# SKDrepaData: A Multisite West African Blood Smear Dataset for Automated Sickle Cell Disease Detection and Classification

Abdourahmane Baldé[1], Ousmane Sall[1], Dieu-Donné Okalas Ossami[2], Mamadou Soumboundou[3], Cheikh Sall[3], Youssou Faye[4], Radhia M'Kacher[5], Avewe Bassene[6], Lamine Faty[4], Papa Seydou Wane[1], Mariama Dione[1], and Oumy Niass[1]

[1] Cheikh Hamidou Kane Digital University, Dakar, Senegal
{abdourahmane.balde, ousmane1.sall, papaseydou.wane, mariama.dione, oumy.niass}@unchk.edu.sn
[2] University of Science and Technology of Masuku, Franceville, Gabon
dieudonne.okalas@univ-masuku.org
[3] Iba Der Thiam University, Thiès, Senegal
{mamadou.soumboundou, cheikh.sall}@univ-thies.sn
[4] Assane Seck University, Ziguinchor, Senegal
{yfaye, l.f}@univ-zig.sn
[5] Cell Environment, Évry-Courcouronnes, France
radhia.mkacher@cell-environment.com
[6] Cheikh Anta Diop University, Dakar, Senegal
avewe.bassene@ucad.edu.sn

**Abstract.** Sickle cell disease (SCD) is one of the most prevalent inherited blood disorders worldwide, with the highest burden occurring in sub-Saharan Africa. Its diagnosis still relies heavily on manual microscopic examination of peripheral blood smears, a time-consuming and observer-dependent procedure, while publicly available microscopy datasets remain scarce and poorly representative of African clinical settings. We present SKDrepaData, a new dataset comprising 1861 full-field blood smear microscopy images collected from more than 100 West African patients across three clinical sites using a standardized May–Grünwald–Giemsa (MGG) staining protocol. Red blood cells are annotated into three morphological categories (normal, sickle-shaped, and other abnormal cells), totaling 33 832 annotated cells with both bounding-box and cell-level labels. Annotation reliability was assessed through a five-expert cross-validation protocol, achieving a mean Intersection over Union (IoU) of 0.85 for object detection and a Fleiss' $\kappa$ coefficient of 0.90 for morphological classification. To establish reproducible baseline benchmarks, we evaluate two previously published deep learning models for red blood cell detection and morphological classification on SKDrepaData. These models achieve a mAP@50 of 90.3 % and a classification accuracy of 97.0 %, respectively. The reported results provide reliable benchmark performances to facilitate future comparisons on this dataset. The dataset is available at https://doi.org/10.5281/zenodo.21365924.

**Keywords:** Sickle Cell Disease, Blood Smear Dataset, Medical Imaging, Object Detection, Vision Transformer, Global Health

## 1 Introduction

Sickle cell disease (SCD) is among the most common genetic disorders worldwide, affecting about 7.74 million people in 2021 and nearly 300 000 newborns annually, predominantly in West and Central Africa [1,2,3,4]. Without early screening, up to 80 % of affected children may die before the age of five [5,6,7]. Caused by a hemoglobin gene mutation, SCD deforms red blood cells (RBCs) into characteristic sickle shapes, leading to vaso-occlusive crises, chronic anemia, and severe complications [8].

Diagnosis relies on laboratory tests such as hemoglobin electrophoresis [9], the Emmel test [10], HPLC [11], or complete blood count [12], which require specialized equipment and trained personnel [13,14]. In many sub-Saharan African settings, these resources remain limited [15], making peripheral blood-smear microscopy an important initial screening step [16]. However, morphological examination is time-consuming, observer-dependent, and complicated by the small size and frequent overlap of RBCs [17].

Deep learning approaches and dedicated datasets have recently been proposed to support SCD analysis [18,19,20][21,22]. However, existing public datasets [23,24,25][26,27,28] are often small, single-site, or heterogeneous. Several provide classification labels only, lack bounding-box annotations, or do not quantify annotation reliability. Moreover, few reflect African clinical settings, despite the continent carrying the greatest disease burden. The available African datasets remain limited in volume, standardization, or task coverage.

