Science Journal of University of Zakho
DOI: https://doi.org/10.25271/sjuoz.2026.14.3.1701

Original Article

SE-CNN: A LIGHTWEIGHT DEEP LEARNING MODEL FOR ACUTE LYMPHOBLASTIC LEUKEMIA CLASSIFICATION

Awaz Mustafa Abbas[1,*], Maiwan Bahjat Abdulrazaq [1], and Adel AL-Zebari [2]

[1]Department of Computer, College of Science, University of Zakho, Zakho, Kurdistan Region, Iraq.
[2]Department of Information Technology, Technical College of Informatics, Akre University for Applied Sciences, Akre, Kurdistan Region, Iraq.

[*]*Corresponding author, E-mail:*awaz.abbas@uoz.edu.krd (Tel: +234-8064459564)

Received
July 14, 2025

Accepted
August 11, 2025

Published
July 11, 2026

ABSTRACT

Leukemia is a potentially fatal hematological cancer that affects people of all ages and remains one of the leading causes of cancer death worldwide. It mostly impacts white blood cells (WBCs), which leads to the aberrant growth of immature lymphocytes that disrupt the bloodstream and bone marrow. Convolutional Neural Networks (CNNs), a type of deep learning, have become a potent instrument in medical image analysis in recent years, offering significant potential for automated and accurate disease classification. In this study, we present a lightweight CNN architecture enhanced with Squeeze-and-Excitation (SE) blocks for the classification of leukemia from peripheral blood smear images. The proposed model effectively captures channel-wise interdependencies to dynamically recalibrate feature responses, thereby improving the discrimination of relevant features. There were two publicly accessible datasets used for the experiments: C-NMC-2019 and Blood Cells Cancer (Acute Lymphoblastic Leukemia – ALL). The proposed model achieved an accuracy of 97.12% on the C-NMC-2019 dataset (binary classification) and 99.79% on the Blood Cells Cancer (ALL) dataset (multi-class classification). The results demonstrate that the proposed architecture achieves a strong trade-off between computational efficiency and classification accuracy, indicating its suitability for real-time, computer-aided leukemia diagnosis in clinical settings.

Keywords: Squeeze and Excitation, Convolutional Neural Network, Medical image classification, Acute Lymphoblastic Leukemia, Peripheral blood smear.

1. INTRODUCTION

Cancer is among the most deadly diseases, impacting human lives across the world and spreading a shadow across the globe at large (Ansari *et al*., 2018; Abioye *et al*., 2024). Among the various forms of cancer, blood cell cancer is considered one of the major types due to its severity and prevalence. Blood cancer deriving from bone marrow is called leukemia, which leads to excessive production of atypical white blood cells. Hence, weakening the immune system and reducing platelet and red blood cell (Ananth *et al*., 2023; Bibi *et al*., 2020; Onciu, 2009; Sampathila *et al*., 2022). The subtypes of leukemia are Chronic Leukemia (CL) and Acute Leukemia (AL). The progression of CL is usually slow. Alternatively, without specialized treatment, the average life expectancy of patients afflicted with AL is only 3 months (Rahman *et al*., 2023).

There are two subcategories of AL as classified by the French-American-British (FAB) classification system: Acute Myeloid Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL). Moreover, Chronic Lymphocytic Leukemia (CLL) and Chronic Myeloid Leukemia (CML) are the two separate forms of CL (Attallah & Omneya, 2024; Bennett *et al*., 1976; Das *et al*., 2020). Since ALL is one of its more aggressive variants, improving survival rates requires a prompt and precise diagnosis (Muhsin *et al*., 2024). Occasionally, the early diagnosis is complicated by various circumstances in clinical practice because bone marrow examination, a time-consuming and specialized test, is required (Chiaretti *et al*., 2014). The performance of automated medical image analysis has improved considerably (Anwar *et al*., 2018). CNNs have shown strong capabilities in feature extraction (Shamama & Afrin, 2020) and classification tasks (Agustin *et al*., 2021) , thus being effective tools in the analysis of blood cell images and the detection of leukemia (Saeed *et al*., 2024). However, numerous models are costly in terms of computation, which restricts their application to real-world systems such as small clinics, mobile health devices, and so on. This research has made the following contributions.

- Development of a lightweight CNN architecture in both multiclass and binary classification tasks, addressing the need for efficient and accurate leukemia detection in medical imaging.
- The developed model is computationally efficient, which balances high classification accuracy with significantly fewer parameters, enabling potential deployment in low-resource clinical settings.
- The integration of SE blocks with SeparableConv2D layers aims to improve the model, effectively focusing on the most relevant features while significantly lowering computational complexity.

Access this article online

DOI:https://doi.org/10.25271/sjuoz.2026.14.3.1701

Science Journal of University of Zakho
Vol. 14, No. 03, pp. 512-522, July -2026

Printed ISSN 2410-7549;
Electronic ISSN 2414-6943

This research is organized as follows. Section 2 reviews previous research on leukemia classification and deep learning techniques. Section 3 details the proposed lightweight CNN model, datasets, and implementation approach. The results are presented and discussed in Section 4, and then the study is concluded and future research directions are outlined in Section 5.

2. 2. LITERATURE REVIEW

Deep learning has substantially advanced the field of medical image analysis (Suzuki, 2017). In recent years, these advancements have been particularly notable in the context of automated leukemia classification from PBS images, demonstrating promising results in clinical decision support (Erten *et al*., 2024). CNNs have achieved remarkable success in diagnosing hematologic malignancies by automatic learning of hierarchical feature representations, thereby circumventing the reliance on manual feature engineering and enhancing the accuracy of classification. Recent developments in this field are reviewed critically below. In binary and multiclass classification settings, researchers have developed a range of lightweight and hybrid approaches that have emerged.

Multiclass Classification Deep Learning Models:

Rahman *et al*., (2023) proposed a hybrid machine and deep learning model for multiclass classification of ALL, integrating pre-trained CNNs InceptionV3, VGG19, Xception, and ResNet50, for feature extraction with classical classifiers such as SVM, logistic regression (LR), and random forest (RF). Feature selection was refined using PCA, LDA, and SVC, further optimized by Particle Swarm Optimization (PSO) and Cat Swarm Optimization (CSO). As the best performing, the optimum configuration, ResNet50 with SVC and LR, was able to reach 99.84% accuracy after training on a set of 3,262 images. Despite using lightweight CNNs (2M–5M parameters), the model's complexity, because of the use of multiple selection and classification algorithms, had problems with scaling and overfitting. Expanding on the necessity for lightweight models, a hybrid approach by Nayak *et al*. (2024) was proposed by employing MobileNet as well as a support vector machine (SVM) to classify ALL subtypes. In this case, MobileNet served as an effective feature extractor and SVM as a classifier to yield a 99.3% accuracy. However, the method had difficulty generalizing across datasets.

Another study by Prakash *et al*., (2024) proposed a YOLOv8-based hybrid model that integrated a Residual Convolutional Block Attention Mechanism (RCBAM) and Depthwise Separable Convolutions (DWSCNN). The resulting architecture, comprising only 4 million parameters, was optimized for computational efficiency and rapid inference, achieving a mean Average Precision (mAP) of 98.4% and an F1-score of 96.2%. The model was trained on 7,535 annotated images representing four stages of leukemia, with data augmentation and gamma correction to improve the contrast and increase generalization. While successful, the model did struggle to some extent with class imbalance and was not very demographically adaptable. Alim *et al*., (2024) presented a fine-tuned ResNet-50 model using five-fold cross-validation and standard data augmentation. The model achieved an accuracy of 99.38%, outperforming VGG-16, DenseNet-121, and EfficientNetB0, however, despite its robust performance without segmentation, its 24.8 million parameters (of which 1.2 million were trainable) posed computational challenges.

To improve the preservation of morphological features, Dubai *et al*., (2025) proposed MCB-HyperNet, a hybrid DL model based on ResNet and a MobileNetV3 with novel Morphological Context Blocks (MCBs). Trained on the ALL dataset, containing 3256 images, the model got an accuracy of 99.69% with a relatively lightweight characteristic of MobileNetV3. While MCB-HyperNet effectively minimized false positives and negatives, its morphological approximations brought complications and some risk of overfitting. In a related effort to improve interpretability, Dutta *et al*., (2025) presented an attention-based CNN model for better interpretability, achieved 99% accuracy, and used visual explanation techniques like LIME and SHAP. While the model demonstrated strong performance, its practical deployment was limited by a high sensitivity to hyperparameter tuning and a dependence on high-quality annotated data factors that pose significant challenges in real-world clinical settings.

