Article

Not peer-reviewed version

A Lightweight Hybrid CNN–CBAM Model for Multistage Acute Lymphoblastic Leukemia Classification from Peripheral Blood Smear Images

Kittipol Wisaeng *

Posted Date: 26 February 2026

doi: 10.20944/preprints202602.1601.v1

Keywords: blood cancer classification; lightweight deep learning; EfficientNetV2-S; Convolutional Block Attention Module; B-cell acute lymphoblastic leukemia (B-ALL)

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026

doi:10.20944/preprints202602.1601.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Lightweight Hybrid CNN–CBAM Model for Multistage Acute Lymphoblastic Leukemia Classification from Peripheral Blood Smear Images

Kittipol Wisaeng

Mahasarakham Business School, Mahasarakham University, Mahasarakham 44150, Thailand;
Kittipol.w@acc.msu.ac.th; Tel.: +66 866393870

Abstract

Accurate and efficient classification of hematological malignancies from peripheral blood smear (PBS) images remains challenging due to the scarcity of annotated datasets, staining variability, and subtle morphological differences among blood cancer subtypes. To address these limitations, this study proposes an Advanced Lightweight Deep Learning (ALDL) framework for the multi-class classification of Acute Lymphoblastic Leukemia (ALL) across four clinically significant stages: Benign, Pro-B, Pre-B, and Early Pre-B. The framework integrates EfficientNetV2-S with Convolutional Block Attention Modules (CBAM) to enhance spatial and channel-wise feature refinement. At the same time, Focal Loss is employed to mitigate class imbalance by prioritizing hard-to-classify samples. A robust preprocessing pipeline, including CLAHE contrast enhancement, Reinhard stain normalization, and data augmentation, improves feature visibility and dataset generalization. Lesion segmentation is performed using RGB-based thresholding and watershed overlay, followed by lesion-level cropping to ensure consistency across inputs. Experimental evaluations on the ALL-DB dataset demonstrate the superior performance of the proposed method, achieving average accuracy of 96.11%, F1-score of 95.99%, and AUC of 0.9875. Comparative analyses against MobileNetV3, ResNet50, DenseNet121, VGG16, and InceptionV3 confirm that the hybrid approach consistently outperforms conventional CNN architectures across both 70:30 and 60:40 train–test splits. Furthermore, a detailed investigation of color spaces (RGB, HSV, LAB, HED) indicates that RGB yields the most reliable segmentation and classification results. At the same time, HED enhances lesion visualization at the expense of higher computational cost. The proposed ALDL framework demonstrates strong potential for real-world application as a computer-aided diagnostic (CAD) system for early leukemia detection, offering improved diagnostic reliability, reduced error rates, and practical scalability for clinical environments.

Keywords: blood cancer classification; lightweight deep learning; EfficientNetV2-S; Convolutional Block Attention Module; B-cell acute lymphoblastic leukemia (B-ALL)

---

1. Introduction

Medical image analysis continues to face substantial challenges, particularly due to the scarcity of high-quality labeled datasets essential for reliable classification. Annotating medical images is a resource-intensive and highly specialized process that requires domain expertise, especially in tasks involving the segmentation and classification of blood cancer cells across multiple stages [1]. The difficulty in curating large, well-annotated datasets constrains the effectiveness of traditional supervised learning approaches, thereby limiting the robustness and generalizability of deep learning models in hematopathology. To address this limitation, researchers have increasingly explored advanced strategies that combine small sets of labeled data with larger pools of unlabeled data, thereby enhancing the models' capacity to learn meaningful representations for blood cancer imaging. A common strategy to mitigate data scarcity is to use transfer learning with supervised

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

pretraining on large-scale datasets such as ImageNet. This approach enables convolutional neural networks (CNNs) to learn general-purpose features, including edges, textures, and object structures, which can be fine-tuned for medical imaging tasks. While such pre-training accelerates convergence and often improves classification performance, its application to blood cancer analysis remains suboptimal due to the substantial domain gap between natural and medical images. Specifically, blood smear images exhibit highly specialized morphological and chromatic patterns that differ markedly from everyday photographs. Subtle visual cues, such as nuclear morphology, cytoplasmic staining, and fine-grained structural variations, are critical for distinguishing between clinically relevant stages of B-cell acute lymphoblastic leukemia (B-ALL), including Benign, Pro-B, Pre-B, and Early Pre-B cells [2,3]. Models pre-trained on natural image datasets often lack the inductive biases needed to capture domain-specific features, thereby reducing their diagnostic reliability.

To overcome these limitations, CNN-based approaches employing self-supervised and weakly supervised learning paradigms have gained traction in medical imaging. Such methods utilize surrogate tasks (e.g., contrastive learning, image inpainting, or context prediction) to uncover latent structures in large quantities of unlabeled medical data [4,5]. By extracting intrinsic and discriminative representations without extensive manual labeling, these methods offer a promising avenue for analyzing blood cancer cells. Nonetheless, their application to hematological image classification remains limited. Existing CNN methodologies are typically optimized for broader medical imaging domains and fail to adequately account for the distinctive morphological complexity, high intra-class variability, and overlapping cellular structures characteristic of microscopic blood smear images. Consequently, there is a pressing need to design deep learning frameworks tailored to the unique visual and clinical attributes of hematological datasets. This requires incorporating domain-specific inductive biases, optimizing network architectures for fine-grained feature learning, and integrating robust training mechanisms that enhance both accuracy and interpretability. Addressing these requirements is pivotal for achieving reliable automated diagnostic systems in hematopathology.

In this study, we propose a novel hybrid framework for classifying B-ALL subtypes that integrates EfficientNetV2-S with a Convolutional Block Attention Module (CBAM) and Focal Loss. The framework is designed to effectively capture discriminative features from peripheral blood smear images while maintaining computational efficiency. Moreover, the approach incorporates a semantic segmentation component to enhance white blood cell (WBC) localization, thereby supporting not only classification but also interpretability in diagnostic workflows. By focusing on both robustness and efficiency, this study advances the development of clinically applicable computer-aided diagnostic (CAD) systems for the early and accurate detection of leukemia.

2. Related Works

The classification of blood cell cancers (BCCs) remains challenging due to the complex morphological and visual characteristics of peripheral blood smear images. Factors such as high intra-class variability among white blood cell (WBC) types, overlapping cellular structures, and subtle staining differences complicate the distinction between malignant and non-malignant cells. Differentiating BCC subtypes poses additional difficulties, as visual differences between stages are often subtle and fine-grained.

To address these challenges, researchers have investigated diverse feature extraction strategies, which can broadly be categorized into: (1) traditional machine learning models using handcrafted features (e.g., histograms, texture descriptors, and morphological properties), and (2) deep learning-based methods that automatically extract high-level features using convolutional neural networks (CNNs) and advanced classifiers such as Support Vector Machines (SVM) or Random Forests (RF). While handcrafted methods offer interpretability and modest computational requirements [6,7], they suffer from limitations in scalability, robustness to noise, and adaptability to diverse datasets [8,9]. In contrast, CNN-based feature extraction enables automated discovery of discriminative patterns but often requires large annotated datasets and careful adaptation to hematopathological image

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

characteristics. Several studies highlight the potential of deep architectures in hematology. Lu et al. [10] proposed a multiscale WBC classification framework that incorporates a residual-block encoder to mitigate vanishing gradients and capture hierarchical representations of blood smear images. Similarly, Roy et al. [11] developed a DeepLabv3C framework with a ResNet-50 backbone, achieving 96.1% accuracy across three datasets by integrating semantic segmentation and residual learning to improve robustness in cancer detection. Other researchers have explored hybrid methods; for instance, Abdurrazzaq et al. [12] applied singular value decomposition (SVD) to vascular pattern detection, while Khomairoh et al. [13] employed Haar cascades for nucleus and cytoplasm segmentation in acute myeloid leukemia (AML) subtypes, achieving accuracies ranging from 71% to 90%. Comparative studies also reveal trade-offs between handcrafted and deep features. Hegde et al. [14] demonstrated that handcrafted descriptors focusing on morphological and chromatic cues achieved performance comparable to AlexNet-derived deep features, with an overall accuracy of 99%, underscoring the value of interpretability in clinical contexts. Feature fusion frameworks further enhance performance; for example, Saleem et al. [15] combined DarkNet-53 and ShuffleNet, achieving 98.6% segmentation accuracy while balancing semantic depth with computational efficiency. Likewise, Ramya et al. [16] employed gray-level co-occurrence matrix (GLCM) descriptors to distinguish healthy from malignant cells, whereas Rad et al. [17] improved segmentation robustness by employing automatic region initialization strategies based on statistical and morphological cues, achieving 96% accuracy. Beyond CNNs, researchers have employed a diverse range of machine learning and optimization techniques. Puigdollers et al. [18] used a Bag-of-Words (BoW) approach, achieving 80% accuracy while providing interpretability. Ruberto et al. [19] applied SVM and k-nearest neighbors (KNN), reporting up to 99% accuracy in vessel counting and classification, while Hussein et al. [20] introduced a platelet-based classification approach enhanced by Bee Colony and Reptile Search Optimization. H. Rai et al. [21] evaluated pancreatic, prostate, colorectal, and leukemia cancers using standard Machine Learning (ML) and Deep Learning (DL) techniques, highlighting the need for improved cancer diagnosis and treatment. Swanson et al. [22] discussed the use of ML methods in analyzing medical imaging and genetic data for cancer applications, advocating strategies to enhance clinical effectiveness. Aly et al. [23] introduced a method for training a Multi-Layer Perceptron (MLP) using the Hunger Games Optimization (HGO) technique, which improved classification accuracy for cervical cancer. Mahesh et al. [24] presented a hybrid of Ant Lion Mutation and Ant Colony Optimization with Particle Swarm Optimization, achieving an SVM prediction accuracy of 87.8% for optimal feature selection. Wais et al. [25] developed a pipeline for categorizing acute lymphoblastic leukemia, using a novel neighbourhood pixel transformation with differential evolution to enhance the clarity of blood cell images. Recent developments in the field increasingly emphasize integrating deep learning architectures with ensemble-based strategies. For instance, Nssibi et al. [26] proposed the iBABC-CGO framework, which synergistically combines bee colony optimization with chaotic game-theoretic principles to enhance exploration within the feature space. Similarly, Vogelbacher et al. [27] utilized a dual deep neural network architecture for avian blood cell recognition, illustrating the cross-domain versatility of convolutional neural networks (CNNs). Although a wide range of automated computer-aided detection systems has been introduced, many studies continue to report performance constraints related to processing efficiency and classification precision [28–30]. Traditional feature extraction techniques such as Speeded-Up Robust Features (SURF) [31–33], Scale-Invariant Feature Transform (SIFT) [34], and Histograms of Oriented Gradients (HoG) [35], remain critical preprocessing components; however, their capability is frequently outperformed by CNN-based feature learning approaches [36–38]. Evidence from [39] demonstrated that HoG descriptors combined with logistic regression can effectively support feature extraction for leukemia prediction using the ALL-IDB dataset. Complementary work in [40] analyzed color and texture features derived from histogram-based thresholding and watershed segmentation, and outlined the influence of different feature-extraction pipelines [30]. Moreover, SURF has shown strong potential for retrieving discriminative features from color-rich query images, and its coupling with Genetic Algorithms (GA) [41] has

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

yielded substantial improvements, achieving feature-extraction accuracy levels of up to 92%. Among classical detectors, SIFT [42], FAST [43], and CenSurE [44] remain notable for their specialized detection capabilities. At the same time, feature-fusion strategies have been shown to further enhance performance by integrating multiple hand-crafted descriptors [45]. Object recognition tasks commonly rely on segmentation-driven pipelines incorporating both quantitative and qualitative feature analysis and matching [46]; however, such conventional approaches often exhibit limited robustness. In contrast, CNN-based methods automate feature extraction and enable superior performance in complex image classification scenarios [47–51]. Within the broader spectrum of deep learning research [52–55], pre-trained architectures such as AlexNet [56], GoogleNet [56], ResNet [57], VGG-16 [58], and Inception-V3 [59] have emerged as highly effective feature extractors and selection mechanisms. These models have also served as foundational components in advanced analytical frameworks such as the Simulation of Electron Spectra for Surface Analysis (SESSA). In this context, Jagadev and Virani [60] proposed an innovative approach to leukemia cell classification that integrates deep feature extraction with a support vector machine (SVM) classifier. At the same time, work in [61] further demonstrated the potential of SVM-based methodologies for detecting acute lymphoblastic leukemia (ALL). Additional studies have explored hybrid pipelines, including the integration of K-means clustering and SVMs to improve the performance of microscopic ALL recognition [62–65]. More recent research trends highlight the growing dominance of deep learning-driven ALL classification, where CNN and custom CNN architectures yield highly competitive results across training and validation scenarios [66]. In particular, [67] introduces a weighted ensemble CNN framework that enhances diagnostic performance for ALL detection in microscopic images. Table 1. Summary of Existing Approaches for Blood Cancer / ALL Classification and Key Research Insights.

Table 1. Summary of Existing Approaches for Blood Cancer / ALL Classification and Key Research Insights.

