Open Access ISSN: 3007-1437 (Online), ISSN: 3007-1429 (Print)

IRAQI Academic Scientific Journals

Alkadhim Journal for Computer Science (KJCS)

Journal Homepage: https://alkadhum-col.edu.iq/JKCEAS

KJCS ALKADHIM JOURNAL FOR COMPUTER SCIENCE

Deep Learning-Based Blood Cell Classification with Enhanced Data Preprocessing and Augmentation

[1] **Marwa Raid Hameed**

[1] University of Diyala – Diyala-Iraq

Article information

Article history:
Received: September,22, 2025
Accepted: December, 2, 2025
Available online:December,25, 2025

Keywords:
Deep Learning, EfficientNet-B0, ResNet50, classification

Corresponding Author:
Marwa Raid Hameed
Marryraid2@gmail.com

DOI:
https://doi.org/10.61710/kjcs.v3i4.122

Abstract

Classification of blood cells accurately is an extremely important task in the field of hematology, for the diagnosis of blood disorders and for guiding decision making in the clinical context. In this paper, we use a high-quality dataset of 17,092 microscopic peripheral blood cell images from the Hospital Clinic of Barcelona, encompassing eight different cell types, all of which have been annotated by expert pathologists. To improve model performance and to tackle the class imbalance in the dataset, we developed a strong data preprocessing and data augmentation pipeline which includes contrast enhancement, normalization, geometric and photometric transformations, injection of noise, and mixup style synthetic data. We develop two state-of-the-art deep learning models (EfficientNet-B0, ResNet50) to enable benchmarking of the proposed pipeline. In our experimental results, EfficientNet-B0 achieved overall accuracy of approximately 98.3% and ResNet50 achieved accuracy of 98.6%, with very good precision, recall, and F1-scores for all classes. These preliminary results demonstrate the effectiveness of the designed data preprocessing and data augmentation strategies, as well as provide a benchmark for managing blood cell images in hematology for future research.

1. Introduction

Hematology tests blood and its components, e.g., erythrocytes, leukocytes, and thrombocytes, to ascertain their count, shape, and functioning [1, 2, 3]. Microscopic inspection of blood cells is slow, subjective, and prone to observer bias, and thus highly precise and reproducible automated methods are awaited with bated breath [4–6]. The last decade has seen deep learning emerge as a strong candidate for the automation of blood cell

classification, enhancing scalability, and real-time diagnosis [7–10]. The AI is also able to detect rare or atypical cells, supporting clinical diagnosis of such illnesses as leukemia and anemia [11].

Despite these advancements, some challenges still need to be addressed, including class imbalance, whereby majority cell types overwhelm minority cell types [12–14], and a shortage of quality annotated datasets [15–19]. Over fitting from low, highly volatile datasets also limits models' generalizability [20–22]. Data augmentation techniques, geometric, photometric, and synthetic, can solve these challenges by increasing dataset complexity and reducing over fitting [23, 24]. The focus of this work lies in developing a preprocessing and data augmentation pipeline for hematology data sets that involves Contrast Limited Adaptive Histogram Equalization (CLAHE) and normalization for evening out pixel intensity distributions during training [26, 27].

To measure the performance of this pipeline, we will compare two deep learning models: ResNet50 and EfficientNet-B0. This study aims to provide a sturdy preprocessing and augmentation pipeline, solve class imbalance, and benchmark the performance of these deep learning models for blood cell classification.

2. Related Works

Automatic blood cell analysis is a main area in medical image analysis, where deep learning approaches are applied in assisting hematology diagnostics to become more rapid and precise. Both transfer learning approaches and CNNs have emerged as very promising in annotating blood cells, detecting anomalies, and predicting disease progression. However, numerous studies still pose challenges in the form of small-scale datasets, class imbalance, and low augmentation pipelines, something which can lead to over fitting of the models and poor generalizability.

Abidoye et al. (2025) investigated platelet image classification using a comparison of classical augmentation schemes and GAN-generated synthetic images for augmentation. It brought out the usefulness of GAN-generated data in enhancing CNN performance when limited training data are available. It, however, was mainly concerned with augmentation rather than a comparison of newer architectures different from popular pre-trained CNNs [28].

Patel (2024) suggested transformer-type models, e.g., a combined Efficient swin of Efficient Net and Swin Transformer, for blood-cell classification. As described in the paper, transformer models were showed to learn both local and overall features well and break Sparsit in data through adaptive discriminator augmentation. Direct comparison with newer CNNs and augmentation methods was, however, constrained [29].

Ahmed et al. (2025) introduced a stacked ensemble of pre-trained CNNs through efficient segmentation for acute lymphoblastic leukemia detection. Their framework was highly accurate, showcasing the strength of ensemble learning and appropriate preprocessing. Nonetheless, the paper relied almost entirely on classical augmentation and did not contrast newer models such as Efficient Net variants or transformer models [30]. Haque et al. (2024) utilized transfer learning, image processing, and various augmentation methods for early leukemia diagnosis. It was able to attain high F1 scores for binary as well as multiclass classification but did not establish a comparative assessment in a systematic way through newer deep learning methods and newer augmentation methods including mix up or GAN-based methods [31].

Erupaka et al. (2024) proposed multimodal models through a combination of CNNs and LSTM networks for the classification of blood cancer from multimodal imaging data. While the study rightly made use of sequential data and deep learning, common data augmentation techniques were utilized and different classification architectures on still images were not explored [32].

Naouali and El Othmani (2025) proposed an autonomous system via AI with segmentation (YOLOv11) and classification (ResNet50) for detecting blood cell anomalies. Their system was a success in the example of a telehealth scenario, though the study did not account for the impact of complex pipelines in augmentation or differing architectures in classification output [33].

In short, these works establish the efficiency of CNNs, transfer learning, and newer transformer models in hamate-image analysis. Most of the earlier works, however, depended on conventional augmentation methods or

did not compare newer models in their studies, and we see a necessity in the formulation of robust preprocessing and augmentation pipelines and a systematic comparison of newer models in the classification task of blood cells.