Binary classification deep learning models:

Extending the ensemble learning techniques, Kasani *et al*. (2020) proposed an ensemble deep model that utilizes NASNetLarge and VGG19, and its performance for the classification of leukemic B-lymphoblast cells was evaluated on C NMC-2019 dataset. The original dataset was composed of 10661 images, and was augmented to 70,000 samples due to the class imbalance. Based on transfer learning, fine-tuning of CNN and ensemble methods was also adopted to improve the classification performance. NASNetLarge+VGG19 ensemble, with around 111 million parameters, achieves the top accuracy of 96.58%. The treatment also remained computationally efficient, with each image taking approximately 0.03 s to be processed. However, it performed poorly in distinguishing healthy and leukemic cells due to the visual similarity among them, suggesting that more discriminative feature extraction is needed. In the spirit of ensemble methods Chen *et al*. (2021) introduced ResNet101-9, an ensemble of nine fine-tuned ResNet-101 models. The model was trained on a total of 12,528 C-NMC images, with the approach of majority voting and the Taguchi optimal method, and obtained an F1-score of 88.94% and an accuracy of 85.11%. Although the method made successful use of transfer learning and had a low number of parameters, its performance relied on careful hyperparameter selection, and it was sensitive to mislabeled samples when learning from the data.

A custom CNN architecture named ALL-NET was proposed by Sampathila *et al*. (2022). Trained on 10,661 images from the same dataset and augmented to 12,000 images to address class imbalance, this network comprised four convolutional layers and three dense layers. It was able to obtain 95.54% accuracy and 95.43% F1 score. Nevertheless, while achieving state-of-the-art results, the proposed model was limited by extensive preprocessing and computation, and was susceptible to noise and overfitting. To address the computational demands of deeper architectures, Jawahar *et al*. (2022) proposed ALNet, a lightweight CNN with only 535,010 parameters. It was trained on ISBI 2019 C-NMC dataset images resized to 224×224×3. The model featured a compact design with clustered convolutional and pooling layers and achieved 91.13% accuracy, 93% precision, 98% recall, and a 96% F1-score. Even though ALNet achieved better performance than many pretrained deep learning models in classification efficiency, it still had a small limitation for generalization to unseen data.

To further improve performance through multi-model integration, Hasanaath *et al*. (2024) presented a two-stage ensemble method. This method extracted deep features using InceptionResNetV2, VGG16, and DenseNet121, then classified using SVM. The model achieved an F1-score of 90.14% and accuracy of 91.63% of 10,661 C-NMC dataset images when preprocesses such as resizing and normalizing were applied. Although the complexity of those approaches was increased by the feature fusion, it introduced a large computational cost and degraded interpretability. In a study by Darmawan and Shabrina (2025), an interpretable deep learning model combining EfficientNetV2B3 with LIME was proposed for the diagnosis of ALL, using the C-NMC-2019 dataset. The dataset comprised 12,528 images, which were resized to 300×300 pixels and augmented to mitigate class imbalance. The system achieved 97.95% accuracy, precision, recall, and F1-score. To ensure interpretability, LIME was employed to visualize image regions influential in the model's predictions. Although having an overall good performance, the model occasionally misclassified the healthy cells, due to morphological similarity with the leukemic ones.

The literature reveals that Recent deep learning models include both multiclass and binary ALL classification models with high accuracies ranging from 91% to 99.84% on the classification of ALL from PBS images. Yet, several challenges in the fields like class imbalance, insufficient generalization, complex pre-processing, and deployment constraints still exist. These limitations underscore the need for lightweight, robust models suitable for real-world clinical applications.

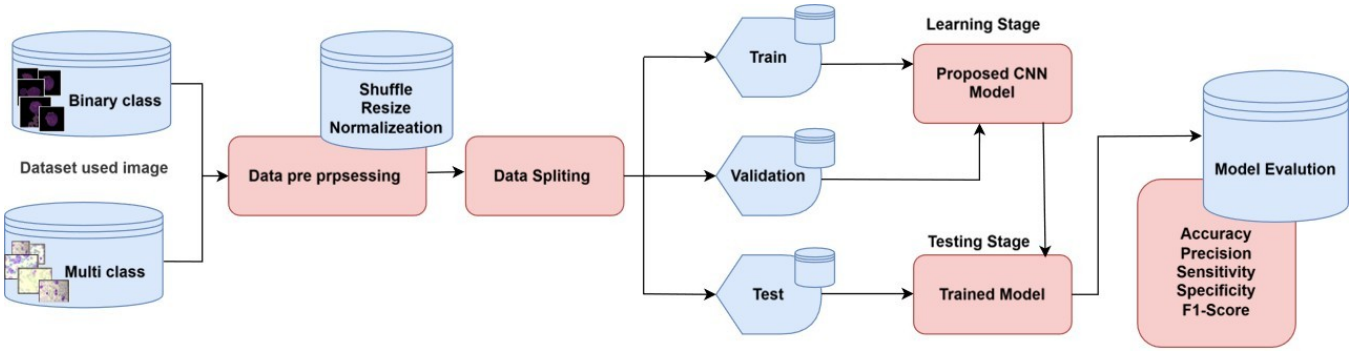


Figure 1: Proposed methodology for this research.

3. MATERIALS AND METHODS

The methodology employed for classifying blood cell cancers involves four key stages: Dataset description, pre processing of the images, proposed CNN structure design, and model evaluation. Figure 1 illustrates the broad model of the proposed method. The following sections present a detailed account of all the stages.

Dataset Description:

The proposed method was evaluated using two publicly available datasets from the Kaggle platform, one designed for multiclass classification and the other for binary classification. A detailed description of each dataset is provided below.

Blood Cells Cancer (ALL):

A dataset comprising of 3,242 images was collected from 89 individuals who were thought to have ALL and was used in this investigation (Hosseini *et al*., 2023). The images were prepared in the Taleqani Hospital's bone marrow lab in Tehran, Iran, where skilled lab workers processed and stained blood samples. The dataset consists of two main classifications: benign (hematogenous) and malignant. The malignant class is further subdivided into three ALL subtypes: Pro-B, Early Pre-B, and Pre-B lymphoblasts. Figure 2(A) shows images captured at 100× magnification using a Zeiss microscope-mounted camera and saved as JPGs. Cell types and also subtypes were definitively classified by a hematology expert using flow cytometry.

C-NMC-2019:

The C-NMC-2019 dataset (Mourya *et al*., 2019) is compiled by the University of Arkansas for Medical Sciences. It consists of 10,661 labeled PBS images collected from 73 subjects, comprising 26 healthy people and 47 people with an ALL diagnosis. The data is split into two groups: ALL (cancerous) and normal (non-cancerous). Among the total images, 7,272 (approximately 68%) correspond to leukemic cells, while 3,389 (approximately 32%) represent normal cells, as shown in Figure 2 (B). The dataset creators applied preprocessing techniques to minimize common issues in microscopic imaging, such as staining variability and illumination inconsistencies, ensuring improved image clarity and reliability.

Image Pre Processing:

Preprocessing of images is important for enhancing model performance and lowering computation overhead. To start with, each image was resized to 256×256 pixels for consistency within the dataset and batch processing. The pixel values were then normalized after resizing by dividing by 255 to rescale RGB color values into the [0, 1] range. This normalization step is required to speed up the convergence of the training of neural network and to have a more stable and efficient learning behavior. After preprocessing and normalization, the dataset was divided with the ratio of 70% to 15% to 15% for training, testing, and validation, as indicated in Table 1. This division minimizes overfitting and enables objective performance evaluation on unobserved data.