Ref. / Author	Method / Model	Key Contribution	Performance	Strengths	Limitations
Lu et al. [10]	Residual multiscale CNN	Hierarchical WBC feature learning	—	Handles gradient issues	Needs a large dataset
Roy et al. [11]	DeepLabv3C + ResNet-50	Segmentation + residual learning	96.1%	Robust detection	High computation
Abdurrazzaq et al. [12]	SVD vascular detection	Enhances vascular features	—	Better representation	Limited classification
Khomairoh et al. [13]	Haar cascade segmentation	AML nucleus/cytoplasm detection	71–90%	Lightweight	Moderate accuracy
Hegde et al. [14]	Morphological vs AlexNet	Handcrafted ≈ deep features	99%	Interpretable	Dataset dependent
Saleem et al. [15]	DarkNet-53 + ShuffleNet	Feature fusion	98.6%	Efficient + deep	Complex design
Ramya et al. [16]	GLCM descriptors	Healthy vs malignant	—	Low cost	Noise sensitive
Rad et al. [17]	Region initialization	Robust segmentation	96%	Stable segmentation	Feature dependent
Puigdollers et al. [18]	Bag-of-Words	Interpretable classification	80%	Simple	Lower accuracy
Ruberto et al. [19]	SVM + KNN	Vessel classification	Up to 99%	High accuracy	Generalization limits
Hussein et al. [20]	Bee + Reptile optimization	Feature search	—	Strong optimization	Complex tuning
Rai et al. [21]	ML + DL comparison	Multi-cancer evaluation	—	Broad analysis	Not specialized
Swanson et al. [22]	ML imaging analysis	Clinical insight	—	Conceptual value	No experiment
Aly et al. [23]	MLP + HGO	Optimized training	—	Better convergence	Dataset specific
Mahesh et al. [24]	ALO + ACO + PSO + SVM	Feature selection	87.8%	Effective search	Computational cost

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

Wais et al. [25]	Pixel transform + DE	Image clarity	—	Enhances visibility	Heavy preprocessing
Nssibi et al. [26]	iBABC-CGO	Feature exploration	—	Strong exploration	Complex model
Vogelbacher et al. [27]	Dual DNN	Cross-domain blood recognition	—	Versatile	Needs training data
Various [31–33]	SURF	Keypoint extraction	Up to 92%	Robust	Inferior to CNN
Lowe [34]	SIFT	Scale-invariant detection	—	Stable	Slow
Dalal & Triggs [35]	HoG	Gradient descriptors	Effective	Simple	Less discriminative
Study [39]	HoG + Logistic Regression	Leukemia prediction	Effective	Efficient	Limited depth
Study [40]	Histogram + Watershed	Color-texture segmentation	—	Simple	Threshold sensitive
Study [41]	SURF + GA	Feature optimization	92%	Improved accuracy	Tuning needed
Various [42–44]	SIFT / FAST / CenSurE	Keypoint detection	—	Specialized	Low semantics
Study [45]	Feature fusion descriptors	Multi-feature integration	—	Complementary	Complexity
Study [46]	Segmentation pipeline	Structured analysis	—	Systematic	Low robustness
Various [47–51]	CNN models	Automated features	High	Strong performance	Data intensive
Various [56–59]	AlexNet / ResNet / VGG / Inception	Pretrained extractors	High	Transfer learning	Heavy models
Jagadev & Virani [60]	Deep features + SVM	DL + ML classifier	—	Generalizable	Multi-stage
Study [61]	SVM detection	ALL detection	—	Effective	Feature dependent
Various [62–65]	K-means + SVM	Improved recognition	—	Better clustering	Initialization sensitive
Study [66]	Custom CNN	End-to-end classification	Competitive	High accuracy	Data demand
Study [67]	Ensemble CNN	Diagnostic improvement	—	Robust prediction	Computational load

3. Research Gap and Contributions

A review of the existing literature on blood cancer analysis reveals that most feature-extraction methods have focused primarily on images of normal or healthy blood cells. Researchers have focused on standard WBC types such as neutrophils, eosinophils, and lymphocytes, mainly because of the greater availability of well-annotated datasets. While these studies have advanced our understanding of cellular morphology, they often overlook the complexities of pathological smear images from cancer patients. A significant challenge is the scarcity of annotated datasets for malignant cells, which typically exhibit considerable morphological variability and irregular staining. This limitation has constrained the development of robust and generalizable classification models.

To address this gap, the present study introduces feature-extraction strategies tailored to pathological blood smear images for the detection and classification of malignant cell types. The focus is placed on capturing morphological, chromatic, and textural alterations associated with malignant transformation, with particular emphasis on the multi-class recognition of B-cell acute lymphoblastic leukemia (B-ALL) subtypes: Benign, Pro-B, Pre-B, and Early Pre-B. By focusing on pathological rather than healthy cells, this study aims to enhance the diagnostic applicability of automated frameworks in hematopathology. The key contributions of this study are as follows:

1. We propose a lightweight hybrid model capable of accurately detecting acute lymphoblastic leukemia (ALL) by differentiating it from healthy samples and distinguishing acute myeloid leukemia (AML) from non-cancerous counterparts.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

2. A carefully designed pipeline integrates contrast-limited adaptive histogram equalization (CLAHE), Reinhard stain normalization, and data augmentation to enhance image quality, improve feature visibility, and increase model generalizability across heterogeneous datasets.
3. We introduce a novel architecture based on MobileNetV3-Small, enhanced with Convolutional Block Attention Modules (CBAM) for fine-grained feature extraction. To improve classification, deep features are fused with an XGBoost classifier, thereby leveraging both neural network representations and the efficiency of gradient-boosting. The model employs Softmax activation for multi-class prediction and integrates categorical cross-entropy with focal loss to effectively address class imbalance.
4. Rigorous performance analysis is conducted using publicly available blood cancer datasets under both 70:30 and 60:40 train–test splits. Each experiment is repeated 10 times to ensure statistical robustness, and the averaged results are compared against baseline methods to validate the superiority of the proposed framework.

4. Methodology

This study proposes a lightweight deep learning framework for efficiently and interpretably classifying blood cancer cell images. The methodological pipeline consists of four sequential stages: (1) data preprocessing, (2) deep representation learning, (3) hybrid feature refinement, and (4) final classification. Each stage is systematically structured to enhance data quality, extract discriminative features, and achieve robust classification performance while maintaining computational efficiency. The overall architecture of the proposed system is depicted in Figure 1, which illustrates the integration of preprocessing techniques, lightweight network architecture, attention mechanisms, and hybrid learning strategies into a unified diagnostic framework.

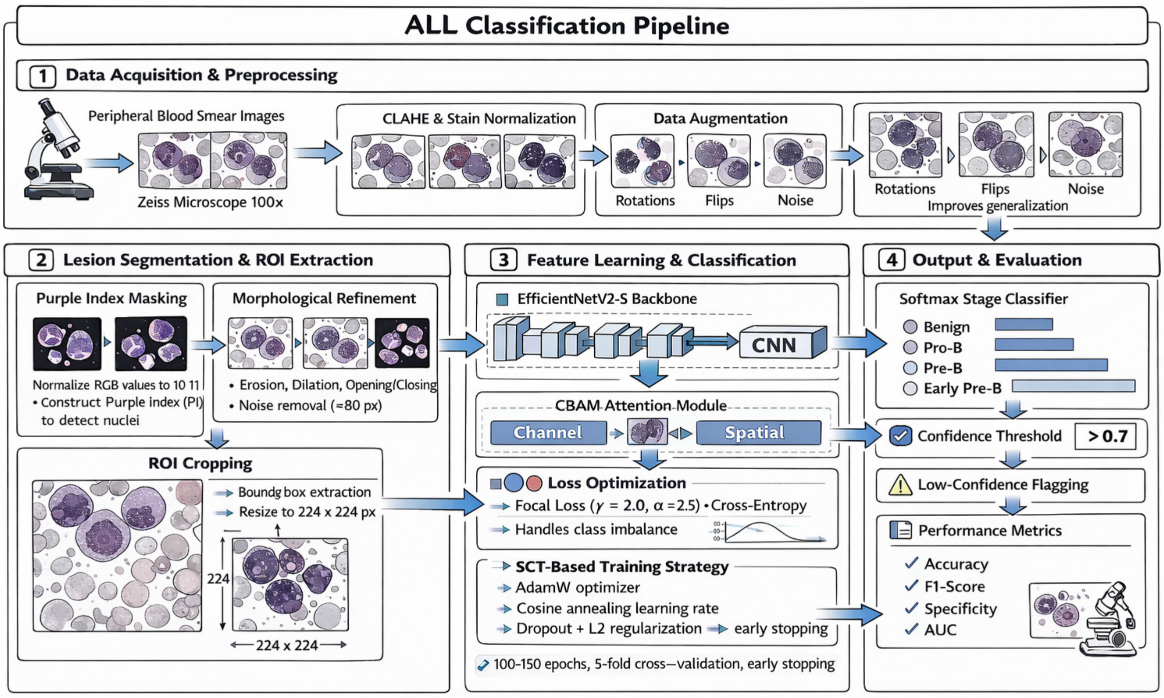


Figure 1. End-to-end workflow of the proposed ALL peripheral blood smear classification pipeline.

The proposed framework begins with data collection and preprocessing of peripheral blood smear images. A series of enhancement techniques is applied, including contrast-limited adaptive histogram equalization (CLAHE) to improve local contrast and Reinhard stain normalization to achieve a consistent color distribution across samples. To increase generalizability and mitigate overfitting, a range of data augmentation strategies is employed. The enhanced images are

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

subsequently processed by a lightweight convolutional neural network (CNN) with a MobileNetV3-Small backbone, augmented with Convolutional Block Attention Modules (CBAM) to emphasize salient spatial and channel-wise features relevant to medical imaging. To further improve interpretability and classification performance, a hybrid feature-refinement strategy is employed. In this stage, deep feature representations extracted from the CNN are integrated with an Extreme Gradient Boosting (XGBoost) classifier, which selectively refines the most informative features for distinguishing among blood cancer subtypes. The final classification stage comprises a fully connected layer with Softmax activation, producing a probability distribution over four clinically significant B-cell acute lymphoblastic leukemia (B-ALL) categories: Benign, Pro-B, Pre-B, and Early Pre-B. To address the inherent class imbalance in the datasets, the training process employs a combination of categorical cross-entropy and focal loss, thereby improving prediction fairness and enhancing the overall robustness and reliability of the diagnostic system.

4.1. Dataset

The definitive diagnosis of acute lymphoblastic leukemia (ALL) is typically complex, requiring invasive, costly, and time-consuming procedures; however, peripheral blood smear (PBS) images provide a non-invasive, cost-effective alternative for preliminary screening, despite the potential diagnostic variability inherent in the nonspecific presentation of ALL. In this study, a dataset comprising 3,242 PBS images from 89 patients suspected of ALL was used. Blood samples were prepared and stained by experienced laboratory personnel according to standard protocols, and images were captured using a Zeiss microscope camera at 100× magnification and stored as JPEG files. The dataset is categorized into benign cases (hematogenous samples without malignancy) and malignant cases, the latter subdivided into three clinically significant B-cell ALL subtypes: Early Pre-B, Pre-B, and Pro-B. To ensure diagnostic reliability, a hematopathology specialist confirmed the cell types and subtypes using flow cytometry, making this dataset a robust resource for developing and evaluating automated classification frameworks [68].

4.2. Blood Cell Lesion Segmentation and Classification

4.2.1. Lesion Mask

In this stage, lesion regions corresponding to the nuclei of leukemic cells are segmented from peripheral blood smear (PBS) images by applying intensity-based thresholding in the RGB color space [69]. Unlike HSV or LAB transformations, which require nonlinear conversions and may distort original intensity distributions, the RGB-based approach retains the native pixel values. This reduces computational complexity, preserves the morphological integrity of leukemic nuclei, and provides a consistent basis for subsequent feature extraction. Let the original color image be denoted as Eq. (1).

$$\mathrm{I}(x,y)=\left[\mathrm{R}(x,y),\mathrm{G}(x,y),\mathrm{B}(x,y)\right],\ \mathrm{R},\mathrm{G},\mathrm{B}\in[0,255] \tag{1}$$

where R(x, y), G(x, y), and B(x, y) represent the intensities of the red, green, and blue channels at pixel coordinates (x, y).

4.2.2. Normalization

Normalization is a crucial preprocessing step in medical image analysis, as it mitigates the influence of illumination variability, staining inconsistencies, and acquisition artifacts, thereby ensuring that the model learns discriminative features rather than noise. In this study, the RGB values of each pixel were normalized to the [0, 1] range, standardizing input intensity levels across all images and improving model convergence during training. For each pixel location (x, y), the normalized red, green, and blue channel intensities are computed using Eq. (2).

$$\mathrm{R}_n(x,y)=\frac{\mathrm{R}(x,y)}{255},\ \mathrm{G}_n(x,y)=\frac{\mathrm{G}(x,y)}{255},\ \mathrm{B}_n(x,y)=\frac{\mathrm{B}(x,y)}{255} \tag{2}$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026

doi:10.20944/preprints202602.1601.v1

where R(x, y), G(x, y), and B(x, y) represent the original channel values, each ranging from 0 to 255, and dividing by 255 scales the values to the range [0, 1], which facilitates numerical stability during backpropagation and prevents saturation in activation functions, the normalized RGB image is then represented as Eq. (3).

$$I(x,y)=\left[R_n(x,y),G_n(x,y),B_n(x,y)\right] \tag{3}$$

4.2.3. Color-Based Segmentation

Accurate segmentation of blood cell lesions is critical for distinguishing malignant nuclei from surrounding cytoplasm and background regions. In hematological imaging, leukemic nuclei frequently exhibit a purple hue, resulting from the combination of high red (R) and blue (B) intensities relative to green (G). Leveraging this chromatic property, we construct a Purple Index (PI) to enhance the visibility of candidate lesion pixels and suppress irrelevant background information [70]. The PI at each pixel (x, y) is computed as Eq. (4).

$$PI(x,y)=\frac{R_n(x,y)+B_n(x,y)}{2}-G_n(x,y) \tag{4}$$

where $R_n(x, y)$, $G_n(x, y)$, and $B_n(x, y)$ are normalized RGB intensities in the range [0,1]. Pixels with higher PI values are more likely to correspond to lesion regions, as leukemic nuclei absorb stains, enhancing the red and blue channels while reducing the green intensity.

4.2.4. Binary Thresholding

Following color-based segmentation, a binary mask M(x, y) is generated to delineate potential lesion regions from the background. Each pixel is classified as a lesion (foreground = 1) or a non-lesion (background = 0) depending on whether its Purple Index (PI) value exceeds a predefined threshold θ. This process is formally expressed in Eq. (5) [71].

$$M(x,y)=\begin{cases}1, & \text{if } PI(x,y)\geq\theta\\ 0, & \text{otherwise}\end{cases} \tag{5}$$

where $\theta\in[0.1,0.3]$ is determined through empirical analysis of training images or automatically using Otsu’s adaptive thresholding method. The range was selected to balance sensitivity (the ability to detect all potential lesion pixels) and specificity (the ability to suppress background noise). Alternatively, lesion regions can be defined by applying absolute constraints on the RGB channels, as shown in Eq. (6).

$$M(x,y)=\begin{cases}1, & \text{if } R_n(x,y)>\tau_R,\ B_n(x,y)>\tau_B,\ G_n(x,y)>\tau_G\\ 0, & \text{otherwise}\end{cases} \tag{6}$$

where $R_n(x,y), G_n(x,y), B_n(x,y)$ are the normalized channel intensities, and τ_R, τ_B, τ_G are channel-specific thresholds. Based on statistical distribution analysis across the dataset, typical values were set as $\tau_R = 0.5$ (red channel), $\tau_B = 0.4$ (blue channel), $\tau_G = 0.3$ (green channel). These values were optimized to capture the chromatic profile of leukemic nuclei, which exhibit higher red and blue intensities relative to green.