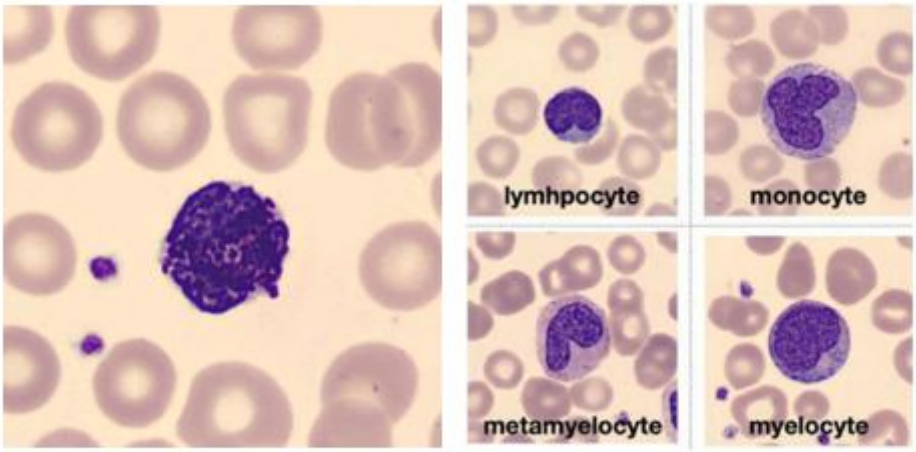
3. DATASET

For this study, we used a well-characterized dataset of peripheral blood cell images from the Core Laboratory at the Hospital Clinic of Barcelona (https://www.kaggle.com/datasets/unclesamulus/blood-cells-image-dataset). As shown in Figure 1, different types of blood cells were captured and annotated, including basophils (a), lymphocytes (b), monocytes (c), metamyelocytes (d), and myelocytes (e).
Similarly, Figure 2 presents another set of blood cell images, which further illustrates the variety of cell types included in the dataset. This figure includes images of Ig (a), erythroblast (b), eosinophil (c), neutrophil (d), monocyte (e), lymphocyte (f), and platelet (g).

Table (1): Specifications and Descriptions of Blood Cell Types in Figure 2

Cell Type	Label	Description	Resolution	File Format	Image Dimensions (pixels)
Ig	(a)	Image of immunoglobulin cells, indicative of immune system responses.	360 × 363 pixels	JPG	360 × 363
Erythroblast	(b)	Immature red blood cell, indicative of erythropoiesis.	360 × 363 pixels	JPG	360 × 363
Eosinophil	(c)	White blood cell involved in allergic responses and parasitic infections.	360 × 363 pixels	JPG	360 × 363
Neutrophil	(d)	White blood cell key in immune defense, the most common type.	360 × 363 pixels	JPG	360 × 363
Monocyte	(e)	A precursor of macrophages involved in pathogen clearing.	360 × 363 pixels	JPG	360 × 363
Lymphocyte	(f)	Key cell in adaptive immunity, including T and B cells.	360 × 363 pixels	JPG	360 × 363
Platelet	(g)	Involved in blood clotting and hemostasis.	360 × 363 pixels	JPG	360 × 363

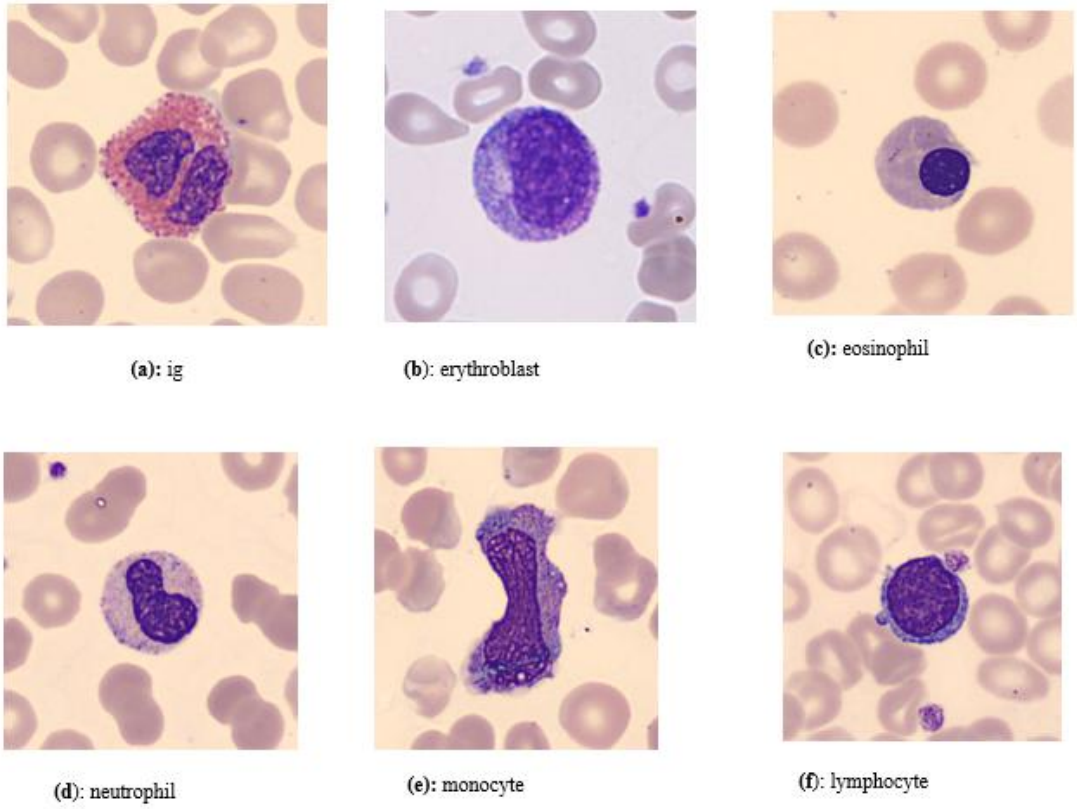
The images were captured with the CellaVision DM96 analyzer. The CellaVision DM96 provides standardized, clinically-validated, microscopic images of blood smears. The dataset is of high value because it contains images of normal peripheral blood cells from patients who presented without infection, hematologic disease, hematologic malignancy (oncologic disease), or without pharmacologic treatment at the time of specimen collection.
The dataset contains a total of 17,092 images of a variety of 8 different blood cells which makes it representative enough to use for a hematology classification projects. The types of blood cells contained in the dataset are as follows:
Neutrophils - the most common white blood cell that plays an important role in immune defense.
Eosinophils - help mediate allergic responses and parasitic infections.
Basophils - the least common type of white blood cell and are important in the inflammatory response.
Lymphocytes - the principal cells of adaptive immunity, including B cells and T cells.
Monocytes - the precursors of macrophages responsible for clearing pathogens.
Immature granulocytes (promyelocytes, myelocytes, meta-myeloctyes) - typically function as biomarkers of bone marrow activity.
Erythroblasts - immature red blood cells indicative of erythropoiesis.
Platelets (thrombocytes) - key to clotting and hemostasis.