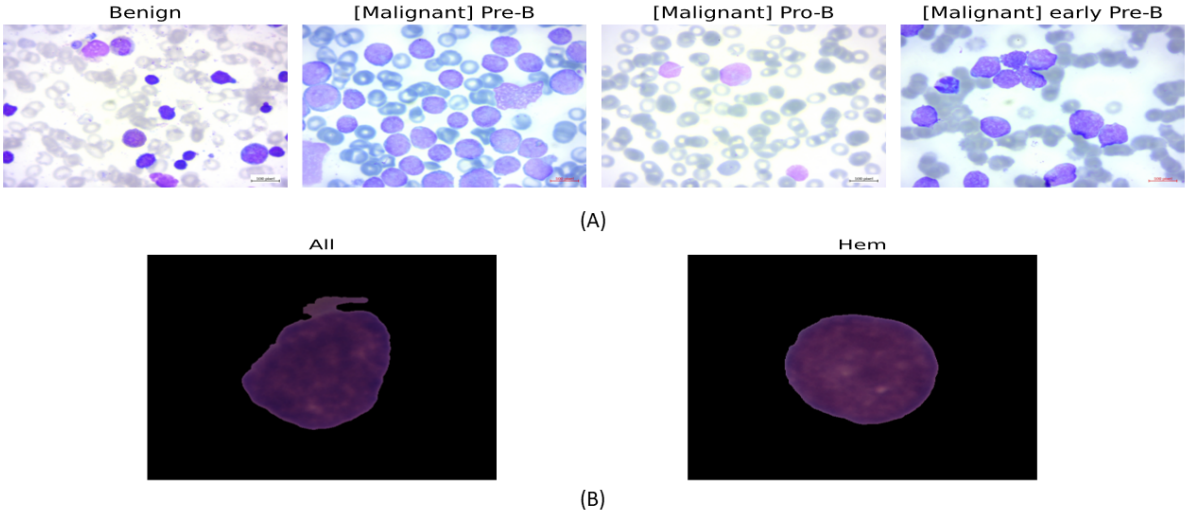


Figure 2: Representative PBS images from the datasets utilized in this study: (A). Sample PBS images of the Blood Cells Cancer (ALL) dataset, (B). Sample PBS images of C-NMC-2019 dataset.

Table 1: Number of images utilized for training, validation, and testing sets in Blood Cell Cancer (ALL) and C-NMC-2019 datasets.

	Class Label	Training	Validation	Testing
Blood Cells Cancer (ALL) (Multi-class)	Benign	358	77	77
	Pre-B	669	143	143
	Pro-B	557	120	119
	Early Pre-B	685	147	147
C-NMC-2019 (Binary)	All	5090	1091	1091
	Hem	2372	509	508

Squeeze and Excitation Blocks:

The proposed architecture incorporates the Squeeze-and-Excitation (SE) block, a lightweight yet effective module designed to boost representational capacity by explicitly modeling inter-channel dependence (Hu *et al*., 2018). As seen in Figure 3, the SE block adaptively recalibrates channel-wise feature responses given an input feature map tensor X□R W×H×C, where W and H stand for spatial dimensions (width and height) and C for the number of channels.

The procedure initiates with a squeeze operation in which global average pooling (GAP) is executed on each channel, generating a descriptor S□R 1×1×C, that encodes the global context of each channel. This descriptor is then fed into two fully connected (FC) layers: the first FC layer reduces the channel dimension by a factor r=8, then a Swish activation, and the second FC layer restores the original dimension C with normalized channel weights in the [0,1] range interval that are obtained by sigmoid activation.

These weights are reshaped and broadcasted to match the dimensions of the original feature map and are used to recalibrate the input via channel-wise multiplication. By means of this attention mechanism, the network can selectively emphasize informative features and suppress irrelevant or noisy ones, thereby improving the feature discriminability in downstream convolutional layers.

Proposed architecture (SE-CNN):

Table 2 summarizes the proposed CNN architecture designed for binary and also multiclass classification of PBS images. The model accepts RGB images resized to 256×256×3 as input. The feature extraction backbone consists of four sequential convolutional blocks.

The first block employs a SeparableConv2D layer to reduce computational complexity while preserving spatial feature extraction capabilities, thereby enabling efficient model performance with fewer parameters compared to a standard convolutional layer. SeparableConv2D layer with 64 filters of size 3×3 and a stride of 1. This is followed by a BatchNormalization (BN) layer, where the momentum parameter set to 0.9 controls the moving average update rate of the mean and variance during training. Subsequently, the Swish activation function is applied to enable smooth and non-monotonic nonlinear transformations.

Then, an SE block with a reduction ratio of 8 is integrated to enhance channel-wise attention (detailed in Section 3.3). Downsampling is achieved using a max pooling layer with a pool size of 2×2, and 'same' padding, as depicted in Figure 4.

The second block follows the same architectural pattern but raises the quantity of filters to 256. The third block continues similarly, employing 186 filters and still holding the SE mechanism. In the fourth block, the number of filters is set to 156, but the SE block is intentionally omitted to reduce computational overhead in deeper layers. Each block maintains a consistent kernel size (3×3) and He-normal initialization to promote stable training.

Following the convolutional stages, a GAP layer is employed to reduce spatial dimensions and feed high-level features into a dense layer with 128 neurons activated by Swish. A dropout layer with a drop rate of 0.3 is used to avoid overfitting. Finally, an FC layer with a single neuron and sigmoid activation is employed for binary classification. In the multiclass classification setup, the sigmoid output layer is replaced by a dense layer with four output neurons activated by a softmax function. This architecture maintained a low parameter count while achieving an optimal trade-off between accuracy and computational efficiency through the integration of separable convolutions and attention mechanisms.

Table 2: Overview of the SE-CNN model architecture

Stage	Type of layer	Param	Shape
Input	Input Layer	0	(None, 256, 256, 3)
Conv Block 1	SeparableConv2D+ BN	539	(None, 256, 256, 64)
	Activation(swish)	0	(None, 256, 256, 64)
	SE	1096	(None, 256, 256, 64)
	MaxPooling2D	0	(None, 128, 128, 64)
Conv Block 2	SeparableConv2D+ BN	18240	(None, 128, 128, 256)
	Activation(swish)	0	(None, 128, 128, 256)
	SE	16672	(None, 128, 128, 256)
	MaxPooling2D	0	(None, 64, 64, 256)
Conv Block 3	SeparableConv2D+ BN	50850	(None, 64, 64, 186)
	Activation(swish)	0	(None, 64, 64, 186)
	SE	8765	(None, 64, 64, 186)
	MaxPooling2D	0	(None, 32, 32, 186)
Conv Block 4	SeparableConv2D+ BN	31470	(None, 32, 32, 156)
	Activation(swish)	0	(None, 32, 32, 156)
	MaxPooling2D	0	(None, 16, 16, 156)
Classification Head	GlobalAveragePooling2	0	(None, 156)
	Dense	20096	(None, 128)
	Dropout	0	(None, 128)
	Dense_1	516 (Multiclass Classification) 129 (Binary Classification)	(None, 4)
Total params (Multiclass Classification)		148244	
Total params (Binary Classification)		147857	

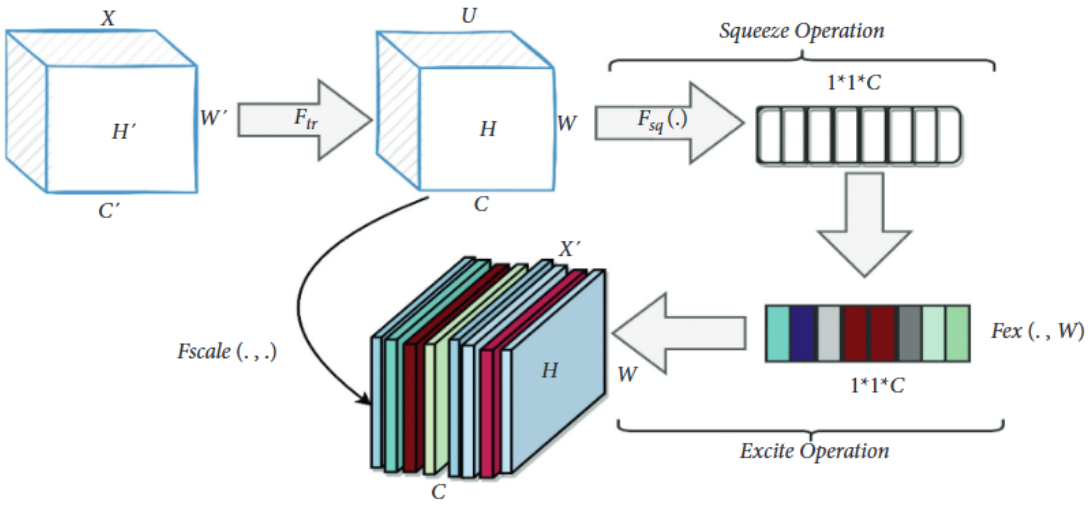


Figure 3: A pictorial representation of the SE operation (Bukhari *et al*., 2022).

