

# AneRBC dataset: a benchmark dataset for computer-aided anemia diagnosis using RBC images

Muhammad Shahzad<sup>1,2</sup>, Syed Hamad Shirazi<sup>1</sup>, Muhammad Yaqoob<sup>1,3</sup>, Zakir Khan<sup>1,2</sup>,  
Assad Rasheed<sup>1</sup>, Israr Ahmed Sheikh<sup>4</sup>, Asad Hayat<sup>4</sup>, Huiyu Zhou<sup>5</sup>

<sup>1</sup>Department of Information Technology, Hazara University Mansehra, Dhodial, Mansehra, Khyber Pakhtunkhwa 21120, Pakistan

<sup>2</sup>Department of Computer Science, National University of Sciences and Technology (NUST), Kach Road, near Sheikh Mohammad Bin Zayed Al Nahyan Cardiac Centre, Quetta, Balochistan (NCB) 87300, Pakistan

<sup>3</sup>School of Physics, Engineering & Computer Science, University of Hertfordshire, College Lane Campus, Hatfield, Hertfordshire AL10 9AB, UK

<sup>4</sup>Department of Pathology, Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH&RC), 7 A Khayaban-e-Firdousi, Block R3 Block R 3 M.A Johar Town, Lahore, Punjab 5400, Pakistan

<sup>5</sup>School of Computing and Mathematical Sciences, University of Leicester, University Road, Leicester LE1 7RH, UK

Corresponding author. Syed Hamad Shirazi, Department of Information Technology, Hazara University Mansehra, Dhodial, Mansehra, Khyber Pakhtunkhwa, 21120, Pakistan. Email: [syedhamad@hu.edu.pk](mailto:syedhamad@hu.edu.pk); Muhammad Yaqoob, School of Physics, Engineering & Computer Science, University of Hertfordshire, College Lane Campus, Hatfield, Hertfordshire, AL10 9AB, UK. Email: [m.yaqoob3@herts.ac.uk](mailto:m.yaqoob3@herts.ac.uk)

Citation details: Shahzad, M., Shirazi, S., Yaqoob, M. et al. AneRBC dataset: a benchmark dataset for computer-aided anemia diagnosis using RBC images. *Database* (2025) Vol. 2025: article ID baae120; DOI: <https://doi.org/10.1093/database/baae120>

## Abstract

Visual analysis of peripheral blood smear slides using medical image analysis is required to diagnose red blood cell (RBC) morphological deformities caused by anemia. The absence of a complete anaemic RBC dataset has hindered the training and testing of deep convolutional neural networks (CNNs) for computer-aided analysis of RBC morphology. We introduce a benchmark RBC image dataset named Anemic RBC (AneRBC) to overcome this problem. This dataset is divided into two versions: AneRBC-I and AneRBC-II. AneRBC-I contains 1000 microscopic images, including 500 healthy and 500 anaemic images with  $1224 \times 960$  pixel resolution, along with manually generated ground truth of each image. Each image contains approximately 1550 RBC elements, including normocytes, microcytes, macrocytes, elliptocytes, and target cells, resulting in a total of approximately 1 550 000 RBC elements. The dataset also includes each image's complete blood count and morphology reports to validate the CNN model results with clinical data. Under the supervision of a team of expert pathologists, the annotation, labeling, and ground truth for each image were generated. Due to the high resolution, each image was divided into 12 subimages with ground truth and incorporated into AneRBC-II. AneRBC-II comprises a total of 12 000 images, comprising 6000 original and 6000 anaemic RBC images. Four state-of-the-art CNN models were applied for segmentation and classification to validate the proposed dataset.

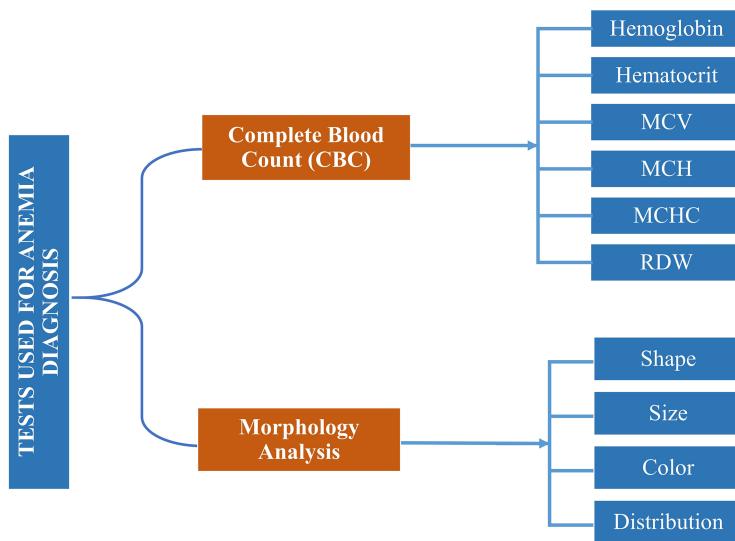
Database URL: <https://data.mendeley.com/datasets/hms3sjzt7f/1>

## Background and summary

A precise and proactive approach reduces the worst results in medical treatment procedures. Most of the worldwide disease rate goes to the chronic stage due to the nonavailability of precise and early diagnostic approaches, especially in developing countries. Anemia is the seventh leading cause of disability in women and 12.8% of maternal deaths [1]. Meanwhile, pathology tests act as bases for treating and diagnosing disease. Approximately 70% of the decisions related to the diagnosis and treatment of disease rely on the pathology test results. In this regard, different blood count and morphology tests are considered a base for diagnosing various diseases. These diseases are related to blood morphology disorders (including anemia, leukemia, polycythemia, etc.) and the immune system (including allergy, autoimmune anemia, etc.) [2]. Blood morphology disorders include several medical conditions that affect the size, shape, structure, and overall

appearance of blood cells. Anemia is the most common type of blood morphology disorder. The chief cause of anemia is the nutritional deficiency in children and infants [3]. Iron deficiency anemia is the most common type worldwide among the different kinds of anemia disease. In South Asian countries like Pakistan, Bangladesh, and India, the economic burden due to anemia is reported as \$4.5 billion annually [4, 5].

In reference [6], authors reported that around 600 million school-aged and preschool children are affected by anemia globally [7]. According to the World Health Organization (WHO) [8], 65.5% of preschool children in Southeast Asia are anemic. Based on mean corpuscular volume (MCV) and mean corpuscular hemoglobin concentration (MCHC) of complete blood count (CBC), anemia can morphologically be categorized into three groups, i.e. microcytic, normocytic, and macrocytic hypochromic anemia [9]. MCV measures the average volume or size of individual red blood cells (RBCs) in a



**Figure 1.** Clinical tests complete-blood-count (CBC) and morphology analysis are used for anemia diagnosis.

blood sample. It is calculated by dividing the total volume of RBCs by their number, while MCHC measures the average concentration of hemoglobin in a given volume of RBCs. Both are key indices used in the CBC to help diagnose and classify different types of anemia and other blood disorders. Based on morphological disorder due to anemia, RBC can be classified into five subgroups, i.e. normal, microcytes, macrocytes, elliptocytes, and target RBCs [10]. Clinical tests used for anemia diagnosis [11] based on CBC and morphology are listed in Fig. 1.

Currently, both manual and automated methods are used to evaluate these tests. Manual methods include analysis under the microscope, while Sysmex XP-300, Nihon Kohden Blood Cell Counter, and DH36 3-Part Auto Hematology Analyzer are used for automated blood cell analysis [2]. Hematology Analyzers is a medical laboratory instrument used to measure and count blood cells along with other components in blood samples. It is based on the principle of light scattering and electrical impedance.

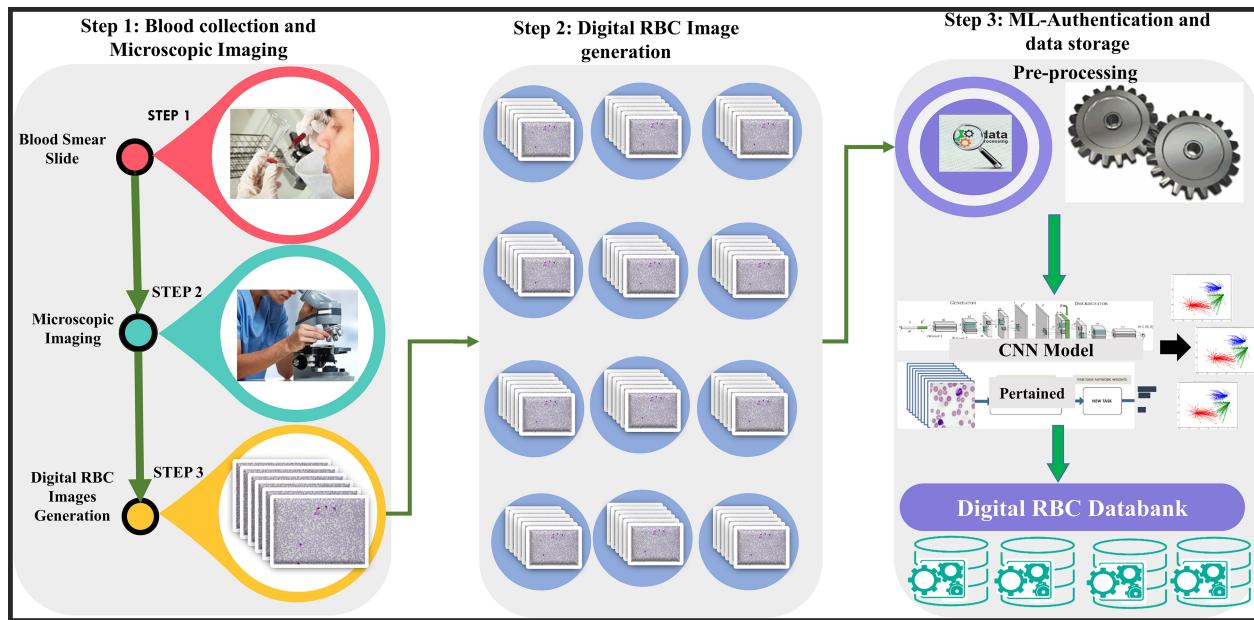
Besides the high cost and resources required for these devices, CBC counting without structural, qualitative, and morphological evaluation is another severe downside. So, after CBC through these hematology analysers, they prepare peripheral blood smear slides to evaluate structural, qualitative, and morphological disorders under the microscope. The microscopic evaluation was carried out by pathologists manually. During manual blood cell analysis, pathologists have faced several issues. Some of them are given below.

- They required outstanding expertise and rich experience for accurate microscopic blood cell analysis.
- Pathologists face challenges when identifying or predicting abnormal cell structures at the pixel level using a simple microscope [11].
- Pathologists also faced difficulties identifying the percentage change of infected cells to the normal using a simple/compound microscope.
- Approximately 5–10 min are required for the examination of one slide.

- Overlapped structures hide the density of blood elements in unit volume and morphological features (shape and size) of abnormal RBCs to normal.
- Pathologists cannot accurately differentiate between normal and abnormal shapes and sizes using a compound microscope, where RBCs show overlapped structures.
- The manual analysis required resources and dedicated concentration to examine the blood sample.
- Manual analysis can be affected due to the errable nature of humans, like tiredness, domestic problems, etc.
- Current blood analyzers perform CBCs but do not facilitate pathologists in examining the blood cell morphological structure at the pixel level.