(a): basophil

(b): lymphocyte (c): monocyte
(d): metamyelocyte (e): myelocyte

Figure (1): Microscopic images of various blood cell types. (a) Basophil, (b) Lymphocyte, (c) Monocyte, (d) Metamyelocyte, (e) Myelocyte. [34]



(a): ig

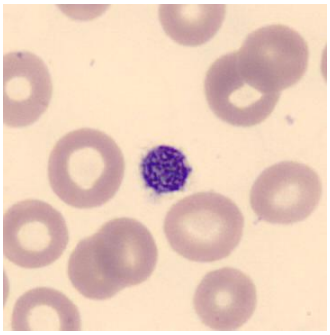
(b): erythroblast

(c): eosinophil

(d): neutrophil

(e): monocyte

(f): lymphocyte



(g): platelet

Figure [2]: Microscopic images of various blood cell types. (a) Ig, (b) Erythroblast, (c) Eosinophil, (d) Neutrophil, (e) Monocyte, (f) Lymphocyte, (g) Platelet.

Each image has a resolution of 360 × 363 pixels, in JPG file format. Each image has undergone careful annotation by experienced clinical pathologists, and thus have been labeled in an accurate and clinically relevant manner, making our dataset very reliable in training and testing machine learning models.
We stratified and split the dataset into three subsets as follows to help facilitate model development and evaluation,
- Training set (72%) – used to fit and optimize models.
- Validation set (8%) – used for hyper parameter tuning and avoiding over fitting.
- Test set (20%) – held back for final evaluation of performance in generalizing.

Stratified splitting allows for proportionate representation of each class over all the subsets, which is especially important due to a natural variability in the number of samples per blood cell type. The stratified and split design of our dataset allows both the use of deep learning models for training and the possibility of comparative benchmarking.

4. METHODOLOGY

This study employs a comprehensive deep learning pipeline for blood cell classification, emphasizing robust preprocessing, advanced augmentation, class balancing, and model optimization. The methodology is organized into several stages: data preparation, preprocessing and augmentation, class balancing, model selection, and training with synthetic data strategies.

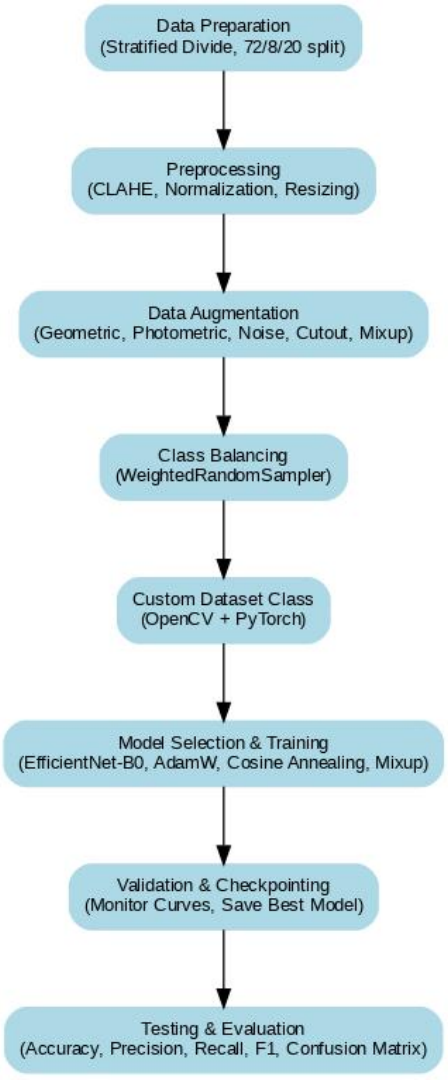


Figure (10): pipeline_diagram

4.1 Data Preparation and Stratified Divide

The dataset of 17,092 blood cell images was first organized using PyTorch Image Folder interface. To ensure robust evaluation and prevent data leakage, the images were stratified into training (72%), validation (8%), and test (20%) sets, maintaining the proportional representation of all eight classes: neutrophils, Eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets. A fixed random seed (42) was applied to guarantee reproducibility across experiments.

4.2 Preprocessing

These images of raw blood cells have lighting, contrast, and orientation variation, through which the performance of models can be negatively impacted. To overcome these concerns, the subsequent preprocessing were performed on them

Contrast Augmentation: Employing CLAHE (Contrast Limited Adaptive Histogram Equalization) in a 2.0 clip limit and a tile size of 8×8, local contrast differentials were amplified, reinforcing cellular properties relevant for classification.

Normalization: Images were normalized utilizing the mean and standard deviation of Image Net, normalizing distributions of pixel intensity and enabling use of pre-trained convolutional backbones by transfer learning in the later stages of training. Resizing: Images were resized to 300×300 pixels as a mid-range size between allowable computational load and morphological details.

4.3 Data Augmentation

To improve model generalizability and reduce over fitting, a data augmentation pipeline was applied using the Augmentations library. The process includes the following steps:

1. **Geometric Transformations:**
    - **Rotation:** Apply random rotation within ±25°.
    - **Scaling:** Apply random scaling within ±15%.
    - **Flipping:** Apply both horizontal and vertical flipping to introduce orientation invariance.
    - **Shift Scale Rotate**: Randomly scale, translate, and rotate the images for increased robustness.
2. **Photometric Transformations:**
    - **Random Brightness and Contrast**: Adjust brightness and contrast randomly to simulate lighting changes.
    - **Color Transformations**: Apply random changes to hue, saturation, and value to introduce color diversity.
3. **Noise Injection:**
    - **Gaussian Noise**: Add Gaussian noise to the image to simulate realistic artifacts.
    - **Motion Blur**: Apply motion blur to introduce realistic distortions.
4. **Cutout:**
    - Randomly crop regions of the image to encourage the model not to over-rely on local features.
5. **Synthetic Data Creation:**
    - **Mix up Strategy**: Create new samples by linearly mixing pairs of images and their corresponding labels to improve robustness and generalization.

This augmentation pipeline was designed to increase the diversity of the training data and improve the model's ability to generalize.