To address this gap, we present SKDrepaData, a multi-site West African blood-smear microscopy dataset collected *in situ* and validated by experts. Images were acquired at three clinical sites using distinct microscopes at $\times 40$ and $\times 100$ magnifications. The dataset combines standardized May–Grünwald–Giemsa (MGG) staining, multi-site acquisition, cell-level annotations, and expert validation. Annotation reliability was assessed through a five-expert protocol, achieving a mean detection IoU of 0.85 and a Fleiss' $\kappa$ of 0.90. Two previously published models are also evaluated to provide reproducible benchmarks for RBC detection and morphological classification.

The main contributions are as follows: (1) a multi-site dataset comprising 1861 full-field MGG-stained images from more than 100 West African patients and 33 832 annotated RBCs across three morphological classes; (2) a quantitative assessment of annotation reliability by five experts; (3) a comparison with existing datasets highlighting its multi-site African acquisition, standardized staining, expert validation, and multi-task annotations; and (4) reproducible baseline benchmarks for RBC detection and morphological classification.

## 2 Related Work

Several microscopy datasets support research on automated RBC analysis for SCD, yet few combine the clinical representativeness, acquisition standardization, annotation quality, and multi-task coverage required for robust automated screening.

ErythrocytesIDB [23] is a widely used SCD resource, providing 196 full-field images across three morphological classes with subsets for classification, segmentation, and detection; however, it originates from a single center, and its small size requires intensive augmentation. Xu *et al.* [24] released nearly 14 000 RBC patches ($100 \times 100$ pixels) across six classes, but from only eight patients, restricted to classification, without detection or segmentation labels, and not publicly available. Tushabe *et al.* [26] introduced an African (Ugandan) SCD dataset with bounding boxes, but it is limited to 422 images and relies on a non-standardized protocol (Field/Leishman staining, smartphone digitization at the eyepiece), which affects image quality and reproducibility. erythroSight [28] is the largest recent resource, with over 300 000 single-cell images from 138 participants in Nepal and Canada, acquired with the Octopi automated microscope; however, it relies on a proprietary device, targets phase-contrast morphological characterization rather than multi-class detection on standard clinical staining, and provides no African grounding. Finally, general-purpose hematology datasets such as ALL-IDB [25] (108 images, ∼39 000 cells) and BCCD [27] (364 images, 4888 objects) provide detection annotations but contain no sickle cells, limiting their relevance for SCD.

Table 1: Comparison of SKDrepaData with existing datasets. Det.: detection; Cls.: classification; Seg.: segmentation.

| **Dataset** | **Images** | **Sickle** | **Multi-site** | **Tasks** | **African** |
|---|---|---|---|---|---|
| ErythrocytesIDB [23] | 196 | ✓ | ✗ | Cls.+Seg.+Det. | ✗ |
| Xu *et al.* [24] | ∼14 000 (p) | ✓ | ✗ | Cls. | ✗ |
| Tushabe *et al.* [26] | 422 | ✓ | ✓ | Det. | ✓ (UG) |
| erythroSight [28] | >300 000 | ✓ | ✓ | Seg./Morph. | ✗ |
| ALL-IDB [25] | 108 | ✗ | ✗ | Cls.+Det. | ✗ |
| BCCD [27] | 364 | ✗ | ✗ | Det. | ✗ |
| **SKDrepaData** | **1 861** | ✓ | **✓ (3)** | **Det.+Cls.** | **✓ (WA)** |

Overall, these resources share recurring limitations: small size, single-site acquisition, non-standardized staining, and the absence of patient metadata; above all, none combines multi-task coverage with an African clinical grounding. As summarized in Table 1, SKDrepaData uniquely combines a substantial multi-site volume, standardized MGG staining, expert-validated annotations, and joint detection/classification support, while being explicitly grounded in a West African clinical context.

