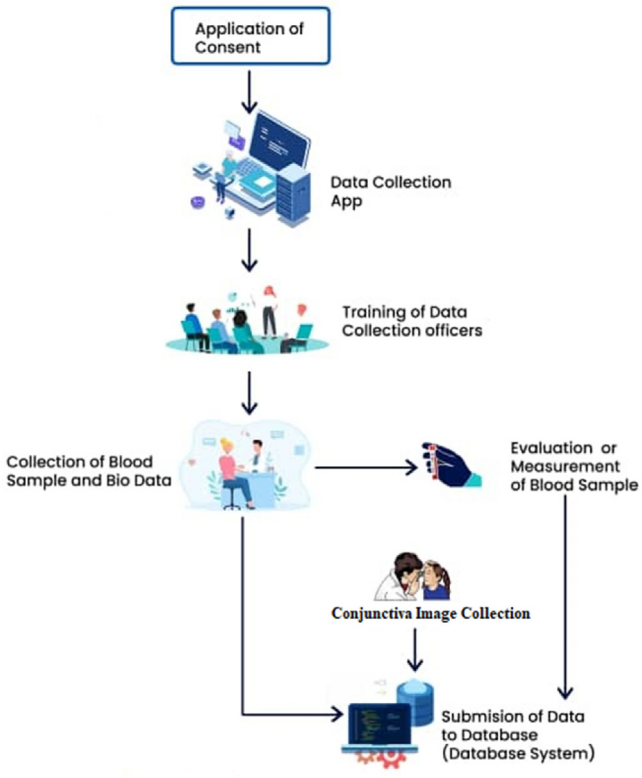

Fig. 2. The pipeline of the data collection for anemia diagnosis.

the conjunctiva, the lower eyelid was gently pushed back with the thumb with the aid of the index finger. All images were taken by laboratory personnel in ambient natural light with a standard camera of 12 MP image resolution mounted on an electronic tablet.

Furthermore, the cameras' spotlights were turned off when capturing the photographs to prevent excessive shine effects created by the picture quality, that is to eradicate ambient light, which dramatically influences detection or classification by the models. This method is a great technique to remove the effect of ambient light on photographs in datasets. The ROI of the conjunctiva of the eyes were extracted after the application of the triangle thresholding algorithm, in combination with the entropy grayscale image algorithm. This data is then stored along with the captured conjunctiva pallor image and the demographic information.

3.2. Data pre-processing

Based on the outcome of image extraction and the development of the dataset, each image's ROI was converted to the CIELAB (also known as CIE L*a*b*) colour space model. The L*a*b colour space is intended to simulate human eyesight or perception. The standard deviation value of the ROI a* components, which is the mean value, is used to express it. A* components are red components (a* $>$ 0) and green components (a* 0). The average value A* components are red components (a* $>$ 0) and green components (a* 0). Previous research in this field shows that there is a strong relationship between a* components and Hb levels when calculated using the Pearson Correlation Index, and various experiments in this domain show that individuals with higher Hb values tend to have an average value of a* greater than 160, while patients with lower Hb values tend to have an average value of a* less than 142 [21].

The datasets are analyzed to get the detection of anemia. To begin, colour characterization of pictures was performed to utilize the CIE L*a*b* colour space. This colour space converts all colours visible to the human eye into a three-dimensional integer space, allowing device-independent digital representation. Because the above-mentioned

components correlate to systemic changes in the alleged colour, the relative perceptual differences between two colours in L*, a*, and b* may be determined. Each colour may be approximated by considering it as a point in three-dimensional space (with three components: L* a*b*) and measuring the Euclidean distance between them. L* (Lightness) indicates the darkest black at zero and the brightest white at one hundred, whereas a* and b* are colour channels. Nonalignment grey is defined by its existence in the Cartesian coordinate system (a*, b*).

3.3. Dataset collection limitations

This study encountered some challenges which occurred before and during the collection of the dataset for the study, even though we had a higher number of a dataset with a balanced anemic and non-anemic classes.

Low Response Rate: Anemia is a condition that is often underreported due to the lack of awareness and education. As a result, it was difficult to obtain a high response rate when collecting the dataset on anemia. In addition, most parents and guardians were reluctant to give consent for their ward(s) to partake in the study, although their reasons were not conveyed. Most parents or guardians however gave their consent after several consultations and detailed explanations of the benefits of the study to the healthcare fraternity.