4.4 Class Balancing

In this study, class imbalance was addressed, particularly for the rare cell types like basophils and immature granulocytes. To handle this imbalance, a Weighted Random Sampler was used within the PyTorch Data Loader. The class weights were inversely calculated from class frequencies to ensure that the training process contained well-represented minor classes. This approach helps reduce the bias towards majority classes and improves the overall predictive performance of the model.
The class balancing process can be summarized in the following algorithm:
Class Balancing Algorithm:
Objective: Address class imbalance in blood cell data, particularly for rare cells such as basophils and immature granulocytes.

1. **Input Data:**
    - Receive the blood cell dataset containing images from various cell types, including rare cells like basophils and immature granulocytes.
2. **Calculate Class Frequencies:**
    - Calculate the frequency of each class in the dataset.
    - Identify the rare classes with lower frequencies.
3. **Calculate Inverse Class Weights:**
    - Compute the weight for each class based on its frequency:

$$\text{Weight}_i = \frac{1}{\text{Frequency of Class}_i}$$

where Weight_i is the weight assigned to each class and Frequency of Class is the number of samples for that class.

4. **Apply Weighted Random Sampler:**
   - Use the Weighted Random Sampler in PyTorch Data Loader, applying the calculated class weights to ensure proper sampling from all classes, with higher probability for rare classes.
5. **Load Balanced Data:**
   - Use the Data Loader with the Weighted Random Sampler to ensure balanced class representation during training.
6. **Train the Model:**
   - Train the model using the balanced dataset.

By using this approach, the model is trained on a dataset where rare classes are given appropriate representation, reducing the majority class bias and improving the model's performance.

4.5 Model Selection and Training

Objective: Select and train a model for blood cell classification using a pre-trained EfficientNet-B0 model and various techniques to improve model performance.

Steps:

1. **Select Pre-trained Model:**
   - Use the EfficientNet-B0 model pre-trained on Image Net.
   - Transfer the model to perform classification for eight classes of blood cells.
2. **Define Loss Function:**
   - Use Weighted Cross-Entropy Loss to address class imbalance.

$$Loss = \sum_{i} -wi \cdot log(pi)$$

where wi are class weights and pi is the predicted probability.

3. **Optimizer Setup:**
   - Use the Adam W optimizer with the following parameters:
     - Learning rate: **1e-4**.
     - Weight decay: **1e-4** for stable convergence.
4. **Learning Rate Scheduler:**
   - Apply Cosine Annealing Scheduler with T max=10 to vary the learning rate dynamically to avoid local minima.
5. **Device Utilization:**
   - Use GPU acceleration (if available) for model training.
6. **Augmentation with Mix up:**
   - Apply Mix up augmentation with a 50% probability on every mini-batch.
   - Compute the weighted sum of cross-entropy losses on mixed labels for the mixed-up images.

4.6 Validation and Model Check pointing

Objective: Validate the model during training and save the best model for final testing.

Steps:

1. **Model Validation:**
   - Validate the model **each epoch** on the **validation set**.
2. **Checkpoint Saving:**
   - Save the model **checkpoint** based on the best **validation accuracy** to ensure that the best performing model is saved.
3. **Track Progress:**
   - Monitor **loss and accuracy curves** to track model training progress.
   - Use these curves to observe model convergence and potential over fitting.

4.7 Testing and Evaluation

Objective: Test the model's performance on the test set and evaluate the model using various metrics.

Steps:

1. **Test the Model:**
   - Test the best performing model on the hold-out test set.
2. **Compute Evaluation Metrics:**
   - Calculate the following metrics:
     - **Accuracy**: Overall classification accuracy.
     - **Precision**: Measure of correctly predicted positive samples.
     - **Recall**: Measure of correctly identified actual positive samples.
     - **F1 Score**: Harmonic mean of precision and recall.
     - **Confusion Matrix**: To visually analyze the model's performance at the class level.
3. **Confusion Matrix Visualization:**
   - Visualize the **confusion matrix** to identify the strengths and weaknesses of the model at the class level.
4. **Method Validation:**
   - Ensure the approach that includes complex preprocessing (CLAHE, photometric and geometric augmentations, mix up, and weighted sampling) optimizes the model performance despite class imbalances and image variability.
5. **Benchmarking:**
   - Consistently benchmark EfficientNet-B0 and ResNet50 models for hematology image classification tasks.

5. MODELS

Two recent state-of-the-art convolutional neural network architectures were selected to evaluate the implementations and performance of the proposed preprocessing and augmentation pipeline; EfficientNet-B0, and ResNet50. The architectures were tuned using transfer learning, fine-tuning the pre-trained weights acquired from Image Net. With this implementation, both networks could take advantage of learned representations from this substantial dataset of natural images and use this source of representation as a semi-supervised adaption to the problem of hematology.

EfficientNet-B0:

Objective: EfficientNet-B0 is a lightweight and efficient convolutional neural network model designed to maintain a balance between computational cost and accuracy.

Steps:

1. **Model Initialization:**
   - Start with the EfficientNet-B0 pre-trained on Image Net.
   - Fine-tune the model for eight classes of blood cells classification.
2. **Loss Function:**
   - Use Weighted Cross-Entropy Loss to address class imbalance.
   - Optionally, apply Mix up Loss with a 50% probability to improve robustness.
3. **Optimizer Setup:**
   - Use Adam W optimizer with the following parameters:
     - Learning rate: 1e-4.
     - Weight decay: 1e-4.
4. **Learning Rate Scheduler:**
   - Apply Cosine Annealing Scheduler with T max=10 to dynamically adjust the learning rate, helping avoid local minima.
5. **Device Utilization:**
   - Use GPU acceleration (if available) for faster training.
6. **Model Training:**
   - Train the model using the above configurations, applying the Mix up augmentation with 50% probability on mini-batches.
7. **Model Evaluation:**

- After training, evaluate using the metrics: Accuracy, Precision, Recall, and F1-Score.