## 3 The SKDrepaData Dataset

### 3.1 Data Acquisition

SKDrepaData was built from blood smears of patients monitored at a pediatric hospital in West Africa, prospectively enrolled between 2022 and 2025 after

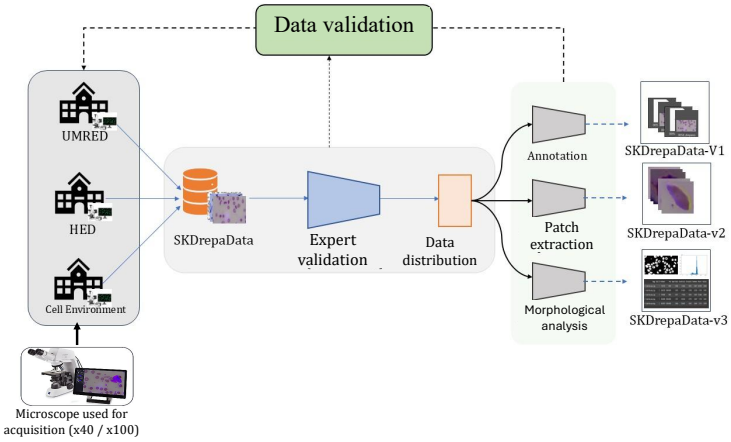


Fig. 1: Construction of SKDrepaData: multi-site acquisition, expert validation, and three subsets for detection, classification, and pixel-level segmentation.

informed consent and ethical approval[7]. Three hemoglobin profiles were included: homozygous (SS), heterozygous carriers (AS), and healthy controls (AA). More than 100 patients were enrolled, with two to five slides per patient. Images were acquired at three complementary clinical sites using distinct optical microscopes (two health facilities at $\times 100$ oil-immersion magnification and one specialized laboratory at $\times 40$), each coupled with a digital camera. This diversity introduces controlled acquisition variability, while a shared protocol preserves visual homogeneity. All slides were uniformly stained using the May–Grünwald–Giemsa (MGG) protocol. Figure 1 summarizes the process from multi-site collection to the construction of the three dataset subsets.

### 3.2 Annotation

Annotations were performed at the cell level, assigning each red blood cell a bounding box and a morphological label. Three classes were defined with hematologists and biologists: normal, sickle, and other abnormalities, including elliptocytes, target cells, and atypical shapes. Each blood smear slide remained associated with its source patient during data collection.

Annotations were initially produced using LabelImg and QuPath and subsequently validated by five experts, including hematologists, biologists, and computer scientists. Disagreements were reviewed and resolved collegially under the supervision of a reference hematologist to establish the final dataset annotations.

### 3.3 Quality Control and Annotation Reliability

Images with major acquisition artifacts, including severe blur, over- or under-exposure, debris, or air bubbles obscuring cells, were discarded during expert

[7] Ethical approval number and institutional details withheld for anonymous review.

Table 2: Inter-annotator agreement on SKDrepaData.

| **Task** | **Metric** | **Score** |
|---|---|---|
| Detection | Mean IoU | 0.85 |
| Classification | Fleiss' $\kappa$ | 0.90 |

quality control. The final corpus nevertheless preserves morphological variability and complex clinical cases, such as overlapping cells and atypical shapes.

Annotation reliability was quantified before the final consensus procedure. For detection, bounding boxes were matched between annotators using an IoU threshold of 0.5, yielding a mean IoU of 0.85. For classification, agreement among the five annotators reached Fleiss' $\kappa = 0.90$, indicating near-perfect agreement [29,30]. Final annotations were subsequently established through expert consensus.

### 3.4 Statistics and Organization

After quality control, SKDrepaData comprises 1861 full-field images and 33 832 annotated red blood cells, distributed across three classes: 11 106 normal (32.8 %), 11 473 sickle (33.9 %), and 11 253 other abnormalities (33.3 %). This balance reflects the dataset construction for machine learning and should not be interpreted as clinical prevalence.