These factors may lead to the incorrect diagnosis and treatment of the disease. However, penetration of bioimage analysis using a deep neural network and image processing techniques improves the structural, qualitative, and morphological analysis even on the pixel level [2]. Medical image analysis penetrates multiple domains, including quantitative microscopy, digital pathology, systems pathology, and computational pathology [12–17]. Image processing techniques rapidly expand bioimage analysis through technological advancement in imaging modalities regarding images' morphological, structural, texture, and pixel-level analysis. Furthermore, publicly available medical image databases and platforms allow the research community to access image datasets and annotations to develop sophisticated algorithms to analyze complex medical images [18].

Over the past few decades, several techniques and datasets have been proposed for the morphological analysis of RBCs and white blood cells (WBCs). In medical image analysis, segmenting the region of interest (ROI) is the initial rudimentary step for image processing [19]. Based on the efficient segmentation of blood cell elements like WBCs, RBCs, and platelets, numerous analyses can be carried out after the segmentation, like feature extraction, classification of blood cell elements, etc. These analyses provide valuable information during morphological analysis and disease diagnostic procedures [20]. In

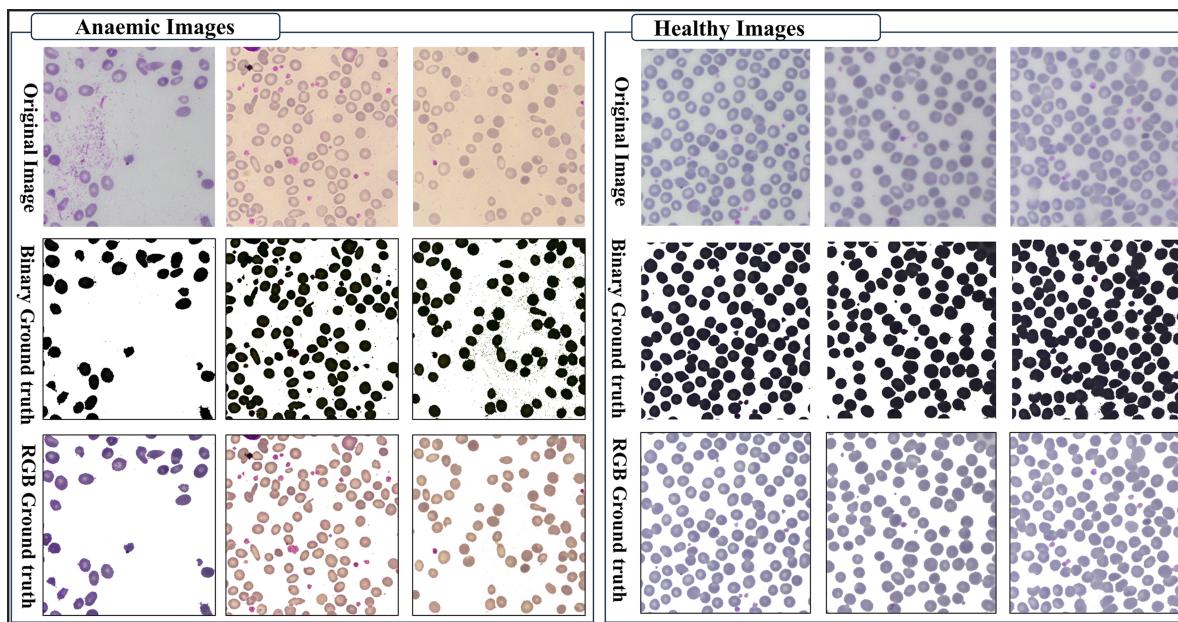


**Figure 2.** The complete schematic overview of blood collection to dataset development.

reference [21], Nikitaev *et al.* introduced a technique for segmenting leukocytes from blood and bone marrow cells. A total of 1018 cell images were formed out of 50 samples. In references [22–24], authors proposed algorithms based on marker-controlled watershed and thresholding, morphological-based techniques for systematically segmenting leukocytes and erythrocytes. Due to their close proximity, blood cells may overlap in microscopic images of blood smears. In reference [25], the authors try to overcome the problem of overlapping cells using image processing techniques and neural networks. Cao *et al.* [26] performed a deep study on the segmentation of leukocytes in peripheral blood smears. They introduced an algorithm SWAM&IVFS (stepwise averaging method) considering RGB and HIS color space. In reference [27], the authors proposed an image-processing algorithm for classifying WBCs and nuclei based on the nuclei features. In reference [28], authors classify RBCs based on their shape. They introduced an outlier tolerant machine learning (ML) technique to determine the blood's slipper-shaped and croissant-shaped cells. The deep neural network used in references [29–31] to segment MRI and CT scan images, while [32, 33] classified the WBCs and RBCs, respectively, to detect sickle cell anemia. Another study [3], explores the diagnosis of three types of anemia, i.e. (i) iron deficiency anemia (IDA), (ii)  $\alpha$ -thalassemia trait, and (iii)  $\beta$ -thalassemia trait, by using WEKA software. The author used two parameters to evaluate the data mining technique, i.e. highest accuracy and lowest mean absolute error. They combined the J48, IBK, and Naïve Bayes algorithms with vote algorithms and got accuracy scores of 96.343 and 96.2169 to diagnose iron deficiency anemia. They analyzed blood smear images of 793 individuals, containing 184 IDA patients (71 males and 118 females), 409 thalassemia patients ( $\alpha$ -thalassemia 206 and  $\beta$ -thalassemia 203), and a control group of 200 healthy individuals (100 males and 100 females). In reference [5], the authors analyzed the classification approaches, i.e. decision tree (KNN, CNN, SVM, logistic regression) and association rule mining for statistical analysis

of anemia [11]. Veterans Affairs Informatics and Computing Infrastructure database were used in reference [34] to find the anemia risk to be within a 95% confidence interval of 74.95%–75.26%. Random prediction (Rp) classification algorithms were used in reference [35] to select anemia in pregnant women.

The experimental data used to develop and authenticate all the above algorithms is microscopic blood images. The availability of state-of-the-art blood image datasets is pivotal in developing deep ML models. We need diverse state-of-the-art image data, annotations, and ground truth to train and test deep ML models to diagnose and classify different diseases. To develop a curated dataset, the generation of ground truth, annotations, and labeling of images should be carried out under the supervision of a team of expert pathologists. Developing a fine-tuned dataset for medical purposes is challenging and time-consuming, but it plays a crucial role in smartening the medical diagnostic process. Unfortunately, limited blood image databases are available for the research community to evaluate blood cell morphology. These databases are ALL-IDB-I, ALL-IDB-II [36], extended ALL-ID [37], BCCD, IUMS-IDB [38], SMC-IDB [39], BS\_DB3 [40], Ash bank, BBBC [41], peripheral blood cell image dataset [42], RBC counting automation [43], malarial-infected blood RGB images [44], and Raabin-WBC [2]. Most of these datasets are not publicly available. The publicly available datasets are composed of a small number of images with short information related to blood cell morphology. As an example, ALL-ID-I contains only 108 images with only 39 000 blood elements. All images have no masks to authenticate segmentation results. All datasets have no ground truth for the authentication of proposed techniques except extended ALL-IDB [37] and Raabin-WBC [2]. Other datasets also have no expert metadata, annotation, or labeling of images that may help to diagnose the disease accurately. The comparative analysis of all these datasets is given in Table 3.



**Figure 3.** Original images along with segmented binary, and RGB ground truth of healthy and anemic samples.

Original Image	Ground Truth	CBC Report	Doctor decision
		<p>Test,Result/Units,Norm. Range</p> <p>WBC,* 19.31 x10.e 3/<math>\mu</math>l,5 --- 15  RBC,* 3.14 x10.e 6/<math>\mu</math>l,4.3 --- 5.3  HGB,* 8.3 g/dL,9.4 --- 13  HCT,* 24 %,28.5 --- 35.5  MCV,* 72.3 L,27 --- 33  MCH,* 25 pg,28 --- 34  MCHC,* 34.6 g/dL,32 --- 36  PLT,* 487 x10.e 3/<math>\mu</math>l,150 --- 450  MPV,10.6 FL,7.2 --- 11</p>	<p>Test: RBC Morphology  Specimen: Peripheral smear  <b>RBC MORPHOLOGY / PERIPHERAL SMEAR:</b> Peripheral smear shows microcytic hypochromic picture of RBCs with anisocytosis, target cell.  <b>WBCs MORPHOLOGY:</b> WBCs increased with lymphocytosis.  <b>PLATELETS:</b> Platelets increased.</p>
		<p>Test, Result/Units, Norm. Range</p> <p>WBC, * 1.95 x10.e, 4 --- 10  RBC, * 2.1 x10.e 6/<math>\mu</math>l, 3.8 --- 4.8  HGB, * 5.9 g/dL, 12 --- 15  HCT, * 17.5 %, 36 --- 46  MCV, 83.3 FL, 76 --- 96  MCH, 28.1 pg, 27 --- 32  MCHC, 33.7 g/dL, 31.5 --- 34.5  PLT, * 106 x10.e 3/<math>\mu</math>l,150 --- 450  MPV, 10.6 FL, 7.2 --- 11</p>	<p>Test : RBC Morphology  Specimen : Peripheral smear  <b>RBC MORPHOLOGY / PERIPHERAL SMEAR:</b> Peripheral smear shows microcytic hypochromic picture of RBCs with polychromasia, and elliptocytes  <b>WBCs MORPHOLOGY:</b> WBCs decreased with neutropenia.  <b>PLATELETS:</b> Platelets reduced with no clumping.</p>

**Figure 4.** Original and RGB segmented ground truth samples of healthy and anemic images including original, ground truth, CBC report, and doctor decision in AneRBC-I dataset.

We introduced a state-of-the-art benchmark AneRBC-I and AneRBC-II datasets to overcome these challenges and automate the morphological analysis of blood cells. We target the morphological disorder of RBCs due to anemia. The Schematic overview of this study is given in Fig. 2.

The key characteristics of the AneRBC dataset are listed below:

#### A large number of images

The AneRBC-I dataset has 1000 original blood images, including 500 normal and 500 anemic images. These 1000 images were further processed, and 2000 manually generated

1000 binary and 1000 RGB segmented ground truths were obtained, as shown in Fig. 3. So, the total number of images included in the AneRBC-I repository are 3000 (1000 original, 1000 binary segment, and 1000 RGB segmented). The size of the image is  $1224 \times 960$ . Each image contains approximately 1550 RBC elements. So, the original 1000 images comprise approximately 1 550 000 RBC elements. These elements include normocytes, microcytes, macrocytes, elliptocytes, and target RBCs. The AneRBC-II dataset has a total of 36 000 images. Out of these, 12 000 are original images, including 6000 normal and 6000 anemic images. The remaining 24 000 included 12 000 binary segmented and 12 000 RGB segmented images.

## Ground truth for result authentication

The AneRBC dataset is equipped with two different types of ground truth for each image, i.e. binary segmented ground truth and RGB segmented ground truth, as shown in Fig. 4. These ground truths were generated under the supervision of a team of expert pathologists at the pathology lab of Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH&RC) in Lahore, Pakistan. These ground truths will help the research community to train and authenticate more generic CNN models. Manually generated ground truth for the ROI plays a crucial role in evaluating segmentation and classification models.

## Abnormal RBC morphology

The images in both the AneRBC-I and II datasets are divided into two subgroups: healthy and anemic images. This classification will assist the research community in training deep learning models to identify morphological disorders in RBC. Furthermore, the research community can also use this dataset to classify Anemic RBC images into microcytes, macrocytes, elliptocytes, and target cells.

## CBC and morphology reports

Another distinct characteristic of the AneRBC datasets is the availability of CBC and morphology reports of each image. Based on our understanding, the AneRBC dataset represents the first instance of a blood image dataset that includes both CBC and morphology reports for each image, enabling the automated anemia diagnosis decision to be compared against clinical test results.