Sample Selection: The dataset collection was tedious, especially taking pictures of the conjunctiva of the eyes. This was due to the reason that the participants were minors (aged 6–59 months) which required extra attention and caution to avoid the finger entering the eyes.

Ethical Consent Clearance: Since the study involved human life, there was a need to acquire ethical clearance from hospitals and the Committee for Human Research and Ethics for the study. This passed through a lot of procedures and took a lot of time. Even though the ethical clearance was approved, there were some rescheduled meetings with hospital administrators due to their busy schedules.

Costs: Collecting the dataset on anemia was expensive due to the need for laboratory tests, training of biomedical scientists, and visits to hospital facilities.

3.4. Annotation

According to the WHO criteria, the anemia in children of the age, six months to 59 months is defined with Hb level (value) $<$ 11 g/dL as anemic and$\geq$11 g/dL as non-anemic. Each image in the CP-AnemiC dataset is annotated with meta-information such as participant's Hb levels, gender, age, remark, and the location of the sample collection.

3.5. Dataset statistics

Table 2 summarizes the CP-AnemiC statistics. The CP-AnemiC dataset includes 710 images with 424 (60%) labelled as anemic patients and 286 (40%) labelled as non-anemic patients. Unlike the existing datasets, anemic.

Table 2
A patient-level characteristics of the dataset.

	Patient Class		
	Anemic	Non-Anemic	Total
Patients	424 (60%)	286 (40%)	710 (100%)
Female	174 (57%)	132 (43%)	306 (43%)
Male	250 (62%)	154 (38%)	404 (57%)
Age (months)	31.04 ± 17.02	32.31 ± 16.46	31.58 ± 16.78
	Anemia Diagnosis for Age 6–59 months		
Anemia Classification	Anemic	Non-Anemic	
Hemoglobin Levels	<11 g/dL	≥ 11 g/dL	

Class covers the largest part our dataset. The mean age of the study participants was 31.58 months (95% confidence interval (CI), 14.8–48.36 years) (range, 6–59 months). Out of the 710 participants, 306 (43%) were female and 404 (57%) were male. Fig. 3 (left) shows the gender wise distribution of the Hemoglobin concentration (g/dL) and Fig. 3 (right) shows the age wise distribution of the Hemoglobin concentration (g/dL) along with anemia classification for each participant.

4. Experiment and evaluation

4.1. Baseline models

The models used in the baseline experiments are VGG16 [28], ResNet50 [29], DenseNet121 [30], and Vision Transformer (ViT) [31], and ConvNeXtBase [32]. The VGG network proposed by Simonyan et al. [28] employs the superposition of multiple 3 $\times$ 3 convolution filters to replace a large convolution filter, which can both increase network depth and reduce total parameters. ResNet50 uses residual blocks to extract more in-depth features without any network degradation. To accomplish this, ResNet50 used a convolution operation on the input, then four residual blocks, and finally a full connection operation to perform classification tasks. Densely connected convolutional network (DenseNet) ensures the advantages of both Highway [33] and ResNet [29] architectures. Furthermore, it allows maximum information flow between layers in the network. The Vision Transformer (ViT) introduced by Dosovitskiy et al. [31] is the first transformer architecture inherited from NLP that can be directly applied to images. ViT treats an image as a sequence of patches and performs favorably to state-of-the-art convolutional networks. At last, ConvNeXtBase, a hierarchical convolutional neural network (CNN) modified from ResNet50 with a pure CNN architecture is used. ConvNeXt combines the intrinsic superiority of transformer to bridge the gap between CNNs and ViTs. Next, as shown in