The dataset comprises three complementary subsets: SKDrepaData-v1 for detection (bounding boxes), SKDrepaData-v2 for morphological classification (cell patches normalized to $224 \times 224$ pixels), and SKDrepaData-v3 for quantitative analysis (pixel-level masks and morphological descriptors).

For benchmarking, SKDrepaData-v1 and SKDrepaData-v2 were split into training, validation, and test sets using class stratification. The benchmark partition is therefore defined at the data level rather than as a strict patient-level partition.

## 4 Benchmark Experiments

To demonstrate the potential of SKDrepaData and provide reference points for future work, we re-evaluate two previously published models: DREPADetect [31], dedicated to red blood cell detection, and Drepa-ViT [32], devoted to morphological classification. Experiments are carried out on two subsets of the dataset: SKDrepaData-v1 for detection and SKDrepaData-v2 for classification. For both tasks, the data were partitioned into training, validation, and test sets using a class-stratified 70/20/10 split to preserve the representation of the three morphological classes across the experimental subsets. The benchmark experiments were conducted using anonymized image- and cell-level data rather than a strict patient-level partition. Accordingly, the reported results are intended as reproducible reference benchmarks for the detection and morphological classification tasks supported by SKDrepaData, rather than as a dedicated evaluation

of patient-independent generalization. Experiments were conducted on a single NVIDIA Tesla T4 GPU using PyTorch, following the same experimental protocols as in the original works. This setup provides a consistent basis for comparing the baseline models under the reported benchmark conditions.

### 4.1 Red Blood Cell Detection

DREPADetect [31] is a lightweight single-stage detector derived from the YOLOv8 architecture [33], optimized for the constraints of smear microscopy: small red blood cells, frequent overlap, and the need for fast inference in resource-limited settings. Its architecture concentrates the computational capacity on small-object scales, significantly reducing model complexity (16.2 M parameters and a latency of 14.1 ms per image, versus 43.6 M and 18.3 ms for YOLOv8) while preserving accuracy. Table 3 reports the performance of DREPADetect compared with two-stage (Faster R-CNN) and single-stage (YOLOv5, YOLOv7, YOLOv8) detectors. DREPADetect achieves the best overall performance, with a precision of 87.6 %, an mAP@50 of 90.3 %, and an mAP@50:95 of 64.6 %, clearly outperforming two-stage detectors (Faster R-CNN: 57.4 % mAP@50) and earlier YOLO versions. YOLOv8 retains a slightly higher recall (88.8 % versus 87.6 %), as DREPADetect favors the reliability of its detections at the cost of a marginally lower recall—a favorable trade-off in clinical screening, where limiting false detections matters as much as missing none.

Table 3: Baseline detection performance on SKDrepaData-v1.

| **Model** | **Prec.** | **Recall** | **mAP@50** | **mAP@50:95** |
|---|---|---|---|---|
| Faster R-CNN [34] | 60.0 | 62.5 | 57.4 | 33.6 |
| YOLOv5 [35] | 65.4 | 72.9 | 72.4 | 46.3 |
| YOLOv7 [36] | 65.0 | 72.6 | 69.6 | 42.4 |
| YOLOv8 [33] | 84.6 | **88.8** | 90.0 | 64.5 |
| **DREPADetect** [31] | **87.6** | 87.6 | **90.3** | **64.6** |

