A Vision Transformer with a Self-Attention Mechanism for High-Accuracy Blood Cell Classification

Noor Ayesha
Center of Excellence in Cyber Security (CYBEX), Prince Sultan University Riyadh, Saudi Arabia
drnayesha@gmail.com (corresponding author)

Humaira Khalidi
College of Medicine, Alfaisal University, Riyadh, Saudi Arabia
hkhalidi@alfaisal.edu

Received: 29 November 2025 | Revised: 28 December 2025 | Accepted: 4 January 2026

 DOI: https://doi.org/10.48084/etasr.16600

ABSTRACT

Blood has several essential functions in the body, including immune defense against foreign elements and the transportation of oxygen, nutrients, and hormones. In this study, an IoT-based learning system was developed for classifying blood cell types, including basophils, erythroblasts, monocytes, myeloblasts, and segmented neutrophils. The study utilized a dataset of five blood cells, which was systematically split into three parts to ensure robust evaluations. A Convolutional Neural Network (CNN) model based on a Vision Transformer (ViT) combined with a Self-Attention Mechanism (SAM) was utilized to extract and learn discriminative features from cell images. To comprehensively evaluate and demonstrate execution, a suite of assessment metrics, classification reports, and Receiver Operating Characteristic (ROC) curves with the Area Under the Curve (AUC) was used in the training, validation, and testing phases. Overall, the proposed model achieved accuracies of 99%, 98%, and 98%, with balanced precision and recall, across the training, validation, and testing samples of five classes. Additionally, 5-fold cross-validation was conducted during the preparation and approval stages to improve generalizability and decrease overfitting. The model consistently demonstrated strong performance and maintained robust discriminative ability on the unseen test set. The proposed framework offers a promising solution to the limitations of conventional blood flow investigation by enabling digital, precise, and adaptable cell classification.

Keywords-medical imaging; vision transformer; classification; cancer cells; health risks

I. INTRODUCTION

Blood plays several essential functions in the body, including immune defense and the transportation of oxygen and nutrients. The blood cells comprise 45% of the entire volume, while the liquid plasma constitutes the remaining 55% [1]. The blood cells can be categorized into three types: Red Blood Cells (RBCs) or erythrocytes, White Blood Cells (WBCs) or leukocytes, and platelets (thrombocytes). Approximately 40%-45% of the blood consists of RBCs, while WBCs represent approximately 1% of the blood [1]. The harmful role of basophils is observed in illnesses such as certain skin diseases, asthma, chronic rhinosinusitis, in the intestinal diseases, and in a few anaphylactic responses. Basophils are responsible for various immune system infections and persistent inflammatory or fibrotic conditions of the lungs, intestine, kidneys, or heart [2]. The RBC helps the tissues in the body to produce energy by delivering the appropriate amount of oxygen. The erythroblasts are the adolescent RBCs that are mostly present in newborn children, typically in the age range of 0-4 months [3].

The Monocytes are WBC that develop from the bone marrow. They are an important part of innate immune reaction and direct cellular homeostasis, particularly during disease and inflammation. Monocytes account for around 5% of circulating nucleated cells in typical adult blood, and the half-life of circulating monocytes is around 1-3 days [4]. Acute Myeloid Leukemia (AML) arises from the transformation of myeloid precursor cells, leading to the uncontrolled proliferation of immature WBC, known as myeloblasts. These explosions are not distinguished by fully functional blood cells, leading to a lack of RBCs, platelets, and mature WBCs. This disorder of normal blood cell production causes symptoms such as anemia, infection, and bleeding [5]. Neutrophils are the main type of granulocyte and include 40% to 70% of all human leukocytes. There are 100-200 billion neutrophils produced daily in humans by hematopoiesis in the bone marrow [6].

Blood cells are used to diagnose diseases. The microscopic imaging of blood cells is only done manually by medical professionals [7]. It has been revealed that a CNN model performs better than traditional Machine Learning (ML) models [8]. The objective of the present study is to develop an IoT-based learning system trained for classifying blood cell types. The study utilized ViT-based Self-Attention Mechanisms (SAM) combined with a CNN model to evaluate the diagnostic effectiveness for the microscopic image dataset of five blood cell types. The research ultimately reduces model bias and variance, and increases generalizability and classification reliability. The main contributions of this study are:

- The ViT-based (SAM) CNN model achieves 99%, 98%, and 98% accuracy with balanced precision and recall across five classes of training, validation, and testing samples, respectively.
- This investigation contributes to the developing field of Artificial Intelligence (AI)-driven advanced pathology and offers guidelines for upgrading interpretability, extending datasets with varying qualities, and validating the framework in real-world clinical IoT conditions.