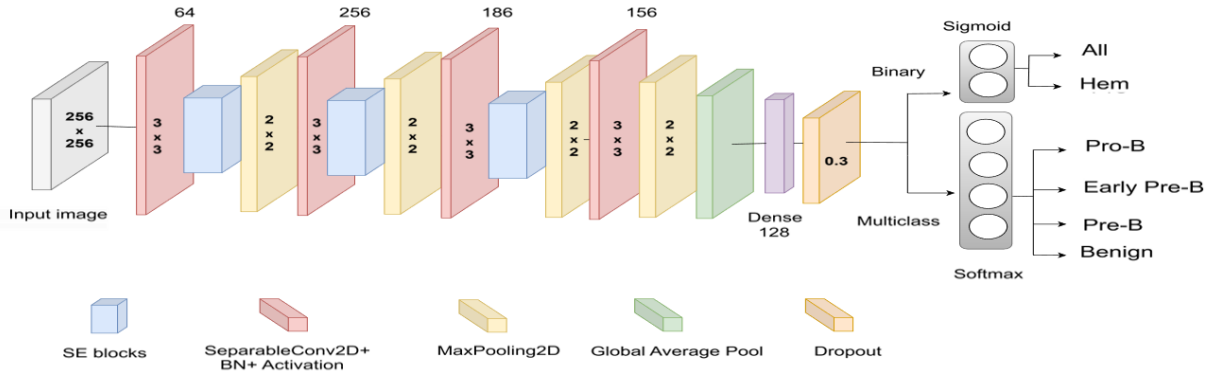


Figure 4: Architecture of the SE-CNN model.

Model training:

The training of the SE-CNN method is carefully designed to achieve a balance between the convergence speed, classification accuracy, and generalization ability such that the learning is efficient and the model performance is optimal. The training process starts with data preprocessing, which is designed to improve the quality and resilience of input samples in order to make the model less sensitive to the input samples. Then, the network parameters are optimized through a well-structured training pipeline that incorporates adaptive optimization techniques and dynamic learning rate scheduling, making the approach suitable for both multiclass as well as binary classification tasks.

The Nadam optimizer is used to train the SE-CNN model, which is the combination of adaptive moment estimation and the momentum of Nesterov Accelerated Gradient optimizer, and is suitable for handling sparse gradients and noisy data (Salehi *et al*., 2020). In addition, a Triangular Cyclical Learning Rate (CLR) schedule is used to help accelerate training and improve convergence. This learning rate technique allows the optimizer to explore the loss surface more effectively, helping to avoid local minima and achieve better performance.

Both datasets exhibited reasonably balanced class distributions, therefore, no specific imbalance-handling techniques were applied. The model is trained with 60 epochs. For binary classification tasks, we use the BinaryCrossentropy loss function, for multi-class classification tasks, we use the Sparse Categorical Crossentropy loss, which allows integer labeled classes. A summary of the complete training configuration can be provided in Table 3.

Table 3: Summary of the respective hyperparameters used for the training and evaluation of the model.

Hyperparameters	Values (Multiclass)	Values (Binary)
Loss Function	sparse_categorical_crossentropy	BinaryCrossentropy
Optimizers	Nadam	Nadam
Epochs	60	60
Batch_Size	32	52
Seed Values	42	42
Triangular Cyclical Learning Rate	initial_learning_rate=1e-6 maximal_learning_rate=1e-3	initial_learning_rate=1e-6 maximal_learning_rate=1e-2

Performance Measures:

This section explains the fundamental metrics required to measure the efficiency of the SE-CNN model. These measures give us the ability to know how well the model distinguishes between leukemic and non-leukemic samples. The assessment consists of four fundamental results:

- True Positive (TP): ALL cases that were correctly identified as ALL.
- True Negative (TN): Healthy cases correctly identified as healthy.
- False Positive (FP): Healthy cases incorrectly predicted as ALL.
- False Positive (FP): Healthy cases incorrectly predicted as ALL.

Accuracy:

Accuracy represents the overall correctness of the model by evaluating the proportion of correctly predicted cases (both ALL and healthy) overall number of cases (Devi *et al*., 2024). It is computed as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

Precision:

Precision evaluates the model's dependability in identifying ALL cases by determining the percentage of true ALL detections out of all cases predicted as ALL (Pushpalatha *et al*., 2024):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

Sensitivity (True Positive Rate/ Recall):

Sensitivity, also referred to as recall or the true positive rate, measures the model's capability detect ALL instances (Brahmaiah *et al*., 2024). It is defined as the proportion of accurately predicted ALL instances to the total number of actual ALL cases, as expressed by the following equation:

$$\text{Sensitivity (Recall)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

Specificity (True Negative Rate):

Specificity assesses the capability of the model to correctly identify healthy individuals by measuring the proportion of non-ALL cases that are accurately classified as negative (Zolfaghari & Sajedi, 2022):

$$\text{Specificity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

F1 Score:

The F1 Score is a single, balanced metric that considers both false positives and false negatives. It is the harmonic average of sensitivity and precision. The F1 Score is especially well suited if there is a class imbalance, as it gives a more meaningful number than accuracy alone (Hasanaath *et al*., 2024), as denoted in the equation:

$$\text{F1 Score} = \frac{2 \times \text{TP}}{(2 \times \text{TP}) + \text{FP} + \text{FN}} \quad (5)$$

4. RESULTS ANALYSIS AND DISCUSSION

The evaluation process of the SE-CNN model was detailed in this section by means of diverse analytical modalities. A preliminary comparison of training and validation accuracy offers a broad overview of the model's performance. Subsequently, the performance measures such as precision, specificity, recall, and F1-score were calculated to provide a detailed evaluation of the classification abilities. This method is also robust and provides fold-wise performance measures as well as classification using the confusion matrix. Finally, the proposed method was compared with several state-of-the-art methods to demonstrate its competitive advantage in the classification of blood cell cancer.

Accuracy and Loss vs. Epoch:

Figure 5 presents the training and validation performance of the SE-CNN model over 60 epochs for both binary and multiclass classification tasks, evaluated in terms of accuracy and loss. The accuracy curves indicate that training accuracy remains consistently high after the initial epochs. Although the validation accuracy exhibits some fluctuations, likely due to dataset complexity and class imbalance, it remains at relatively high levels throughout, demonstrating the model's robust generalization capabilities on unseen data across both tasks.

The corresponding loss curves further support these observations. Training loss steadily decreases and stabilizes at a low value, indicating efficient learning. In contrast, the validation loss shows greater variability, with intermittent spikes that may be attributed to sample-level noise or more challenging instances within the validation set. Nonetheless, the absence of a sustained upward trend in validation loss suggests that the SE-CNN model effectively avoids overfitting and maintains a stable, generalizable learning trajectory.