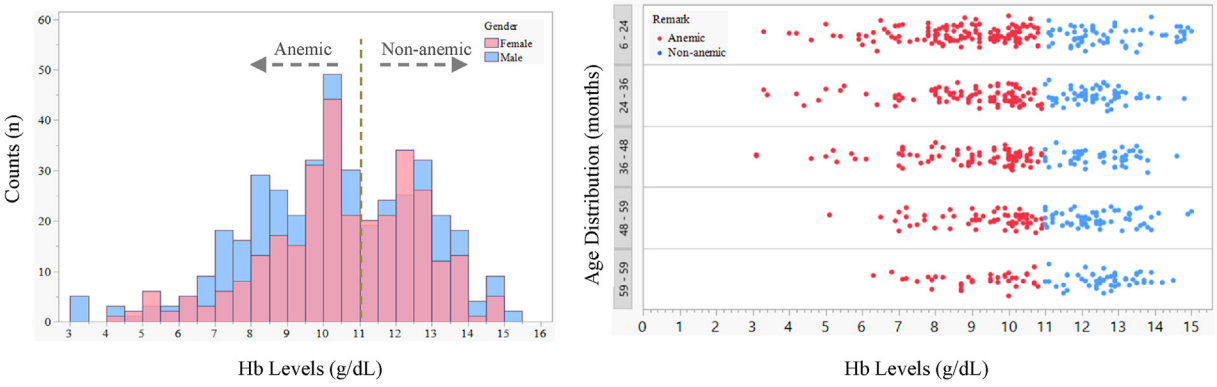

Fig. 3. The gender wise distribution of the Haemoglobin concentration (g/dL) (left). The age wise distribution of the Haemoglobin concentration (g/dL) along with remarks of each participant (right).

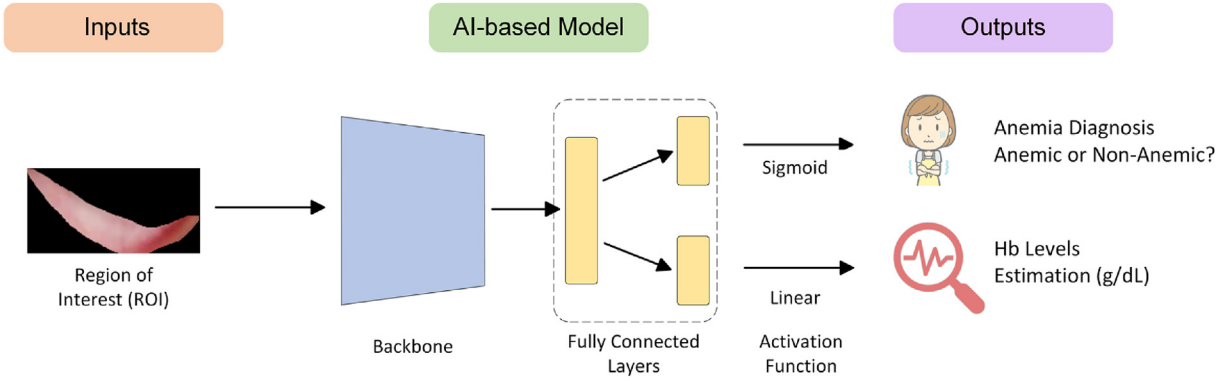

Fig. 4. Overview of the proposed deep neural network based joint classification and regression learning framework.

Fig. 4, we modify the fully connected layer of each model to get one output for Hb levels estimation via linear activation function and another output for anemia diagnosis task using sigmoid activation function. The network is mathematically described as follows.

Consider N data points, $X = \{x_1, x_2, x_3, \ldots. x_N\}$ where $x_n \in X$ represents the conjunctiva pallor image belonging to n^{th} participant. Here, $y^c = \{y_n^c\}_{n=1}^N (c = 0, 1)$ denotes the anemic/non-anemic labels annotated by healthcare professionals. $z^s = \{z_n\}_{n=1}^N (s \geq 0)$ denotes the Hb levels estimated using blood sample hemoglobin screening. The anemic labels and Hb levels are used in a back-propagation procedure to update the network weights in the model layers and learn the most important features in the fully connected layers. The network aims to learn a nonlinear mapping $\Psi: X \rightarrow \{\{y^c\}_{c=0}^1, z^s\}$ from the input space to both spaces of the anemic label and Hb levels. We equally treat the tasks of anemia classification and Hb levels regression, with the objective function defined as follows in Eq. (1):

$$\underset{w}{\arg\min} \frac{1}{N} \sum_{n=1}^{N} \{y_n \log(\hat{y}_n) + (1 - y_n)\log(1 - \hat{y}_n)\} + \frac{1}{N} \sum_{n=1}^{N} (z_n - \hat{z}_n)^2 \quad (1)$$

where the first term is the cross-entropy loss for the classification of anemia, and the second one is the mean squared loss for regression to evaluate the difference between the predicted Hb levels and the ground truth Hb levels. Here, $\hat{y}_n = P(y_n^c = c | x_n; w)$ denotes the probability map obtained by sigmoid on the activation values and $\hat{z}_n = f(x_n; w)$ is the regression score obtained by the model for Hb levels.