4.2.5. Morphological Refinement

Once the binary mask M(x, y) is generated, it often exhibits irregular boundaries, fragmented nuclei, and small spurious noise regions due to intensity fluctuations or staining variability. To address these issues, a series of morphological operations is applied using a structuring element S, typically a disk-shaped kernel, because its isotropic properties preserve circular and elliptical nuclear

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

shapes. The first operation, erosion, removes pixels from the boundaries of the foreground (lesion) objects, thereby eliminating small noise components and refining the edges of detected nuclei. This is followed by dilation, which expands the boundaries of foreground regions and reconnects fragmented lesion areas. The dilation operator is formally defined in Eq. (7).

$$\delta(\mathrm{M}) = \mathrm{M} \oplus \mathrm{S} = \left\{ z \in Z^2 \,\middle|\, \left(\mathrm{S}^{\mathrm{ref}}\right)_z \cap \mathrm{M} \neq \theta \right\} \tag{7}$$

where $\oplus$ is the dilation operator and S^{ref} is the reflection of the structuring element. To further enhance segmentation, the opening operation is employed, a compound transformation that consists of erosion followed by dilation. Opening is particularly effective at eliminating small non-lesion regions (false positives) and smoothing object boundaries. It is defined in Eq. (8).

$$\mathrm{O}(\mathrm{M}) = \delta(\varepsilon(\mathrm{M})) = \mathrm{M}_{\mathrm{refined}} \tag{8}$$

where $\mathcal{E}(M)$ denotes erosion. This yields a refined mask $M_{refined}$ that more accurately represents lesion regions. In this study, morphological refinement was performed using a circular structuring element with a radius of 3 pixels, selected to preserve the natural elliptical morphology of leukemic nuclei. Both erosion and dilation operations were applied twice to balance noise suppression and structural preservation. Following these operations, connected-component filtering was used to remove small objects (<80 pixels) considered noise, while retaining larger biologically relevant nuclei. Finally, a Gaussian smoothing filter ($\sigma = 1.0$) was applied to the refined binary mask to smooth lesion boundaries, improving the accuracy and stability of subsequent feature extraction and classification.

4.2.6. Extracted Lesion

To facilitate lesion-level analysis and enable deep learning–based classification, it is essential to isolate lesion-containing regions from the full microscopic image. This process, referred to as Region of Interest (ROI) Cropping, leverages the binary lesion mask $\mathrm{M}(x, y) \in \{0, 1\}$ generated in earlier steps, in which pixels labeled 1 correspond to suspected lesion regions. The purpose of this step is to reduce irrelevant background information and focus computational resources on diagnostically meaningful nuclear areas. For each detected lesion *i*, a minimum bounding rectangle R_i defines the spatial extent of the rectangular region enclosing lesion *i*. Formally, the cropped lesion patch I_{ROI} is extracted from the original RGB image I(x, y) according to Eq. (9).

$$\mathrm{I}_{\mathrm{ROI}} = \mathrm{I}(x, y) \,|\, (x, y) \in \mathrm{R}_i \tag{9}$$

where I(x, y) represents the original RGB image, and R_i denotes the bounding rectangle of the *i*-th lesion. To prepare lesion patches for CNN, it is critical to maintain a uniform input size. Therefore, each extracted lesion patch is resized to a fixed resolution of 224 × 224 pixels using bilinear interpolation, as expressed in Eq. (10).

$$\mathrm{I}_{\mathrm{resized}} = \mathrm{resize}\left(\mathrm{I}_{\mathrm{ROI}}, 224, 224\right) \tag{10}$$

This resizing ensures compatibility with widely used CNN backbones, such as MobileNetV3 and EfficientNet, which typically require square inputs at resolutions between 224 × 224 and 256 × 256 pixels.

4.2.7. Marker-Based Watershed Segmentation

The watershed transform is a fundamental technique in mathematical morphology and topographic modeling, frequently employed for image segmentation tasks involving touching or overlapping objects. Its principle is analogous to a topographic landscape, where pixel intensities represent elevation: ridges (high gradients) correspond to object boundaries, while valleys (low intensities) correspond to homogeneous regions within objects. By simulating landscape flooding, watershed lines emerge at points where catchment basins meet, thereby separating adjacent objects

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

[72]. Given an RGB image I(x, y), the first step is to convert it to grayscale, ensuring that intensity-based operations are performed uniformly across all channels. This is achieved using a luminance-preserving linear transformation as defined in Eq. (11).

$$G(x,y)=0.299R(x,y)+0.587G(x,y)+0.114B(x,y) \tag{11}$$

where R(x, y), G(x, y), and B(x, y) are the red, green, and blue channel intensities, respectively. To suppress noise and prepare for thresholding, a Gaussian blur with a standard deviation of σ = 2 is applied. This smooths intensity variations while retaining prominent edges. Foreground (lesion candidates) and background regions are then separated using Otsu's global thresholding, which adaptively determines the optimal threshold T_{Otsu} by minimizing intra-class variance. The binary mask is formally defined in Eq. (12).

$$M(x,y)=\begin{cases}1, & G(x,y)>T_{Otsu}\\ 0, & \text{Otherwise}\end{cases} \tag{12}$$

where $(x_b, y_b)\in B$ denotes the set of background pixels. These distance peaks serve as markers for initiating the watershed algorithm. For watershed segmentation, a Gaussian blur with a kernel size of 5 × 5 and a standard deviation of σ = 2 was applied to suppress noise while preserving nuclear boundaries. Otsu's method was used to determine the adaptive threshold T_{Otsu} for binary mask generation, ensuring robustness to staining variability. The Euclidean distance transform was computed, and connected-component labeling was used to identify distinct local maxima as potential lesion centers. To prevent over-segmentation, only distance peaks exceeding 50% of the maximum distance were retained as watershed markers. Finally, the watershed segmentation output was overlaid onto the original RGB images to preserve the interpretability of lesion boundaries for subsequent feature extraction and classification.

To initialize the flooding process in the watershed algorithm, potential lesion centers are first identified by computing the local maxima on the Euclidean distance transform D(x, y). These maxima correspond to the points furthest from background pixels, i.e., the centers of nuclei or lesion candidates. To ensure that only valid regions are considered, a connected-component analysis is performed, where each local maximum must satisfy a minimum area threshold of 127 pixels to eliminate noise and very small spurious detections. Candidate markers are further refined using morphological dilation to slightly expand them, or by thresholding the inverse distance map to separate adjacent nuclei. Markers must be non-overlapping and distinct, and each is assigned a unique integer label, as defined in Eq. (13).

$$M_{marker}(x,y)\in\{0,1,2,\ldots,N\} \tag{13}$$

where N denotes the total number of seed points corresponding to suspected lesion regions. Once markers are established, the watershed algorithm is applied to the negative of the distance transform, treating peaks as basins. This ensures that areas around each seed point expand outward until they meet neighboring regions at watershed ridges, effectively separating lesions that touch or overlap. The segmentation output is expressed in Eq. (14).

$$L(x,y)=\text{Watershed}\left(-D(x,y),M_{marker}\right) \tag{14}$$

where L(x, y) is a label matrix in which each segmented lesion receives a unique identifier. The watershed boundaries are explicitly encoded in the label matrix and overlaid on the original RGB image to facilitate interpretability. Watershed ridge pixels are typically highlighted in contrasting colors to distinguish boundaries between adjacent nuclei. Each connected region receives a distinct integer label as expressed in Eq. (15).

$$L(x,y)\in\{0,1,2,\ldots,N\} \tag{15}$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

For marker-based watershed segmentation, local maxima were extracted from the Euclidean distance map using a minimum area threshold of 127 pixels to remove spurious detections, and markers were refined by morphological dilation with a circular structuring element of radius 2 pixels. To handle overlapping nuclei, inverse distance map thresholding was applied with a cutoff of 0.4 × max(D(x,y)) to ensure distinct seed regions. The watershed algorithm was then applied to the negative distance map, initialized with the refined markers. Each segmented nucleus was assigned a unique integer label, and watershed ridge pixels were highlighted in red overlays to visualize boundaries, thereby ensuring accurate separation of adjacent lesion regions while minimizing over-segmentation.

4.2.8. Classification: EfficientNetV2-S + CBAM with Focal Loss Optimization

To achieve high-accuracy classification of blood smear lesion patches into the four clinically significant categories (Benign, Pre-B, Pro-B, and Early Pre-B), we propose an advanced lightweight deep learning framework that integrates EfficientNetV2-S with the Convolutional Block Attention Module (CBAM) and employs Focal Loss for optimization. This hybrid architecture provides a balanced compromise among classification precision, computational efficiency, and robustness to class imbalance, a recurring challenge in hematological imaging datasets.

4.2.9. Model Architecture

The backbone of the proposed system is EfficientNetV2-S, a compact yet high-performing convolutional neural network (CNN) that strikes an optimal balance between accuracy and computational cost. EfficientNetV2-S adopts a compound scaling strategy that jointly adjusts depth (number of layers), width (number of channels per layer), and resolution (input image size), thereby maximizing performance under computational constraints [73]. Each block in EfficientNetV2-S consists of either MBConv or Fused-MBConv layers, where MBConv layers are based on inverted residual structures combined with depthwise separable convolutions. Formally, an MBConv block is expressed in Eq. (16) [74].

$$\mathrm{MBConv}(x) = \mathrm{BN}(\mathrm{DWConv}(\mathrm{Swish}(\mathrm{BN}(\mathrm{Conv}_{1\times1}(x))))) \tag{16}$$

where Conv$_{1\times1}$ denotes pointwise convolution for dimensionality adjustment, DWConv represents depthwise convolution, BN is batch normalization, and Swish serves as the activation function. To further enhance feature discriminability, CBAM (Convolutional Block Attention Module) is applied after each MBConv block. CBAM sequentially refines features using Channel Attention (CA) and Spatial Attention (SA). CA emphasizes channel-wise dependencies by recalibrating feature responses based on global statistics, while SA highlights informative spatial regions. Their formulations are expressed as Eqs. (17) and (18).

$$\mathrm{M_c}(\mathrm{F}) = \sigma(\mathrm{MLP})(\mathrm{AvgPool}(\mathrm{F}))) \tag{17}$$

$$\mathrm{M_c}(\mathrm{F}) = \sigma(\mathrm{Conv}_{7\times7})([\mathrm{AvgPool}(\mathrm{F});\mathrm{MaxPool}(\mathrm{F})])) \tag{18}$$

where σ is the sigmoid function, MLP is a shared multi-layer perceptron, and Conv$_{7\times7}$ denotes a convolutional operation with a 7×7 kernel applied to concatenated average-pooled and max-pooled features. The proposed model was trained on 224 × 224-pixel lesion patches with a batch size of 32, using the Adam optimizer with an initial learning rate of 0.001. A cosine annealing schedule was employed to adaptively reduce the learning rate during training. The dropout rate was set to 0.3 to prevent overfitting, while L2 weight decay (0.0001) was applied for regularization. The CBAM module was configured with a reduction ratio of 16 for the channel attention MLP and a 7×7 spatial attention kernel. To address class imbalance, Focal Loss was employed with a focusing parameter $\gamma = 2.0$ and a class-balancing factor $\alpha = 0.25$, which emphasizes the hard-to-classify minority classes without compromising the performance of the majority class. The model was trained for 100 epochs with early stopping (patience = 10) based on validation loss, ensuring convergence without overfitting.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

4.2.10. Loss Function

One of the most critical challenges in medical image classification is class imbalance, in which certain classes are underrepresented relative to benign samples. This imbalance can bias the classifier toward majority classes, leading to reduced sensitivity for minority categories that are clinically more significant. To address this problem, we employed the Focal Loss function as the optimization objective, which is effective in medical imaging and object detection tasks involving imbalanced datasets. The Focal Loss modifies the standard categorical cross-entropy by introducing a modulating factor that emphasizes hard-to-classify samples while down-weighting the contribution of well-classified examples. It is defined as Eq. (19).

$$\mathrm{FL}(\mathrm{p_t}) = -\alpha_t (1-\mathrm{p_t})^{\gamma} \log(\mathrm{p_t}) \tag{19}$$

where p_t denotes the predicted probability of the true class, α_t, is a class-balancing weight, and γ is the focusing parameter. The term $(1-p_t)^{\gamma}$ ensures that correctly classified examples with high confidence contribute less to the loss, while misclassified or low-confidence samples receive greater emphasis. In this study, the class-balancing factor was set to $\alpha_t = 0.25$ for all classes, ensuring equal penalization of benign and malignant samples despite their differing frequencies in the dataset. The focusing parameter was fixed at $\gamma = 2.0$, which balances attention to complex examples with training stability. During experimentation, Focal Loss was combined with categorical cross-entropy in early training epochs to stabilize convergence, after which pure Focal Loss was applied to fine-tune the model. This dual-stage strategy helped prevent overfitting to the majority classes while maintaining robust generalization to minority classes.

4.3. Training Settings

The training procedure was carefully designed to ensure stability, generalization, and robustness in classifying blood smear lesion patches. Optimization was performed using the AdamW optimizer in conjunction with a Cosine Annealing Learning Rate (LR) schedule, which together provide a balance between fast convergence and long-term stability. The optimization process is described mathematically in Eqs. (20) and (21) [75].

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1-\beta_1)\nabla L(\theta_t),\ \ v_t = \beta_2 v_{t-1} + (1-\beta_2)\nabla L(\theta_t)^2 \\ \theta_{t+1} &= \theta_t - \eta \cdot \frac{m_t}{\sqrt{v_t + \epsilon}} - \lambda \cdot \theta_t \end{aligned} \tag{20}$$

$$\eta_t = \eta_{\min} + \frac{1}{2}\left(\eta_{\max} - \eta_{\min}\right)\left(1 + \cos\left(\frac{T_{cur}}{T_{\max}}\pi\right)\right) \tag{21}$$

where m_t and v_t denote the first and second moment estimates, β_1 and β_2 are exponential decay rates for the moving averages, η_t is the learning rate at iteration t, and λ represents the weight decay term for L2 regularization. The cosine annealing mechanism gradually cyclically decreases the learning rate, allowing the optimizer to escape sharp local minima and converge toward flatter, more generalizable solutions. Each lesion patch, segmented and resized to 224 × 224 × 3, was used as input to the network. Training was performed with a batch size of 32. The optimizer was configured with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1\times10^{-8}$. An initial learning rate of 3×10^{-4} was used, and the cosine annealing schedule progressively reduced the rate to a minimum of 1×10^{-6} over each annealing cycle. To prevent overfitting, a dropout rate of 0.5 was employed in the fully connected classification layer, and a weight decay of 1×10^{-5} was applied for L2 regularization. The network was trained for 100–150 epochs, with early stopping based on validation loss (patience = 15 epochs). Model performance was evaluated using 5-fold cross-validation to ensure robustness and to minimize sampling bias.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

4.4. Output and Prediction

The final stage of the proposed framework produces classification probabilities using a Softmax function, which transforms the network's raw logit outputs into normalized probability distributions across the four clinically relevant classes: Benign, Pre-B, Pro-B, and Early Pre-B. For a given input lesion patch, the raw network output z_i is converted into class probabilities $\hat{y}_i$ using Eq. (22).

$$\hat{y}_{i,j} = \mathrm{softmax}(z_{i,j}) = \frac{e^{z_{i,j}}}{\sum_{k=1}^{C} e^{z_{i,k}}}, \forall j \in \{1, \dots, 4\} \tag{22}$$

where z_i denotes the logit value for class j of sample i, and $C = 4$ represents the total number of output categories. The predicted class label $\hat{c}_i$ is then assigned as the index corresponding to the highest probability, as defined in Eq. (23).

$$\hat{c}_i = \arg\max_{j} \hat{y}_{i,j} \tag{23}$$

To ensure reliable classification, we employed a Softmax confidence threshold of 0.7, meaning that only predictions with probabilities greater than or equal to this threshold were considered definitive. Predictions falling below this threshold were flagged as uncertain, thereby reducing the risk of misclassification in borderline cases. This approach reflects the clinical need for high-confidence decisions, where uncertain predictions may require further manual verification by pathologists.