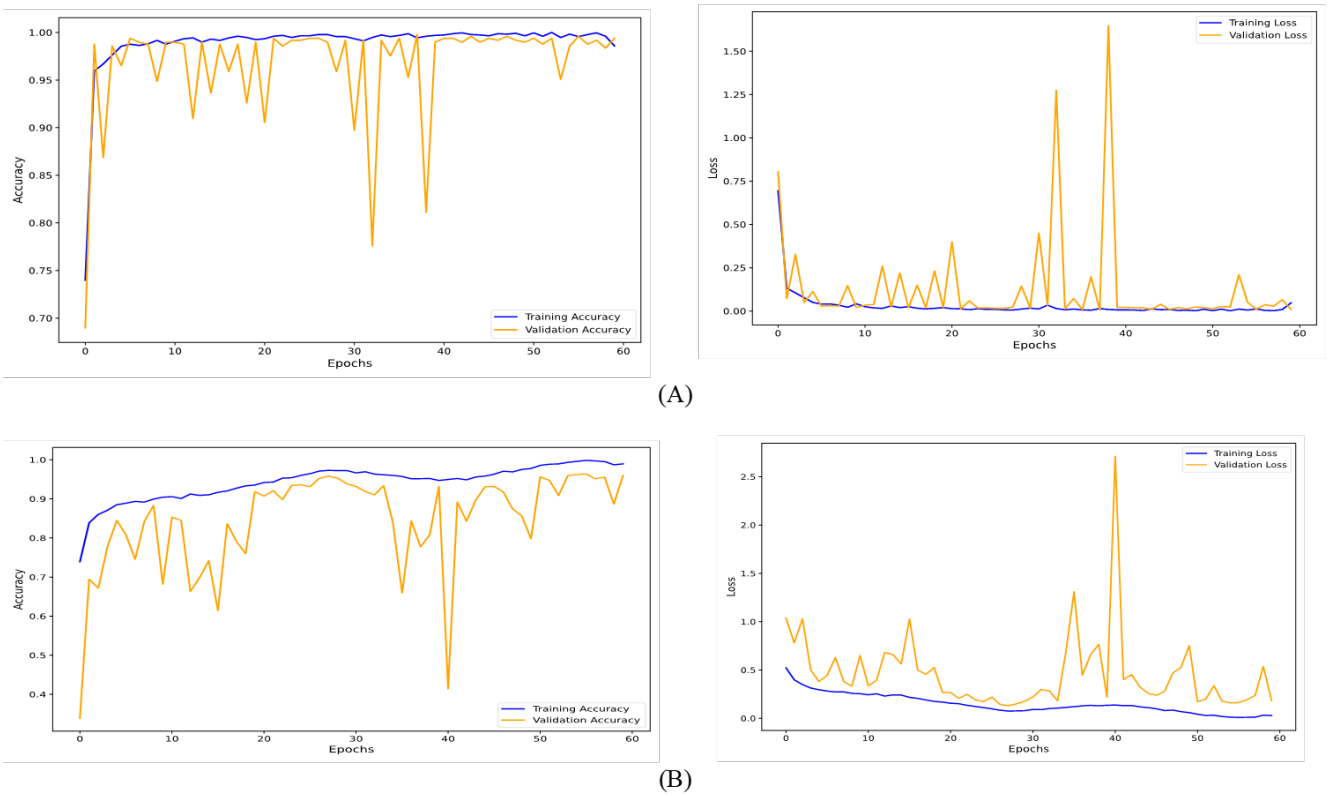


Figure 5: Training and validation accuracy and loss curves: (A) multiclass classification, (B) binary classification.

4.1 Confusion Matrix:

The multiclass and binary class confusion matrices in Figures 6 and 7 provide a more complete view of model performance. In the multiclass task, the model presents high accuracy across training, validation, and also testing datasets, with minimal misclassifications among the four categories: Benign, Early Pre-B, Pro-B, and Pre-B. For instance, in the testing set, only one Benign sample was misclassified as Early Pre-B, while the remaining categories exhibited strong predictive consistency. These results indicate the model effectively distinguishes between closely related ALL subtypes.

In the binary classification, ALL and Healthy, the model achieved high precision and recall with insignificant errors. Specifically, only 1, 20, and 21 ALL samples were misclassified as Healthy in the training, validation, and testing sets. On the other hand, 5, 38, and 25 Healthy samples were incorrectly labeled as ALL. These minor errors correspond to exceptionally low misclassification rates, indicating strong generalization and robust feature learning. Overall, the proposed model performs excellently in both classification tasks, highlighting its capability to distinguish between subtle morphological differences in hematological images.

4.2 Performance Evaluation of SE-CNN Model:

The SE-CNN model was evaluated under both binary (2-class) and multiclass (4-class) classification settings. In the multi-class scenario, the model obtained a test accuracy of 99.79%, with validation and training accuracies of 99.79 % and 99.91%, respectively. Figure 8 illustrates the 4-class model's classification Performance.

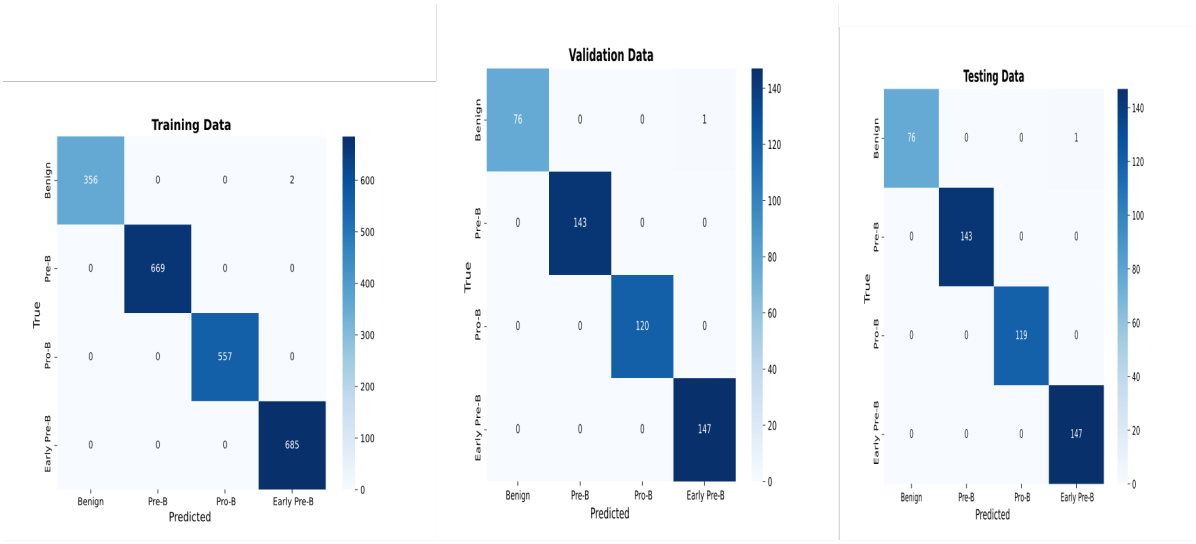


Figure 6: Confusion matrices for SE-CNN multi-class classification.

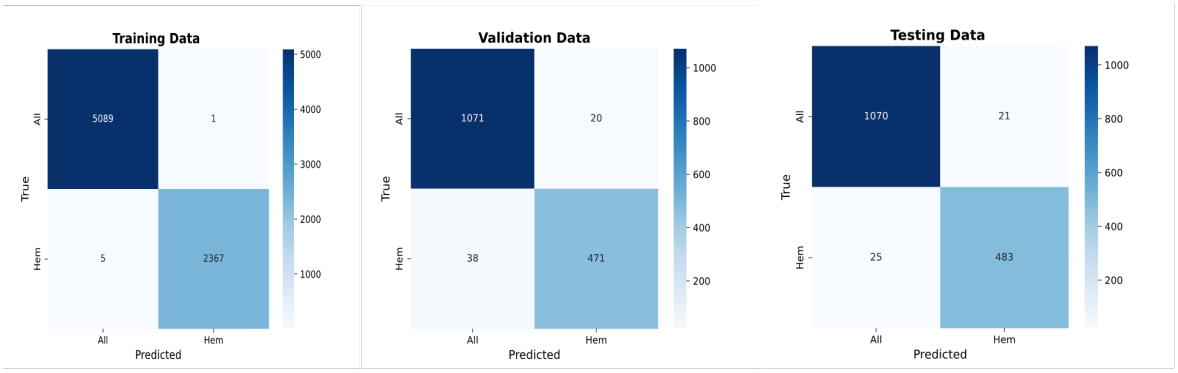


Figure 7: Confusion matrices for SE-CNN binary classification.

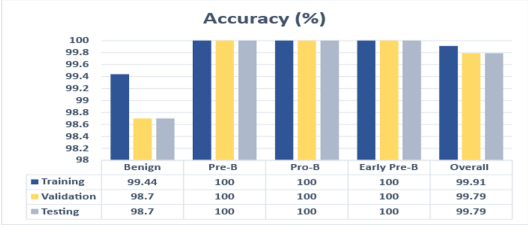


Figure 8: Class-wise and overall accuracy performance of the proposed multi-class classification model.

Figure 9 illustrates that for the binary classification, the model attained a test accuracy of 97.12%, alongside 99.92% training accuracy and 96.37% validation accuracy.

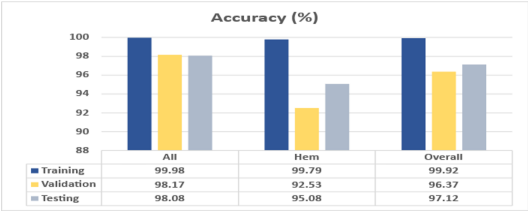


Figure 9: Class-wise and overall accuracy performance of the proposed binary classification model.