4.2. Experiment setting

To increase the data size of training images, all the models used various data augmentation strategies such as random horizontal flip, random rotation, random shifts, and random scaling. We apply a 5-fold cross-validation technique where four folds are reserved for training and another fold is used for testing. Furthermore, the weights of each model were randomly initialized prior to training with a batch size of 32 on the NVIDIA Quadro RTX 6000 GPU. The Cross-Entropy loss function is used with the Adam Optimizer and the learning rate is set to 10^{-4}. The joint deep neural network-based framework, which consists of a backbone and fully connected layers, is fine-tuned end-to-end and converges after 150 epochs.

4.3. Results

Evaluation of Anemia Diagnosis. As shown in Table 3, we compare the performance of each model using metrics such as classification accuracy, precision, recall, and F1-Score. Here, ViT and ResNet50 have similar performances irrespective of the completely different design choices. We notice that DenseNet121 and ConvNeXtBase achieve the lowest accuracy among the five models. Despite the higher number of network parameters in VGG16, the performance is still comparable to DenseNet121. As depicted in Fig. 5, we also show the precision-recall curves of the above results. Here, DenseNet121 achieved an AUC of 0.793 with a precision of 0.790 and recall of 0.788, ResNet50 has an AUC of 0.835 with a precision of 0.852 and recall of 0.830, VGG16 achieved AUC of 0.791 with a precision of 0.807 and recall of 0.785, ViT achieved AUC of 0.841 with a precision of 0.833 and recall of 0.835, and ConvNeXtBase achieved an AUC of 0.745 with a precision of 0.749 and recall of 0.739. The areas under the receiver operating characteristic curve (AUC) is highest for ViT among the different models. The lowest AUC value of 0.744 was achieved on ConvNeXtBase.

Table 3
Anemia Detection result on CP-AnemiC using different state-of-the-art algorithms.

Backbone	Precision	Recall	F1-Score	AUC	Accuracy
ResNet50	0.852	0.830	0.837	0.835	84.79%
VGG16	0.807	0.785	0.799	0.791	80.56%
DenseNet121	0.790	0.788	0.786	0.793	79.58%
ViT	0.833	0.835	0.833	0.841	84.08%
ConvNeXtBase	0.749	0.739	0.740	0.744	75.63%

Evaluation of Hb levels Estimation. As shown in Table 4, we assess the performance of regression task by the models using the Mean Average Error (MAE), and the Mean Squared Error (MSE). In addition, a goodness-of-fit plot between the actual and the estimated Hb levels is shown in Fig. 6. The blue broken line shows the bias values plotted in the Bland Altman plots with red broken lines visualizing the distribution of values within 95% limits of agreement. It should be noted that the higher

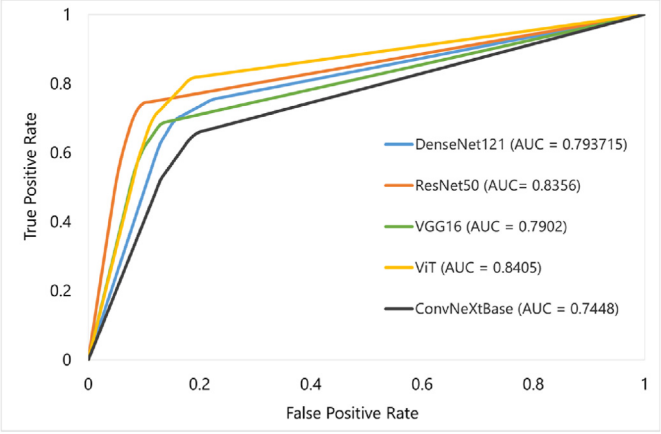

Fig. 5. ROC curves for detecting anemia from the conjunctiva pallor.

Table 4
Hemoglobin levels prediction result on CP-AnemiC using different state-of-the-art algorithms.