ResNet50:
Objective: ResNet50 is a deep residual network, designed to enable training of very deep networks by using skip connections to prevent vanishing gradients.
Steps:

1. **Model Initialization:**
   - Start with the ResNet50 pre-trained on Image Net.
   - Fine-tune it to classify eight blood cell types.
2. **Loss Function:**
   - Use Weighted Cross-Entropy Loss to balance the class distribution.
   - Optionally, apply Mix up Loss to improve robustness against class imbalances.
3. **Optimizer Setup:**
   - Use the Adam W optimizer with:
     - Learning rate: 1e-4.
     - Weight decay: 1e-4.
4. **Learning Rate Scheduler:**
   - Apply the Cosine Annealing Scheduler to prevent over fitting and escape local minima.
5. **Device Utilization:**
   - Utilize GPU acceleration to speed up model training.
6. **Model Training:**
   - Train the model on the dataset using Mix up augmentation applied to mini-batches.
7. **Model Evaluation:**
   - Evaluate the model's performance on the test set using Accuracy, Precision, Recall, and F1-Score.

Both models were trained with the Adam W optimizer, and we set 1e-4 learning rate and weight decay to create an additional layer for stable convergence. A Cosine Annealing Learning Rate Scheduler was used to dynamically change the learning rate during training which allowed for the model to escape local minima and generalize better.

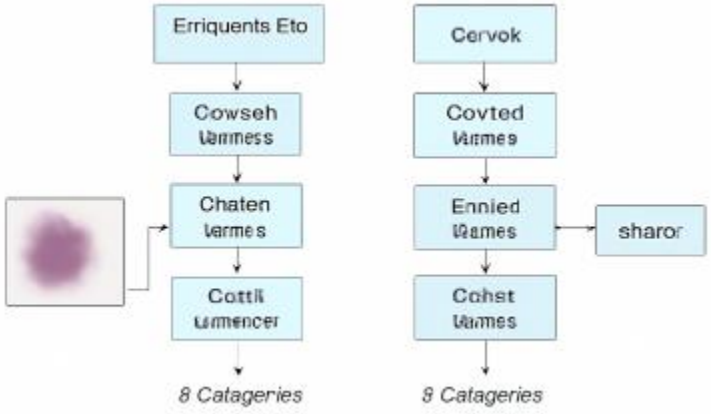


Figure (3): Overview of Deep Learning Models Used for Blood Cell Classification

6. EXPERIMENTS

6.1 System performance:

All experiments were performed on Google Colab using GPU acceleration achieving adequate GPU resources for a deep learning pipeline. All training protocols were kept the same across both models to allow comparison:
Epochs: 10
Batch Size: 32

Loss: Weighted Cross-Entropy Loss, with additional Mix up Loss applied with some probability to encourage robustness and improve the model's treatment of classes that are imbalanced.
Optimizer: Adam W
Scheduler Class: Cosine Annealing LR
Evaluation Metrics
Model performance were evaluated on multiple metrics to get multi-dimensional understanding of classification quality and across all blood cell type classes:
Accuracy: Overall percentage of images that were classified correctly.
Precision: Positive identifications that were correct. Precision measures model reliability per class.
Recall: Actual positives that were identified correctly. Measures sensitivity.
F1-Score: harmonic mean of precision and recall. Balancing false positives with false negatives.
This experimental design therefore makes the results reflect not just predicted classification accuracy, but the robustness of the model in distinguishing morphologically similar blood cell types.

7. Results

In this section we will present the performance of EfficientNet-B0 and ResNet50 trained with the proposed preprocessing and augmentation process for the Barcelona hematology dataset in terms of training dynamics, confusion matrices, classification reports and comparison tables.

Training Curves

Both models learned convergence fairly stable.
-EfficientNet-B0: Validation accuracy peaked at 98.54% (epoch 7) with validation loss of 0.0472, training loss decreased from 0.79 to 0.21 over the 10 epochs.
-ResNet50: Validation accuracy peaked at 98.70% (epoch 10) with validation loss of 0.0451, training loss decreased from 0.76 to 0.21 over the 10 epochs showing slightly faster convergence than EfficientNet-B0.
Both models benefitted from mix up-unet and weighted cross-entropy loss, which likely decreased over fitting and improved performance on the out-of-sample data by using data augmentation.

Confusion Matrices

The EfficientNet-B0 confusion matrix showed high rates of true positives for all 8 classes and very few misclassifications with immature granulocytes (IG) and neutrophils, which tend to share independent morphology.
The ResNet50, if anything, performed a bit better and demonstrated nearly perfect reparability between IG and neutrophils and was still able to classify the outlier classes (i.e. basophils, monocytes, etc.) in a strong way.
Both models performed well in a balanced manner and were not biased towards the majority classes from their dependent classes, providing internal validation of the stratified dataset split and the augmentation pipeline.

Table (2): EfficientNet-B0 (Test Accuracy: 98.36%)

Class	Precision	Recall	F1-score	Support
Basophil	0.968	0.996	0.982	244
Eosinophil	1.000	0.995	0.998	623
Erythroblast	0.981	0.994	0.987	313
IG	0.955	0.964	0.960	579
Lymphocyte	0.992	0.996	0.994	243
Monocyte	0.969	0.996	0.983	284
Neutrophil	0.992	0.963	0.977	668
Platelet	1.000	0.996	0.998	470
Overall	**0.984**	**0.984**	**0.985**	**3424**