Table 4 presents a thorough analysis of additional performance indicators, such as sensitivity, precision, F1-score, and specificity, which further validates the model's efficacy in both classification tasks.

To evaluate the discriminative performance of the SE-CNN model, Receiver Operating Characteristic (ROC) curves were generated for both multiclass and binary classification tasks. As shown in Figure 10(A), the multiclass model obtained an Area Under the Curve (AUC) of 1.0 during the training, validation, and testing phases. This near perfect performance can be attributed to the well structured nature of the Blood Cells Cancer (ALL) dataset, which exhibits clear visual distinctions among the different leukemia subtypes, thereby facilitating effective class separation. Additionally, the incorporation of SE blocks and the application of strong regularization techniques contributed to improved feature recalibration and reduced overfitting, further enhancing the model's class discrimination capability.

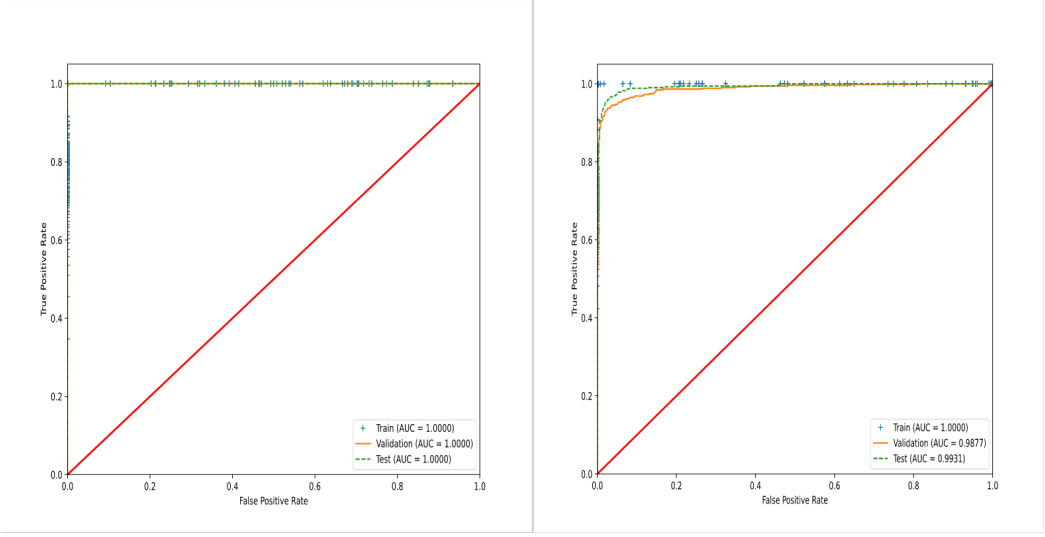
Similarly, Figure 10(B) presents the binary classification ROC curves, with AUC values of 1.0, 0.99, and 0.99 for training, validation, and testing, respectively. These results demonstrate the robustness and generalization capability of the SE-CNN architecture, showing high performance in both binary and multi-class classification of blood cell images.

Comparative Analysis and Computational Efficiency:

The computational complexity of deep learning models plays a pivotal role in their practical deployment, particularly in medical diagnostic tasks (Shah & Bhavsar, 2022). Table 5 indicates that the SE-CNN model has a few parameters and strong performance at the same time, hence it is an efficient model computationally.

Table 4: Summary of evaluation metrics for Binary and Multiclass classification tasks.

Metrics	Training	Validation	Testing	Training	Validation	Testing
	Binary (2-class)			Multiclass (4-class)		
Precision (%)	99.96	95.93	95.83	99.93	99.83	99.83
Sensitivity (%)	99.79	92.53	95.08	99.86	99.68	99.68
Specificity (%)	99.98	98.17	98.08	99.97	99.93	99.93
F1-score (%)	99.87	94.20	95.45	99.89	99.75	99.75



(A) (B)

Figure 10: ROC curve of the proposed SE-CNN model (A) multiclass classification, (B) binary classification.

All experiments were conducted on the Kaggle cloud platform, utilizing an NVIDIA Tesla P100 GPU. The computational efficiency of the SE-CNN model has been evaluated in terms of parameters number, training time, and inference speed. The multiclass classification variant has 146,920 trainable parameters and ca. 573.91 KB of memory, in particular computationally efficient. The average time for training a model per epoch was around 21 seconds. At inference time, it produced a fast and accurate test by processing 486 PBS test images in 3 seconds, achieving a high test accuracy of 99.79% at a very low final loss of 0.0043.

In contrast, the binary classification model has a total of 146,533 trainable parameters (approximately 572.39 KB) and can be considered as a lightweight architecture design for computational efficiency. It took roughly 52 seconds per epoch for training and could classify 1599 PBS test image samples in 4 seconds of computing time with a test accuracy of 97.12% and a final loss of 0.114. In addition, the Computational complexity of the SE-CNN model was analyzed using the Floating Point Operations (FLOPs). A concrete function of the model was created on the model's computation graph, and the amount of floating-point arithmetic was approximated from a static graph profile. From the analysis, the model needs (about 0.56G) FLOPs for a single forward pass.

To further assess the efficiency and effectiveness of the proposed architecture, Table 5 presents a comparative evaluation of three architectural configurations, each designed to isolate and quantify the individual contributions of the SE block and SeparableConv2D layers. The evaluation was conducted on both binary and multiclass leukemia classification tasks, using a consistent image input size across all variants to ensure fair comparison.

The first configuration omits the SE block, serving as a baseline to evaluate the model's performance without channel-wise feature recalibration. The second configuration retains the SE block but replaces SeparableConv2D layers with standard Conv2D layers, enabling an assessment of the trade-offs between computational cost and representational capacity. configuration The third corresponds to the proposed SE-CNN model, which integrates both the SE block and SeparableConv2D layers to maximize both accuracy and computational efficiency.

The SE-CNN model achieves the best performance. Although the no-SE model attained competitive results in a multiclass setup, its performance reduced significantly on binary classification, thus demonstrating the necessity of the channel reuse mechanism in discriminating among fine-grained classes. On the other extreme end, the Conv2D based model had perfect accuracy on a multiclass task but with the high computational complexity of 888K parameters and 4.58G FLOPs. The SE-CNN provides an effective balance, obtaining high accuracy at a substantially lower parameter number and FLOPs, and low train and test times.

Table 5: Comparison of different architectural variants of the proposed model on binary and multiclass classification tasks.

Model Variant	Task	Memory	Training Time (s/epoch)	Testing Time (s)	FLOPs	Loss	Accuracy (%)	Trainable params
Baseline CNN without SE block	Multiclass	470.26 KB	18	3	0.55 G	0.0031	99.79	120387
	Binary	468.75 KB	42	7		0.2011	95.3	120000
CNN with SE block and Conv2D	Multiclass	3.39 MB	22	3	4.58 G	2.9069	100	888003
	Binary	3.39 MB	59	5		0.1633	95.99	887616
Proposed model SE-CNN	Multiclass	573.91 KB	21	3	0.56 G	0.0043	99.79	146920
	Binary	572.39 KB	52	4		0.1148	97.12	146533

Table 6: Comparative analysis of multiclass SE-CNN model with current methods in terms of dataset size and accuracy.

Reference	Method	Sample size	Accuracy (%)
Rahman *et al*., (2023)	ResNet50 with LR and SVC Feature Selector and CSO	3262	99.84
Prakash *et al*., (2024)	Lightweight Attention-YOLOv8	7535	98.4
Nayak *et al*., (2024)	MobileNet with SVM	Not specified	99.3
Alim *et al*., (2024)	ResNet-50	3242	99.38
Dubai *et al*., (2025)	MCB-HyperNet	3256	99.69
Dutta *et al*., (2025)	Attention-augmented (LEU3)	3242	99
	Proposed model (SE-CNN)	3242	99.79

Table 7: Comparative analysis of binary SE-CNN model with current methods in terms of dataset size and accuracy.