II. LITERATURE REVIEW

AI is widely employed in medical image processing, the agriculture field, and healthcare. ML and the Deep Learning (DL) models are employed for the automated classification of blood cells. The blood cell problem is divided into two parts: classification and segmentation [9]. The segmentation process, including thresholding, morphological operations, and ML, enhances the classification performance by removing the non-important parts of the images. Some authors have employed traditional ML techniques, while others have utilized DL approaches such as custom CNNs, transfer learning, and hybrid models for blood cell classification [10]. ML models, such as K-Nearest Neighbor, Support Vector Machine (SVM), Naïve Bayes, Decision Tree, and Artificial Neural Networks, have been applied to classify the WBC using the microscopic blood smear images [11].

Most DL models are implemented to classify the WBC into five types, such as neutrophils, eosinophils, basophils, monocytes, and lymphocytes. Different transfer learning methods are also deployed to classify the blood cells and leukemia. While various CNN architectures have been applied to detect and classify WBC, the VGG16 architecture has been trained and has achieved better results [12]. Authors in [13] applied CNN-based different architectures for feature selection and classification of WBC, achieving an accuracy of 97.95% for four different cell types [13]. Authors in [14] applied the ResNet50 model to 10,868 microscopic images of six classes. This model demonstrated high precision and recall for each class, achieving an overall accuracy of 94% [14].

Authors in [15] applied a CNN-based AlexNet model to 17,000 blood smear microscopic images to identify various blood cells. Their model achieved a minimal Quadratic Loss of 0.0049, with an accuracy of 95.08%. Authors in [16] demonstrated that a custom CNN model used to classify the eight classes of blood cells utilizing microscopic images achieved an accuracy of 92.60%, surpassing VGG16, Res-Net101, and EfficientNet variants, as well as other models. Overall, most DL-based studies have identified blood cells using microscopic images of different types of cells. However, the present study utilizes five types of blood cells to classify them by deploying the proposed model. Additionally, the YOLOv8 model is used to detect Acute Lymphoblastic Leukemia. The performance of the proposed model was compared with other models, including SVM, ResNet-50, a hybrid model that integrates ResNet-50 with the SVM classifier, and DenseNet-121. The YOLOv8 model achieved an accuracy of 95% and 94% on the C-NMC and ALL-IDB2 datasets, respectively [17]. Additionally, the present study proposes a ViT-based SAM integrated with a CNN model for diagnosing five types of blood cells using a microscopic image dataset.

III. PROPOSED FRAMEWORK

The study uses a publicly available blood cell dataset [18] for leukemia detection. The dataset consists of 1024×1024 pixel images, with Wright-Giemsa minimum staining, 100-fold magnification, and 24-bit RGB colour. Multiple focal planes per sample are used for the diagnostic process. The dataset comprised 5,000 images, with 1,000 images from each of the five classes, distributed evenly across them, as shown in Figure 1. The detailed proposed framework is illustrated in Figure 2.

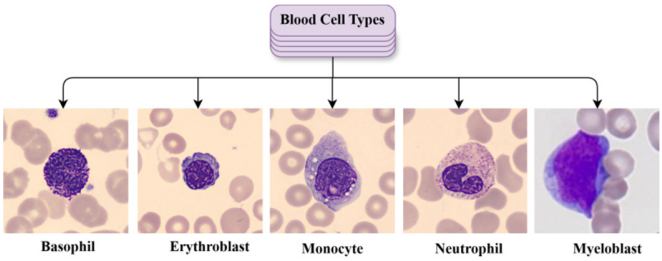


Fig. 1. Graphical representation of the 5 blood cell types.

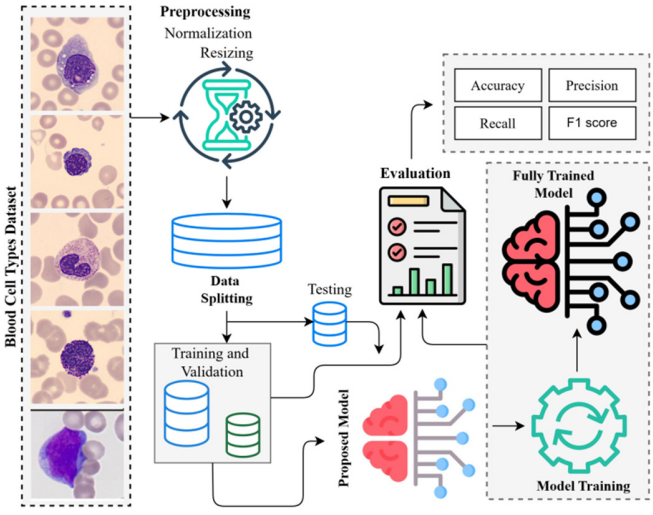


Fig. 2. Proposed research framework.