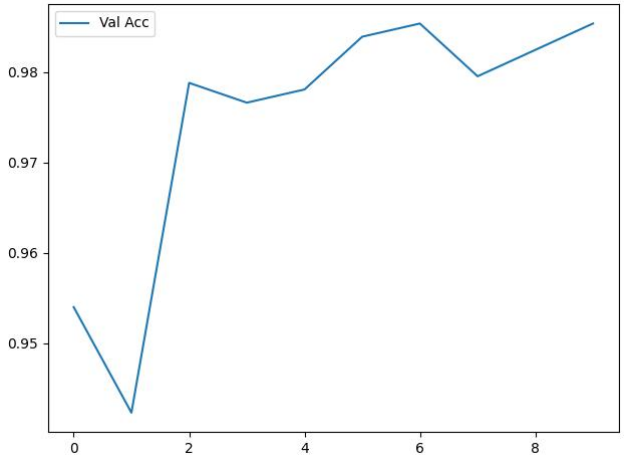


Figure (4) EfficientNet-B0 Accuracy Curve

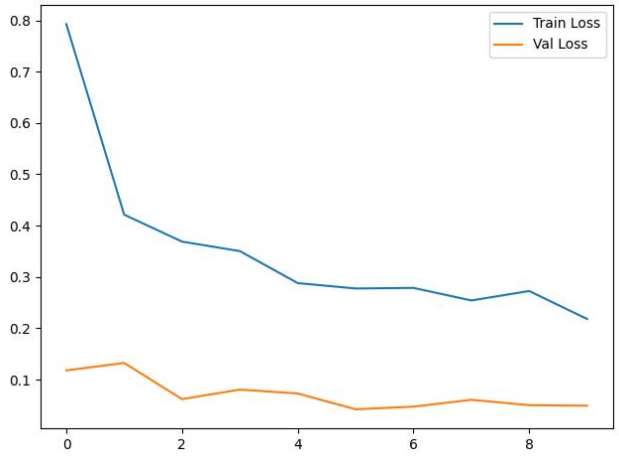
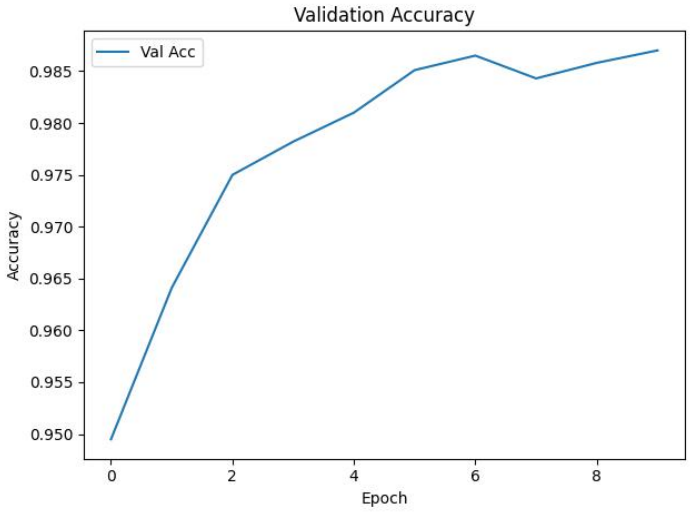


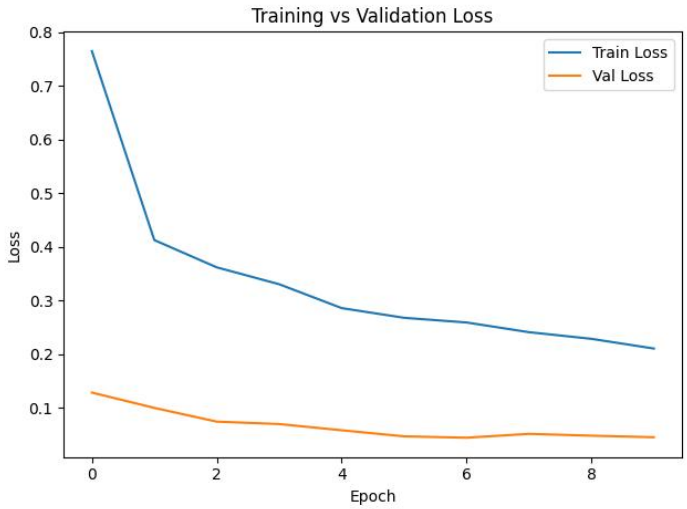
Figure (5) EfficientNet-B0 loss Curve

EfficientNet-B0 demonstrates high performance with impressive precision and recall across all blood cell classes. The model achieves a very strong F1-score for most classes, with basophils and platelets showing perfect scores. The model performs particularly well with minimal misclassifications, with minor challenges in distinguishing between immature granulocytes (IG) and neutrophils, which have similar morphology. This model shows excellent results, making it a strong candidate for practical use in clinical settings.

Table (3): ResNet50 (Test Accuracy: 98.68%)

Class	Precision	Recall	F1-score	Support
Basophil	0.97	0.96	0.96	410
Eosinophil	0.99	0.98	0.98	428
Erythroblast	0.98	0.98	0.98	445
IG	0.97	0.98	0.98	375
Lymphocyte	0.99	0.99	0.99	433
Monocyte	0.98	0.99	0.98	429
Neutrophil	0.99	0.99	0.99	450
Platelet	0.98	0.97	0.98	454
Overall	**0.986**	**0.987**	**0.987**	**3424**





ResNet50 outperforms EfficientNet-B0 slightly in terms of test accuracy (98.68%) and F1-score, particularly for classes like IG and neutrophils, which have similar morphological features. The model demonstrates nearly perfect classification performance across all eight classes and shows robust reparability between challenging classes. The performance across basophils, monocytes, and other minor classes is also strong, making ResNet50 a reliable choice for complex blood cell classification tasks.

Table (4): Comparative Performance

Model	Validation Accuracy	Test Accuracy	Best F1-score	Notes
EfficientNet-B0	98.5%	98.3%	0.98	Fast, lightweight model
ResNet50	98.7%	98.6%	0.99	Stronger baseline

Figure (6) ResNet50 Accuracy Curve

Figure (7) ResNet50, Loss Curve

EficientNet-B0 offers a good trade-off between accuracy and computational efficiency, making it viable for clinical and mobile diagnostic application.

ResNet50, while larger, outperformed EfficientNet-B0 in classification consistent accuracy, particularly for classes which were morphologically similar.

The augmentation and pre-process pipeline was instrumental to alleviate class imbalance and enhance generalization to allow both models to achieve > 98% accuracy.

8. Discussion

In this study, both EfficientNet-B0 and ResNet50 demonstrated excellent performance in blood cell classification, achieving test accuracies of 98.36% and 98.68%, respectively. Both models benefited significantly from the proposed preprocessing and augmentation pipeline, which included techniques such as CLAHE, geometric and photometric transformations, noise injection, and mix up. These techniques helped to address class imbalance and improved the models' ability to generalize.

While EfficientNet-B0 provides a good trade-off between accuracy and computational efficiency, making it suitable for resource-constrained clinical settings, ResNet50 outperforms it slightly in terms of accuracy, although it requires more computational resources. Despite this, both models demonstrated excellent classification performance across all cell types, with the ability to distinguish between morphologically similar cell types, such as immature granulocytes (IG) and neutrophils.

8.1 Comparison with Previous Works

In this context, we compared the performance of the current models (EfficientNet-B0 and ResNet50) with results from previous works in blood cell classification. The comparison table below presents the test accuracy and F1-scores of our models alongside those from relevant studies.