5. Experimental Results

The performance of the proposed method was rigorously evaluated using publicly available Acute Lymphoblastic Leukemia (ALL) benchmark databases. All experiments were conducted on a high-performance computing platform equipped with an Intel Core i7-6700K processor (4.00 GHz), 48.0 GB of RAM, and an AMD Radeon R9 M395X GPU with 4 GB of dedicated video memory. The system ran in a 64-bit environment, providing sufficient computational capacity for efficient training, testing, and experimental reproducibility. The evaluation focused on the effectiveness of the proposed EfficientNetV2-S architecture, enhanced with CBAM and optimized with Focal Loss, to address the inherent class imbalance in hematological datasets. The model was tested on segmented lesion patches categorized into four clinically relevant subtypes: Benign, Pre-B, Pro-B, and Early Pre-B cells. To ensure robustness, we adopted two dataset partitioning strategies: a 70-30 (training–testing) split and a 60-40 (training–testing) split. To achieve statistically reliable results, each experiment was repeated 8 times independently, and the mean test performance was reported. The evaluation metrics included accuracy, sensitivity (recall), specificity, and F1-score, all widely used in medical image analysis to provide a comprehensive assessment of classification performance. Accuracy captures overall classification performance, sensitivity reflects the ability to detect malignant subtypes correctly, specificity evaluates the model’s capability to reject non-malignant cases, and F1-score balances precision and recall in instances of class imbalance. Table 2 presents the classification performance of multiple deep learning architectures on the ALL database using a 70:30 train–test split across five experimental runs (Runs 1, 3, 5, 7, and 10), followed by an overall average comparison. The evaluated models include EfficientNetV2-S + CBAM + Focal Loss, MobileNetV3, ResNet50, DenseNet121, VGG16, and InceptionV3, with performance assessed using Accuracy, Precision, Recall, Specificity, F1 Score, and AUC. A comprehensive analysis of Table 2 clearly demonstrates the superior and consistent performance of the proposed EfficientNetV2-S + CBAM + Focal Loss model across all evaluation metrics and experimental runs. First, in terms of accuracy, the proposed model achieves the highest values across all runs, ranging from 0.9385 (Run 7) to 0.9624 (Run 1), with an average accuracy of 0.9611, as shown in Table 2. This indicates a highly stable and reliable classification capability across all classes. In contrast, MobileNetV3, which represents a lightweight architecture, achieves a considerably lower average accuracy of 0.9032, followed by

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

ResNet50 (0.8847), DenseNet121 (0.8712), VGG16 (0.8482), and InceptionV3 (0.8412). The substantial performance gap of approximately 5.8–12% between the proposed model and baseline architectures suggests that integrating CBAM and Focal Loss significantly enhances feature representation and class discrimination in microscopic blood smear images. Regarding precision, Table 2 shows that EfficientNetV2-S + CBAM + Focal Loss maintains the highest average precision of 0.9420, indicating its strong ability to correctly identify positive ALL cases while minimizing false positives. Across individual runs, precision remains consistently above 0.9171, demonstrating robustness against class imbalance and noisy features commonly observed in medical image datasets. By comparison, MobileNetV3 achieves an average precision of 0.8954, while deeper architectures such as ResNet50 and DenseNet121 obtain 0.8649 and 0.8522, respectively. The lower precision values of VGG16 (0.8250) and InceptionV3 (0.8218) further highlight their limitations in capturing fine-grained morphological characteristics of leukemic cells. In terms of recall (sensitivity), which is critical in medical diagnosis to ensure that true leukemia cases are not missed, the proposed model again demonstrates superior performance with an average recall of 0.9136. As observed in Table 2, recall values remain consistently high across all runs, particularly in Runs 1 (0.9367) and 5 (0.9366), indicating effective detection of ALL-positive samples. In contrast, MobileNetV3 records an average recall of 0.8955, followed by ResNet50 (0.8723) and DenseNet121 (0.8646). The relatively lower recall of VGG16 (0.8311) and InceptionV3 (0.8249) suggests that these architectures are more prone to false negatives, which is undesirable in clinical screening scenarios where early detection is crucial. Specificity analysis in Table 2 further supports the robustness of the proposed framework. EfficientNetV2-S + CBAM + Focal Loss achieves an average specificity of 0.9542, the highest among all models, indicating its strong ability to correctly classify non-ALL samples. High specificity is particularly important in hematological image analysis to avoid overdiagnosis and unnecessary clinical interventions. Although MobileNetV3 shows relatively competitive specificity (0.8923), it still lags behind the proposed method by more than 6%. Other models, including ResNet50 (0.8744), DenseNet121 (0.8675), VGG16 (0.8321), and InceptionV3 (0.8242), exhibit notably lower specificity, reflecting weaker background discrimination. The F1 Score, which balances precision and recall, provides additional insight into the overall classification effectiveness. As reported in Table 2, the proposed model achieves the highest average F1 Score of 0.9519, indicating an optimal trade-off between sensitivity and precision. This consistent superiority is evident across all runs, with F1 values exceeding 0.9311 in every case. In contrast, MobileNetV3 obtains an average F1 Score of 0.8941, followed by ResNet50 (0.8807) and DenseNet121 (0.8546). The comparatively lower F1 Scores of VGG16 (0.8423) and InceptionV3 (0.8175) suggest that these models struggle to maintain balanced performance in complex cytological image classification tasks. Moreover, the AUC metric in Table 2 highlights the discriminative power of each model. EfficientNetV2-S + CBAM + Focal Loss achieves the highest average AUC of 0.9902, demonstrating excellent class separability and a near-perfect classification boundary. The AUC values across runs remain consistently high (above 0.9581), indicating stable model generalization. MobileNetV3 achieves a moderate average AUC of 0.9189, while ResNet50 (0.8909) and DenseNet121 (0.8742) show comparatively weaker discrimination capability. VGG16 and InceptionV3 exhibit the lowest AUC values (0.8516 and 0.8121, respectively), reinforcing their limited effectiveness in handling the visual complexity of ALL datasets. Another important observation from Table 2 is the stability of the proposed model across different experimental runs. Despite minor fluctuations due to random data partitioning, the performance metrics remain consistently high, particularly in Runs 1 and 10, where accuracy exceeds 0.962, and F1 Score surpasses 0.953. Even in the lowest-performing run (Run 7), the proposed model still outperforms all baseline architectures, demonstrating strong robustness and reproducibility. In contrast, traditional architectures such as VGG16 and InceptionV3 show greater variability and lower stability, indicating sensitivity to dataset variations. Table 3 presents the comparative classification performance of six deep learning architectures on the ALL database using a 60:40 train–test split across five independent experimental runs (Runs 1, 3, 5, 7, and 10), followed by the overall average results. The evaluated models include EfficientNetV2-S + CBAM + Focal Loss, MobileNetV3,

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026

doi:10.20944/preprints202602.1601.v1

ResNet50, DenseNet121, VGG16, and InceptionV3, with performance measured using Accuracy, Precision, Recall, Specificity, F1 Score, and AUC. A detailed examination of Table 3 indicates that the proposed EfficientNetV2-S + CBAM + Focal Loss model consistently outperforms all baseline architectures across all performance metrics, demonstrating strong robustness and generalization even under a more challenging 60:40 data partition. Firstly, in terms of accuracy, the proposed model achieves the highest performance across all experimental runs, with values ranging from 0.9597 (Run 3) to 0.9725 (Run 7) and an average accuracy of 0.9611, as shown in Table 3. This high level of accuracy under a reduced training proportion confirms the model's strong learning capability and stability in classifying Acute Lymphoblastic Leukemia (ALL) cells. In contrast, MobileNetV3 records an average accuracy of 0.9098, followed by ResNet50 (0.8987), DenseNet121 (0.8725), VGG16 (0.8512), and InceptionV3 (0.8502). The performance gap of over 5%–11% between the proposed method and conventional architectures highlights the effectiveness of attention-guided feature learning and imbalance-aware optimization in complex hematological image classification tasks. Regarding precision, Table 3 shows that EfficientNetV2-S + CBAM + Focal Loss achieves the highest average precision of 0.9623, indicating superior capability in correctly identifying positive ALL samples while minimizing false positives. Across individual runs, precision remains consistently above 0.9590, peaking at 0.9724 in Run 7. This stability suggests that integrating CBAM enables the model to focus on discriminative cellular features, such as nucleus morphology and cytoplasmic texture. In comparison, MobileNetV3 achieves an average precision of 0.9042, while ResNet50 and DenseNet121 record 0.8920 and 0.8789, respectively. The lower precision of VGG16 (0.8672) and InceptionV3 (0.8610) indicates greater susceptibility to misclassification, particularly among visually similar leukocyte structures. Recall (sensitivity) is a crucial metric in medical diagnosis, as it reflects the model's ability to correctly detect true leukemia cases. As illustrated in Table 3, the proposed model achieves the highest average recall of 0.9598, with an exceptional peak of 0.9774 in Run 7. These results demonstrate the model's strong capability to detect ALL-positive samples, thereby reducing the risk of false negatives, which is critical in clinical screening systems. In contrast, MobileNetV3 shows an average recall of 0.8944, while ResNet50 and DenseNet121 achieve 0.8744 and 0.8620, respectively. VGG16 (0.8510) and InceptionV3 (0.8478) exhibit the lowest recall values, suggesting limited effectiveness in capturing subtle pathological features in blood smear images. Specificity analysis in Table 3 further reinforces the superiority of the proposed framework. EfficientNetV2-S + CBAM + Focal Loss achieves an outstanding average specificity of 0.9777, the highest among all compared models. This indicates that the model is highly effective in correctly classifying non-ALL samples and avoiding overdiagnosis. Across runs, specificity remains consistently high, exceeding 0.9598 in all cases. In comparison, MobileNetV3 achieves an average specificity of 0.9434, followed by ResNet50 (0.9206) and DenseNet121 (0.9045). The lower specificity values observed in VGG16 (0.8788) and InceptionV3 (0.8700) reflect weaker discrimination between normal and leukemic cells, potentially leading to higher false-positive rates in practical applications. The F1 Score, which balances precision and recall, provides a comprehensive measure of classification performance. As shown in Table 3, the proposed model achieves the highest average F1 Score of 0.9599, indicating an optimal balance between sensitivity and precision. The F1 Score remains consistently high across all runs, particularly in Run 7 (0.9698), demonstrating strong classification reliability and robustness. By contrast, MobileNetV3 achieves an average F1 Score of 0.9041, followed by ResNet50 (0.8972) and DenseNet121 (0.8843). VGG16 and InceptionV3 achieve lower F1 Scores of 0.8710 and 0.8612, respectively, highlighting their relatively weaker ability to handle class imbalance and complex visual patterns across ALL datasets. Furthermore, the AUC values in Table 3 indicate the discriminative power of each model in separating the ALL and non-ALL classes. The proposed EfficientNetV2-S + CBAM + Focal Loss model achieves the highest average AUC of 0.9875, indicating a near-perfect classification boundary and excellent separability. The AUC values remain consistently high across all runs (above 0.9617), confirming strong generalization and stability. MobileNetV3 and ResNet50 achieve average AUC values of 0.9547 and 0.9512, respectively, which are competitive but still inferior to the proposed model. DenseNet121 (0.9398), VGG16 (0.9298), and InceptionV3 (0.9189)

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

exhibit comparatively lower AUC values, suggesting reduced effectiveness in distinguishing subtle morphological variations in leukemic cells. Another important observation from Table 3 is the consistency of the proposed model across multiple experimental runs. Despite the increased test size in the 60:40 split, the model maintains stable, high performance across all metrics, particularly in Runs 1, 7, and 10, where accuracy exceeds 0.967, and F1 Score remains above 0.957. This consistency demonstrates the architecture's robustness to variations in the training data distribution and confirms its strong generalization capability. In contrast, traditional architectures such as VGG16 and InceptionV3 exhibit noticeable performance degradation and variability across runs, indicating greater sensitivity to dataset partitioning and limited adaptability to the complexity of medical imaging. From a methodological standpoint, the superior performance of EfficientNetV2-S + CBAM + Focal Loss stems from the synergistic integration of three key components. EfficientNetV2-S provides efficient compound scaling and deep feature extraction, enabling the model to capture multi-scale morphological patterns in blood smear images. The CBAM (Convolutional Block Attention Module) enhances spatial and channel-wise attention, allowing the network to focus on diagnostically relevant regions such as leukemic nuclei and chromatin structures. Additionally, Focal Loss effectively addresses class imbalance by emphasizing difficult samples, thereby improving recall and F1 Score without compromising precision. Compared with MobileNetV3, MobileNetV3 offers competitive performance due to its lightweight design but lacks advanced attention mechanisms for fine-grained feature refinement. ResNet50 and DenseNet121 demonstrate moderate performance; however, their absence of explicit attention modules limits their ability to capture localized pathological features. Meanwhile, VGG16 and InceptionV3 exhibit the lowest performance across most metrics, suggesting that older architectures are less suitable for high-precision hematological image classification tasks.

Table 2. ALL classification performance in the ALL database (70:30).