A. Imaging Data Preprocessing and Partition

Multiple data preparation methods were used to build a robust image classifier, including data resizing and normalization. Before conducting any analysis, it is important to preprocess the dataset images. Preprocessing was conducted by resizing all 5000 microscopic images to 100×100×3 RGB pixels for better feature selection. The resized images were normalized after dividing them by 255.0 for a more balanced presentation. The preprocessed data were systematically split into three parts: training, validation, and testing (holdout).

B. Convolutional Neural Network

This study employs the ViT-based SAM by integrating it with a CNN model, along with the introduction of a lightweight model for diagnosing the 5 types of blood cells using the microscopic images. As CNN models are weak in operating on the local neighbors and miss global information, the custom CNN model was adopted in this study for its efficient performance on microscopic images.

C. Self-Attention Mechanism

SAMs are capable of addressing the problem of global information and capturing long-range interactions. They are also a key component of the ViT and are utilized to enhance neural network classification performance [19]. The proposed model integrates a CNN with a dual attention network to detect the leukocytes.

D. Proposed Model

The ViT-based SAM, integrated with a CNN model, was considered in this study to detect the five blood cell types using microscopic images. Authors in [20] demonstrated that integrating SAM with the CNN model achieves outstanding classification performance. The proposed model integrates local spatial features with global contextual dependencies using a CNN and a ViT-based SAM. This hybrid design is a key modification, enabling both local and global feature representation, which enhances classification reliability compared to conventional CNN-only models. The CNN architecture was integrated with the SAM layer. The proposed model consists of CNN-based feature extraction, max pooling, SAM, flattening, and fully connected layers, as well as the final model composition.

1) Input and CNN Block

The proposed model processes the input image $I \in \mathbb{R}^{H*W*C}$, where *H* is the height of the image, *w* is the width of the image, and *C* is the number of channels (RGB). The fifth convolution layer was used for the proposed model, with the mathematical expression given in:

$$X^{(l)} = ReLU\left(Conv2D^{(l)}(X^{(l-1)}; K^{(l)})\right) \quad (1)$$

where $X^{(l)}$ is the output of the l^{th} convolutional block, $K^{(l)} \in \mathbb{R}^{(3x3)*C_{in}^{(l)}*C_{out}^{(l)}}$ is the kernel, and $l = 1, 2, 3, 4, 5$. The kernel size is the same for each block: (3, 3), with the same padding and ReLU activation function.

2) Max Pooling

Max pooling layers were applied after each convolutional layer, with the mathematical expression given in (2). The number of filters increases the depth with $C_{out}^{(1)}$=16, $C_{out}^{(2)} = 32$, $C_{out}^{(3)}$=32, $C_{out}^{(4)} = 64$, $C_{out}^{(5)} = 64$.

$$X^{(l)} = MaxPooling\left(X^{(l)}\right) \quad (2)$$

3) Self-Attention

The output of the final CNN layer is $X \in \mathbb{R}^{H' \times W' \times C}$, where $C = 64$. The tensor shape was reshaped to $X_r \in \mathbb{R}^{S \times C}$, where $S = H' \times W'$applied to SAM. The values, query, and key of SAM were projected using (3), where h = 1, 2,…, H heads. With ${W_h}^Q$, ${W_h}^K$, ${W_h}^V \in \mathbb{R}^{C \times d_k}$, were learnable weights metrics, $X_r$were the feature vectors, and d_k is the number of key dimensions, and *H* is the number of heads.

$$Q_h = X_r \times {W_h}^Q, K_h = X_r \times {W_h}^K, V_h = V_r \times {W_h}^V \quad (3)$$

The scaled dot product attention score was computed using (4), where $Q_h K_h^T$ is the pairwise attention score between all tokens, divided by $\sqrt{d_k}$ to prevent the gradients from being too large:

$$Attention_h(X_r) = softmax\left(\frac{Q_h K_h^T}{\sqrt{d_k}}\right) \times V_h \quad (4)$$

Similarly, the outputs of all heads were concatenated and projected with (5), where $W^0 \in \mathbb{R}^{H \times d_k \times C}$ is a learned linear projection. The final output of SAM was achieved using (6) by pulsing the residual connection and layer normalization:

$$SAM(X_r) = Concat(\, Attention_1, \ldots Attention_H) W^0 \quad (5)$$

$$Z = LayerNorm\big(X_r + SAM(X_r)\big) \quad (6)$$

4) Flatten and Fully Connected Layers

The output of attention or SAM $Z \in \mathbb{R}^{S \times C}$ was further flattened into a $z \in \mathbb{R}^{S \times C}$ vector that was passed through the fully connected layer. The first dense layer, h_1 was defined as follows, where $W_1 \in \mathbb{R}^{64x(S \times C)}$. The dropout was set to 0.5 or 50% of the activations. The second dense layer , h_2 were defined as follows, where $W - 2 \in \mathbb{R}\text{^}128 \times 64$. Similarly, the output layer $y^{\wedge}$ was added with 5 classes for classification purposes, where $W_3 \in \mathbb{R}^{C_{classes}x128}$. The detailed graphical presentation of the proposed model is presented in Figure 3.