### 4.2 Morphological Classification

Drepa-ViT [32] is a hybrid CNN–Transformer classifier that refines the detection outputs by assigning each cropped red blood cell to one of the three morphological classes. It combines a ResNet-50 [37] backbone for local feature extraction with a Vision Transformer [38] for global context modeling. Although the three classes are globally balanced, they present unequal learning difficulty: reversible sickle forms and atypical abnormalities are morphologically ambiguous. A *Focal Loss* ($\gamma = 2$), complemented by weighted sampling, is used to focus training on these difficult cases. Table 4 compares Drepa-ViT with eight reference architectures. Drepa-ViT achieves the best performance across all metrics, with an accuracy of 97.00 %, an F1-score of 96.92 %, and an AUC of 0.9979, while maintaining a moderate computational cost (26.39 M parameters, 9.29 GFLOPs). It surpasses

the best convolutional backbones (ResNet-50: 96.46 % F1) and, above all, the standalone Vision Transformer (ViT/16: 77.75 % F1), confirming that the fusion of local (CNN) and global (Transformer) representations is decisive: the pure ViT, lacking a convolutional inductive bias, generalizes poorly on cells with fine morphological differences. The Grad-CAM visualizations and UMAP projections obtained during the experiments show that predictions rely on clinically relevant morphological cues, thereby reinforcing the interpretability of the model.

Table 4: Baseline classification performance on SKDrepaData-v2.

| Model | Acc. | Prec. | Recall | F1 | AUC |
|---|---|---|---|---|---|
| MobileNetV2 [39] | 0.9326 | 0.9353 | 0.9326 | 0.9328 | 0.9899 |
| VGG19 [40] | 0.9468 | 0.9480 | 0.9468 | 0.9466 | 0.9947 |
| VGG16 [40] | 0.9610 | 0.9640 | 0.9610 | 0.9612 | 0.9930 |
| Xception [41] | 0.9255 | 0.9275 | 0.9255 | 0.9249 | 0.9911 |
| ResNet-50 [37] | 0.9645 | 0.9654 | 0.9645 | 0.9646 | 0.9973 |
| EfficientNet-B0 [42] | 0.9433 | 0.9448 | 0.9433 | 0.9428 | 0.9956 |
| InceptionV3 [43] | 0.8191 | 0.8176 | 0.8191 | 0.8167 | 0.9362 |
| ViT/16 [38] | 0.7778 | 0.8378 | 0.7778 | 0.7775 | 0.9252 |
| **Drepa-ViT** [32] | **0.9700** | **0.9872** | **0.9691** | **0.9692** | **0.9979** |

## 5 Discussion

SKDrepaData addresses a well-identified gap: the lack of representative, standardized, and multi-task blood smear datasets for sickle cell disease in African clinical settings. Its multi-center acquisition, uniform MGG staining, two-level annotation, and quantified inter-annotator agreement (IoU 0.85, Fleiss' $\kappa$ 0.90) make it a reliable resource for red blood cell detection, morphological classification, and pixel-level segmentation. The reference results obtained with two previously published models show that accurate red blood cell detection and robust morphological classification can be achieved on SKDrepaData, thereby providing reproducible comparison points for future work.

Several directions could further strengthen and extend this work. The current benchmark experiments use a class-stratified training/validation/test split that preserves the distribution of the three morphological classes. Future evaluations based on patient-level partitions will provide a complementary assessment of generalization to previously unseen patients. Likewise, the current benchmark results should not be interpreted as a dedicated assessment of cross-site or cross-device generalization. Future work could therefore evaluate model performance across the different acquisition sites and settings, including leave-one-site-out protocols, to further assess robustness to variations in microscopes and magnification conditions. In addition to the detection and classification benchmarks presented in this work, SKDrepaData-v3 provides pixel-level segmentation masks that can support future studies on segmentation and the extraction of finer morphometric biomarkers. The *other abnormalities* category also groups heterogeneous morphologies, including elliptocytes, target cells, and atypical shapes, whose internal diversity could, in the longer term, justify a more detailed taxonomy. Finally,

external validation on cohorts from other countries remains necessary to further assess cross-population generalizability.