Backbone	MAE	MSE
ResNet50	1.54	4.15
VGG16	1.60	4.26
DenseNet121	1.62	4.51
ViT	1.50	3.82
ConvNeXtBase	1.70	4.76

performance of a model in the anemia classification does not imply higher performance in the Hb levels estimation. ViT model obtained best results with lowest MAE and MSE scores of 1.50 and 3.82 respectively. ResNet50 which reported the best results for anemia classification, reported MAE and MSE scores of 1.54 and 4.15 respectively. Furthermore, we use Scatter plots and Bland-Altman plots to depict the degree of agreement for the data. As shown in Fig. 6 (left), Scatter plots of the ground truth Hb levels (g/dL) and predicted Hb levels (g/dL) are calculated for the five models. As depicted in Fig. 6 (right), Bland-Altman plots report the differences in Hb levels (g/dL) measured by the two methods

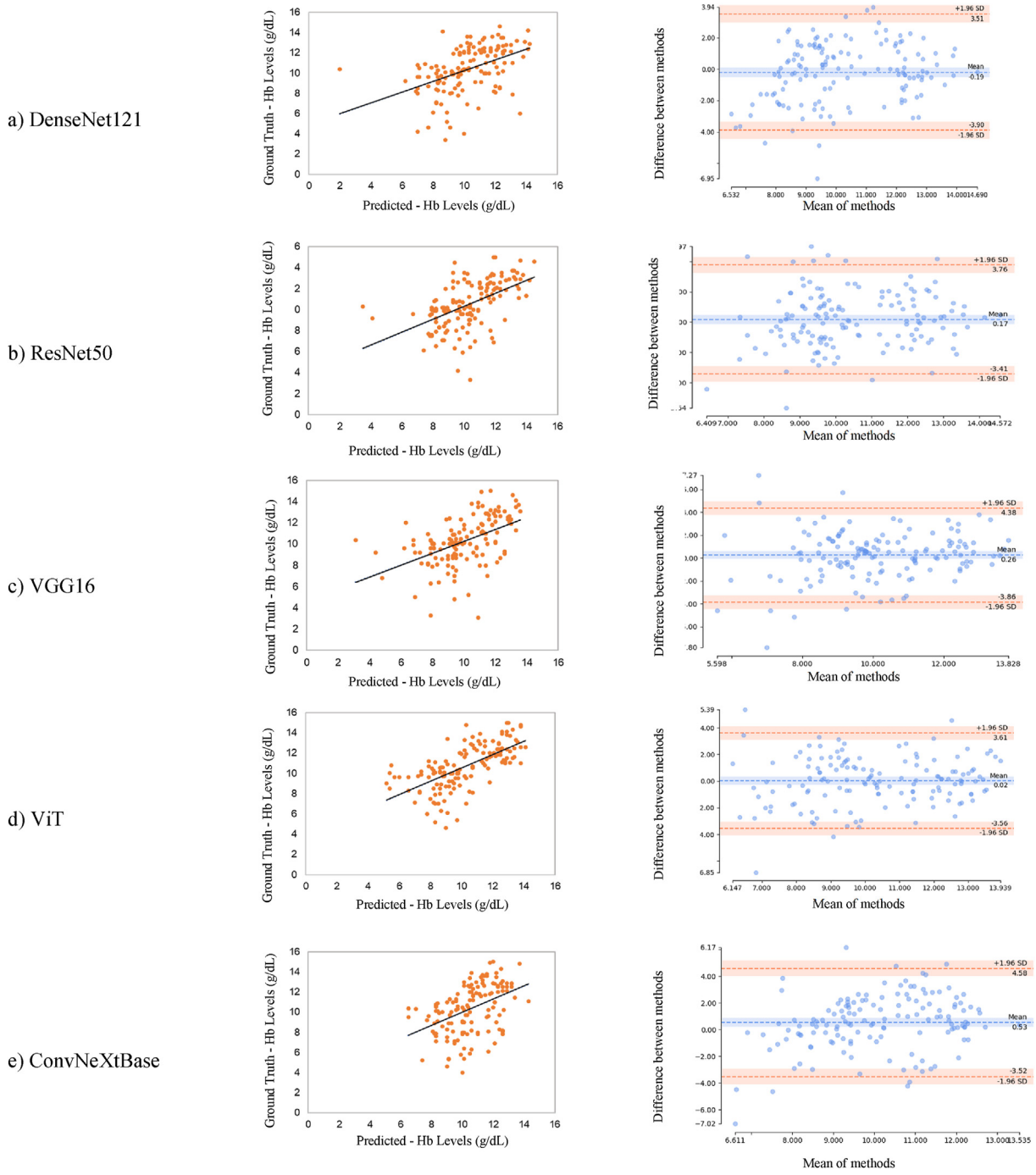

Fig. 6. Scatter plots of the predicted Hb Levels (g/dL) vs. the real Hb Levels (g/dL) achieved by using five different models (Left). Bland–Altman plot for predicted and measured Hb levels. Each blue dot represents the difference between the ground truth Hb value and the predicted Hb value. The blue broken line represents the mean of the difference, and red broken lines represent 95% limits of agreement (mean ± 1.96 × s.d.). (Right).

against the average of the two methods. Here, ViT reports the best results with the bias of 0.02 and the limit of agreement range (LoA) from −3.56 to 3.61 (see Fig. 6d). ResNet50 also performed good with bias of 0.17 and LoA range from −3.41 to 3.76 (see Fig. 6b). The Bland-Altman plots for all the five models showed a random nature of the spread with no clear pattern in the differences of the paired results.