Reference	Method	Sample size	Accuracy (%)
Kasani *et al*., (2020)	NASNetLarge with VGG19	10661	96.58
Chen *et al*., (2021)	ResNet101-9	12528	85.11
Sampathila *et al*., (2022)	ALL-NET	10661	95.54
Jawahar *et al*., (2022)	ALNet	10661	91.13
Hasanaath *et al*., (2024)	Multiple deep CNN models + SVM	10661	91.63
Darmawan & Shabrina, (2025)	EfficientNetV2B3 and LIME	12528	97.95
	Proposed model (SE-CNN)	10661	97.12

Comparison with previous studies:

The SE-CNN model's low parameter count and high classification accuracy on both tasks demonstrate their efficiency and practicality for real-time blood cell cancer detection. Comparisons with the state-of-the-art leukemia classification methods by PBS images present in Tables 6 and 7 validate the performance of the proposed model.

While direct comparisons are difficult given differences in dataset sizes, preprocessing steps, class distributions, and evaluation metrics, this analysis offers valuable insight into the relative performance of the proposed architecture. In spite of these complexities, the SE-CNN model achieved improved classification accuracies when compared with the various existing approaches with a significantly lower number of learnable parameters. While some competing models may achieve marginally higher accuracy, SE-CNN distinguishes itself through its lightweight design and computational efficiency. This balance of performance and efficiency makes it particularly suitable for resource constrained clinical environments.

Moreover, a comprehensive evaluation encompassing F1 score, sensitivity, and precision, as illustrated in Figures 11 and 12, further highlights the practical deployment potential and real-world applicability of the SE-CNN model, particularly when compared to more complex and parameter-intensive alternatives.

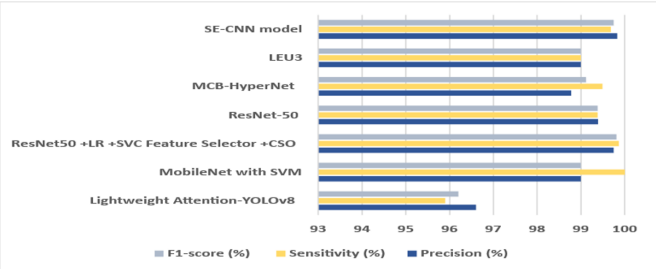


Figure 11: Comparison of F1 Score, sensitivity, and specificity results for the proposed multiclass model versus recently published methods in the literature.

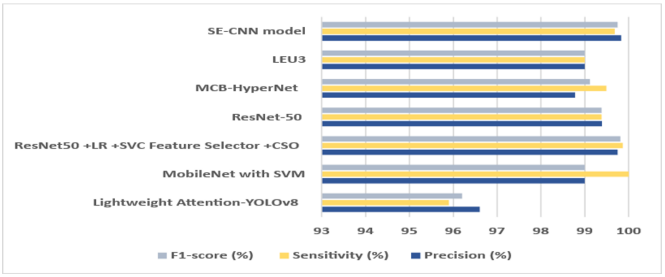


Figure 12: Comparison of F1 Score, sensitivity, and specificity results for the proposed binary model versus recently published methods in the literature.

Limitations and Future Directions:

Despite the promising performance of the SE-CNN model, several limitations should be acknowledged. The model was evaluated on only two datasets, which may not fully capture the variability encountered in real-world clinical settings, including differences in patient demographics, imaging devices, and laboratory protocols. Additionally, cross-institutional validation was not performed, limiting assessment of the model's robustness across diverse clinical environments. Future work will focus on expanding validation to multiple datasets from different hospitals and imaging systems to improve generalizability. Furthermore, integrating Explainable AI (XAI) techniques, such as Grad-CAM, will be prioritized to increase the interpretability and trustworthiness of model predictions, thereby facilitating clinical adoption. Moreover, the use of k-fold cross-validation will be incorporated in future evaluations to provide a more reliable and comprehensive assessment of model performance.

5. CONCLUSION

This study introduced the SE-CNN model, designed for efficient and accurate classification of leukemia using PBS images. The architecture was evaluated on two publicly available datasets that addressed binary and multiclass classification tasks. By incorporating an integrated attention mechanism, SE-CNN significantly reduced the number of trainable parameters while preserving high classification accuracy. Experimental results show that the proposed model outperforms or remains competitive using current cutting-edge methods, offering an optimal trade-off between classification performance and computational efficiency. These results underscore the usefulness of lightweight deep learning models, especially for use in real-time

clinical settings where prompt and precise diagnostic assistance is essential. Real-time evaluations hold significant importance in clinical practice, as they facilitate timely decision-making in emergency scenarios and resource-limited settings, leading to better patient outcomes and faster diagnostic processes. To improve the model's resilience and also generalizability, future research will concentrate on using model compression approaches and carrying out cross-institutional validations. In addition, it is planned to evaluate the proposed method on diverse PBC datasets in real-time scenarios to further assess its practical applicability in clinical environments.

Acknowledgment:

This publication's work is supported by the University of Zakho, Kurdistan Region, Iraq.

Ethical Statement:

This study has used existing online blood cell cancer datasets. No experiments have been done to humans.

Conflict of Interests:

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not for profit sectors. The study was self funded by the authors.

Author Contributions:

A.M.A.: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing. M. B. A.: Supervision, Guidance, Writing – Review & Editing. A. Z.: Supervision, Guidance, Writing – Review & Editing. All authors have reviewed the final version to be published and have agreed to be accountable for all aspects of the work.

6. REFERENCES

Abioye, O. A., Evwiekpaefe, A. E., & Awujoola, A. J. (2024). Performance Evaluation of Efficientnetv2 Models on the Classification of Histopathological Benign Breast Cancer Images. *Science Journal of University of Zakho*, 12(2), 208–214. DOI:https://doi.org/10.25271/sjuoz.2024.12.2.1261

Agustin, R. I., Arif, A., & Sukorini, U. (2021). Classification of immature white blood cells in acute lymphoblastic leukemia L1 using neural networks particle swarm optimization. *Neural Computing and Applications*, 33(17), 10869–10880. DOI:https://doi.org/10.1007/s00521-021-06245-7

Alim, M. S., Bappon, S. D., Sabuj, S. M., Islam, M. J., Tarek, M. M., Azam, M. S., & Islam, M. M. (2024). Integrating convolutional neural networks for microscopic image analysis in acute lymphoblastic leukemia classification: A deep learning approach for enhanced diagnostic precision. *Systems and Soft Computing*, 6(July), 200121. DOI:https://doi.org/10.1016/j.sasc.2024.200121

Ananth, C., Tamilselvi, P., Joshy, S. A., & Kumar, T. A. (2023). Blood Cancer Detection with Microscopic Images Using Machine Learning. *Lecture Notes in Networks and Systems*, 498, 45–54. DOI:https://doi.org/10.1007/978-981-19-5090-2_4

Ansari, M. O., Ahmad, M. F., Shadab, G. G. H. A., & Siddique, H. R. (2018). Superparamagnetic iron oxide nanoparticles based cancer theranostics: A double edge sword to fight against cancer. *Journal of Drug Delivery Science and Technology*, 45, 177–183. DOI:https://doi.org/10.1016/j.jddst.2018.03.017

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical Image Analysis using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, 42(11). DOI:https://doi.org/10.1007/s10916-018-1088-1

Attallah, & Omneya. (2024). Acute lymphocytic leukemia detection and subtype classification via extended wavelet pooling based-CNNs and statistical-texture features. *Image and Vision Computing*, 147(August 2023), 105064. DOI:https://doi.org/10.1016/j.imavis.2024.105064

Bennett, J. M., Catovsky, D., Daniel, M. -T, Flandrin, G., Galton, D. A. G., Gralnick, H. R., & Sultan, C. (1976). Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. *British Journal of Haematology*, 33(4), 451–458. DOI:https://doi.org/10.1111/j.1365-2141.1976.tb03563.x

Bibi, N., Sikandar, M., Din, I. U., Almogren, A., & Ali, S. (2020). IOMT-based automated detection and classification of leukemia using deep learning. *Journal of Healthcare Engineering*, 2020. DOI:https://doi.org/10.1155/2020/6648574

Brahmaiah, O. V., Raju, M. S. N., Jahnavi, V., & Varshini, M. (2024). Dense Net-Based Acute Lymphoblastic Leukemia Classification and Interpretation through Gradient-Weighted Class Activation Mapping. *2024 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, INCOS 2024 - Proceedings*, 1–7. DOI:https://doi.org/10.1109/INCOS59338.2024.10527599