$$h_1 = ReLU(W_{1z} + b_1) \quad (7)$$

$$h_2 = ReLU(W_2 h_1 + b_2) \quad (8)$$

$$y^{\wedge} = softmax(W_3 h_1 + b_3) \quad (9)$$

5) Final Model Composition

The forward pass of the proposed model is defined in (10), where *X* represents the output of the last convolution block:

$$y^{\wedge} = softmax\big(W_3 \times ReLU(W_2 \times ReLU(W_1 \times Flatten\big(SelfAttention(X)\big) + b_1\big) + b_2 + b_3) \quad (10)$$

6) Performance Metrics

Standard metrics were utilized to evaluate the performance of the proposed model for multi-class classification. Their mathematical equations are:

$$Accuracy = \frac{\sum_{i=1}^{5} TP_i}{\sum_{i=1}^{5}(TP_i+FP_i+FN_i+TN_i)} \quad (11)$$

$$Precision\ C_i = \frac{TP_i}{TP_i+FP_i},\ \text{for } i = 1,2,3,4,5 \quad (12)$$

$$Recall\ C_i = \frac{TP_i}{TP_i+FN_i} \quad (13)$$

$$F1\ C_i = 2 \times \frac{Precision\ C_i \times Recall\ C_i}{Precision\ C_i + Recall\ C_i} \quad (14)$$

$$Precision_{macro} = \frac{1}{5}\sum_{i=1}^{5} Precision\ C_i \quad (15)$$

$$Recall_{macro} = \frac{1}{5}\sum_{i=1}^{5} Recall\ C_i \quad (16)$$

$$F1_{macro} = \frac{1}{5}\sum_{i=1}^{5} F1\ C_i \quad (17)$$

where TP_i represents the True Positives for class i, FP_i represents the False Positives for class i, TN_i represents the true negatives for class i, and FN_i represents the False Negatives for class i.

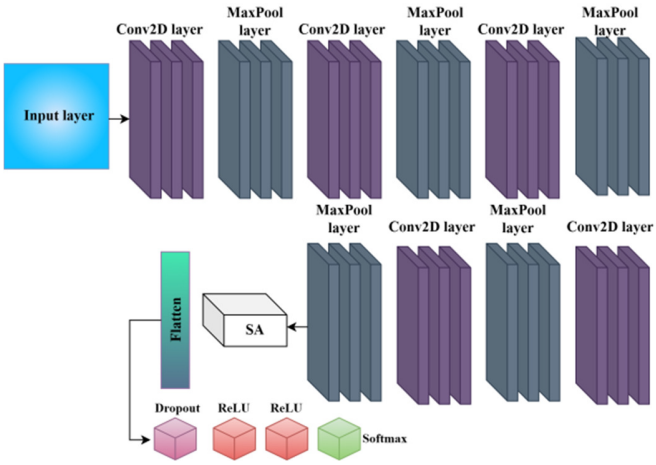


Fig. 3. Architecture of the proposed model for blood cell classification.

IV. EXPERIMENTAL RESULTS

A. Accuracy and Losses

In the training of the proposed model, the standard practice was employed to evaluate performance by plotting losses and accuracies against 25 epochs. These were measured separately for both training and validation datasets to assess both learning and generalization capabilities. A balanced decrease in training and validation losses over the 25 epochs suggests that the model is gradually reducing the error, indicating effective learning and generalization. The detailed presentation is shown in Figure 4, where the "o" marker with a red solid line represents the accuracy of the training data, and the "o" marker with a red dashed line represents the training data losses.

B. Classification Report

The classification report of the validation dataset is displayed in Table I. The results include the precision, recall, F1-score, macro average, weighted average, and support of each class of blood cell. These measures provide a comprehensive view of the model's ability to correctly classify the instances. The validation samples achieved a macro and weighted average precision of 0.98 for each blood cell, indicating a low false positive rate for the validation samples.