5. Conclusion

This paper first introduces CP-AnemiC dataset towards two representative tasks in anemia diagnosis in children, i.e., Hb levels predictions and anemia label classification. The CP-AnemiC is the largest publicly available conjunctiva pallor-based dataset with 710 annotated images of balanced anemic and non-anemic classes. Further, this paper introduces a deep joint deep neural network framework for simultaneous anemia classification and Hb levels (g/dL) regression and benchmarked it for our dataset. The experimental results show that the anemia diagnosis can be greatly benefited from our large, clean and class-balanced CP-AnemiC dataset. It is noted that while the proposed dataset collected from diverse sources in Ghana, the ethnicity is highly biased towards people of one ethnic group. In the future, we aim to create a diverse dataset of different ethnic groups for more robustness and generalization of the proposed deep joint deep neural network framework.

Author contribution

Peter Appiahene: Conceptualization, validation, data curation, writing—review and editing supervision, visualization. Kunal Chaturvedi: Conceptualization, methodology, software, validation, writing—original draft preparation, writing— review and editing. Justice Williams Asare: Conceptualization, software, validation, data curation, writing—original draft preparation, writing—review and editing. Emmanuel Timmy Donkoh: validation, data curation, writing—review and editing, visualization. Mukesh Prasad: methodology, validation, writing—review and editing, supervision, visualization.

Funding source

This research received no external funding.

Ethical approval and informed consent (if applicable)

The ethics and consent Committee for Human Research of the University of Energy and Natural Resources, Sunyani, Ghana approved the commencement of this work.

Patient consent (if applicable)

Before we began this study, the ethical committee for human research of the University of Energy and Natural Resources, Ghana approved the collection of the datasets to be used for this work. Furthermore, because the participants (patients) in the study were minors, ethical consent was obtained from their parent(s) or guardian(s), and the purpose and objectives of the study were explained to them, along with the benefits it would be to the health services. Also, patients' or participants' names and faces were not shown or exposed during the capturing of images, which makes their identities unidentified.