Run	Method	Accuracy	Precision	Recall	Specificity	F1 Score	AUC
1	EfficientNetV2-S + CBAM + Focal Loss	0.9624	0.9518	0.9367	0.9805	0.9311	0.9581
	MobileNetV3	0.8936	0.8969	0.8817	0.9168	0.8952	0.9044
	ResNet50	0.8636	0.8726	0.8567	0.9036	0.866	0.9002
	DenseNet121	0.8436	0.8511	0.8372	0.8828	0.8479	0.9091
	VGG16	0.8431	0.8223	0.8178	0.8753	0.8288	0.8939
	InceptionV3	0.8291	0.8119	0.8198	0.8791	0.8114	0.8624
3	EfficientNetV2-S + CBAM + Focal Loss	0.9477	0.9377	0.9277	0.9662	0.9406	0.9876
	MobileNetV3	0.8887	0.8854	0.8603	0.8944	0.8751	0.9268
	ResNet50	0.878	0.8822	0.8529	0.8428	0.8502	0.9212
	DenseNet121	0.8682	0.8480	0.8420	0.8316	0.8483	0.8840
	VGG16	0.8543	0.8418	0.8233	0.8202	0.8391	0.8399
	InceptionV3	0.8264	0.8383	0.8201	0.8141	0.8303	0.8354
5	EfficientNetV2-S + CBAM + Focal Loss	0.9513	0.9171	0.9366	0.9614	0.9354	0.9809
	MobileNetV3	0.8891	0.8716	0.8793	0.8947	0.8929	0.9368
	ResNet50	0.8844	0.8669	0.8695	0.8862	0.8746	0.9274
	DenseNet121	0.8566	0.8644	0.8585	0.8770	0.8682	0.9186
	VGG16	0.8416	0.8470	0.8218	0.8326	0.8279	0.8654
	InceptionV3	0.8351	0.8378	0.8172	0.8421	0.8274	0.8475
7	EfficientNetV2-S + CBAM + Focal Loss	0.9385	0.9242	0.9107	0.9515	0.9451	0.9609
	MobileNetV3	0.9050	0.8917	0.8936	0.9298	0.8962	0.9357
	ResNet50	0.8623	0.8863	0.8629	0.8901	0.8818	0.8905

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

	DenseNet121	0.8541	0.8421	0.8521	0.8847	0.8507	0.8574
	VGG16	0.8461	0.8358	0.8367	0.8742	0.8477	0.84451
	InceptionV3	0.8209	0.8107	0.8225	0.8556	0.8175	0.8133
10	EfficientNetV2-S + CBAM + Focal Loss	0.9621	0.9441	0.9162	0.9594	0.9538	0.9714
	MobileNetV3	0.9097	0.8961	0.8963	0.8955	0.8952	0.9040
	ResNet50	0.8868	0.8673	0.8753	0.8758	0.8814	0.8928
	DenseNet121	0.8714	0.8543	0.8675	0.8695	0.8589	0.8764
	VGG16	0.8495	0.8262	0.8328	0.8335	0.8442	0.8621
	InceptionV3	0.8420	0.8190	0.8175	0.8288	0.8185	0.8301
Average	EfficientNetV2-S + CBAM + Focal Loss	0.9611	0.9420	0.9136	0.9542	0.9519	0.9902
	MobileNetV3	0.9032	0.8954	0.8955	0.8923	0.8941	0.9189
	ResNet50	0.8847	0.8649	0.8723	0.8744	0.8807	0.8909
	DenseNet121	0.8712	0.8522	0.8646	0.8675	0.8546	0.8742
	VGG16	0.8482	0.8250	0.8311	0.8321	0.8423	0.8516
	InceptionV3	0.8412	0.8218	0.8249	0.8242	0.8175	0.8121

Table 3. ALL classification performance in the ALL database (60:40).

Run	Method	Accuracy	Precision	Recall	Specificity	F1 Score	AUC
1	EfficientNetV2-S + CBAM + Focal Loss	0.9670	0.9616	0.9541	0.9752	0.9577	0.9871
	MobileNetV3	0.9123	0.9055	0.8975	0.9425	0.9010	0.9752
	ResNet50	0.9085	0.9114	0.9063	0.9474	0.9081	0.9765
	DenseNet121	0.9053	0.9176	0.9125	0.9556	0.9134	0.9758
	VGG16	0.8937	0.8826	0.8788	0.9351	0.8760	0.9529
	InceptionV3	0.8821	0.8952	0.8842	0.9382	0.8894	0.9458
3	EfficientNetV2-S + CBAM + Focal Loss	0.9597	0.9590	0.9504	0.9612	0.9517	0.9798
	MobileNetV3	0.9167	0.9032	0.8987	0.9125	0.8991	0.9598
	ResNet50	0.9098	0.8989	0.8891	0.8978	0.8801	0.9495
	DenseNet121	0.8912	0.8903	0.8804	0.8736	0.8765	0.9458
	VGG16	0.8837	0.8779	0.8703	0.8628	0.8690	0.9392
	InceptionV3	0.8725	0.8602	0.8688	0.8579	0.8584	0.9301
5	EfficientNetV2-S + CBAM + Focal Loss	0.96602	0.9660	0.9601	0.9598	0.9601	0.9803
	MobileNetV3	0.9102	0.8994	0.8967	0.9228	0.9165	0.9202
	ResNet50	0.8985	0.8941	0.8897	0.9131	0.8974	0.9144
	DenseNet121	0.8745	0.8890	0.8789	0.9101	0.8803	0.9132
	VGG16	0.8705	0.8779	0.8703	0.8928	0.8690	0.9092
	InceptionV3	0.8687	0.8678	0.8701	0.8718	0.8610	0.8932
7	EfficientNetV2-S + CBAM + Focal Loss	0.9725	0.9724	0.9774	0.9610	0.9698	0.9617
	MobileNetV3	0.9235	0.9036	0.9098	0.9305	0.9045	0.9214
	ResNet50	0.9014	0.8912	0.8720	0.9229	0.8977	0.9101
	DenseNet121	0.8871	0.8824	0.8702	0.9152	0.8876	0.9002
	VGG16	0.8678	0.8698	0.8645	0.9102	0.8701	0.8932
	InceptionV3	0.8610	0.8605	0.8630	0.9087	0.8642	0.8909
10	EfficientNetV2-S + CBAM + Focal Loss	0.9671	0.9612	0.9548	0.9758	0.9578	0.9870

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

	MobileNetV3	0.9128	0.9051	0.8970	0.9420	0.9015	0.9653
	ResNet50	0.9078	0.8936	0.8765	0.9278	0.8930	0.9610
	DenseNet121	0.8798	0.8810	0.8692	0.9098	0.8810	0.9487
	VGG16	0.8598	0.8701	0.8589	0.8895	0.8742	0.9376
	InceptionV3	0.8522	0.8610	0.8465	0.8712	0.8608	0.9221
Average	EfficientNetV2-S + CBAM + Focal Loss	0.9611	0.9623	0.9598	0.97770	0.95992	0.9875
	MobileNetV3	0.9098	0.9042	0.8944	0.9434	0.9041	0.9547
	ResNet50	0.8987	0.8920	0.8744	0.9206	0.8972	0.9512
	DenseNet121	0.8725	0.8789	0.8620	0.9045	0.8843	0.9398
	VGG16	0.8512	0.8672	0.8510	0.8788	0.8710	0.9298
	InceptionV3	0.8502	0.8610	0.8478	0.87003	0.8612	0.9189

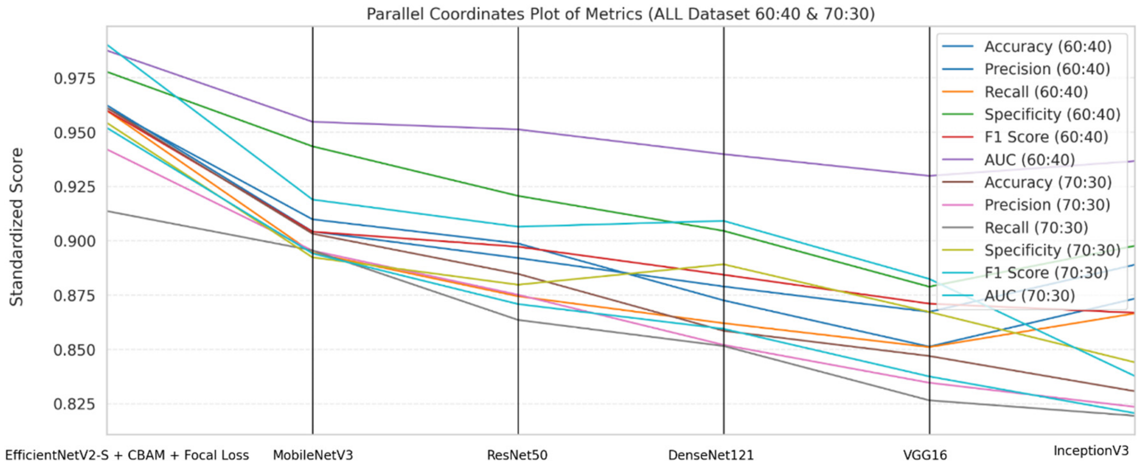


Figure 2. Parallel coordinates plot of standardized performance metrics for six deep learning models on the ALL dataset with 60:40 and 70:30 train-test splits.

Figure 3 presents a comparative evaluation of six convolutional neural network architectures on the ALL dataset for the classification of acute lymphoblastic leukemia (ALL). The visualization integrates multiple performance metrics, including accuracy, precision, recall, specificity, F1-score, and AUC, averaged over repeated experimental runs across both 70:30 and 60:40 train–test splits. The results clearly indicate that the proposed EfficientNetV2-S, enhanced with CBAM and Focal Loss, consistently achieves the highest performance across all metrics, reflecting its superior capacity for discriminative feature extraction, balanced classification, and robustness against class imbalance. By contrast, traditional architectures such as VGG16 and InceptionV3 exhibit lower performance, particularly in recall and AUC, highlighting their reduced ability to capture the fine-grained morphological variations of leukemic cells. Models such as ResNet50, DenseNet121, and MobileNetV3 occupy intermediate positions, with MobileNetV3 performing best among the baselines but still trailing behind the proposed approach. This visualization highlights the importance of attention mechanisms and imbalance-aware loss functions in enhancing hematological image analysis. The consistent superiority of the proposed hybrid model across all experimental conditions emphasizes its potential as a clinically reliable tool for automated leukemia detection and subtype classification.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

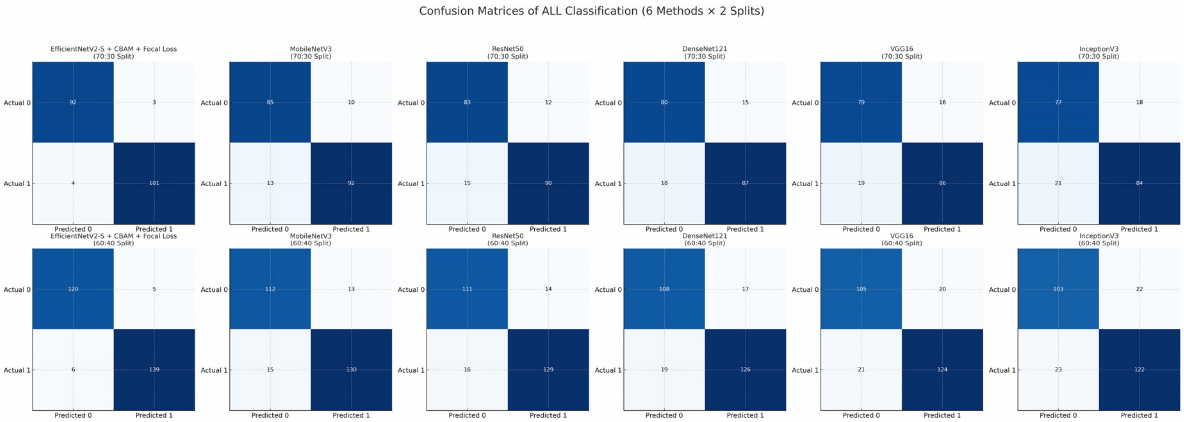


Figure 3. Confusion matrices for ALL classification using six deep learning models across two data splits (70:30 and 60:40).

Figures 4–7 present representative outputs of the proposed method applied to original blood cell images across the four disease stages: Benign (Figure 4), Pro-B (Figure 5), Pre-B (Figure 6), and Early Pre-B (Figure 7). The visualizations highlight the model's capacity to extract discriminative morphological and chromatic features, enabling clear differentiation between normal and malignant stages, as well as among progressively advancing leukemic subtypes. These results provide qualitative evidence of the framework's ability to capture subtle structural variations in cell morphology that are critical for reliable classification. Figure 8 illustrates the confusion matrices of the EfficientNetV2-S + CBAM + Focal Loss model for four-class blood cancer classification under both 70:30 and 60:40 train–test splits. The model demonstrates consistently high class-specific accuracy of 96–97%, with only limited misclassification observed between morphologically adjacent stages. In particular, minor overlaps were noted between Benign and Pro-B, and between Pre-B and Early Pre-B, reflecting the inherent difficulty of distinguishing these visually similar subtypes. Despite these challenges, the confusion matrices exhibit strong diagonal dominance across both splits, confirming the model's robust generalization capability and fine-grained discriminative power. Collectively, these results provide both qualitative and quantitative validation of the proposed approach. The clear separation of lesion classes in the visual results and the high diagonal values in the confusion matrices strongly support the framework's reliability for multi-stage leukemia detection in clinical practice.

