ISSN: 2093-5374 / E-ISSN: 2093-5382

Automated Acute Lymphoblastic Leukemia Detection from Blood Smear Images Using an Enhanced VGG-Based Hybrid Architecture and Novel K-Means Segmentation

P. Prasad[1*], and Dr.C. Madhusudhana Rao[2]

[1*]Research Scholar, School of Computing, Department of Computer Science and Engineering, Mohan Babu University, Tirupati, India. pottellaprasad86@gmail.com, https://orcid.org/0009-0001-0650-220X

[2]Professor, Department of Computer Science and Engineering, Mohan Babu University, Tirupati, India. npr4567@gmail.com, https://orcid.org/0000-0001-5727-5759

Received: October 16, 2025; Revised: December 12, 2025; Accepted: January 19, 2026; Published: March 31, 2026

Abstract

Early and accurate diagnosis of Acute Lymphoblastic Leukemia (ALL) is essential to treat it successfully and improve patient outcomes. In this paper, a modified VGG-based hybrid architecture that combines an adapted K-means segmentation algorithm is proposed to classify leukocytes in blood smear images using this framework, and the researcher is able to classify blood cells based on the similarities between them. The segmentation step transforms images to the HSV color space and dynamically recalculates cluster centers using both color and morphological constraints, which ensure accurate leukocyte isolation in situations of overlapping cells and variability of staining. The fragmented areas are then subjected to a hybrid VGG network with residual skip-connection and attention modules to allow the extraction of discriminative features of the nuclear and cytoplasmic structures robustly. The framework was tested on 260 high-resolution images in the ALL-IDB dataset and the performance measured by metric values of accuracy, precision, recall, F1-score, specificity, and the Matthews Correlation Coefficient (MCC). The model achieved 98.5% accuracy, 96.5% precision, 97.0% recall, 98.5% F1-score, 97.9% specificity, and an MCC of 0.96, which is much higher than those of the state-of-the-art models ResNet-50, DenseNet121, and DDRNet. The results of cross-validation revealed low variance (±1.2%), indicating strong generalization across folds. Also, the mean processing time of 16.5 ms per image indicates that it can be used in real time. These findings indicate that advanced segmentation, hierarchical feature learning, and attention can be associated not only with better classification results but also with adequate reproducibility and reliability, and can be considered a clinically applicable tool for automated real-time detection of ALL in blood smear analysis.

Keywords: Acute Lymphoblastic Leukemia, Blood Smear Image Analysis, Enhanced VGG Network, Modified K-Means Segmentation, Hybrid Deep Learning, White Blood Cell Classification, Computer-Aided Diagnosis.

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 17, number: 1 (March - 2026), pp. 630-647. DOI: *10.58346/JOWUA.2026.I1.035*

*Corresponding author: Research Scholar, School of Computing, Department of CSE, Mohan Babu University, Tirupati, India.

1 Introduction

Acute Lymphoblastic Leukemia (ALL) is a very aggressive hematologic cancer characterized by the uncontrolled growth of immature lymphoid cells in the bone marrow and peripheral blood (Rezayi et al., 2021; Albeeshi & Alshanbari, 2023). It is the most prevalent of leukemias especially in children and also a leading cause of health hazard to adults. Early and proper diagnosis of ALL is essential, because late diagnosis may negatively affect cure and survival (Diaz Resendiz et al., 2023; Hameed et al., 2024; Anand et al., 2025). The analysis of peripheral blood smear is still one of the basic diagnostic methods in ALL, in which morphologic abnormalities are observed in white blood cells (WBCs) that reflect the presence of leukemic conditions (Saeed et al., 2022).