Data Availability

The dataset used for this study have been published on the repository "CP-AnemiC (A Conjunctival Pallor) Dataset from Ghana", Mendeley Data, V1, doi: 10.17632/m53vz6b7fx.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] Chaparro CM, Suchdev PS. Anemia epidemiology, pathophysiology, and etiology in low- and middle-income countries. Ann N Y Acad Sci 2019:14092. https://doi.org/10.1111/nyas.14092. nyas.
[2] Anaemia, World Heal. Organ. (n.d.). https://www.who.int/health-topics/anaemia#tab=tab_1 (accessed December 24, 2022).
[3] Kapil U, Bhavna A. Adverse effects of poor micronutrient status during childhood and adolescence. Nutr Rev 2002;60:S84. https://doi.org/10.1301/00296640260130803. –S90.
[4] Cusick S, Georgieff M, Rao R. Approaches for reducing the risk of early-life iron deficiency-induced brain dysfunction in children. Nutrients 2018;10:227. https://doi.org/10.3390/nu10020227.
[5] Mitani A, et al. Detection of anaemia from retinal fundus images via deep learning. Nat. Biomed. Eng. 2020;4:18–27. https://doi.org/10.1038/s41551-019-0487-z.
[6] Rice BD, Tomkins SE, Ncube FM. Sharp truth: health care workers remain at risk of bloodborne infection. Occup Med 2015;65:210–4. https://doi.org/10.1093/occmed/kqu206.
[7] 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;29:2649–62. https://doi.org/10.3906/elk-2103-77.
[8] Badarinath D, et al. Study of clinical staging and classification of retinal images for retinopathy of prematurity (ROP) screening. In: 2018 int. Jt. Conf. Neural networks. IEEE; 2018. p. 1–6. https://doi.org/10.1109/IJCNN.2018.8489491.
[9] Padmanabha AGA, Appaji MA, Prasad M, Lu H, Joshi S. Classification of diabetic retinopathy using textural features in retinal color fundus image. In: 2017 12th int. Conf. Intell. Syst. Knowl. Eng. IEEE; 2017. p. 1–5. https://doi.org/10.1109/ISKE.2017.8258754.
[10] Prasad M, et al. Fusion based en-FEC transfer learning approach for automobile parts recognition system. In: 2018 IEEE symp. Ser. Comput. Intell. IEEE; 2018. p. 2193–9. https://doi.org/10.1109/SSCI.2018.8628789.
[11] Chou K-P, et al. Robust facial alignment for face recognition. 497–504, https://doi.org/10.1007/978-3-319-70090-8_51; 2017.
[12] Li D-L, Prasad M, Hsu S-C, Hong C-T, Lin C-T. Face recognition using nonparametric-weighted Fisherfaces. EURASIP J Appl Signal Process 2012:92. https://doi.org/10.1186/1687-6180-2012-92. 2012.
[13] Chen Y-M, Miaou S-G, 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.
[14] Chen Y-M, Miaou S-G. A kalman filtering and nonlinear penalty regression approach for noninvasive anemia detection with palpebral conjunctiva images. J. Healthc. Eng. 2017:1–11. https://doi.org/10.1155/2017/9580385. 2017.
[15] 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.
[16] Suner S, et al. Prediction of anemia and estimation of hemoglobin concentration using a smartphone camera. PLoS One 2021;16:e0253495. https://doi.org/10.1371/journal.pone.0253495.
[17] 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.
[18] Mannino RG, et al. Smartphone app for non-invasive detection of anemia using only patient-sourced photos. Nat Commun 2018;9:4924. https://doi.org/10.1038/s41467-018-07262-2.
[19] 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;4:1. https://doi.org/10.1155/2014/543860. 2014.
[20] 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. 2023. p. 1–28. https://doi.org/10.1002/eng2.12667.
[21] 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;16:2. https://doi.org/10.1186/s13040-023-00319-z.
[22] Wei H, Shen H, Li J, Zhao R, Chen Z. AneNet: a lightweight network for the real-time anemia screening from retinal vessel optical coherence tomography images. Opt Laser Technol 2021;136:106773. https://doi.org/10.1016/j.optlastec.2020.106773.
[23] Constantino BT. The red cell histogram and the dimorphic red cell population. Lab Med 2011;42:300–8. https://doi.org/10.1309/LMF1UY85HEKBMIWO.
[24] Elsalamony HA. Healthy and unhealthy red blood cell detection in human blood smears using neural networks. Micron 2016;83:32–41. https://doi.org/10.1016/j.micron.2016.01.008.
[25] Elsalamony HA. Anaemia cells detection based on shape signature using neural networks. Measurement 2017;104:50–9. https://doi.org/10.1016/j.measurement.2017.03.012.

[26] Bauskar S, Jain P, Gyanchandani M. A noninvasive computerized technique to detect anemia using images of eye conjunctiva. Pattern Recogn Image Anal 2019; 29:438–46. https://doi.org/10.1134/S1054661819030027.
[27] Bin Noor N, Anwar MS, Dey M. Comparative study between decision tree, SVM and KNN to predict anaemic condition. In: 2019 IEEE int. Conf. Biomed. Eng. Comput. Inf. Technol. Heal. IEEE; 2019. p. 24–8. https://doi.org/10.1109/BECITHCON48839.2019.9063188.
[28] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. http://arxiv.org/abs/1409.1556; 2014.
[29] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. http://arxiv.org/abs/1512.03385; 2015.
[30] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: 2017 IEEE conf. Comput. Vis. Pattern recognit. IEEE; 2017. p. 2261–9. https://doi.org/10.1109/CVPR.2017.243.
[31] Dosovitskiy A, et al. An image is worth 16x16 words: transformers for image recognition at scale. http://arxiv.org/abs/2010.11929; 2020.
[32] Liu Z, et al. A ConvNet for the 2020s. In: 2022 IEEE/CVF conf. Comput. Vis. Pattern recognit. IEEE; 2022. https://doi.org/10.1109/CVPR52688.2022.01167. 11966–11976.
[33] Srivastava RK, Greff K, Schmidhuber J. Training very deep networks. In: Adv. Neural inf. Process. Syst.; 2015. p. 1–9. https://doi.org/10.5555/2969442.2969505. 28 (NIPS 2015.