## Evaluation of deep neural network models

AneRBC-I dataset is equipped with 1000 original microscopic images, including 1 550 000 healthy/anemic RBC elements, and ground truth, CBC, and morphology report of each image. Furthermore, AneRBC-II comprises 12 000 original RBC images. This data repository enables the research community to evaluate the performance of deep neural networks for the precise detection of morphological disorders in blood cells.

## Diversity of lighting effects for generalized ML models

All the images were collected using a single microscope but with different lighting effects. Images with heterogeneous background lighting effects help to evaluate the performance of generalized ML models.

## Methods

The methodology adopted for the development of this dataset comprises on two phases. (i) Patient selection and peripheral blood smear collection at SKMCH&RC lab, and (ii) manual slide analysis and image capturing with heterogeneous lighting intensity. The research conducted in this study follows the ethical guidelines and has received approval from the Internal Review Board (IRB) of Shaukat Khanum Hospital, Lahore, Pakistan, and Hazara University Mansehra. The IRB carefully reviewed and endorsed the research methodology, ensuring that it fulfills ethical standards and protects the rights and well-being of the participants. All participants involved in this

**Table 1.** Patient population description of both anemic and nonanemic

Category	Total patients	Male	Female	Age range
Anemic	500	250	250	
Nonanemic	500	250	250	
Total	1000	500	500	30–50

study provided informed consent before participating, and they were explicitly informed about the nature of the research, its objectives, and the intended use of their data. Furthermore, participants were guaranteed that their identities, like name, affiliation, age, gender, etc., would remain confidential, and steps were taken to protect their privacy throughout the study. Particularly, each participant agreed to the open publication of the data, understanding that it would contribute to advancing knowledge in the respective field.

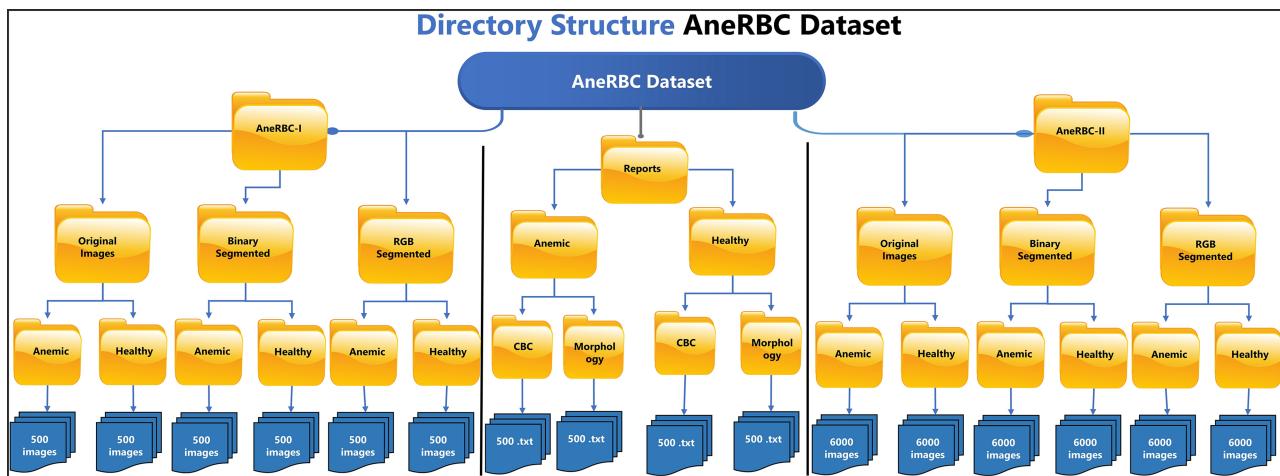
## Patient's selection and peripheral blood smear collection

After obtaining approval from the ethical committee of Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH&RC) Lahore, Pakistan, 1000 participants were selected from the pathology lab at SKMCH&RC to provide peripheral blood smears for the study. Out of which, 500 were healthy, and 500 were anemic participants. To reduce gender inequality, we have selected 50% male and 50% female from each group, i.e. healthy and anemic. The age range for both males and females was 30–50 years. The summary of the population is given in Table 1. Peripheral blood smear collection and slide preparation were done under the supervision of a team of expert pathologists from the SKMCH&RC laboratory. The volume of the blood that was taken from each patient was 3.0 ml. If sample transportation is required, 24:00 h were given with 2°C–8°C temperature. The blood sample was preserved in a Lavender cap EDTA tube.

The sample was received at the laboratory, and clots were checked before further processing. After that, these samples were processed on a Sysmex xn9000 analyzer to generate CBC and morphology parameters. Slides of each sample were prepared with Wright-Giemsa stain for microscopic analysis. Olympus Dp27 8.9-megapixel CMOS sensor camera captured the RBC images with a 4M resolution. Sample images from healthy and anemic blood elements are shown in Fig. 3. The binary ground truth represents a black-and-white segmented version, where the foreground (RBC elements) and the background are distinctly marked. The RGB ground truth includes color segmented images with only foreground (i.e. RBC elements) without the background. Both the binary and RGB ground truths were generated manually using segmentation software under the supervision of expert pathologists from Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH&RC). The binary ground truth was created by selecting the RBC regions with a high-precision tool, while the RGB ground truth used color coding for detailed cell morphology classification.

## Manual slide analysis and image capturing with heterogeneous lighting intensity

After preparing blood smear slides, pathologists from SKMCH&RC lab examined the slides under an Olympus



**Figure 5.** The directory structure of the AneRBC Dataset.

Dp27 8.9-megapixel CMOS microscope. The study was supervised by a team of two senior consultant pathologists and three assistant consultant pathologists. The senior consultants brought extensive experience and knowledge in the field of pathology, particularly in the interpretation of blood cell morphology and the diagnosis of hematological disorders. Their supervision ensured the quality and accuracy of the pathological assessments conducted as part of the study. Additionally, the assistant consultant pathologists, while perhaps less experienced than their senior counterparts, provided valuable support in data collection, analysis, and review. Together, the combined expertise of the pathologist team ensured rigorous and thorough examination of the blood cell images, contributing to the credibility and reliability of the study outcomes.

The pathologists examined all the slides and gave appropriate interpretations of morphological disorders in the cells. The pathologist's interpretation was also stored in the database with morphology reports, as shown in Fig. 4. So, in this way, we have generated 1000 CBC and morphology reports of all images and pathologists' interpretations for ML Models performance evaluation. After that, images were captured with  $1224 \times 960$  pixels resolution and heterogeneous lighting intensity to develop a more diverse blood image dataset. All images were stored on storage media in PNG format.

### Ground truth generation

The AneRBC dataset includes two distinct types of ground truth for each image: binary segmented ground truth and RGB segmented ground truth, as shown in Fig. 4. Initially, these ground truths were generated using Photoshop. We used the magic tool for the selection of the background with a 20% tolerance. The manually generated ground truth was authenticated under the supervision of a team of expert pathologists at the pathology lab of Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH&RC) in Lahore, Pakistan. These annotations serve as valuable references for the research community for the training and validation of deep learning CNN models. Manually generated ground truth for the ROI plays a crucial role in evaluating segmentation and classification models. The detail of both Binary and RGB ground truth is given below.

### Binary ground truth

Binary ground truth annotations are typically used to denote the presence or absence of specific features within an image. In the case of the AneRBC dataset, binary annotations likely indicate the boundaries and presence of RBCs against the background. These annotations are usually in black and white (binary), where white (or one) represents the background and black (or zero) represents the ROI. Binary annotations are often created using image segmentation tools that allow pathologists or trained technicians to outline the cell boundaries manually or semi-automatically. For this purpose, we have used Photoshop software with magic tool for the selection of the background with a 20% tolerance. This method ensures precise delineation of cells, which is crucial for segmentation tasks in ML.

### RGB ground truth

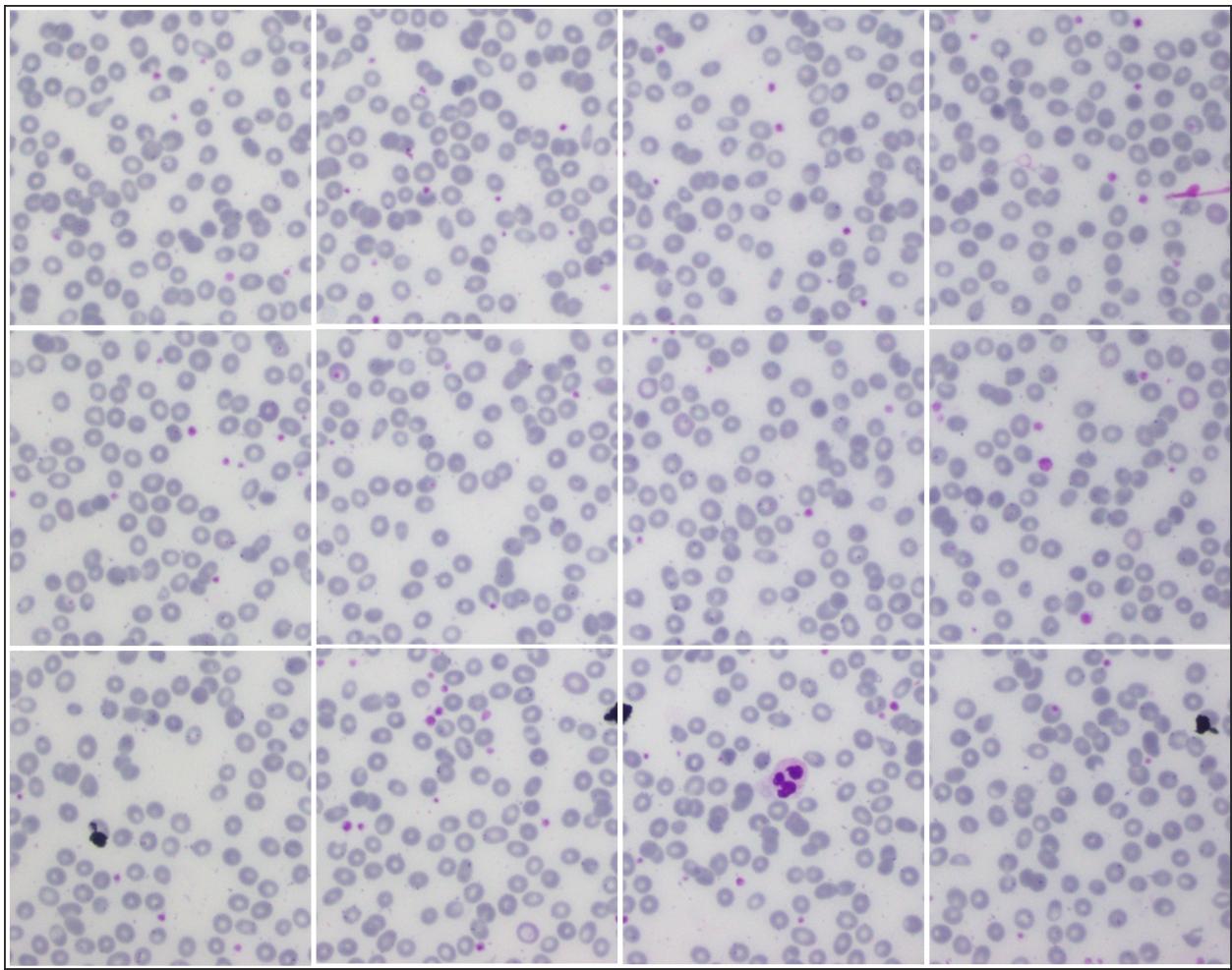
RGB ground truth involves more detailed annotations that include color information. This type of annotation is used to differentiate various types of cells or cell states based on additional attributes like color and texture. In the context of anemia, different colors in the RGB annotations might represent various types of RBCs such as normocytes, microcytes, macrocytes, elliptocytes, and target cells.