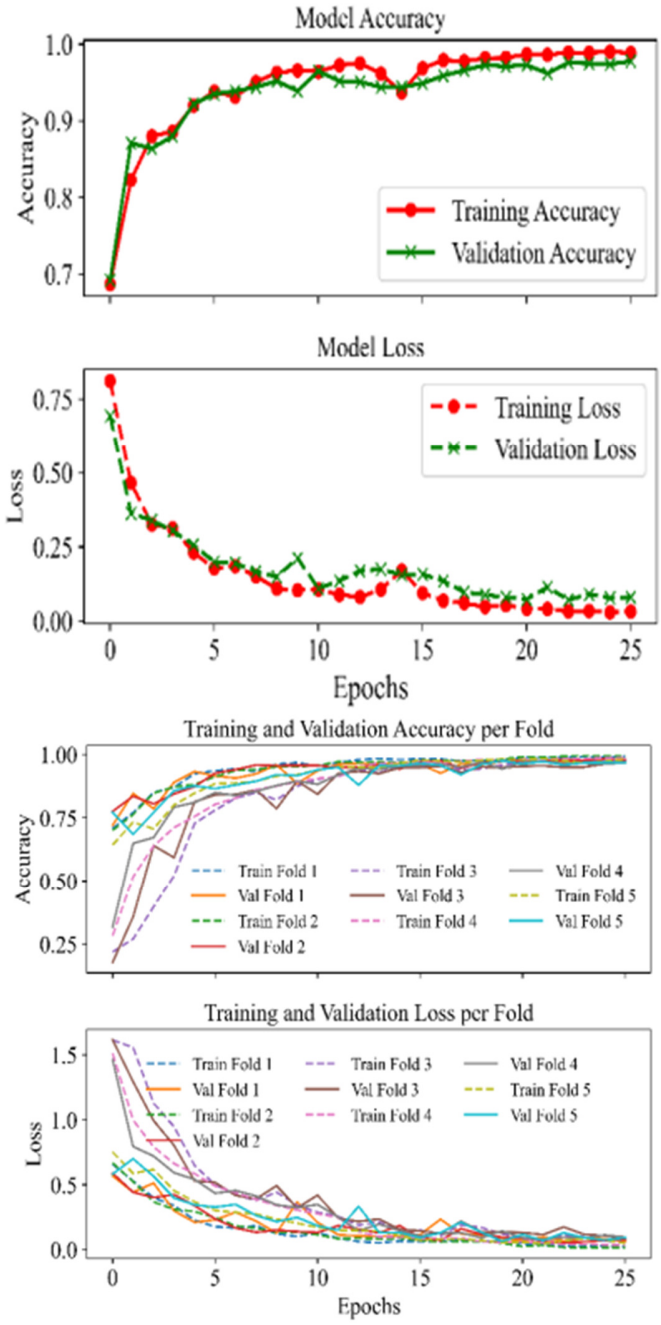


Fig. 4. Accuracy and loss of the proposed model from training validation, k-fold data.

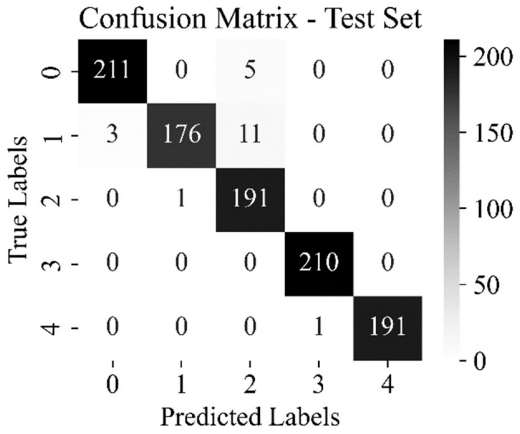
C. Confusion Matrices

Confusion matrices were also calculated for both testing and validation samples, providing a detailed breakdown of the proposed model's performance. A total of 211 microscopic images were correctly classified for class 0 (Basophil class), and 176 microscopic images were correctly classified for class 1 (Erythroblast class). A total of 191 microscopic images were correctly classified for class 2 (Monocyte class). Similarly, 185 microscopic images were correctly classified for class 0, and 208 microscopic images were correctly classified for class 1. The 5×5 confusion matrix of testing and validation samples is given in Figure 5.

TABLE I. CLASSIFICATION RESULTS ON VALIDATION DATA

	Precision	**Recall**	**F1-score**	**Support**
Basophil	0.97	0.99	0.98	186
Erythroblast	1.00	0.95	0.97	220
Monocyte	0.96	0.99	0.98	223
Myeloblast	0.96	1.00	0.98	178
Segmented neutrophil	1.00	0.96	0.98	193
Accuracy			0.98	1000
Macro average	0.98	0.98	0.98	1000
Weighted average	0.98	0.98	0.98	1000

(a)



(b)

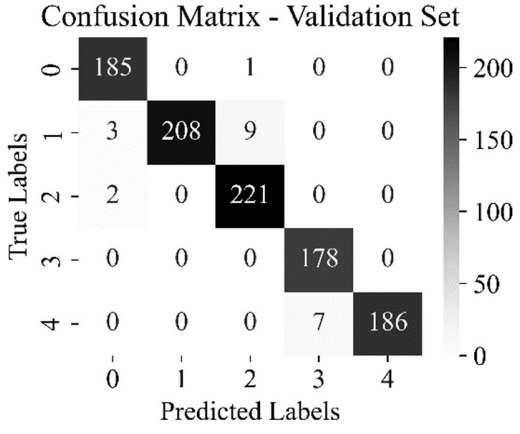


Fig. 5. Confusion matrix for: (a) test set and (b) validation set, where 0: Basophil, 1: Erythroblast, 2: Monocyte, 3: Myeloblast, and 4: segmented neutrophil.