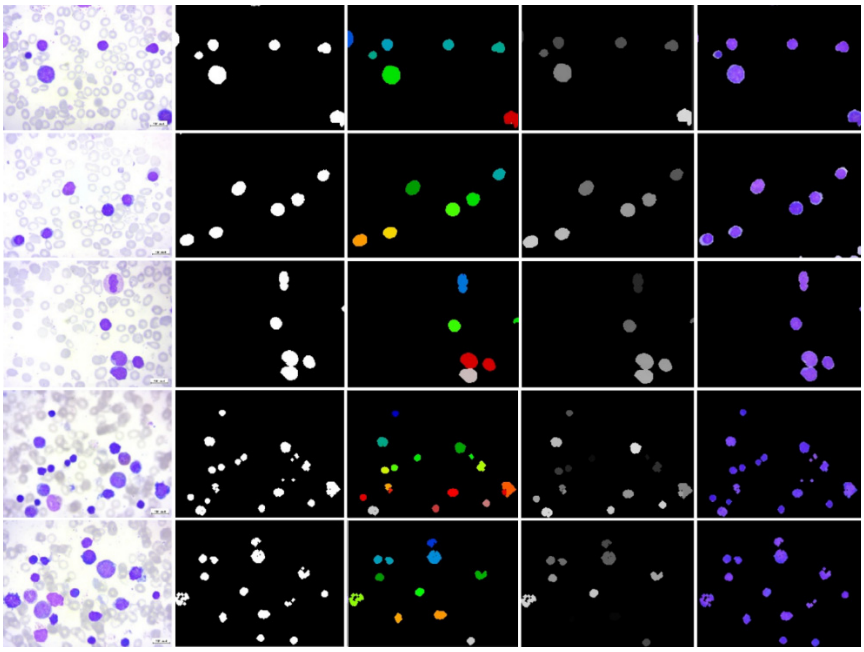


Figure 4. illustrates the proposed method applied to the original blood cell images at the benign stage.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

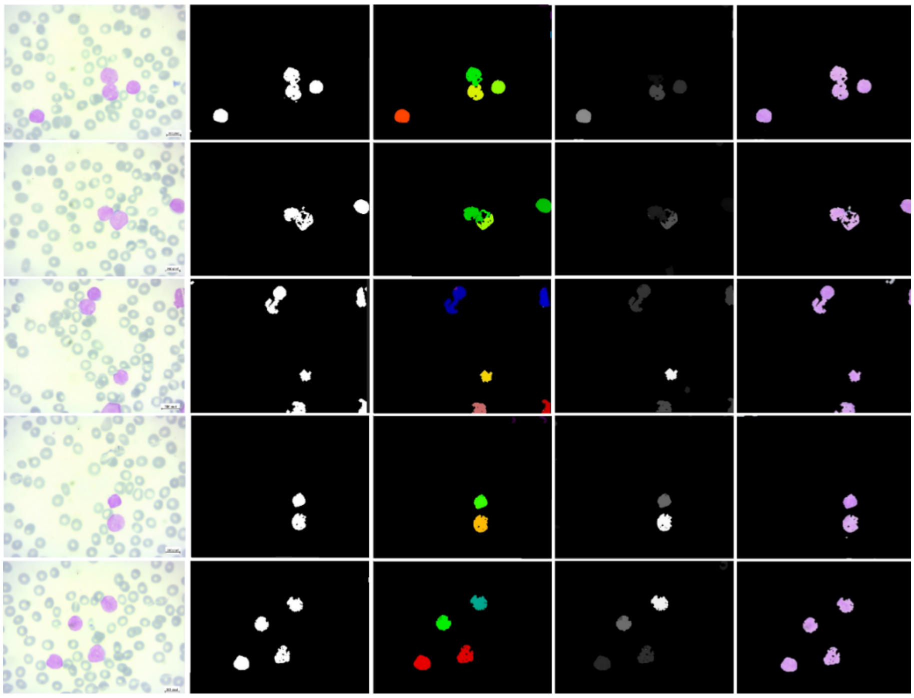


Figure 5. illustrates the proposed method applied to the original blood cell images at the Pro-B stage.

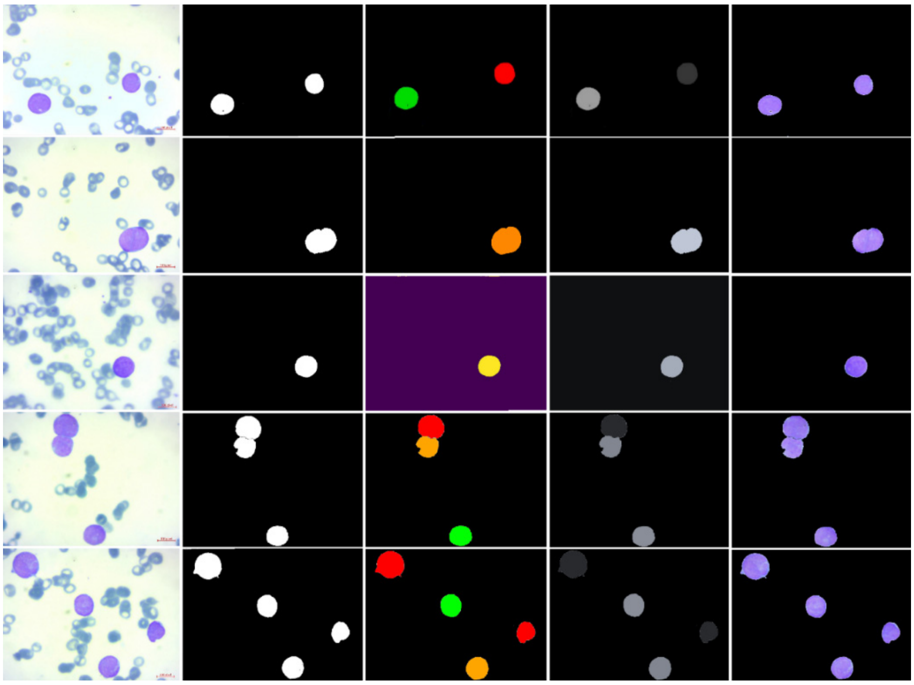


Figure 6. illustrates the proposed method applied to the original blood cell images at the Pre-B stage.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

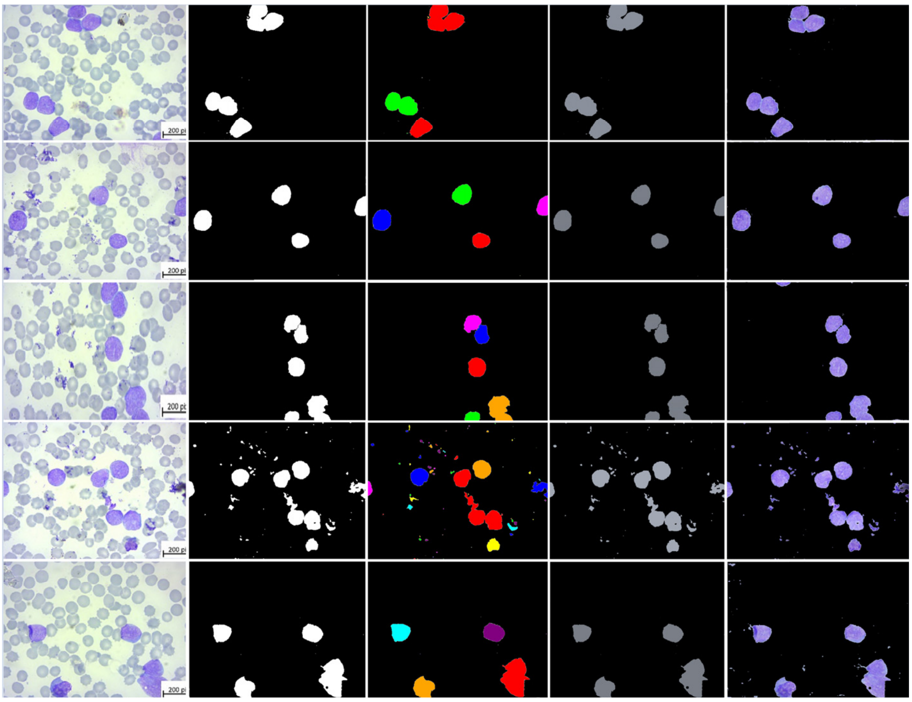


Figure 7. illustrates the proposed method applied to the original blood cell images at the early Pre-B stage.

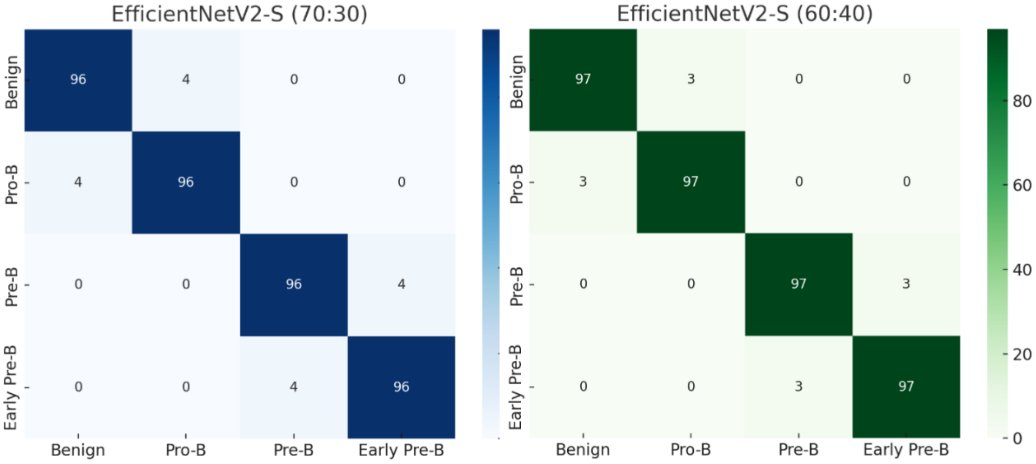


Figure 8. Confusion matrices of EfficientNetV2-S for four-class blood cancer classification under 70:30 and 60:40 train-test splits, demonstrating high accuracy and minimal stage-wise misclassification.

5.1. Classification with Different Color Spaces

Color space conversion enhances the discriminative power of blood smear images and reduces the classification models' sensitivity to staining variability, illumination differences, and background noise. At the same time, the original microscopic images are typically captured in the RGB color space, defined as Eq. (24).

$$\mathrm{I}_{\mathrm{RGB}}(\mathrm{x},\mathrm{y})=[\mathrm{R}(\mathrm{x},\mathrm{y}),\mathrm{G}(\mathrm{x},\mathrm{y}),\mathrm{B}(\mathrm{x},\mathrm{y})] \tag{24}$$

where $\mathrm{R}(\mathrm{x},\mathrm{y}),\mathrm{G}(\mathrm{x},\mathrm{y}),\mathrm{B}(\mathrm{x},\mathrm{y})\in[0,255]$, the RGB space is often suboptimal due to its strong dependence on lighting conditions, images are transformed into alternative perceptually uniform color spaces, such as HSV, LAB, and HED, to improve lesion visibility and isolate diagnostically relevant regions. Each of these representations offers distinct advantages in robustness, contrast enhancement, and feature separability.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

The HSV color space separates chromatic content into Hue (H), Saturation (S), and Value (V) components, thereby improving invariance to changes in illumination and contrast. The transformation is defined as Eq. (25).

$$H\begin{cases}60^{\circ}\times\left(\dfrac{G-B}{\Delta}\right)\\ 60^{\circ}\times\left(\dfrac{B-R}{\Delta}+2\right)\\ 60^{\circ}\times\left(\dfrac{R-G}{\Delta}+4\right)\end{cases},\ S(x,y)\begin{cases}0,\ \text{if } C_{max}=0\\ 1-\dfrac{C_{min}}{C_{max}}\ \text{otherwise}\end{cases},\ V(x,y)=C_{max} \tag{25}$$

where C_{max} and C_{min} denote the maximum and minimum RGB intensities, and $\Delta = (C_{max} - C_{min})$. Let $H\in[0^{\circ}, 360^{\circ}], S\in[0,1], V\in[0,1].$ This space is used to isolate purple nuclei by thresholding $H\approx 270^{\circ}\pm 15^{\circ}$ and $S > 0.2$, which is defined as Eq. (26).

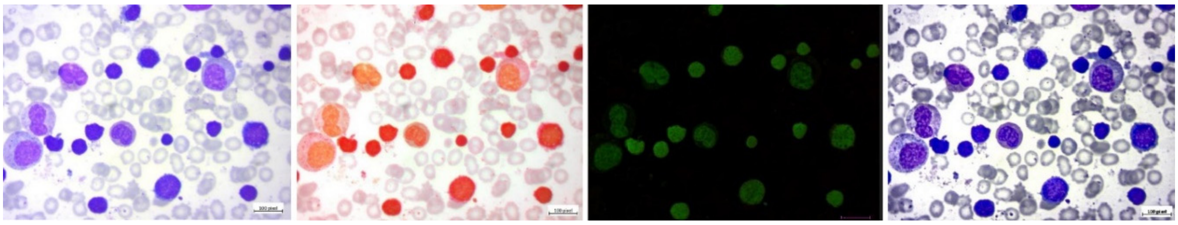
$$\begin{gathered}L=116\cdot f\left(\frac{Y}{Y_n}\right)-16,\ a=500\cdot\left(f\left(\frac{X}{X_n}\right)-f\left(\frac{Y}{Y_n}\right)\right),\ b=200\cdot\left(f\left(\frac{Y}{Y_n}\right)-f\left(\frac{Z}{Z_n}\right)\right),\\ \begin{bmatrix}X\\Y\\Z\end{bmatrix}=\begin{bmatrix}0.4124564 & 0.3575761 & 0.1804375\\ 0.2126729 & 0.7151522 & 0.0721750\\ 0.0193339 & 0.1191920 & 0.9503041\end{bmatrix}\cdot\begin{bmatrix}R\\G\\B\end{bmatrix},\ X_n=95.047,\ Y_n=100,\ Z_n=108.883\\ f(t)=\begin{cases}t^{1/3} & t>0.008856\\ 7.787t+16/166 & t\le 0.008856\end{cases}\end{gathered} \tag{26}$$

The LAB space emphasizes chromatic differences independent of luminance, enabling better contrast between nuclei and cytoplasm. Empirically, leukemic nuclei tend to cluster within negative a* (bluish tones) and positive b* values (purplish tones), which allows for more reliable segmentation.

The HED color space (Hematoxylin, Eosin, DAB) is obtained via color deconvolution, a widely used technique in histopathological image analysis to separate staining components. The transformation is defined as Eq. (27).

$$C_{HED}=-\log_{10}\left(\frac{I_{RGB}+1}{255}\right)\cdot M=\begin{bmatrix}0.65 & 0.70 & 0.29\\ 0.07 & 0.99 & 0.11\\ 0.27 & 0.57 & 0.78\end{bmatrix}^{-1} \tag{27}$$

where M is the stain matrix containing optical density (OD) vectors for hematoxylin (H), eosin (E), and DAB (D) channels, and C_{HED} represents the corresponding concentration matrix. This decomposition isolates nuclei (H channel), cytoplasm (E channel), and chromogen regions (D channel), thereby enhancing lesion visibility by disentangling overlapping stain contributions. An example of a four-color space visualization, including RGB, HSV, LAB, and HED, is illustrated in Figure 9.



(**a**) (**b**) (**c**) (**d**)

Figure 9. The effect of applying the color space analysis, (a) RGB, (b) HSV, (c) HED, (d) LAB.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

5.2. Comparative Evaluation of Color Spaces

The segmentation and classification performance of EfficientNetV2-S + CBAM + Focal Loss was systematically evaluated across four color spaces (RGB, HSV, LAB, and HED) under both 70:30 and 60:40 train–test splits. Metrics, including Accuracy, Precision, Recall, Specificity, F1-Score, AUC, and average running time per image, were computed (see Table 4).

Table 4. Segmentation Performance of EfficientNetV2-S + CBAM + Focal Loss Across Color Spaces.