### Dataset description

The AneRBC dataset is divided into two distinct versions based on image resolution and type: (i) AneRBC-i and (ii) AneRBC-ii. The director structure and the number of images in each directory of the AneRBC dataset are shown in Fig. 5.

### AneRBC-I

AneRBC-I consists of 1000 original images (500 healthy and 500 anemic images) with a resolution of  $1224 \times 960$  pixels. These 1000 original images are further processed, and 2000 images are generated with 1000 binary and 1000 RGB ground truths, as shown in Figs 3 and 4, respectively. The CBC and morphology reports of all 1000 images with the pathologist's interpretation are also included in the database. Resultantly, 3000 images (1000 original, 1000 binary, and 1000 RGB ground-truth) were stored in the AneRBC-I data repository



**Figure 6.** Samples images from AneRBC-II dataset, where  $1224 \times 960$  is divided into 12 subimages with size  $306 \times 320$ .

**Table 2.** Detailed description of both versions of the AneRBC datasets

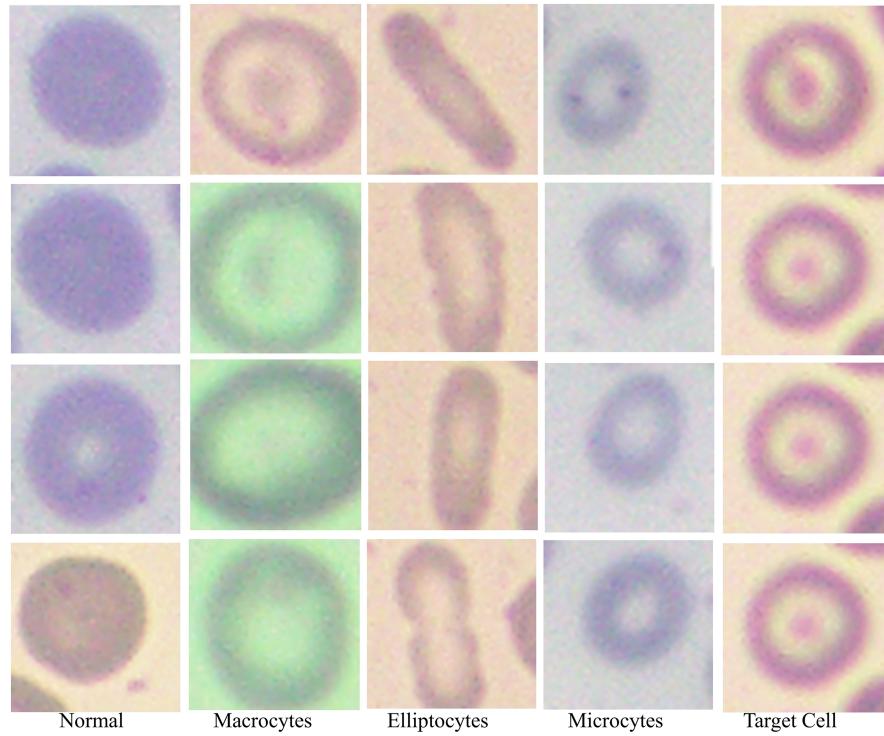
Type	Image acquisition setup	
Camera	Olympus Dp27 8.9-megapixel CMOS	
Magnification of microscope	40×	
Image format	PNG	
Color Depth	3 (RGB)	
Dataset	AneRBC-I	AneRBC-II
Total no. of images	3000	36 000
Original	1000	12 000
Binary ground truth	1000	12 000
RGB ground truth	1000	12 000
Resolution	$1224 \times 960$	$306 \times 320$
Image type	50% anemic and 50% healthy	50% anemic and 50% healthy
Ground truth	Yes	Yes
CBC reports	Yes	Yes
Morphology and doctor decision reports	Yes	Yes

along with CBC, morphology, and doctor decision report of each image.

### AneRBC-II

Due to high resolution, most ML Models cannot process these images in the current size, i.e.  $1224 \times 960$ . So, each image

is cropped and divided into 12 images with a resolution of  $306 \times 320$  pixels. As a result, AneRBC-II comprises 12 000 original images (6000 healthy and 6000 anemic images) along with 12 000 binary and 12 000 RGB ground truths. Resultantly, AneRBC-II stored a total of 36 000 (12 000 original, 12 000 binary, and 12 000 RGB ground truth) images, as



**Figure 7.** Red blood cells in the anemic images can be classified as normal, macrocytes, elliptocytes, microcytes, and target cells.

shown in Fig. 6. A short description of both datasets is given in Table 2.

### Image name representation

The AneRBC-I images are named with annotations xxx\_a.png and xxx\_h.png. Here xxx is a three-digit integer counter, “a” represents images with anemic RBC elements, and “h” represents healthy elements. The naming annotation for AneRBC-II is xxxx\_yy\_a.png and xxxx\_yy\_h.png. The AneRBC-II dataset is the subdivision of AneRBC-I. Each image in AneRBC-I is divided into 12 subimages. So, here “xxxx” represents the 4-digit counter for the total number of images, while “yy” represents the counter for each image’s subdivision. The naming representation of relevant manually segmented images is also the same, but the segmented images are placed in a separate directory named “segmented.”

### Blood cell element types

Based on RBC morphology, the AneRBC dataset is divided into healthy RBC images and anemic RBC images. The anemic RBC images can be further categorized into five classes, i.e. normal, microcytes, macrocytes, elliptocytes, and target RBCs, as shown in Fig. 7. Each image in the AneRBC-I dataset contains approximately 1550 RBC elements, including five classes. So, our dataset comprises approximately 1 550 000 RBC elements. Each image in the dataset, either original or ground truth, is provided in .PNG format in AneRBC-I and II.

### Dataset comparisons

The comparative analysis of proposed datasets with previously developed blood cell datasets like ALL-IDB-I, ALL-IDB-II [36], extended ALL-ID [37], BCCD, IUMS-IDB [38],

SMC-IDB [39], BS\_DB3 [40], Ash bank, BBBC [41] is given in Table 3.

### Technical validation

CNNs remain a preferred choice due to their abilities to capture spatial hierarchies in image data. We have conducted experiments to authenticate the proposed dataset by segmenting and classifying the RBC elements and then validating the results against the ground truth. To set the baseline for deep learning-based instances and semantic segmentation, we compared the segmented ROI of the RBC element with manually generated ground truth. Furthermore, three types of technical validation were carried out on the proposed dataset: clinical validation, segmentation validation, and classification validation. Alternative ML approaches such as support vector machines (SVMs) and decision trees offer several advantages in certain contexts, but they generally lack the spatial hierarchies in image data that CNNs effectively capture. Therefore, tasks involving complex image analysis like medical image segmentation, CNNs remain preferred choice due to automatically detect patterns and features directly from image pixels.

### Clinical validation

Initially, the proposed dataset was clinically verified by the consultant pathologists. The ground truth of each healthy and anaemic image was generated during this verification under the pathologist’s supervision. In parallel, each image’s morphology and decision report were also added to this dataset, which was generated using the CBC report by the consulted pathologists.

**Table 3.** Comparative analysis of proposed AneRBC-I, and AneRBC-II datasets with previously developed blood cell datasets

Dataset	No. of images	Blood elements	Resolution	Ground truth	CBC reports	Morphology reports
ALL-IDB-I [36]	109	WBC & RBC	2592 × 1944	✗	✗	✗
ALL-IDB-II [36]	260	WBC	257 × 257	✗	✗	✗
SMC-IDB [39]	367	WBC	Nil	✗	✗	✗
IUMS-IDB [38]	196	WBC	Nil	✗	✗	✗
Malarial dataset	848	RBC	Nil	✗	✗	✗
AneRBC-I Dataset	1000	RBC	1224 × 960	1000: binary 1000: RGB	✓	✓
AneRBC-II Dataset	12 000	RBC	306 × 320	12 000: binary 12 000: RGB	✓	✓

**Table 4.** The loss, accuracy, Dice Coefficient, IoU, precision, recall, and specificity for UNet, LinkNet, and AttenUNet models, demonstrating the technical effectiveness of AneRBC-I dataset for the medical image segmentation task

Model	Loss	Accuracy	Dice coefficient	IoU	Precision	Recall	Specificity
Unet	0.2503	0.9174	0.9414	0.8916	0.9332	0.9448	0.8982
LinkNet	0.2018	0.9426	0.9497	0.9083	0.9474	0.9557	0.9232
Atten_Unet	0.2719	0.9780	0.9829	0.9665	0.9833	0.9825	0.9700

**Table 5.** Loss, accuracy, Dice Coefficient, IoU, precision, recall, and specificity for each CNN model. The results validate the AneRBC-II dataset's robustness and applicability for developing and testing advanced segmentation models in the medical image processing domain

Model	Loss	Accuracy	Dice coefficient	IoU	Precision	Recall	Specificity
Unet	0.1358	0.9585	0.9693	0.9407	0.9698	0.9689	0.9368
Linknet	0.1851	0.9447	0.9535	0.9129	0.9462	0.9629	0.9221
Atten_Unet	0.1660	0.9812	0.9743	0.9634	0.9415	0.9264	0.8924

## Segmentation validation

The technical validation of the AneRBC-I and AneRBC-II datasets is critical in establishing their utility for the research community, particularly for segmentation tasks in medical imaging. The application of standard convolutional neural network (CNN) models, including UNet, LinkNet, and AttenUNet, along with a generative adversarial network (GAN), provides a robust framework for evaluating the datasets' quality and the feasibility of their use in developing advanced segmentation algorithms. We have applied four state-of-the-art CNN models to validate both datasets concerning segmentation, including GAN, UNet, LinkNet, and AttenUNet models. A detailed description of the results obtained from each model is given below.

### UNet model performance

#### *AneRBC-I*

The standard UNet model demonstrated an impactful performance on the AneRBC-I dataset, with a Dice Coefficient of 0.9414 and an intersection over union (IoU) of 0.8916, indicating a high degree of overlap between the predicted segmentation and the ground truth. The accuracy of 0.9174 is commendable; however, it is not the highest among the tested models. The model's precision and recall are closely matched (0.9332 and 0.9448, respectively), suggesting a balanced segmentation capability. The specificity of 0.8982 indicates that UNet identifies true negatives effectively using the proposed dataset, which is crucial for medical image segmentation tasks.

#### *AneRBC-II*

On the AneRBC-II dataset, the UNet model's performance improved across all metrics, with a prominent increase in the Dice Coefficient to 0.9693 and IoU to 0.9407. The accuracy also raised to 0.9585. This evaluation metric enhancement could be attributed to the diverse dataset's characteristics to generalize the given data better. The precision and recall remained high, with values above 0.96, reinforcing the model's reliability in segmenting the AneRBC-II dataset.