## 6 Impact in Remote and Community Settings

SKDrepaData targets remote and community settings (RCS), where the SCD burden is highest yet annotated data and diagnostic resources are scarcest. Collected *in situ* in West African laboratories with clinical-grade staining, it captures the real imaging conditions of these settings rather than those of well-equipped reference centers, making models trained on it more likely to transfer where they are most needed. As an openly released, expert-validated, multi-task resource, it also lowers the entry barrier for local researchers, who have lacked representative data to develop and benchmark tools adapted to African cohorts. By enabling automated smear analysis, it can reduce the dependence on scarce hematologists and costly confirmatory tests (hemoglobin electrophoresis, HPLC), supporting earlier point-of-care screening in underserved areas. The accompanying lightweight baselines illustrate this potential, running in a few milliseconds on modest hardware (portable microscopes, low-cost workstations, embedded devices). By releasing a representative dataset from within the region most affected by SCD, this work aims to help rebalance the geographic distribution of medical-AI resources toward the communities that need them most.

## 7 Conclusion

We presented SKDrepaData, a new multi-site blood smear dataset for the development and evaluation of AI methods for automated sickle cell disease screening. It comprises 1861 full-field images from more than 100 West African patients, totaling 33 832 red blood cells annotated into three morphological classes, with a standardized acquisition protocol, expert-validated annotations, and quantified reliability (mean IoU 0.85, Fleiss' $\kappa$ 0.90). We further established reference results using two previously published models covering detection and morphological classification. The dataset will be publicly released[8], supporting the development of reproducible, robust, and clinically relevant AI methods for SCD screening in low-resource settings.

## References

1. M. Ally and E. Balandya. Current challenges and new approaches to implementing optimal management of sickle cell disease in sub-saharan africa. *Semin. Hematol.*, 60(4):192–199, 2023.
2. GBD 2021 Sickle Cell Disease Collaborators. Global, regional, and national prevalence and mortality burden of sickle cell disease, 2000–2021. *Lancet Haematol.*, 10(8):e585–e599, 2023.

[8] Anonymized data link provided for review; public DOI released upon acceptance.

3. WHO Regional Office for Africa. Regional fact sheet on sickle cell disease, 2025.
4. A. Kumar and S. Bhattacharya. Sickle cell disease: a comparative perspective on global and national initiatives. *Front. Hematol.*, 3:1457158, 2024.
5. O. P. Adigwe, S. O. Onoja, and G. Onavbavba. A critical review of sickle cell disease burden and challenges in sub-saharan africa. *J. Blood Med.*, 14:367–376, 2023.
6. G. Onavbavba, O. P. Adigwe, and S. O. Onoja. Improving access to healthcare services for sickle cell disease patients in nigeria. *Front. Health Serv.*, 5, 2025.
7. L. Chirande et al. Building capacity for pediatric hematological diseases in sub-saharan africa. *Blood Adv.*, 9:939–947, 2025.
8. I. M. Idris, E. A. Botchwey, and H. I. Hyacinth. Sickle cell disease as an accelerated aging syndrome. *Exp. Biol. Med.*, 247(4):368–374, 2022.
9. F. Cotton, F. Wolff, and B. Gulbis. Automated capillary electrophoresis in the screening for hemoglobinopathies. In *Capillary Electrophoresis of Biomolecules*, pages 227–235. Humana Press, 2013.
10. M. Assoumanou, D. Issifou, and A. Akpona. Le test de précipitation en milieu réduit: une alternative au test d'emmel. *Int. J. Biol. Chem. Sci.*, 4(2), 2010.
11. P. T. McGann, A. G. Hernandez, and R. E. Ware. Sickle cell anemia in sub-saharan africa: advancing the clinical paradigm. *Blood*, 129:155–161, 2017.
12. D. N. Greene, C. P. Vaughn, B. O. Crews, and A. M. Agarwal. Advances in detection of hemoglobinopathies. *Clin. Chim. Acta*, 439:50–57, 2015.
13. C. Patgiri and A. Ganguly. Red blood cell and sickle cell detection from microscopic blood images. In *WiSPNET*, pages 474–478, 2019.
14. B. Aygun and I. Odame. A global perspective on sickle cell disease. *Pediatr. Blood Cancer*, 59(2):386–390, 2012.
15. M. Ally and E. Balandya. Current challenges and new approaches to implementing optimal management of sickle cell disease in sub-saharan africa. *Semin. Hematol.*, 60(4):192–199, 2023.
16. K. Esoh, E. Wonkam-Tingang, and A. Wonkam. Emerging point-of-care technologies for sickle cell disease screening and monitoring. *Expert Rev. Med. Devices*, 13(12):1073–1093, 2016.
17. Y. Alapan et al. Sickle cell disease in sub-saharan africa: transferable strategies for prevention and care. *Lancet Haematol.*, 2021.
18. L. Alzubaidi, M. A. Fadhel, O. Al-Shamma, J. Zhang, and Y. Duan. Deep learning models for classification of red blood cells in microscopy images. *Electronics*, 9, 2020.
19. K. Simon et al. Comparison of deep learning techniques in detection of sickle cell disease. *Artif. Intell. Appl.*, 1, 2023.
20. S. Aruna Kumar, K. Ganesan, and B. Bharathi Kannan. A novel deep learning approach for sickle cell anemia detection in human rbcs. *Biomed. Tech.*, 68:175–185, 2023.
21. G. M. K. B. Karunasena, K. Herath, H. Priyankara, and A. Madhusanka. Sickle cell disease identification using r-cnn and digital image processing. *Int. J. Sci. Res. Eng. Trends*, 1:1–8, 2023.
22. I. ul Haq, S. Anwar, M. T. Khan, and U. Sadique. An intelligent approach for blood cell detection employing faster r-cnn. *Pak. J. Eng. Technol.*, 6:1–6, 2023.
23. ErythrocytesIDB. ErythrocytesIDB dataset, 2015.
24. M. Xu, D. P. Papageorgiou, S. Z. Abidi, M. Dao, H. Zhao, and G. E. Karniadakis. A deep convolutional neural network for classification of red blood cells in sickle cell anemia. *PLOS Comput. Biol.*, 13:e1005746, 2017.