Bukhari, M., Yasmin, S., Sammad, S., & Abd El-Latif, A. A. (2022). A Deep Learning Framework for Leukemia Cancer Detection in Microscopic Blood Samples Using Squeeze and Excitation Learning. *Mathematical Problems in Engineering*, 2022. DOI:https://doi.org/10.1155/2022/2801227

Chen, Y. M., Chou, F. I., Ho, W. H., & Tsai, J. T. (2021). Classifying microscopic images as acute lymphoblastic leukemia by Resnet ensemble model and Taguchi method. *BMC Bioinformatics*, 22, 1–20. DOI:https://doi.org/10.1186/s12859-022-04558-5

Chiaretti, S., Zini, G., & Bassan, R. (2014). Diagnosis and subclassification of acute lymphoblastic leukemia. *Mediterranean Journal of Hematology and Infectious Diseases*, 6(1). DOI:https://doi.org/10.4084/mjhid.2014.073

Darmawan, S. T. S., & Shabrina, N. H. (2025). Explainable deep learning for diagnosing acute lymphocytic leukemia using blood smear images. *Bulletin of Electrical Engineering and Informatics*, 14(2), 1298–1307. DOI:https://doi.org/10.11591/eei.v14i2.9073

Das, P. K., Jadoun, P., & Meher, S. (2020). Detection and Classification of Acute Lymphocytic Leukemia. *Proceedings of 2020 IEEE-HYDCON International Conference on Engineering in the 4th Industrial Revolution, HYDCON 2020*. DOI:https://doi.org/10.1109/HYDCON48903.2020.9242745

Devi, J. R., Kadiyala, P. S., Lavu, S., Kasturi, N., & Kosuri, L. (2024). Enhancing Acute Lymphoblastic Leukemia Classification with a Rapid and Effective CNN Model. *3rd IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics, ICDCECE 2024*. DOI:https://doi.org/10.1109/ICDCECE60827.2024.10549348

Dubai, N. J., Kadhim, O. N., & Najjar, F. H. (2025). A Morphological Context Blocks Hybrid CNN for Efficient Acute Lymphoblastic Leukemia Classification. *International Journal of Robotics and Control Systems*, 5(2), 1102–1119. DOI:https://doi.org/10.31763/ijrcs.v5i2.1824

Dutta, M., Mojumdar, M. U., Kabir, A., Chakraborty, N. R., Siddiquee, S. T., & Abdullah, S. (2025). LEU3 : An Attention Augmented-Based Model for Acute Lymphoblastic Leukemia Classification. *IEEE Access*, 13(February), 31630–31642. DOI:https://doi.org/10.1109/ACCESS.2025.3542609

Erten, M., Barua, P. D., Dogan, S., Tuncer, T., Tan, R. S., & Acharya, U. R. (2024). ConcatNeXt: An automated blood cell classification with a new deep convolutional neural network. *Multimedia Tools and Applications*. DOI:https://doi.org/10.1007/s11042-024-19899-x

Hasanaath, A. A., Mohammed, A. S., Latif, G., Abdelhamid, S. E., Alghazo, J., & Hussain, A. A. (2024). Acute lymphoblastic leukemia detection using ensemble features from multiple deep CNN models. *Electronic Research Archive*, 32(4), 2407–2423. DOI:https://doi.org/10.3934/ERA.2024110

Hosseini, A., Eshraghi, M. A., Taami, T., Sadeghsalehi, H., Hoseinzadeh, Z., Ghaderzadeh, M., & Rafiee, M. (2023). A mobile application based on efficient lightweight CNN model for classification of B-ALL cancer from non-cancerous cells: A design and implementation study. *Informatics in Medicine Unlocked*, 39, 101244. DOI:https://doi.org/10.1016/j.imu.2023.101244

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7132–7141. DOI:https://doi.org/10.1109/QCE60285.2024.10249

Jawahar, M., H, S., L, J. A., & Gandomi, A. H. (2022). ALNett: A cluster layer deep convolutional neural network for acute lymphoblastic leukemia classification. *Computers in Biology and Medicine*, 148, 105894. DOI:https://doi.org/10.1016/j.compbiomed.2022.105894

Kasani, P. H., Park, S. W., & Jang, J. W. (2020). An aggregated-based deep learning method for leukemic B-lymphoblast classification. *Diagnostics*, 10(12). DOI:https://doi.org/10.3390/diagnostics10121064

Mourya, S., Kant, S., Kumar, P., Gupta, A., & Gupta, R. (2019). ALL challenge dataset of ISBI 2019 (C-NMC-2019) (version 1) [dataset]. *The Cancer Imaging Archive*. DOI:https://doi.org/10.7937/tcia.2019.dc64i46r

Muhsin, A. A., Atrushi, A. M., & Al-Doski, A. A. (2024). the Significance of Minimal Residual Disease in Acute Lymphoblastic Leukaemia: a Single Centre Study. *Science Journal of University of Zakho*, 12(2), 144–148. DOI:https://doi.org/10.25271/sjuoz.2024.12.2.1240

Nayak, R., Bekal, A., Suvarna, M., & Sathish, D. (2024). Identifying Subtypes of Acute Lymphoblastic Leukemia Using Blood Smear Images: A Hybrid Learning Approach. *Journal of The Institution of Engineers (India): Series B*. DOI:https://doi.org/10.1007/s40031-024-01069-0

Onciu, & Mihaela. (2009). Acute Lymphoblastic Leukemia. *Hematology/Oncology Clinics of North America*, 23(4), 655–674. DOI:https://doi.org/10.1016/j.hoc.2009.04.009

Prakash, K. D., Khan, J., & Kim, K. (2024). Lightweight and Efficient YOLOv8 with Residual Attention Mechanism for Precise Leukemia Detection and Classification. *IEEE Access*, 12(September), 159395–159413. DOI:https://doi.org/10.1109/ACCESS.2024.3484933

Pushpalatha, M. P., S, J. B., S, M., R, N. B., & K, A. (2024). Detection and Classification of Acute Lymphoblastic Leukemia using CNN. *International Journal of Advanced Research in Computer Science*, 15(4), 35–43. DOI:https://doi.org/10.26483/ijarcs.v15i4.7114

Rahman, W., Faruque, M. G. G., Roksana, K., Sadi, A. H. M. S., Rahman, M. M., & Azad, M. M. (2023). Multiclass blood cancer classification using deep CNN with optimized features. *Array*, 18(May), 100292. DOI:https://doi.org/10.1016/j.array.2023.100292

Saeed, U., Kumar, K., Khuhro, M. A., Laghari, A. A., Shaikh, A. A., & Rai, A. (2024). DeepLeukNet—A CNN based microscopy adaptation model for acute lymphoblastic leukemia classification. *Multimedia Tools and Applications*, 83(7), 21019–21043. DOI:https://doi.org/10.1007/s11042-023-16191-2

Salehi, H. S., Barchini, M., & Mahdian, M. (2020). Optimization methods for deep neural networks classifying OCT images to detect dental caries. *Lasers in Dentistry XXVI*, 11217, 16. DOI:https://doi.org/10.1117/12.2545421

Sampathila, N., Chadaga, K., Goswami, N., Chadaga, R. P., Pandya, M., Prabhu, S., Bairy, M. G., Katta, S. S., Bhat, D., & Upadya, S. P. (2022). Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images. *Healthcare (Switzerland)*, 10(10). DOI:https://doi.org/10.3390/healthcare10101812

Shah, B., & Bhavsar, H. (2022). Time Complexity in Deep Learning Models. *Procedia Computer Science*, 215(2022), 202–210. DOI:https://doi.org/10.1016/j.procs.2022.12.023

Shamama, A., & Afrin, A. (2020). A convolutional neural network–based learning approach to acute lymphoblastic leukaemia detection with automated feature extraction. *Medical and Biological Engineering and Computing*, 58(12), 3113–3121. DOI:https://doi.org/10.1007/s11517-020-02282-5

Suzuki, K. (2017). Overview of deep learning in medical imaging. *Radiological Physics and Technology*, 10(3), 257–273. DOI:https://doi.org/10.1007/s12194-017-0406-5

Zolfaghari, M., & Sajedi, H. (2022). A survey on automated detection and classification of acute leukemia and WBCs in microscopic blood cells. *Multimedia Tools and Applications*, 81(5), 6723–6753. DOI:https://doi.org/10.1007/s11512-022-12108-7