Color Space	Split	Accuracy	Precision	Recall	Specificity	F1 Score	AUC	Running Time (s)/image
RGB	70:30	0.9611	0.9420	0.9136	0.9542	0.9519	0.9902	4.25
	60:40	0.9611	0.9623	0.9598	0.9777	0.9599	0.9875	4.32
HSV	70:30	0.9556	0.9408	0.9202	0.9482	0.9301	0.9825	4.58
	60:40	0.9589	0.9510	0.9389	0.9655	0.9448	0.9837	4.61
LAB	70:30	0.9511	0.9457	0.9308	0.9546	0.9381	0.9862	4.86
	60:40	0.9535	0.9458	0.9420	0.9697	0.9488	0.9871	4.90
HED	70:30	0.9562	0.9523	0.9402	0.9591	0.9462	0.9802	5.23
	60:40	0.9598	0.9514	0.9531	0.9541	0.9572	0.9815	5.31

Table 4 summarizes the segmentation performance of the proposed EfficientNetV2-S + CBAM + Focal Loss framework across four color spaces, evaluated with 70:30 and 60:40 train–test splits. Performance was measured using Accuracy, Precision, Recall, Specificity, F1-Score, AUC, and per-image running time. The results clearly demonstrate that the RGB color space provides the strongest overall performance, with an accuracy of 0.9611, F1-score of 0.9599, and an AUC of 0.9902 under the 70:30 split, while achieving the lowest average running time (4.25–4.32 s/image). This indicates that RGB, despite its sensitivity to illumination variations, remains the most computationally efficient and reliable representation for blood smear classification when combined with advanced deep learning frameworks. The HSV color space achieves competitive performance, with recall values consistently above 0.92, highlighting its robustness to illumination changes. However, its slightly lower F1-scores (0.9301–0.9448) compared to RGB suggest challenges in balancing precision and recall. The LAB color space shows strong generalization, particularly under the 60:40 split, with a recall of 0.9420 and a specificity of 0.9697, outperforming HSV in lesion boundary detection. Its slightly higher runtime (4.86–4.90 s/image) reflects the additional computational overhead introduced by the RGB–XYZ–LAB conversion. The HED color space, optimized for stain separation, achieves high recall (0.9402–0.9531) and specificity (~0.95), but its overall performance is slightly lower than RGB and LAB, with AUC values of 0.9802–0.9815. Importantly, HED incurs the highest computational cost (5.23–5.31 s/image), primarily due to the need for stain deconvolution and matrix inversion. Overall, the findings indicate that while RGB-based segmentation yields the best trade-off among accuracy, efficiency, and clinical reliability, alternative spaces, such as LAB and HED, provide complementary benefits by enhancing boundary detection and stain invariance. The 60:40 split consistently outperforms 70:30 across all color spaces in recall and F1-score, indicating improved generalization with larger test sets.

6. Conclusion

This study introduces an Advanced Lightweight Deep Learning (ALDL) framework for the multi-class classification of blood cancers from peripheral blood smear (PBS) images, targeting four clinically relevant stages: Benign, Pro-B, Pre-B, and Early Pre-B. The proposed approach leverages EfficientNetV2-S integrated with CBAM and optimized with Focal Loss, thereby achieving a robust balance among high predictive accuracy, strong generalization, and computational efficiency. A comprehensive preprocessing pipeline, comprising CLAHE for contrast enhancement, Reinhard stain normalization for chromatic consistency, and data augmentation for improved generalizability, ensured high-quality input data. Furthermore, lesion-level segmentation was performed using RGB-based intensity thresholding and watershed overlay, enhancing lesion visibility and facilitating morphological feature extraction. These steps significantly contributed to the reliability of

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

downstream classification. Experimental evaluations conducted on ALL-DB datasets validated the superiority of the proposed method. The model consistently achieved an average accuracy of 96.11%, an F1-score of 95.99%, and an AUC of 0.9875, outperforming benchmark architectures including MobileNetV3, ResNet50, DenseNet121, VGG16, and InceptionV3 across both 70:30 and 60:40 train–test splits. This consistent improvement underscores the framework's robustness and generalization capabilities. Additionally, a comparative analysis across multiple color spaces (RGB, HSV, LAB, and HED) revealed that RGB provides the most clinically reliable segmentation and classification outcomes, due to its computational simplicity and strong performance across evaluation metrics. While the HED color space enhanced lesion visualization through stain decomposition, it incurred a significantly higher computational cost, highlighting the trade-off between segmentation interpretability and efficiency. The findings demonstrate strong potential for real-world adoption as a computer-aided diagnostic (CAD) system for early detection of leukemia, thereby supporting hematopathologists in decision-making and reducing the likelihood of diagnostic errors.

7. Future Work

While the proposed ALDL framework demonstrates significant improvements in accuracy, robustness, and computational efficiency for multi-class blood cancer classification, several avenues remain open for future research. First, validation should be extended to larger, more diverse clinical datasets across multiple institutions to enhance generalizability and reduce potential dataset-specific biases. Second, integrating multimodal data sources, such as genetic, molecular, and clinical records, with peripheral blood smear images could enhance diagnostic precision and provide a more comprehensive understanding of disease progression. Third, incorporating semi-supervised and self-supervised learning techniques may reduce reliance on extensive annotated datasets, which remain a significant bottleneck in medical imaging. Additionally, optimizing the framework for real-time clinical deployment on resource-constrained hardware, such as mobile or edge devices, will be crucial for achieving broad adoption in low-resource healthcare settings. Finally, exploring explainable AI (XAI) approaches to improve model interpretability could foster greater trust and acceptance among clinicians by providing transparent insights into classification decisions.

Author Contributions: Conceptualization, K.W. methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision, project administration, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research project was financially supported by Mahasarakham Business School, Mahasarakham University, Thailand.

Informed Consent Statement: Patient informed consent was waived for this study because it exclusively utilized secondary data that had been fully anonymized prior to analysis. The research did not involve direct patient interaction, clinical interventions, or the collection of identifiable personal health information. Instead, the study confined its scope to the operational workflow of treatment planning. Accordingly, only non-clinical, activity-based information was analyzed, including the type of activity performed, execution dates, start and end times, resources used, and the roles of personnel involved in each process stage. All data were handled in accordance with ethical research standards, ensuring confidentiality, privacy protection, and compliance with institutional and academic research guidelines.

Acknowledgments: This research project was financially supported by Mahasarakham Business School, Mahasarakham University, Thailand. The authors would also like to thank the laboratory staff and medical experts who contributed to the preparation, staining, and annotation of the peripheral blood smear datasets used in this research. Their expertise and commitment were essential in ensuring the quality and reliability of the experimental data. In addition, the authors acknowledge the contributions of the open-source scientific community, whose tools and frameworks were indispensable in implementing the proposed methodology.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

Constructive comments from anonymous reviewers are also appreciated, as they helped to improve the quality and clarity of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Disclosure statement: No potential conflict of interest was reported by all authors.

References

1. Dulaimi, K.A.L.; Banks, J.; Nugyen, K.; Al-Sabaawi, A.; Reyes, T.I.; Chandran, V. Segmentation of white blood cell, nucleus and cytoplasm in digital haematology microscope images: A Review-challenges, current and future potential techniques, *IEEE Rev. Biomed. Eng.* **2021**, 14, 290-306, http://doi.org/10.1109/RBME.2020.3004639.
2. Joshi, U.; Khanal, S.; Bhetuwal, U.; Bhattarai, A.; Dhakal, P.; Bhatt, V. R. Impact of insurance on overall survival in acute lymphoblastic leukemia: A SEER database study, Clin. Lymphoma Myeloma Leukemia, **2022**, 22, (7) 477–484, http://doi.org/10.1016/j.clml.2022.01.001.
3. Frey, N.V. Approval of brexucabtagene autoleucel for adults with relapsed and refractory acute lymphocytic leukemia, *Blood* **2022**, 140 (1) 11–15, http://doi.org/10.1182/blood.2021014892.
4. Yazdan, S.A.; Ahmad, R.; Iqbal, N.; Rizwan, A.; Khan, A.N.; Kim, D.H. An efficient multi-scale convolutional neural network based multiclass brain MRI classification for SaMD, *Tomography* **2022**, 8 (4) 1905–1927, https://doi.org/10.3390/tomography8040161
5. Shahzad, T.; Iqbal, K.; Khan, M.A.; Iqbal, N. Role of zoning in facial expression using deep learning, *IEEE Access*, **2023**, 11, 16493–16508, http://doi.org/10.1109/ACCESS.2023.3243850.
6. Agustin, R.I.; Arif, A.; Sukorini, U. Classification of immature white blood cells in acute lymphoblastic leukemia l1 using neural networks particle swarm optimization, *Neural Comput. Appl.*, **2021**,33 (17) 10869-10880, http://doi.org/10.1007/s00521-021-06245-7.
7. Parab, M.A.; Mehendale, N.D. Red blood cell classification using image processing and CNN, *Social Netw. Comput. Sci.* **2021**, 2 (2) 1-10, https://doi.org/10.1007/s42979-021-00458-2.
8. Pansombut, T.; Wikaisuksakul, S.; Khongkraphan, K.; Phon-on, A. Convolutional neural networks for recognition of lymphoblast cell images, *Comput. Intell. Neurosci.* **2019**, 1-12, https://doi.org/10.1155/2019/7519603.
9. Nguyen, D.T.; Pham, T.D.; Baek, N.R.; Park, K.R. Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors, *Sensors* **2018**, 18 (3) 1-29, https://doi.org/10.3390/s18030699.
10. Lu, Y.; Qin, X.; Fan, H.; Lai, T.; Li, Z. WBC-Net: A white blood cell segmentation network based on UNetCC and Resnet, **Appl. Soft Comput. 2021**, 101, 107006, 1-11, https://doi.org/10.1016/j.asoc.2020.107006.
11. Roy R.M.; and Ameer, A.P.M., Segmentation of leukocyte by semantic segmentation model: A deep learning approach, *Biomed. Signal Process. Control* **2021**, 65 (102385) 1-13, https://doi.org/10.1016/j.bspc.2020.102385.
12. Abdurrazzaq, A.; Junoh, A.K.; Yahya, Z.; Mohd, I. New white blood cell detection technique by using singular value decomposition concept, *Multimedia Tools Appl.* **2021**, 80 (3) 4627-4638, https://doi.org/10.1016/j.bspc.2020.102385.
13. Khomairoh, N.; Sigit, R.; Harsono, T.; Hernaningsih, Y.; A. Anwar, Segmentation system of acute myeloid leukemia (AML) subtypes on microscopic blood smear image, in Proc. Int. Electron. Symp. (IES), Sep., Surabaya, Indonesia, 2020, pp. 565-570. https://doi.org/10.1109/IES50839.2020.9231651.
14. Hegde, R.B., Prasad, K., Hebbar, H.; Singh, B.M.K. Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: A study, *Australas. Phys. Eng. Sci. Med.* **2019**, 42, (2), 627-638, https://doi.org/10.1007/s13246-019-00742-9.
15. Saleem, S.; Amin, J.; Sharif, M.; Anjum, M. A.; Iqbal, M.; Wang, S.H. A deep network designed for segmentation and classification of leukemia using fusion of the transfer learning models, *Complex Intell. Syst*. **2022**, 8, 3105–3120, https://doi.org/10.1007/s40747-021-00473-z.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

16. Ramya, V.J.; Lakshmi, S. Acute myelogenous leukemia detection using optimal neural network based on fractional black-widow model, *Signal, Image Video Process*. **2022**, 16, (1) 229-238, https://doi.org/10.1007/s11760-021-01976-5.
17. Rad, A.E.; Rahim, M.S.M.; Kolivand, H.; Amin, I.B.M. Morphological region-based initial contour algorithm for level set methods in image segmentation, *Multimedia Tools Appl*. **2017**, 76, 2, 2185-2201, https://doi.org/10.1007/s11042-015-3196-y.
18. Puigdollers, D.L.; Traver, V.J.; Pla, F. Recognizing white blood cells with local image descriptors, *Expert Syst. Appl*. **2019**, 115, 695-708, https://doi.org/10.1016/j.eswa.2018.08.029.
19. Ruberto, C.D., Loddo, A.; Putzu, L. A multiple classifier learning by sampling system for white blood cells segmentation, in Computer Analysis of Images and Patterns. Cham, Switzerland: Springer, 2015, pp 415–425, https://doi.org/10.1007/978-3-319-23117-4_36.
20. Hussein, A.I.; Saleh M.A.; Aly, R.H.M. Bee Colony-Reptile Search Optimization Technique for Blood Cell Cancer Detection, 2025 17th International Conference on Computer and Automation Engineering (ICCAE), Perth, Australia, 2025, pp. 292-297, https://doi: 10.1109/ICCAE64891.2025.10980547.
21. Rai, H.; Yoo, J.; Razaque, A. Comparative analysis of machine learning and deep learning models for improved cancer detection: A comprehensive review of recent advancements in diagnostic techniques, *Expert Sys. with App*. **2024**, 225, (124838), https://doi.org/10.1016/j.eswa.2024.124838.
22. Swanson, K.; Wu, E.; Zhang, A.; Zou, J. From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment, **Cell** 2023, 186, (8), 1772-1791, https://doi.org/10.1016/j.cell.2023.01.035.
23. Aly, R.; Hussein, A.; Youssef, R. Accurate classification of cervical cancer based on multi-layer perceptron hunger games search optimization technique, 2024 21st Learning and Technology Conference (L&T), IEEE, Jeddah, Saudi Arabia, 2024, https://doi.org/10.1109/LT60077.2024.10468761.
24. Mahesh, R.; Santhakumar, D.; Balajee, A.; Shreenidhi, S.; Annand, R. Hybrid ant lion mutated ant colony optimizer technique with particle swarm optimization for leukemia prediction using microarray gene data, *IEEE Access* **2024**, 12, 10910-10919, https://doi.org/10.1109/ACCESS.2024.3351871.
25. Awais, M.; Abdal, N.; Akram, T.; Alasiry, A.; Masood, A. An efficient decision support system for leukemia identification utilizing nature-inspired deep feature optimization, *Frontiers in Oncology* **2024**, 14, (1328200), https://doi.org/10.3389/fonc.2024.1328200.
26. Nssibi, M.; Manita, G.; Chhabra, A.; Mirjalili, S.; Korbaa, O. Gene selection for high dimensional biological datasets using hybrid island binary artificial bee colony with chaos game optimization, *Artificial Intelligence Review* **2024**, 57, (3), 51, 1-74, https://doi.org/10.1007/s10462-023-10675-1.
27. Vogelbacher, M.; Strehmann, F.; Bellafkir, H.; Mühling, M.; Freisleben, B. Identifying and counting avian blood cells in whole slide images via deep learning, *Birds* **2024**, 5(1), 48-66, https://doi.org/10.3390/birds5010004.
28. Amin, J.; Sharif, M.; Anjum, M.A.; Siddiqa, A.; Kadry, S.; Nam, Y.; Raza, M. 3D semantic deep learning networks for leukemia detection, Tech Sci. Press, Henderson, NV, USA, Tech. Rep., 2021, https://doi.org/10.32604/cmc.2021.015249.
29. Amin, J., Anjum, M.A.; Sharif, M.; Kadry, S.; Nam, Y.; Wang, S. Convolutional bi-LSTM based human gait recognition using video sequences, *Comput., Mater. Continua* 2021, 68, (2), 2693–2709, https://doi.org/10.32604/cmc.2021.016871.
30. Saleem, S.; Amin, J.; Sharif, M.; Anjum, M.A.; Iqbal, M.; Wang, S.H. A deep network designed for segmentation and classification of leukemia using fusion of the transfer learning models, *Complex Intell. Syst*. **2022**, 8, (4), 3105–3120, https://doi.org/10.1007/s40747-021-00473-z.
31. Zhang, Q., Sun, L.; Chen, J.; Zhou, M.; Hu, M.; Wen, Y.; Li, Q. Speeded-up robust features-based image mosaic method for large-scale microscopic hyperspectral pathological imaging, *Meas. Sci. Technol.* **2021**, 32, (3), 1-17, https://doi.org/10.1088/1361-6501/abc868.
32. Noor, A.M.; Zakaria, Z.; Noor, A.M.; Norali, A.N. Classification of white blood cells based on surf feature, *Suranaree J. Sci. Technol.* **2021**, 28, (1), 1-6.
33. Faria, L.C.; Rodrigues, L.F.; Mari, J.F. Cell classification using handcrafted features and bag of visual words, in Proc. Anais do XIV Workshop de Visão Computacional, 2018 pp. 68-75.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