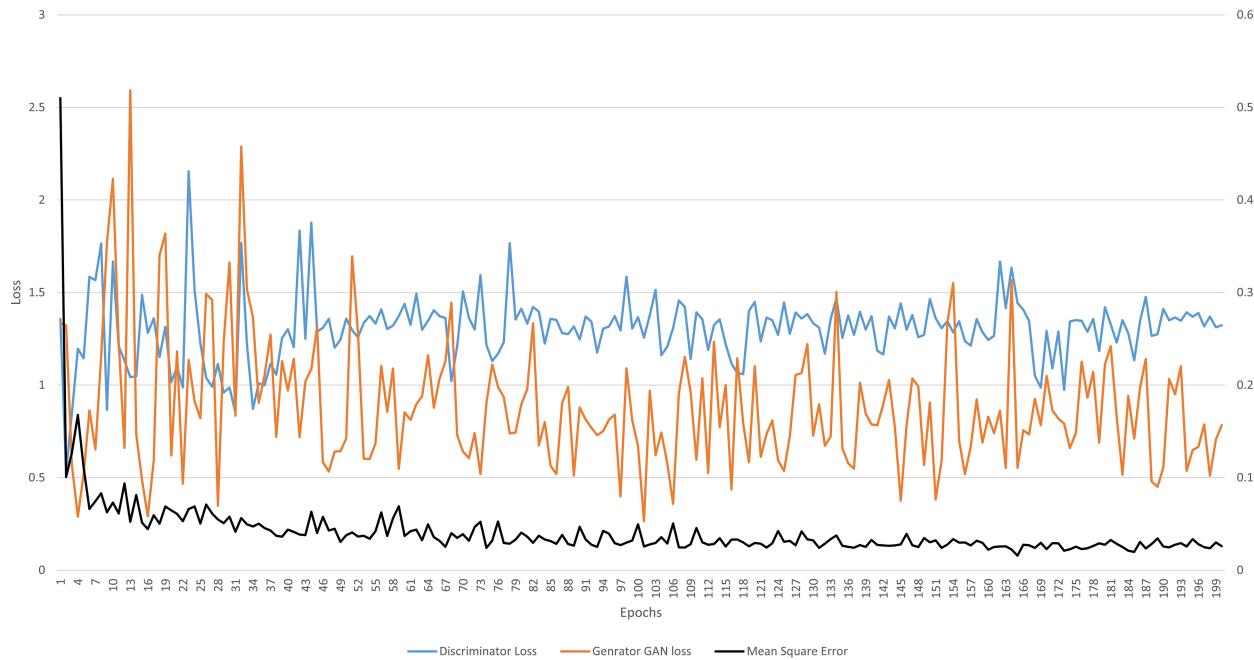
### LinkNet model performance

#### *AneRBC-I*

LinkNet outperformed UNet in the AneRBC-I dataset, achieving a higher accuracy of 0.9426 and the best Dice Coefficient (0.9497) and IoU (0.9083) among the models. The precision and recall were the highest at 0.9474 and 0.9557, respectively, indicating a superior segmentation quality with the AneRBC-I dataset. The specificity of 0.9232 was also the highest, showing that LinkNet was most adept at correctly identifying negative cases.

#### *AneRBC-II*

In the AneRBC-II dataset, the performance of LinkNet slightly increased as compared to its results on AneRBC-I for all metrics except precision and specificity as shown in Tables 4 and 5. The accuracy increased slightly from 0.9426 to 0.9447, Dice Coefficient from 0.9497 to 0.9535, IoU from 0.9083 to 0.9129, and recall from 0.9557 to 0.9629. The model maintained a high Dice Coefficient of 0.9535 and an IoU of 0.9129, but both UNet and AttenUNet outperformed it in most metrics.



**Figure 8.** Graph of discriminator loss, generator loss, and mean square error.

### AttenUNet model performance

#### AneRBC-I

AttenUNet showed better results on the AneRBC-I dataset, with the highest accuracy of 0.9780 and specificity of 0.9700 among the three models. The Dice Coefficient and IoU were outstanding at 0.9829 and 0.9665, respectively. These results suggest that the attention mechanisms integrated into AttenUNet boost the identification of relevant features for segmentation.

#### AneRBC-II

Interestingly, while still excellent, AttenUNet's performance on the AneRBC-II dataset was not as dominant as on the AneRBC-I dataset. The accuracy remained high at 0.9812, but there was a decrease in the Dice Coefficient to 0.9743 and IoU to 0.9634. The precision and recall show a little decline, with recall dropping to 0.9264, which was the lowest among the models for this dataset.

### Conditional Generative Adversarial Network

The Conditional Generative Adversarial Network (cGAN) network comprises a generator and a discriminator. The generator generates a predicted image against each input image, while the discriminator discriminates between the predicted image and ground truth. Unlike traditional GANs that generate images purely from random noise inputs, the cGAN used in our research takes an actual image as input and produces a segmented image as output, aligning closely with the goals of medical image segmentation. The results related to the discriminator, generator losses, and mean square error are shown in Fig. 8. This figure shows that after 15 epochs, the predicted image does not look realistic. An epoch is one complete pass through the entire training set by the learning algorithm. The predicted image looks more natural as epochs increase from 10 to 50. As the epochs increased upto 80, we got

optimized results. This phenomenon indicates that the proposed AneRBC-II dataset is technically valid for the research community. The results of the GAN network were evaluated using the different loss functions. The visual representation of segmentation is shown in Fig. 9.

The reason for GAN network is used for the segmentation task of the proposed datasets is twofold: (i). Using GANs alongside traditional UNet architectures for segmentation allows for the evaluation of the dataset's utility; (ii) Successful segmentation with both UNet and GAN architectures acknowledges the credibility and applicability of the dataset for a broader range of image analysis tasks.

The generator loss determines how accurately the ROI is segmented in the predicted image near the ground truth. Generator loss can be determined using the following formula.

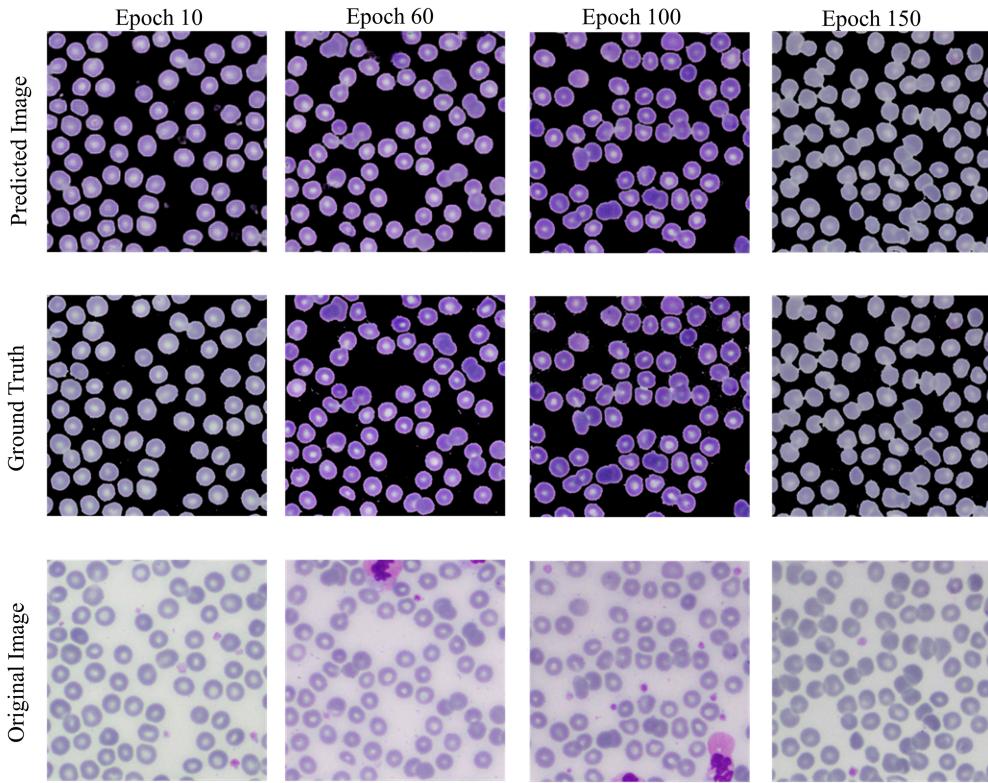
$$G_L = E_x [\log(D(x))] + E_z [\log(1 - D(G(z)))] \quad 1$$

where  $G$  and  $D$  represent the generator and discriminator function of GAN, respectively.  $G(z)$  and  $D(x)$  are the output of generator and discriminator probability estimation, respectively. While  $[\log(D(x))]$  is the output of discriminator on real data. The generator minimizes its loss in this function to generate a more realistic image near the ground truth. Figure 8 shows that the highest generator loss during execution was 2.155 at epoch no 23, while it reduced to 1.322 at epoch no 200.

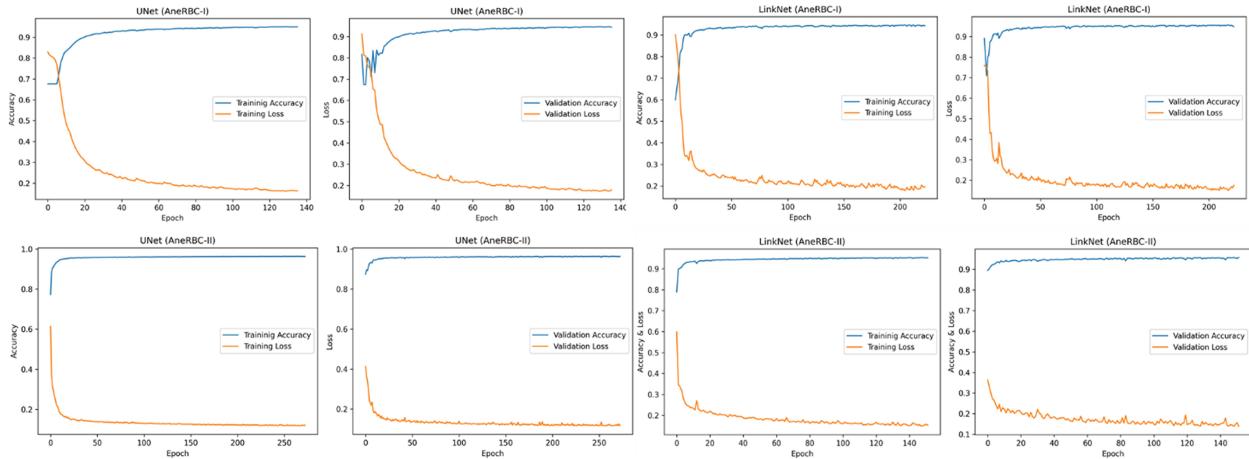
The discriminator loss maximizes the prediction probability related to the generated image as fake. Mathematically, it is represented as:

$$D_L = \text{maximize} \log D(x) + \log(1 - D(G(z))) \quad 2$$

where  $D(G(z))$  represents the discriminator's output when evaluating fake data. The term  $\log(D(x))$  shows the log of



**Figure 9.** Visual results of image segmentation using the GAN network at different epochs.



**Figure 10.** Training and validation graph for Unet and LinkNet models.

discriminator identification. Whereas the term  $\log(1-D(G(z)))$  expresses the discriminator correctly identifying generator-produced data as fake. Figure 8 shows that the highest discriminator loss during execution was 2.592 at epoch no 13, while it reduced to 0.763 at epoch no 198.

The mean squared error (MSE) between generated image ( $g$ ) and ground truth ( $g_0$ ) is calculated as given below.

$$MSE = \frac{1}{n} \sum_{i=1}^n (g_0 - g)^2 \quad 3$$

where the term  $\sum_{i=1}^n (g_0 - g)^2$  shows the squared differences from

$i = 1$  to  $n$ . MSE measured for the AneRBC dataset is shown in Fig. 8. The maximum MSE of 0.510 was calculated at the 1st epoch and reduced to 0.0158 at epoch no 165.

The high-performance metrics achieved by the models across both datasets emphasize their robustness and the technical soundness of the data. The UNet model's improved performance on the AneRBC-II dataset, as compared to the AneRBC-I dataset, as shown in Tables 4 and 5, suggests that the datasets are diverse and can help to train models that generalize well. The graphical visualization regarding the segmentation validation of all datasets is shown in Fig. 10. The other graphical results related to segmentation are given in Appendix (Figs A1(a–c)).