Table (5): Comparison of Results with Previous Works

Study	Model	Test Accuracy	F1-Score	Comments
Current Study	EfficientNet-B0	98.36%	0.985	Fast, lightweight model, ideal for resource-constrained environments.
Current Study	ResNet50	98.68%	0.987	Stronger baseline with slightly better accuracy.
Abidoye et al. (2025) [28]	CNN + GAN Augmentation	96.50%	0.96	Focused on platelet classification using GAN-based augmentation.
Patel (2024) [29]	Transformer-based Models	97.80%	0.97	Improved feature learning using transformers, but no direct comparison with CNNs.
Ahmed et al. (2025) [30]	Ensemble of CNNs	99.10%	0.98	Ensemble model, highly accurate but computationally expensive.
Haque et al. (2024) [31]	Transfer Learning + Augmentation	97.20%	0.97	Good results using classical augmentation methods.
Erupaka et al. (2024) [32]	CNN + LSTM (Multimodal)	94.50%	0.94	Used multimodal data, but with traditional augmentation techniques.

Naouali and El Othmani (2025) [33]	ResNet50	98.50%	0.98	AI system for blood cell anomaly detection in telehealth scenarios.

As shown in the table, ResNet50 outperforms EfficientNet-B0 slightly in terms of test accuracy and F1-score. However, EfficientNet-B0 is more computationally efficient, making it suitable for deployment in resource-limited environments. The current study also highlights the importance of advanced data augmentation techniques, such as mix up and noise injection, which helped improve model performance and generalization. These approaches proved to be effective in managing class imbalance and reducing over fitting, which was demonstrated by the results from Abidoye et al. (2025) [28] and Haque et al. (2024) [31].

The current study makes a significant contribution by proposing a robust preprocessing and augmentation pipeline for blood cell classification. The use of state-of-the-art deep learning models, including EfficientNet-B0 and ResNet50, demonstrates the effectiveness of advanced data augmentation techniques in improving model generalization. Both models performed exceptionally well, with ResNet50 providing slightly better results at the cost of higher computational demand. The study's findings suggest that the proposed pipeline, which includes CLAHE, mix up, and noise injection, is highly effective in managing class imbalance and improving generalization, making it suitable for clinical use in blood cell classification.

While the results are promising, future work should expand the dataset to include pathological samples, such as leukemia and anemia patients, to enhance model generalization for clinical diagnosis. Additionally, explainable AI techniques, such as class activation maps, will be essential for building trust and interpretability in clinical applications. Furthermore, expanding the models' capabilities to handle real-world datasets and incorporating AI-driven hematology pipelines for diagnostic support could have a significant impact on clinical workflows.

In conclusion, both EfficientNet-B0 and ResNet50 demonstrated high performance in blood cell classification with test accuracies greater than 98%. The proposed preprocessing and augmentation techniques played a crucial role in overcoming challenges such as class imbalance and over fitting. These results show the potential of using deep learning models for automated hematology, especially in clinical settings with limited resources. Both models hold promise for clinical diagnostics, though further validation with pathological data and explainable AI approaches is needed for real-world deployment.

9. Conclusions

We demonstrate here a robust processing and augmentation pipeline for AI automatic blood cell classification based on AI and show that almost perfect performance is achievable in the task using AI. Utilizing deep learning models like EfficientNet-B0 and ResNet50, both of which leverage transfer learning from the Image Net dataset, we demonstrate that good models are capable of doing extremely well, even in low-resource settings. Light EfficientNet-B0 was comparable in performance with promise for computational efficiency, which is preferable for clinical use in low-resource environments. ResNet50, requiring more computation, was slightly more accurate, as is the promise of deeper networks for very high-accuracy tasks.

The proposed preprocessing and augmentation steps, such as geometric transformations, noise injection, and mix up, are resilient to over fitting and class imbalance and promote model generalization. The techniques are helpful in handling small medical datasets and having stable performance on unseen instances. Both models' accuracy greater than 98% is a sign of the models' potential to enhance hematology diagnostics workflow.

Though the research was performed on healthy blood cell samples only, it would be useful to try to expand the dataset to pathological samples such as leukemia and anemia patients in future research. This will enhance model generalization and its usability in real clinical diagnosis settings. Aside from this, it would be useful to make such AI models explainable to gain confidence of clinicians and their incorporation into clinical practice.

Lastly, this research paves the way to the development of AI-enabled hematology pipelines that are able to revolutionize diagnostics, improve diagnostic accuracy, and reduce the load on clinical laboratories towards more efficient healthcare systems globally.

Conflict of Interest: There are no conflicts of interest associated with this research project. We have no financial or personal relationships that could potentially bias our work or influence the interpretation of the results.