D. Sensitivity Analysis

ROC with AUC was also used to measure the performance of the proposed model in the case of multiclass classification for both training and validation samples. The AUC of the training samples and validation were given in solid and dashed blue, green, red, black, and brown lines, respectively, as portrayed in Figure 7. The model achieved an excellent classification performance for both training and validation samples.

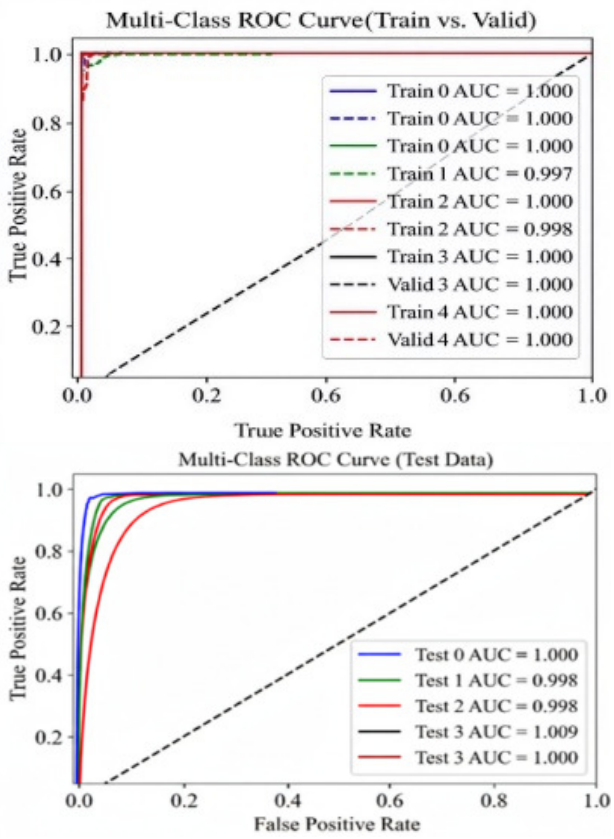


Fig. 6. ROC with AUC of training, validation, and test data, where 0: Basophil, 1: Erythroblast, 2: Monocyte, 3: Myeloblast, and 4: segmented neutrophil.

E. Five-Fold Cross-Validation

The five-fold cross-validation was also applied to evaluate the proposed model's performance and generalizability. It provided a balanced, more stable, and less computationally intensive method than the leave-one-out method. Accuracies of 99.25%, 99.54%, 98.25%, 98.67%, and 98.33% were achieved by the proposed model with 5-fold cross-validation using the training samples. Similarly, accuracies of 97.7%, 97.9%, 97.7%, 97.5%, and 98.1% were achieved by the proposed model with 5 folds, respectively, using the validation samples. The average accuracy of the 5-fold cross-validation using training samples was 98.80%. The average loss of the 5-fold cross-validation using training samples was 0.0353. Similarly, the average loss of the 5-fold cross-validation using validation samples was 0.0837, as presented in Table II.

TABLE II. FIVE-FOLD CROSS-VALIDATION RESULTS

Folds	Training accuracy	Validation accuracy	Training loss	Validation loss
1	0.9925	0.977	0.0241	0.0667
2	0.9954	0.979	0.0152	0.0757
3	0.9825	0.977	0.0529	0.0883
4	0.9867	0.975	0.0349	0.0963
5	0.9833	0.981	0.0496	0.0916
Mean	0.9880	0.977	0.0353	0.0837

The classification results for the testing dataset are displayed in Table III. The testing samples achieved a macro and weighted average precision of 0.98 for each blood cell, indicating a low false-positive rate for the testing samples. Testing samples achieved a macro and weighted average recall of 0.98 for each blood cell, indicating a low false-negative rate for testing samples.

TABLE III. CLASSIFICATION RESULTS OF TESTING DATA

	Precision	Recall	F1-score	Support
Basophil	0.99	0.98	0.98	216
Erythroblast	0.99	0.93	0.96	190
Monocyte	0.92	0.99	0.96	192
Myeloblast	1.00	1.00	1.00	210
Segmented neutrophil	1.00	0.99	1.00	192
Accuracy			0.98	1000
Macro average	0.98	0.98	0.98	1000
Weighted average	0.98	0.98	0.98	1000

Table IV presents the comparison between the proposed model and different existing methods.