25. R. D. Labati, V. Piuri, and F. Scotti. All-idb: The acute lymphoblastic leukemia image database for image processing. In *ICIP*, pages 2045–2048, 2011.
26. F. Tushabe et al. Sickle cell disease dataset, 2024. Kaggle.
27. M. Noukhez. Blood cell count and detection (bccd) dataset, 2024. Kaggle.
28. P. Shrestha et al. erythrosight: Open-access image dataset of blood cells in sickle cell disease and beta-thalassemia, 2024. Federated Research Data Repository.
29. J. L. Fleiss. Measuring nominal scale agreement among many raters. *Psychol. Bull.*, 76(5):378–382, 1971.
30. J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.
31. Anonymous. DREPADetect: An optimized YOLOv8 model for high-precision detection of sickle cells. In *IEEE Int. Conf. Big Data Service and Machine Learning Applications*, pages 203–209, 2025.
32. Anonymous. Drepa-ViT: An interpretable cnn-transformer hybrid for red blood cell classification. In *Int. Conf. Artificial Intelligence, Computer, Data Sciences and Applications*, pages 1–7, 2026.
33. Ultralytics. YOLOv8, 2023.
34. S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In *NeurIPS*, pages 91–99, 2015.
35. G. Jocher. YOLOv5 by ultralytics, 2020.
36. C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *CVPR*, pages 7464–7475, 2023.
37. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, pages 770–778, 2016.
38. A. Dosovitskiy et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2021.
39. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. MobileNetV2: Inverted residuals and linear bottlenecks. In *CVPR*, pages 4510–4520, 2018.
40. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015.
41. F. Chollet. Xception: Deep learning with depthwise separable convolutions. In *CVPR*, pages 1251–1258, 2017.
42. M. Tan and Q. V. Le. EfficientNet: Rethinking model scaling for convolutional neural networks. In *ICML*, pages 6105–6114, 2019.
43. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *CVPR*, pages 2818–2826, 2016.