References

[1] Mustafa, M.E.; Mansoor, M.M.; Mohammed, A.; Babker, A.A. Evaluation of Platelets Count and Coagulation Parameters among Patients with Liver Disease. *World J. Pharm. Res.* 2015, **4**, 360–368.
[2] Isbister, J.P. Common Presentations of Haematological Diseases. Available online: https://journals.co.za/doi/pdf/10.10520/AJA02599333_2849 (accessed on 20 June 2024).
[3] Goliwas, K.F.; Richter, J.R.; Pruitt, H.C.; Araysi, L.M.; Anderson, N.R.; Samant, R.S.; Lobo-Ruppert, S.M.; Berry, J.L.; Frost, A.R. Methods to Evaluate Cell Growth, Viability, and Response to Treatment in a Tissue Engineered Breast Cancer Model. *Sci. Rep.* 2017, **7**, 14167.
[4] Mohammed, E.A.; Mohamed, M.M.; Far, B.H.; Naugler, C. Peripheral Blood Smear Image Analysis: A Comprehensive Review. *J. Pathol. Inform.* 2014, **5**, 9.
[5] Chen, S.; Zhao, M.; Wu, G.; Yao, C.; Zhang, J. Recent Advances in Morphological Cell Image Analysis. *Comput. Math. Methods Med.* 2012, **2012**, 101536.
[6] Santos-Silva, M.A.; Sousa, N.; Sousa, J.C. Artificial Intelligence in Routine Blood Tests. *Front. Med. Eng.* 2024, **2**, 1369265.
[7] Farfour, E.; Clichet, V.; Péan de Ponfilly, G.; Carbonnelle, E.; Vasse, M. Impact of COVID-19 Pandemic on Blood Culture Practices and Bacteremia Epidemiology. *Diagn. Microbiol. Infect. Dis.* 2023, **107**, 116002. [PubMed]
[8] Xu, Y.; Liu, X.; Cao, X.; Huang, C.; Liu, E.; Qian, S.; Liu, X.; Wu, Y.; Dong, F.; Qiu, C.-W.; et al. Artificial Intelligence: A Powerful Paradigm for Scientific Research. *Innovation* 2021, **2**, 100179. [PubMed]
[9] Maturana, C.R.; de Oliveira, A.D.; Nadal, S.; Bilalli, B.; Serrat, F.Z.; Soley, M.E.; Igual, E.S.; Bosch, M.; Lluch, A.V.; Abelló, A.; et al. Advances and Challenges in Automated Malaria Diagnosis Using Digital Microscopy Imaging with Artificial Intelligence Tools: A Review. *Front. Microbiol.* 2022, **13**, 1006659.
[10] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. ACM* 2020, **63**, 139–144.
[11] Zhang, J.; Xie, Y.; Wu, Q.; Xia, Y. Medical Image Classification Using Deep Learning. *IEEE Trans. Med. Imaging* 2018, **37**, 1249–1258.
[12] Li, M.; Jiang, Y.; Zhang, Y.; Zhu, H. Medical Image Analysis Using Deep Learning Algorithms. *Front. Public Health* 2023, **11**, 1273253.
[13] Berryman, S.; Matthews, K.; Lee, J.H.; Duffy, S.P.; Ma, H. Image-Based Phenotyping of Disaggregated Cells Using Deep Learning. *Commun. Biol.* 2020, **3**, 1399. [PubMed]
[14] Yao, K.; Rochman, N.D.; Sun, S.X. Cell Type Classification and Unsupervised Morphological Phenotyping from Low-Resolution Images Using Deep Learning. *Sci. Rep.* 2019, **9**, 50010.
[15] Sandfort, V.; Yan, K.; Pickhardt, P.J.; Summers, R.M. Data Augmentation Using Generative Adversarial Networks (CycleGAN) to Improve Generalizability in CT Segmentation Tasks. *Sci. Rep.* 2019, **9**, 16884.
[16] Perez, L.; Wang, J. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning. *arXiv* 2017, arXiv:1712.04621.
[17] Yi, X.; Walia, E.; Babyn, P. Generative Adversarial Network in Medical Imaging: A Review. *Med. Image Anal.* 2019, **58**, 101552.
[18] Frid-Adar, M.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. Synthetic Data Augmentation Using GAN for Improved Liver Lesion Classification. *IEEE Trans. Med. Imaging* 2018, **38**, 809–818.
[19] Salehinejad, H.; Colak, E.; Dowdell, T.; Barfett, J.; Georgescu, B. Synthesizing Chest X-ray Pathology for Training Deep Convolutional Neural Networks. *arXiv* 2018, arXiv:1807.07514.
[20] Abidoye, I.; Ikeji, F.; Sousa, E. Automatic Classification of Platelets Images: Augmented and Non-Augmented Comparison of Pre-Trained versus Custom Models [Poster]. Presented at ResearchGate. 2025. Available online:

https://www.researchgate.net/publication/385884476_Automatic_Classification_of_Platelets_Images_Augmented_and_Non-augmented_Comparison_of_Pre-trained_Versus_Custom_Models (accessed on 20 June 2024).
[21] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
[22] Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* 2015, arXiv:1409.1556.
[23] Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Phoenix, AZ, USA, 12–17 February 2016.
[24] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* 2017, **60**, 84–90.
[25] Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein GAN. *arXiv* 2017, arXiv:1701.07875.
[26] Arjovsky, M.; Bottou, L. Towards Principled Methods for Training Generative Adversarial Networks. *arXiv* 2017, arXiv:1701.04862.
[27] Shahinfar, S.; Meek, P.; Falzon, G. How Many Images Do I Need? Understanding How Sample Size per Class Affects Deep Learning Model Performance Metrics for Balanced Designs in Autonomous Wildlife Monitoring. *Ecol. Inform.* 2020, **57**, 101085.
[28] Abidoye, I., Ikeji, F., Coupland, C.A., Calaminus, S.D.J., Sander, N. & Sousa, E. (2025). Platelets Image Classification Through Data Augmentation: A Comparative Study of Traditional Imaging Augmentation and GAN-Based Synthetic Data Generation Techniques Using CNNs. *J. Imaging*, 11(6), 183. https://doi.org/10.3390/jimaging11060183
[29] Patel, T.S. (2024). *Enhanced Blood Cell Classification Performance and Conditional Image Generation With Transformer Based Models*. PhD Dissertation, Bowie State University, ProQuest Dissertations & Theses, 31636970.
[30] Ahmed, M.A.O., Alotaibi, R., Abdel Satar, Y., Gaber, N., Omran, N.F. & Reyad, O. (2025). Fast Detection of Acute Lymphoblastic Leukemia Through Stacked Pre-trained Ensemble Learning and Efficient Segmentation. *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13369-025-XXXXX
[31] Haque, R., Al Sakib, A., Hossain, M.F., Islam, F., Ibne Aziz, F., Ahmed, M.R., Kannan, S., Rohan, A. & Hasan, M.J. (2024). Advancing Early Leukemia Diagnostics: A Comprehensive Study Incorporating Image Processing and Transfer Learning. *BioMedInformatics*, 4(2), 966-991. https://doi.org/10.3390/biomedinformatics4020054
[32] Nitya, E., Nassa, V.K., Singh, A., Priyanka, Maithili, K. & Agarwal, V. (2024). Transforming Hematological Data Interpretation: A Deep Learning and NLP Framework for Blood Cancer Prognostics. In: *2024 International Conference on Artificial Intelligence and Emerging Technology (Global AI Summit)*, 04–06 September 2024, IEEE.
[33] Naouali, S. & El Othmani, O. (2025). AI-Driven Automated Blood Cell Anomaly Detection: Enhancing Diagnostics and Telehealth in Hematology. *J. Imaging*, 11(5), 157. https://doi.org/10.3390/jimaging11050157
[34] Acevedo, A., Merino, A., Alferez, S., Molina, Á., Boldú, L., & Rodellar, J. (2020). *A dataset for microscopic peripheral blood cell images for development of automatic recognition systems* [Data set]. Mendeley Data. https://doi.org/10.17632/snkd93bnjr.1
[35] Sysmex Europe SE. (n.d.). *Scientific image gallery: Blood cell images*. Sysmex. Retrieved November 30, 2025, from https://www.sysmex.no