TABLE IV. COMPARISON OF THE PROPOSED MODEL WITH EXISTING METHODS

Model	Classes	Accuracy	Ref.
DL	Neutrophils, eosinophils, basophils, lymphocytes, monocytes, and immature granulocytes	91.37% to 94.72%	[4]
DL	Neutrophils, eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets	98.79 % to 99.91 % (Best)	[21]
MobileNetV2, custom CNN	Neutrophil, Lymphocyte, Monocyte, Eosinophil, and Basophil	0.9986	[22]
CNN Ensemble	Basophil, Eosinophil, Lymphocyte, Monocyte, and Neutrophil	94.7% to 97%	[23]
Hybrid with RF,	Healthy cells, lymphoblasts, and myeloblasts	88% (Best)	[24]
Proposed model	Basophil, Erythroblast, Monocyte, Myeloblast, and Segmented neutrophil	98% to 99%	

V. CONCLUSION

This study evaluated an IoT-based model designed for classifying blood cell images. The five-fold cross-validation approach was applied to the training and validation phases to enhance the generalizability and reliability of the proposed model. The latter's performance was evaluated through testing data, further demonstrating its practical utility. The model's classification ability was demonstrated using a microscopic image dataset. Overall, the proposed model achieved accuracies of 99%, 98%, and 98% with balanced precision and recall across five classes of training, validation, and testing samples. These results also matched those obtained from a five-fold cross-validation approach, which demonstrates the classification reliability of the proposed model. The consistent results across training, validation, and testing phases confirm the effectiveness of the proposed IoT-based classification approach. However, the study has some limitations, including the reduction of resolution from 1024×1024 to 100×100, which may cause loss of some features.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

CONFLICT OF INTEREST

The authors declare no conflicts of interest to report regarding the present study.

ACKNOWLEDGMENT

The authors are thankful to CYBEX and Prince Sultan University, Riyadh Saudi Arabia for APC support.

REFERENCES

[1] N. Abbas *et al.*, "*Plasmodium* Species Aware Based Quantification of Malaria Parasitemia in Light Microscopy Thin Blood Smear," *Microscopy Research and Technique*, vol. 82, no. 7, pp. 1198–1214, Jul. 2019, https://doi.org/10.1002/jemt.23269.

[2] A. Rehman, T. Mahmood, and T. Saba, "Robust Kidney Carcinoma Prognosis and Characterization Using Swin-ViT and DeepLabv3+ with Multi-model Transfer Learning," *Applied Soft Computing*, vol. 170, p. 112518, Feb. 2025, https://doi.org/10.1016/j.asoc.2024.112518.

[3] N. Abbas, Z. Mohamad, H. Abdullah, and A. Altameem, "Clustered Red Blood Cells Splitting via Boundary Analysis in Microscopic Thin Blood Smear Digital Images," *International Journal of Technology*, vol. 6, no. 3, p. 306, Jul. 2015, https://doi.org/10.14716/ijtech.v6i3.522.

[4] V. E. Espinoza and P. D. Emmady, *Histology, Monocytes*. Treasure Island, FL, USA: StatPearls Publishing, 2023.

[5] A. Achir, I. Debbarh, N. Zoubir, I. Battas, H. Medromi, and F. Moutaouakkil, "Advances in Leukemia Detection and Classification: A Systematic Review of AI and Image Processing Techniques," *F1000Research*, vol. 13, p. 1536, Dec. 2024, https://doi.org/10.12688/f1000research.159318.1.

[6] M. Karimi *et al.*, "Feature Selection Methods in Big Medical Databases: A Comprehensive Survey," *International Journal of Theoretical & Applied Computational Intelligence*, pp. 181–209, 2025, https://doi.org/10.65278/IJTACI.2025.21.

[7] A. Rehman, N. Abbas, T. Saba, Z. Mehmood, T. Mahmood, and K. T. Ahmed, "Microscopic Malaria Parasitemia Diagnosis and Grading on Benchmark Datasets," *Microscopy Research and Technique*, vol. 81, no. 9, pp. 1042–1058, Sept. 2018, https://doi.org/10.1002/jemt.23071.

[8] S. Nasir, M. Bilal, and H. Khalidi, "Detection and Classification of Skin Cancer by Using CNN-Enabled Cloud Storage Data Access Control Algorithm Based on Blockchain Technology," *International Journal of Theoretical & Applied Computational Intelligence*, pp. 146–159, 2025, https://doi.org/10.65278/IJTACI.2025.31.

[9] T. Saba, A. Sameh, F. Khan, S. A. Shad, and M. Sharif, "Lung Nodule Detection Based on Ensemble of Hand Crafted and Deep Features,"

Journal of Medical Systems, vol. 43, no. 12, p. 332, Dec. 2019, https://doi.org/10.1007/s10916-019-1455-6.

[10] B. Shehu Aliyu, J. Isuwa, A. Abdulrahim, M. Abdullahi, I. Hayatu Hassan, and T. Sikder Momi, "Enhanced Feature Selection for Imbalanced Microarray Cancer Gene Classification Using Chaotic Salp Swarm Algorithm," *International Journal of Theoretical & Applied Computational Intelligence*, pp. 259–283, 2025, https://doi.org/10.65278/IJTACI.2025.16.