34. Gheisari, S.; Catchpoole, D.; Charlton, A.; Melegh, Z.; Gradhand, E.; Kennedy, P. Computer-aided classification of neuroblastoma histological images using scale invariant feature transform with feature encoding, *Diagnostics* **2018**, 8, (3) (56), 1-18, https://doi.org/10.3390/diagnostics8030056.
35. Abhishek, A.; Jha, R.K.; Sinha, R.; Jha, K. Automated classification of acute leukemia on a heterogeneous dataset using machine learning and deep learning techniques, *Biomed. Signal Process. Control* **2022**, 72, (103341), https://doi.org/10.1016/j.bspc.2021.103341.
36. Mishra, S.; Majhi, B.; Sa, P.K.; Sharma, L. Gray level co-occurrence matrix and random forest-based acute lymphoblastic leukemia detection, *Biomed. Signal Process. Control* **2017**, 33 272–280, https://doi.org/10.1016/j.bspc.2016.11.021.
37. Sunny, S.P.; Khan, A.I.; Rangarajan, M.; Hariharan, A.; Birur, P.; Shah, N.; Kuriakose, M.A.; Suresh, A. Oral epithelial cell segmentation from fluorescent multichannel cytology images using deep learning, *Comput. Methods Programs Biomed.* **2022**, 227, (107205), https://doi.org/10.1016/j.cmpb.2022.107205.
38. Deng, Z.; Jia, W.; He, Y.; Zhang, L.; Rose, K. Adaptive Upscaling Filter Selection with Wiener Filter Compression Techniques in Reference Picture Resampling, 2025 IEEE International Symposium on Circuits and Systems (ISCAS), London, United Kingdom, **2025**, pp. 1-5, https://doi: 10.1109/ISCAS56072.2025.11043367.
39. Abedy, H., Ahmed, F., Bhuiyan, M.N.Q.; Islam, M.; Ali, N.Y.; Shamsujjoha, M. Leukemia prediction from microscopic images of human blood cell using HOG feature descriptor and logistic regression, in Proc. 16th Int. Conf. ICT Knowl. Eng., Nov., Bangkok, Thailand, **2018**, pp. 1–6, https://doi.org/10.1109/ICTKE.2018.8612303.
40. Molina, A.; Alférez, S.; Boldu, L.; Acevedo, A.; Rodellar, J.; Merino, A. Sequential classification system for recognition of malaria infection using peripheral blood cell images, *J. Clin. Pathol*. **2020**, 73, (10), 665–670, https://doi.org/10.1136/jclinpath-2019-206419.
41. Wasson, V. An efficient content-based image retrieval based on speeded up robust features (SURF) with optimization technique, in Proc. 2nd IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT), May, Bangalore, India, **2017**, pp. 730–735, https://doi.org/10.1109/RTEICT.2017.8256693.
42. Li, Y.; Li, Q.; Liu, Y.; Xie, W. A spatial–spectral SIFT for hyperspectral image matching and classification, *Pattern Recognit. Lett*. **2019**, 127, 18–26, https://doi.org/10.1016/j.patrec.2018.08.032.
43. Li, S.; Wang, Z.; Zhu, Q. A research of ORB feature matching algorithm based on fusion descriptor, in Proc. IEEE 5th Inf. Technol. Mechatronics Eng. Conf. (ITOEC), Jun. **2020**, pp. 417–420, https://doi.org/10.1109/ITOEC49072.2020.9141770.
44. Saleem, S.; Amin, J.; Sharif, M.; Mallah, G.A.; Kadry, S.; Gandomi, A.H. Leukemia segmentation and classification: A comprehensive survey, *Comput. Biol. Med*. **2022**, 150, (106028), https://doi.org/10.1016/j.compbiomed.2022.106028.
45. Amin, J.; Sharif, M.; Raza, M.; Saba, T.; Rehman, A. Brain tumor classification: Feature fusion, in Proc. Int. Conf. Comput. Inf. Sci. (ICCIS), Sakaka, Saudi Arabia, **2019**, pp. 1–6, https://doi.org/10.1109/ICCISci.2019.8716449.
46. J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, Image matching from handcrafted to deep features: A survey, *Int. J. Comput. Vis.* **2020**, (129), 23–79, https://doi.org/10.1007/s11263-020-01359-2.
47. Naqi, S.M.; Sharif, M.; Yasmin, M.; Fernandes S.L. Lung nodule detection using polygon approximation and hybrid features from CT images, *Current Med. Imag. Rev*. **2017**, 14, (1), 108–117, https://doi.org/10.2174/1573405613666170306114320.
48. Shi, F.; Wang, J.; Shi, J.; Wu, Z.; Wang, Q.; Tang, Z.; He, K.; Shi, Y.; Shen, D. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19, *IEEE Rev. Biomed. Eng.* **2020**, 14, 4–15, https://doi.org/10.1109/RBME.2020.2987975.
49. Claro, M. L.; Veras, R.D.M.S.; Santana, A.M.; Vogado, L.H.S.; Junior, G.B.; Medeiros, F.N.S.D.; Tavares, J.M.R.S. Assessing the impact of data augmentation and a combination of CNNs on leukemia classification, *Inf. Sci.* **2022**, 609, pp. 1010–1029, https://doi.org/10.1016/j.ins.2022.07.059.
50. Kanumuri, C.; Madhavi, C.R. A survey: Brain tumor detection using MRI image with deep learning techniques, in Smart and Sustainable Approaches for Optimizing Performance of Wireless Networks: Real-time Applications, Hoboken, NJ, USA: Wiley, **2022**, pp. 125–138, https://doi.org/10.1002/9781119682554.ch6.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 February 2026 doi:10.20944/preprints202602.1601.v1

51. Malik, S.; Amin, J.; Sharif, M.; Yasmin, M.; Kadry, S.; Anjum, S. Fractured elbow classification using hand-crafted and deep feature fusion and selection based on whale optimization approach, *Mathematics* **2022** 10, (18), 1-13, https://doi.org/10.3390/math10183291.
52. Amin, J.; Sharif, M.; Yasmin, M.; Fernandes, S.L. A distinctive approach in brain tumor detection and classification using MRI, *Pattern Recognit. Lett*. **2020**, 139, 118–127, https://doi.org/10.1016/j.patrec.2017.10.036.
53. Amin, J.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Big data analysis for brain tumor detection: Deep convolutional neural networks, Future Gener. *Comput. Syst*. **2018**, 87, 290–297, https://doi.org/10.1016/j.future.2018.04.065.
54. Saba, T., Mohamed, A.S., Affendi, M.E., Amin, J.; Sharif, M. Brain tumor detection using fusion of hand crafted and deep learning features, *Cogn. Syst. Res.* **2020**, 59, 221–230, https://doi.org/10.1016/j.cogsys.2019.09.007.
55. Sharif, M.I.; Li, J.P.; Amin, J.; Sharif, A. An improved framework for brain tumor analysis using MRI based on YOLOv2 and convolutional neural network, *Complex Intell. Syst.* **2021**, 7, (4), 2023–2036, https://doi.org/10.1007/s40747-021-00310-3.
56. Anand, R.; Shanthi, T., Nithish, M.; Lakshman, S. Face recognition and classification using GoogleNET architecture, in Soft Computing for Problem Solving, Berlin, Germany: Springer, **2020**, pp. 261–269, https://doi.org/10.1007/978-981-15-0035-0_20.
57. Amin, J., Sharif, M., Gul, E., Nayak, R.S. 3D-semantic segmentation and classification of stomach infections using uncertainty aware deep neural networks, *Complex Intell. Syst.* **2022**, 8, 4, 3041–3057, https://doi.org/10.1007/s40747-021-00328-7.
58. Shahzad, M., Umar, A.I.; Khan, M.A.; Shirazi, S.H.; Khan, Z.; Yousaf, W. Robust method for semantic segmentation of whole-slide blood cell microscopic images, *Comput. Math. Methods Med.* **2020**, 1–13, https://doi.org/10.1155/2020/4015323.
59. Meenakshi, A.; Ruth, J.A.; Kanagavalli, V.; Uma, R. Automatic classification of white blood cells using deep features based convolutional neural network, *Multimedia Tools Appl.* **2022**, 81, 21, 30121–30142, https://doi.org/10.1007/s11042-022-12539-2.
60. Jagadev P.; Virani, H.G. Detection of leukemia and its types using image processing and machine learning, in Proc. Int. Conf. Trends Electron. Informat. (ICEI), May, Tirunelveli, India, 2017, pp. 522–526, https://doi.org/10.1109/ICOEI.2017.8300983.
61. Osman H.M.; Yaba, S.P. Automated segmentation of acute lymphocytic leukemia (ALL) subtypes by the combination of color space conversion and K-means cluster, *Zanco J. Pure Appl. Sci.* **2022**, 34, 3, 11–20, https://doi.org/10.21271/ZJPAS.34.3.2.
62. Leng, B.; Wang, C.; Leng, M.; Ge, M., Dong, W. Deep learning detection network for peripheral blood leukocytes based on improved detection transformer, *Biomed. Signal Process. Control*, **2023**, 82, 104518, https://doi.org/10.1016/j.bspc.2022.104518.
63. Raji, H.; Tayyab, M.; Sui, J.; Mahmoodi, S.R.; Javanmard, M. Biosensors and machine learning for enhanced detection, stratification, and classification of cells: A review, *Biomed. Microdevices* **2022**, 24, 3, 1–20, https://doi.org/10.1007/s10544-022-00627-x.
64. Billah, M.E.; Javed, F. Bayesian convolutional neural network-based models for diagnosis of blood cancer, *Appl. Artif. Intell.* **2022**, 36, 1, 2011688, https://doi.org/10.1080/08839514.2021.2011688.
65. Devi, T.G.; Patil, N.; Rai, S.; Philipose, C.S. Survey of leukemia cancer cell detection using image processing, in Proc. Int. Conf. Comput. Vis. Image Process. Cham, Switzerland: Springer, **2022**, pp. 468–488, https://doi.org/10.1007/978-3-031-11346-8_41.
66. Atteia, G.; Alhussan, A.; Samee, N. BO-ALLCNN: Bayesian-based optimized CNN for acute lymphoblastic leukemia detection in microscopic blood smear images, *Sensors* **2022**, 22, 15, 5520, 1-20, https://doi.org/10.3390/s22155520.
67. Saeed, A.; Shoukat, S.; Shehzad, K.; Ahmad, I.; Eshmawi, A.A.; Amin, A.H.; Tag-Eldin, E. A deep learning-based approach for the diagnosis of acute lymphoblastic leukemia, *Electronics* **2022**, 11, 19, 3168, https://doi.org/10.3390/electronics11193168.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

68. Hosseini, A.; Eshraghi, M.A.; Taami, T.; Sadeghsalehi, H.; Hoseinzadeh, Z.; Ghaderzadeh, M.; Rafiee, M. 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* **2023**, 39, 101244, 1-9, https://doi.org/10.1016/j.imu.2023.101244.
69. Togban, E.; Ziou, D. Improved image display by identifying the RGB family color space, *Displays* **2025**, 90, (103106) (2025), 1-6, https://doi.org/10.1016/j.displa.2025.103106.
70. Xue, W.; Liu, Y.; Zhuang, Y. A weight-sharing based RGB-T image semantic segmentation network with hierarchical feature enhancement and progressive feature fusion, *Neurocomputing* **2025**, 652, 131023, https://doi.org/10.1016/j.neucom.2025.131023.
71. Durom, E.; Yang, C.; Mozaffaripour, A.; Matheson, A.M.; Eddy, R.L.; Svenningsen, S.; Parraga, G. Quantification of 129Xe MRI Ventilation-defect-percent Using Binary-threshold, Gaussian Linear-Binning and K-means Methods: Differences in Asthma and COPD, *Academic Radiology* **2025**, 32, 8, 4893-4902, https://doi.org/10.1016/j.acra.2025.04.030.
72. Mohammadi, S.; Ghaderi, S.; Ghaderi, K.; Mohammadi, M.; Pourasl, M.H. Automated segmentation of meningioma from contrast-enhanced T1-weighted MRI images in a case series using a marker-controlled watershed segmentation and fuzzy C-means clustering machine learning algorithm, *Inter. Journal of Surgery Case Reports* **2023**, 111, 108818, https://doi.org/10.1016/j.ijscr.2023.108818.
73. Zeng, Z.; Liu, J.; Zheng, B.; Yi, S.; Yuan, X.; Liu, Q. A Pneumonia Recognition Model Based on Multiscale Attention Improved EfficientNetV2, *Computers, Materials and Continua* **2025**, 84, 1, 513-536, https://doi.org/10.32604/cmc.2025.063257.
74. Chen, L.; Yao, H.; Fu, J.; Ng, C.T. The classification and localization of crack using lightweight convolutional neural network with CBAM, *Engineering Structures* **2025**, 275, Part B, 115291, https://doi.org/10.1016/j.engstruct.2022.115291.
75. AbuKaraki, A.; Alrawashdeh, T.; Abusaleh, S.; Alksasbeh, M.Z.; Alqudah, B.; Alemerien, K.; Alshamaseen, H. Pulmonary edema and pleural effusion detection using efficientNet-V1-B4 architecture and AdamW optimizer from chest X-rays images, *Computers, Materials and Continua* **2024**, 80, 1, 1055-1073, https://doi.org/10.32604/cmc.2024.051420.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.