**Table 6.** Classification results of AneRBC-I Datasets on key metrics, including training accuracy, training loss, validation accuracy, validation loss, test accuracy, precision, recall, and F1 score for AneRBC-I dataset

	Matrices	MobileNetV2	Resnet152V2	VGG16	InceptionV3
Without transfer learning	Training accuracy	0.98	0.98	0.98	0.97
	Training loss	0.05	0.05	0.05	0.10
	Validation accuracy	0.89	0.91	0.94	0.90
	Validation loss	0.32	0.47	0.35	0.31
	Test accuracy	0.83	0.77	0.91	0.89
	Precision	0.95	0.70	0.91	0.77
	Recall	0.68	0.95	0.96	0.84
	F1 score	0.78	0.86	0.93	0.80
With transfer learning	Training accuracy	0.94	0.71	0.88	0.85
	Training loss	0.18	5.82	0.28	0.51
	Validation accuracy	0.84	0.70	0.88	0.78
	Validation loss	0.40	3.65	0.78	0.45
	Test accuracy	0.71	0.70	0.86	0.45
	Precision	0.73	0.69	0.88	0.51
	Recall	0.52	0.86	0.78	0.56
	F1 score	0.60	0.77	0.92	0.58

**Table 7.** Classification results of AneRBC-II Datasets on key metrics, including training accuracy, training loss, validation accuracy, validation loss, test accuracy, precision, recall, and F1 score for AneRBC-II dataset

	Matrices	MobileNetV2	Resnet152V2	VGG16	InceptionV3
Without transfer learning	Training accuracy	0.98	0.97	0.97	0.98
	Training loss	0.04	0.52	0.05	0.03
	Validation accuracy	0.95	0.92	0.95	0.97
	Validation loss	0.23	0.53	0.34	0.55
	Test accuracy	0.88	0.82	0.94	0.89
	Precision	0.92	0.86	0.95	0.95
	Recall	0.93	0.93	0.95	0.98
	F1 score	0.90	0.96	0.95	0.91
With transfer learning	Training accuracy	0.89	0.73	0.86	0.73
	Training loss	0.31	8.68	0.42	0.70
	Validation accuracy	0.86	0.67	0.78	0.73
	Validation loss	0.34	4.11	0.41	0.66
	Test accuracy	0.84	0.62	0.76	0.70
	Precision	0.84	0.85	0.84	0.75
	Recall	0.86	0.91	0.94	0.94
	F1 score	0.87	0.87	0.86	0.86

## Classification validation

Four state-of-the-art CNN models were applied to the AneRBC datasets to validate the classification problem. We have performed classification validation on the original anemic and healthy images. We used the original image for training and testing during this classification without providing manually generated ground truth. The classification results for the AneRBC-I and AneRBC-II datasets, as obtained from the MobileNetV2, ResNet152v2 [45], VGG16, and Inceptionv3 [46] models, provide an understanding of the datasets' utility for binary classification tasks in medical imaging. The models were executed twice, once without transfer learning to establish a baseline and then with transfer learning to check the impact of pretrained weights on performance. The detailed discussion of classification results on both datasets is given below.

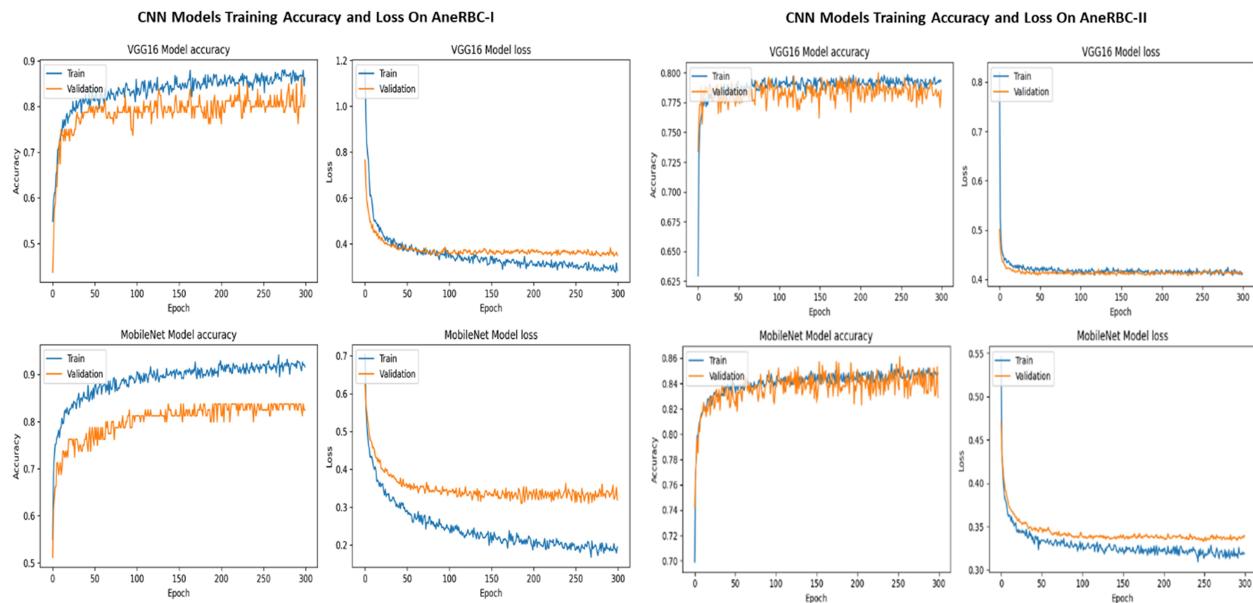
### AneRBC-I dataset classification

During classification, we explored the performance of four prominent CNNs, i.e. MobileNetV2, Resnet152V2, VGG16, and InceptionV3 on the AneRBC-I dataset. We have performed analysis on the AneRBC-I dataset both with and

without the application of transfer learning. During without transfer learning experiment, all models outperformed in their training phases and achieved accuracies close to 98%. The higher training accuracies show robust feature-learning capabilities from the dataset. Among them, VGG16 outperformed with the highest validation accuracy of 94%. It also obtained a balanced precision, recall and F1 score which indicate its reliable performance across various metrics. When models were trained with transfer learning parameters, we observed a general dip in training accuracies, with Resnet152V2 achieving up to 71% accuracy from its previous high. This model also got significant increase in training and validation losses, suggesting difficulties in adapting the pretrained weights to our specific medical dataset. Meanwhile, VGG16 continued to maintain its performance with test accuracy at 86% and an F1 score of 0.92. The detail summary of all CNN model results is given in Table 6.

### AneRBC-II dataset classification

For the AneRBC-II dataset, same models were executed with and without the application of transfer learning. All models show high training accuracies, with InceptionV3 and



**Figure 11.** Training and validation accuracy and loss for both datasets regarding the classification of VGG16 and MobileNet.

MobileNetV2 leading at 0.98. MobileNetV2 also showed the lowest training loss (0.04), indicating a very efficient learning process. In contrast, Resnet152V2 shows higher loss at 0.52. The validation accuracies were improved across the models, particularly for InceptionV3, which peaked at 0.97. In test scenarios, VGG16 achieved highest test accuracy at 0.94 and an F1 score of 0.95 which aligns closely with its precision and recall. Despite a lower test accuracy, Resnet152V2 managed an F1 score of 0.96, showing its ability to balance the precision and recall on AneRBC-II dataset. When applying transfer learning to the AneRBC-II dataset, all models show decreased training accuracies and high loss rates. VGG16 still perform better and achieved test accuracy of 0.76, showing its capability to utilize features learned in different contexts. The recall rates for models like InceptionV3 and VGG16 remained high at 0.94, showing their effectiveness in identifying true positives. The detail summary of classification results on AneRBC-II dataset is given in Table 7.

The consistent improvement in model performance across both datasets provides technical proof of their validity. The high precision, recall, and F1 scores achieved without transfer learning, particularly with VGG16 and InceptionV3 on the AneRBC-I dataset and InceptionV3 on the AneRBC-II dataset, demonstrate the datasets' suitability for training classification models in the medical image domain. The fact that all models performed better without transfer learning also suggests that the datasets contain features representative of the broader medical imaging domain. The visual results for training and testing are shown in Fig. 11. Intermediate results on segmentation and classification are given in the Appendix (see Figs A1(a–c), and A2(a–c)).

## Conclusion

This study introduced the AneRBC dataset, a benchmark dataset for computer-aided anemia diagnosis using RBC images. The dataset is divided into two versions, i.e. AneRBC-I and AneRBC-II. Each contains high-resolution images with

corresponding binary and RGB ground truths generated under the supervision of an expert hematologist. To validate both datasets, we have applied state-of-the-art CNN models such as UNet, LinkNet, and AttenUNet, and a cGAN architecture for segmentation. The proposed datasets were also validated for classification using MobileNetv2, ResNet152V2, VGG16, and InceptionV3. Our results demonstrate that the proposed AneRBC datasets have the potential to promote the development of automated anemia diagnostic tools by providing robust image data for training and evaluating deep learning models. Future work will focus on further annotating individual RBC classes to expand the utility of the datasets for specialized diagnostic applications.

## Acknowledgements

I would like to express my sincere gratitude and appreciation to Shaukat Khanum Hospital, Lahore, Pakistan, for their invaluable contribution in providing the RBC image dataset for this research. The dataset provided by Shaukat Khanum Hospital will be instrumental in enabling the future analysis and diagnosis of anemia. Furthermore, I would like to acknowledge the Higher Education Commission (HEC) Pakistan for their support through the National Research Program for Universities (NRPU) project (ID 16017).

## Conflict of interest

None declared.

## Funding

None declared.

## Data availability

The complete image data of this dataset is available at

<https://data.mendeley.com/datasets/hms3sjzt7f/1>

The code used for segmentation and classification is available at

[https://github.com/shahzadmscs/AneRBC\\_Segmentation\\_Classification\\_code](https://github.com/shahzadmscs/AneRBC_Segmentation_Classification_code).