[11] H. Ramoser, V. Laurain, H. Bischof, and R. Ecker, "Leukocyte Segmentation and Classification in Blood-Smear Images," in *IEEE Engineering in Medicine and Biology 27th Annual Conference*, Shanghai, China, 2005, pp. 3371–3374, https://doi.org/10.1109/IEMBS.2005.1617200.

[12] G. M. Devi and V. Neelambary, "Computer-Aided Diagnosis of White Blood Cell Leukemia Using VGG16 Convolution Neural Network," in *2022 4th International Conference on Inventive Research in Computing Applications*, Coimbatore, India, Sept. 2022, pp. 1064–1068, https://doi.org/10.1109/ICIRCA54612.2022.9985611.

[13] M. Toğaçar, B. Ergen, and Z. Cömert, "Classification of White Blood Cells Using Deep Features Obtained from Convolutional Neural Network Models Based on the Combination of Feature Selection Methods," *Applied Soft Computing*, vol. 97, p. 106810, Dec. 2020, https://doi.org/10.1016/j.asoc.2020.106810.

[14] A. Bhattacherjee, V. Gupta, M. Kaur, S. Mehta, E. Jain, and P. Kaushik, "Blood Cell Cancer Classification Using ResNet50 and Transfer Learning: Enhancing Diagnostic Accuracy in Hematology," in *2024 International Conference on Decision Aid Sciences and Applications*, Manama, Bahrain, Dec. 2024, pp. 1–6, https://doi.org/10.1109/DASA63652.2024.10836321.

[15] Md. I. H. Abir, S. D. Anik, and H. I. Peyal, "Explainable Multiclass Blood Cell Classification: Combining Custom CNNs with SHAP," in *2025 International Conference on Electrical, Computer and Communication Engineering*, Chittagong, Bangladesh, Feb. 2025, pp. 1–4, https://doi.org/10.1109/ECCE64574.2025.11012986.

[16] T. Saba, S. T. F. Bokhari, M. Sharif, M. Yasmin, and M. Raza, "Fundus Image Classification Methods for the Detection of Glaucoma: A Review," *Microscopy Research and Technique*, vol. 81, no. 10, pp. 1105–1121, Oct. 2018, https://doi.org/10.1002/jemt.23094.

[17] R. Baluabid, H. Alnasri, R. Alowaybidi, R. Hafiz, A. Alsini, and M. Alharbi, "Detecting Acute Lymphocytic Leukemia in Individual Blood Cell Smear Images," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19167–19173, Feb. 2025, https://doi.org/10.48084/etasr.9123.

[18] S. S. Kothwal, "Blood Cell Images for Cancer Detection." Kaggle, Jan. 2025, [Online]. Available: https://www.kaggle.com/datasets/sumithsingh/blood-cell-images-for-cancer-detection.

[19] S. Iftikhar *et al.*, "An Evolution Based Hybrid Approach for Heart Diseases Classification and Associated Risk Factors Identification," *Biomedical Research*, vol. 28, no. 8, pp. 3451–3455, 2017.

[20] K. T. Navya, K. Prasad, and B. M. K. Singh, "Analysis of Red Blood Cells from Peripheral Blood Smear Images for Anemia Detection: A Methodological Review," *Medical & Biological Engineering & Computing*, vol. 60, no. 9, pp. 2445–2462, Sept. 2022, https://doi.org/10.1007/s11517-022-02614-z.

[21] R. Asghar, S. Kumar, and P. Hynds, "Automatic Classification of 10 Blood Cell Subtypes Using Transfer Learning via Pre-trained Convolutional Neural Networks," *Informatics in Medicine Unlocked*, vol. 49, p. 101542, 2024, https://doi.org/10.1016/j.imu.2024.101542.

[22] O. Saidani *et al.*, "White Blood Cells Classification Using Multi-Fold Pre-processing and Optimized CNN Model," *Scientific Reports*, vol. 14, no. 1, p. 3570, Feb. 2024, https://doi.org/10.1038/s41598-024-52880-0.

[23] X. Wang, G. Pan, Z. Hu, and A. Ge, "A Two Stage Blood Cell Detection and Classification Algorithm Based on Improved YOLOv7 and EfficientNetv2," *Scientific Reports*, vol. 15, no. 1, p. 8427, Mar. 2025, https://doi.org/10.1038/s41598-025-91720-7.

[24] S. Kasim *et al.*, "Multiclass Leukemia Cell Classification Using Hybrid Deep Learning and Machine Learning with CNN-Based Feature Extraction," *Scientific Reports*, vol. 15, no. 1, p. 23782, Jul. 2025, https://doi.org/10.1038/s41598-025-05585-x.