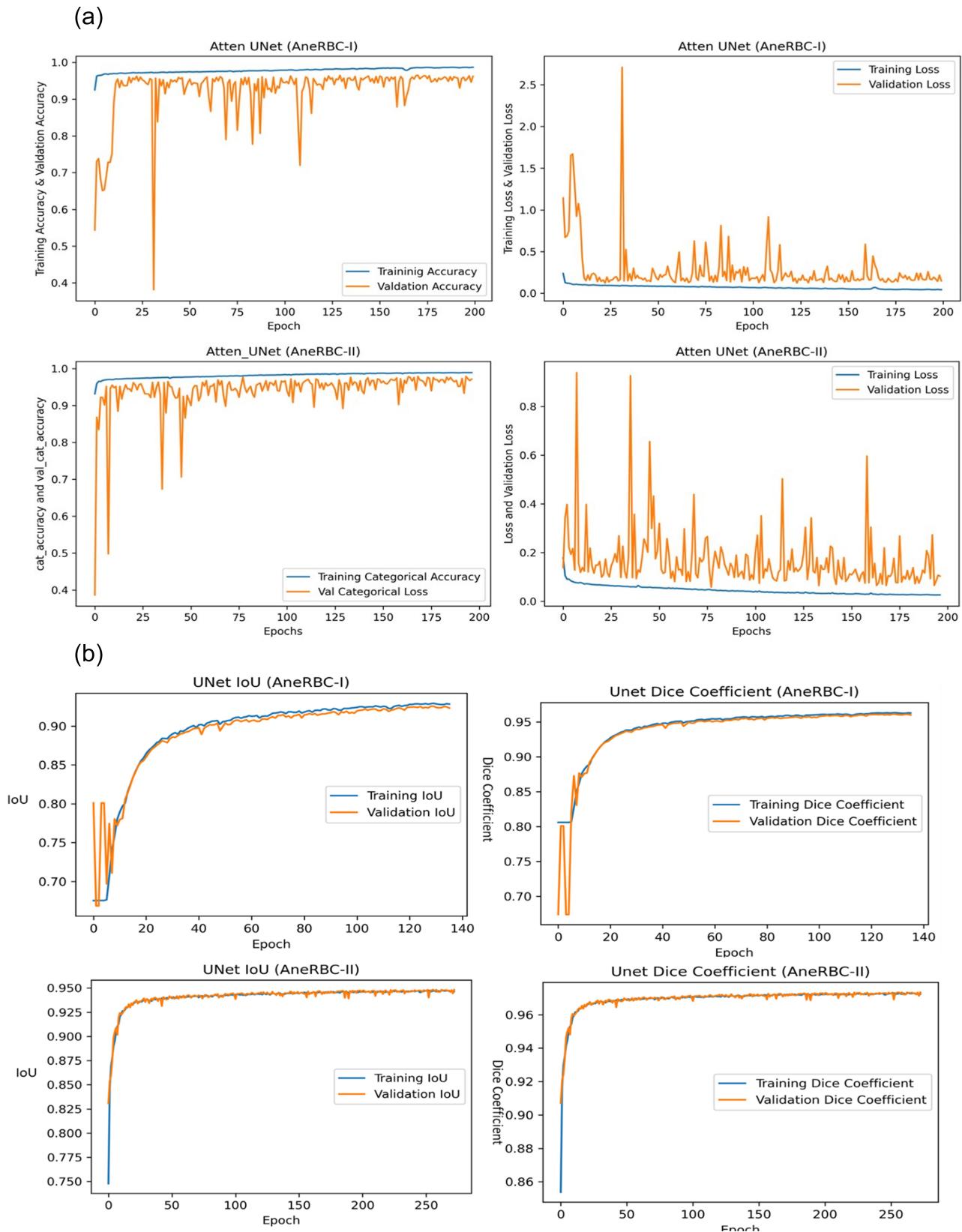
The code for uploading image data along with CBC and morphology reports for the training of the CNN model is

available at

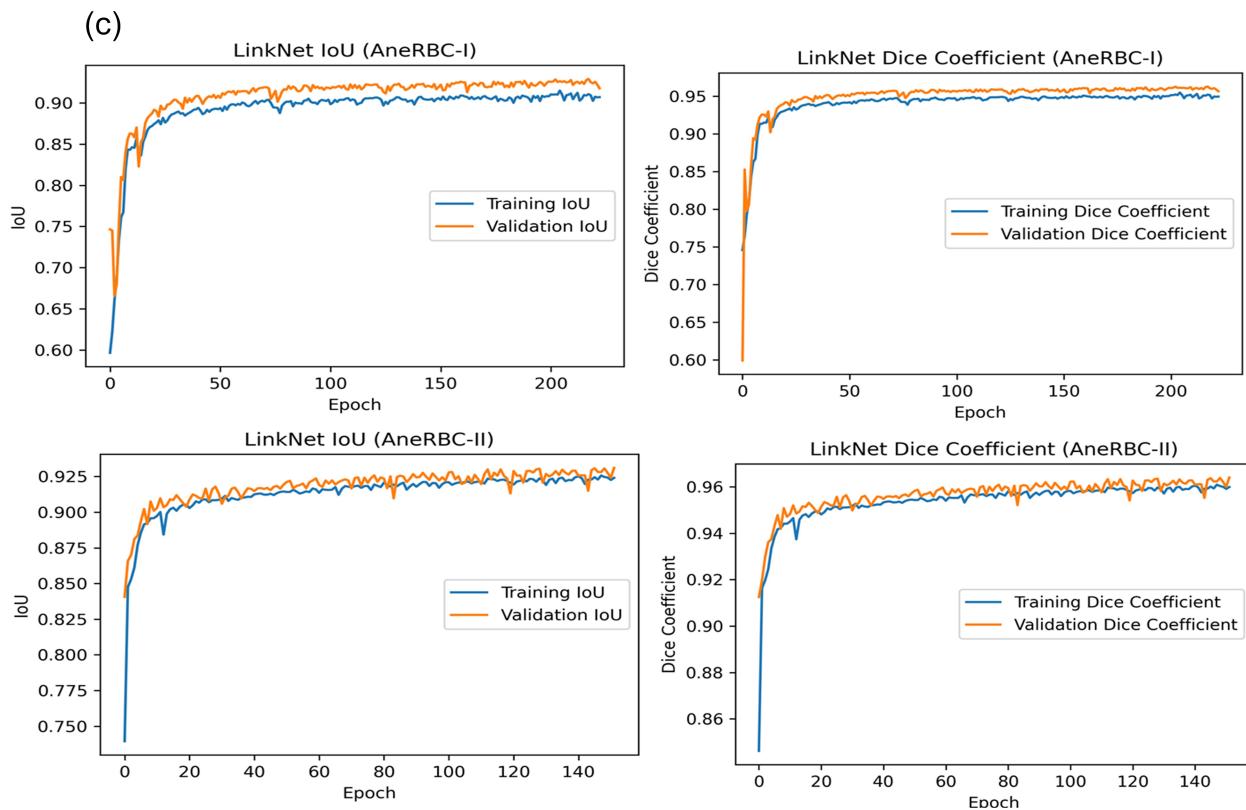
[https://github.com/shahzadmscs/AneRBC\\_Segmentation\\_Classification\\_code/blob/main/upload\\_data\\_for\\_training.ipynb](https://github.com/shahzadmscs/AneRBC_Segmentation_Classification_code/blob/main/upload_data_for_training.ipynb)

## Appendix

### Segmentation results of CNN models

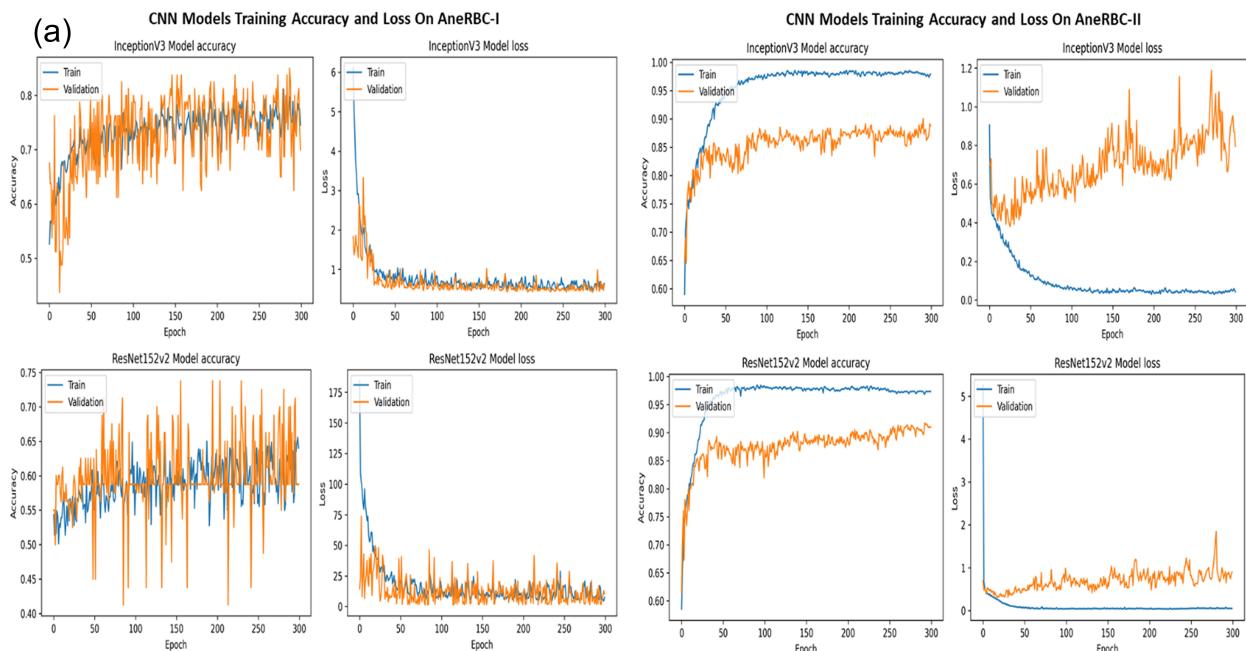


**Figure A1.** (a) Training progress of Atten\_UNet on AneRBC-I and AneRBC-II. (b) IoU and Dice-Coefficient of UNet on AneRBC-I and AneRBC-II. (c) IoU and Dice-Coefficient of LinkNet on AneRBC-I and AneRBC-II.

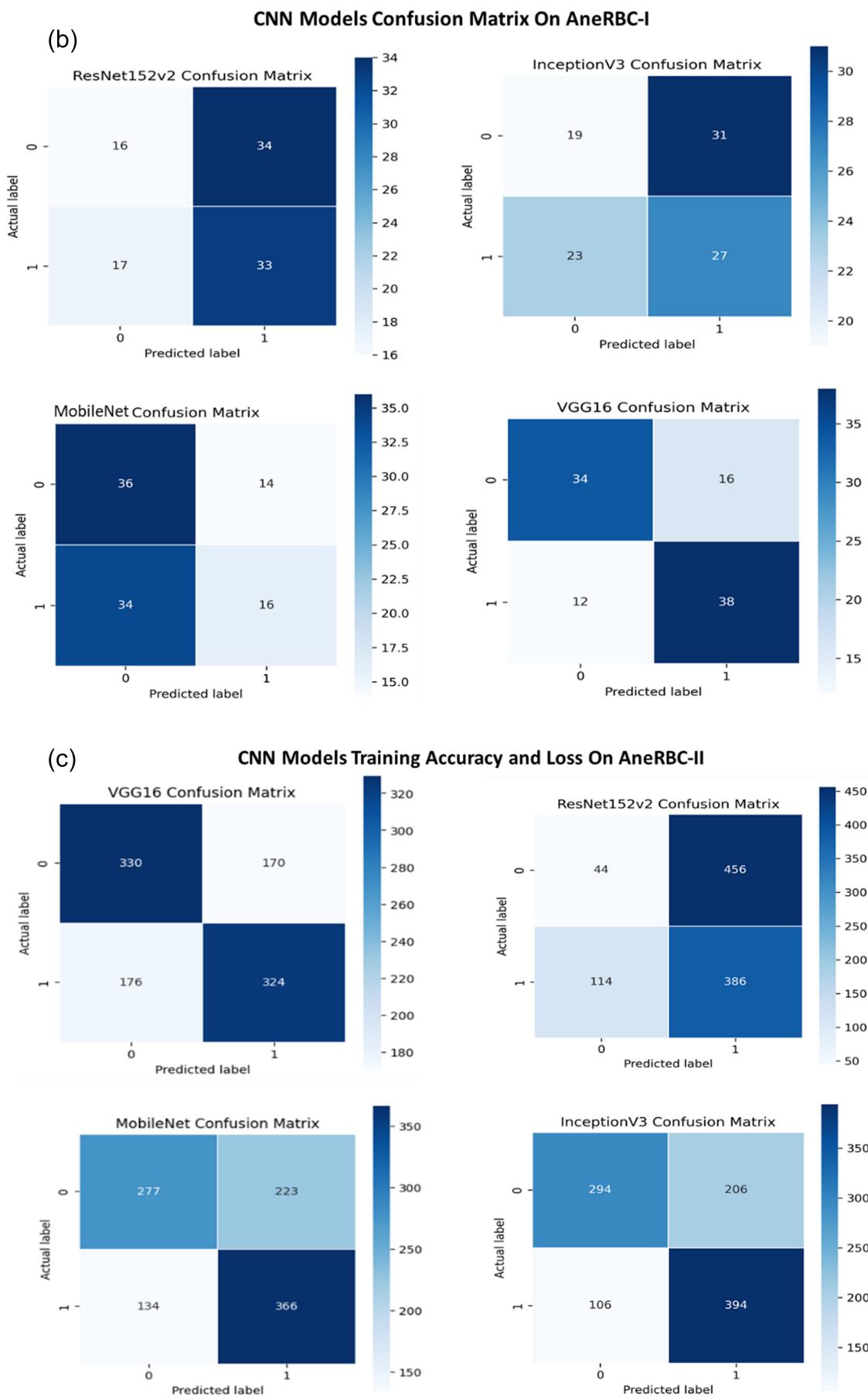


**Figure A1.** continued

### Classification results of CNN models



**Figure A2.** (a) Training accuracy and loss for InceptionV3 and ResNet152v2 for both datasets. (b) Confusion Matrix for InceptionV3 and ResNet152v2, VGG16, and InceptionV3 for AneRBC-I dataset. (c) Confusion Matrix for InceptionV3 and ResNet152v2, VGG16, and InceptionV3 for AneRBC-II dataset.

**Figure A2.** continued

## References

- WHO. *Strategies to Prevent Anaemia: Recommendations from an Expert Group Consultation*. World Health Organization, 2016:5–6. [http://www.searo.who.int/entity/nutrition/documents/recommendations\\_of\\_the\\_expert\\_group\\_on\\_anaemia.pdf?ua=1](http://www.searo.who.int/entity/nutrition/documents/recommendations_of_the_expert_group_on_anaemia.pdf?ua=1) (20 April 2024, date last accessed).
- Kouzehkanan ZM, Saghari S, Tavakoli E et al. Raabin-WBC: a large free access dataset of white blood cells from normal peripheral blood. 2021. <https://doi.org/10.1101/2021.05.02.442287>.
- Hasani M, Hanani A. Automated diagnosis of iron deficiency anemia and thalassemia by data mining techniques. *Int J Comput Sci Network Secur (IJCSNS)* 2017;17:326–31.
- Freire WB, Kahn GS, McGuire J et al. Anemia prevention and control. *What Works*, 2003.
- Meena K, Tayal DK, Gupta V. 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>
- Shanita SN, Siti Hanisa A, Noor Afifah AR et al. Prevalence of anaemia and iron deficiency among primary schoolchildren in Malaysia. *Int J Environ Res Public Health* 2018;15:2332. <https://doi.org/10.3390/ijerph15112332>
- De-Regil LM, Jefferds ME, Sylvestsky A et al. Intermittent iron supplementation for improving nutrition and development in children under 12 years of age. *Cochrane Database of Systematic reviews* 2011;2011:CD009085. <https://doi.org/10.1002/14651858.CD009085.pub2>
- World Health Organization (WHO). *The Global Prevalence of Anaemia in 2011*. 2011. <https://www.who.int/publications/item/9789241564960> (20 April 2024, date last accessed).
- Ivanova N, Gugleva V, Dobreva M et al. Laboratory approach to anemia. *InTech* 2016;ii:13.
- Abdullah M, Al-Asmari S. Anemia types prediction based on data mining classification algorithms. In: *Communication, Management and Information Technology - Proceedings of the International Conference on Communication, Management and Information Technology, ICCMIT 2016*, CRC Press, 26–29 April 2016. pp. 615–21. Cosenza, Italy, 2016.
- Shahzad M, Umar AI, Shirazi SH et al. Identification of anemia and its severity level in a peripheral blood smear using 3-tier deep neural network. *Appl Sci* 2022;12:1–22. <https://doi.org/10.3390/app12105030>
- Meijering E, Carpenter AE, Peng H et al. Imagining the future of bioimage analysis. *Nat Biotechnol* 2016;34:1250–55. <https://doi.org/10.1038/nbt.3722>
- Ching T, Himmelstein DS, Beaulieu-Jones BK et al. Opportunities and obstacles for deep learning in biology and medicine. *J Royal Soc Interface* 2018;15:20170387. <https://doi.org/10.1098/rsif.2017.0387>
- Blom S, Paavolainen L, Bychkov D et al. Systems pathology by multiplexed immunohistochemistry and whole-slide digital image analysis. *Sci Rep* 2017;7:1–13. <https://doi.org/10.1038/s41598-017-15798-4>
- Waters JC. Accuracy and precision in quantitative fluorescence microscopy. *J Cell Biol* 2009;185:1135–48. <https://doi.org/10.1083/jcb.200903097>
- Ronneberger O, Baddeley D, Scheipl F et al. Spatial quantitative analysis of fluorescently labeled nuclear structures: problems, methods, pitfalls. *Chromosome Res* 2008;16:523–62. <https://doi.org/10.1007/s10577-008-1236-4>
- Shahzad M, Umar AI, Shirazi SH et al. Semantic segmentation of anaemic RBCs using multilevel deep convolutional encoder-decoder network. *IEEE Access* 2021;9:161326–41. <https://doi.org/10.1109/ACCESS.2021.3131768>
- Florian K, Boz E, Fikret R et al. An annotated fluorescence image dataset for training nuclear segmentation methods. *Scientific Data* 2020;7:262. <https://doi.org/10.1038/s41597-020-00608-w>
- Caicedo JC, Roth J, Goodman A et al. Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytom Part A* 2019;95:952–65. <https://doi.org/10.1002/cyto.a.23863>
- Liu Y, Long F. Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning. In: *ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging - Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI)*, Venice, Italy, 8–11 April, 2019, pp. 113–21. 2019. [https://doi.org/10.1007/978-981-15-0798-4\\_12](https://doi.org/10.1007/978-981-15-0798-4_12)
- Nikitaei VG, Nagornov OV, Pronichev AN et al. Research of the leukocytes segmentation method in the blood cells recognition systems. *KnE Energy Phys* 2018;3:350–54. <https://doi.org/10.18502/ken.v3i2.1835>
- Miao H, Xiao C. Simultaneous segmentation of leukocyte and erythrocyte in microscopic images using a marker-controlled watershed algorithm. *Comput Math Methods Med* 2018;2018:7235795. <https://doi.org/10.1155/2018/7235795>
- Alomari YM, Norul S, Sheikh H et al. Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection algorithm. *Comput Math Methods Med* 2014;2014:979302. <https://doi.org/10.1155/2014/979302>
- Carolina A, Monteiro B, Iano Y et al. 7.1 Introduction. 2020:165–86. <https://doi.org/10.1016/B978-0-12-819061-6.00007-0>
- Shirazi SH, Umar AI, Naz S et al. Efficient leukocyte segmentation and recognition in peripheral blood image. *Technol Health Care* 2016;24:335–47. <https://doi.org/10.3233/THC-161133>
- Cao H, Liu H, Song E. A novel algorithm for segmentation of leukocytes in peripheral blood. *Biomed Signal Process Control* 2018;45:10–21. <https://doi.org/10.1016/j.bspc.2018.05.010>
- Hegde RB, Prasad K, Hebbar H et al. Development of a robust algorithm for detection of nuclei and classification of white blood cells in peripheral blood smear images. *J Med Syst* 2018;42:1–8. <https://doi.org/10.1007/s10916-018-0962-1>
- Kihm A, Kaestner L, Wagner C et al. Classification of red blood cell shapes in flow using outlier tolerant machine learning. *PLoS Comput Biol* 2018;14:1–15. <https://doi.org/10.1371/journal.pcbi.1006278>
- Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 2016;30:108–19. <https://doi.org/10.1016/j.media.2016.01.005>
- Oktay O, Schlemper J, Le Folgoc L et al. Attention U-Net: learning where to look for the pancreas. *MIDL*, 2018. <http://arxiv.org/abs/1804.03999>
- Roth HR, Lu L, Farag A et al. Deeporgan: multi-level deep convolutional networks for automated pancreas segmentation. *Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 2015;9349:556–64. [https://doi.org/10.1007/978-3-319-24553-9\\_68](https://doi.org/10.1007/978-3-319-24553-9_68)
- Shahin AI, Guo Y, Amin KM et al. White blood cells identification system based on convolutional deep neural learning networks. *Comput Methods Programs Biomed* 2019;168:69–80. <https://doi.org/10.1016/j.cmpb.2017.11.015>
- Xu M, Papageorgiou DP, Abidi SZ et al. A deep convolutional neural network for classification of red blood cells in sickle cell anemia. *PLoS Comput Biol* 2017;13:1–27. <https://doi.org/10.1371/journal.pcbi.1005746>
- Tuck MG, Alemi F, Shortle JF et al. A comprehensive index for predicting risk of anemia from patients' diagnoses. *Big Data* 2017;5:42–52. <https://doi.org/10.1089/big.2016.0073>
- Dithy MD, Krishnapriya V. Anemia selection in pregnant women by using random prediction (Rp) classification algorithm. *Int J Recent Technol Eng* 2019;8:2623–30. <https://doi.org/10.35940/ijrte.B3016.078219>
- Labati RD, Piuri V, Scotti F. ALL-IDB: the acute lymphoblastic leukemia image database for image processing. In: *18th IEEE International Conference on Image Processing*, Brussels, Belgium, pp. 2045–2048. IEEE, 2011. doi: [10.1109/ICIP.2011.6115881](https://doi.org/10.1109/ICIP.2011.6115881)
- Shahzad M, Umar AI, Khan MA et al. Robust method for semantic segmentation of whole-slide blood cell microscopic images. *Comput Math Methods Med* 2020;2020:1–13. <https://doi.org/10.1155/2020/4015323>

38. Sarrafzadeh O, Rabbani H, Talebi A *et al.* Selection of the best features for leukocytes classification in blood smear microscopic images. *Med Imaging 2014 Digit Pathol* 2014;9041:90410. <https://doi.org/10.1117/12.2043605>
39. Mohamed M, Far B, Guaily A. An efficient technique for white blood cells nuclei automatic segmentation. In: *Conference Proceeding - IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Seoul, South Korea, 14-17 October 2012. pp. 220–25. 2012. <https://doi.org/10.1109/ICSMC.2012.6377703>
40. Mohamed MMA, Far B. A fast technique for white blood cells nuclei automatic segmentation based on gram-schmidt orthogonalization. *Proc - Int Conf Tools Artif Intell ICTAI* 2012;1:947–52. <https://doi.org/10.1109/ICTAI.2012.133>
41. Ljosa V, Sokolnicki KL, Carpenter AE. Annotated high-throughput microscopy image sets for validation. *Nat Methods* 2012;9:637. <https://doi.org/10.1038/nmeth.2083>
42. Acevedo A, Merino A, Alférez S *et al.* A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data Brief* 2020;30:105474. <https://doi.org/10.1016/j.dib.2020.105474>
43. Abbas S. Microscopic images dataset for automation of RBCs counting. *Data Brief* 2015;5:35–40. <https://doi.org/10.1016/j.dib.2015.08.006>
44. Somasekar J, Ramesh G, Ramu G *et al.* A dataset for automatic contrast enhancement of microscopic malaria infected blood RGB images. *Data Brief* 2019;27:104643. <https://doi.org/10.1016/j.dib.2019.104643>
45. He K, Zhang X, Ren S *et al.* Deep residual learning for image recognition. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016;2016:770–78. <https://doi.org/10.1109/CVPR.2016.90>
46. Szegedy C, Vanhoucke V, Ioffe S *et al.* Rethinking the inception architecture for computer vision. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016;2016:2818–26. <https://doi.org/10.1109/CVPR.2016.30>