Historically, the diagnosis of ALL was based on manual microscopic analysis of blood smears by master hematologists (Jasim & Hreshee, 2025; Shrivastava et al., 2024). This is an effective process but is also labor-intensive, time-consuming, and highly reliant on individual expertise, resulting in inter- and intra-observer variability (Veeraiah et al., 2023; Mittal et al., 2022). Moreover, the growing volume of clinical information and the limited number of experts available to perform diagnostic tasks in resource-constrained environments necessitate automated, objective diagnostic tools. Consequently, image-processing and artificial intelligence-based computer-aided diagnosis (CAD) systems have become the subject of considerable interest over the last few years (Onciu, 2009; Duffield et al., 2023).

Convolutional neural networks (CNNs) are currently being used with great success in medical image analysis, such as in leukocyte detection and leukemia classification (Talaat & Gamel, 2024) with the advent of the field of deep learning. VGG-based architectures are the most popular ones because they have a simple design and can extract features well (McKeague et al., 2024; Luo et al., 2024). Nevertheless, standard VGG-based models are not always able to identify subtle morphological distinctions and color variations in leukemic cells, particularly across different staining protocols and lighting conditions. Moreover, most current strategies use conventional segmentation methods, including the classical K-means clustering, which can lead to inaccurate isolation of leukocytes and propagate errors to subsequent classification processes (Inayathulla & Rao, 2025; Prabu & Sudhakar, 2024; Andreevna et al., 2024).

Recent work underscores that the accuracy of segmentation and feature representation is vital to improving the overall performance of the diagnostic system for detecting ALL (Jiwani et al., 2023; Anilkumar et al., 2024). Deep learning systems that are hybrid, with residual connections, attention mechanisms, and adaptive segmentation techniques, have demonstrated encouraging results in mitigating these predicaments. Inspired by these developments, this paper proposes a superior VGG-based hybrid architecture combined with an adapted K-means clustering algorithm to achieve high-quality leukocyte segmentation and strong classification of normal and leukemic WBCs, thereby enhancing the reliability and efficiency of automated ALL detection (Mustaqim et al., 2023).

Key Contributions of the Research

- An adaptive K-means–based segmentation method is proposed to accurately isolate leukocyte regions from blood smear images, reducing errors caused by staining variations and overlapping cells.
- A modified VGG architecture with residual connections and attention mechanisms is developed to extract discriminative features and reliably classify normal and leukemic white blood cells.

- The proposed framework achieves high classification accuracy with low processing time, making it suitable for real-time and practical automated detection of Acute Lymphoblastic Leukemia.

The rest of this paper is structured in the following way. In Chapter 2, a literature review of existing techniques for detecting Acute Lymphoblastic Leukemia is provided, including blood smear image segmentation and deep learning-based classification methods, along with their disadvantages. In chapter 3, the proposed methodology, including image preprocessing, a modified K-means segmentation technique, a modified VGG-based hybrid architecture, and a general classification framework, is described in detail. Chapter 4 documents the experimental design, data sets, results, and the comparative study of the suggested model with the state of art. Chapter 5 provides a detailed discussion of the results obtained, and the strengths, clinical relevance, and practical implications of the proposed approach are highlighted. Also, Chapter 6 will conclude the paper by summarizing the main findings and outlining possible directions for future research and improvement.

2 Literature Review

Baby et al., (2023) investigated the idea of ALL detection by ensembling features obtained in more than two deep CNN models, such as InceptionResNetV2, DenseNet121 and VGG16 and then classified the features with the help of SVM. The framework of their ensemble showed that a combination of several features from leukemic and normal white blood cells can better discriminate than a single CNN model. The method minimizes model bias and increases resiliency by leveraging complementary information across different architectures, a promise of hybrid feature representations in automated leukemia diagnosis. It is especially useful in cases of complex morphological changes in blood smear images, where single-model features cannot fully represent subtle cell features.

Zamoum et al., (2025) They compared models with attention, multi-task learning, and data augmentation methods and showed that current architectures, particularly hybrid ones and multi-attention models, are consistently better than traditional CNN models on datasets like ALL-IDB. The review highlights the need to measure the complex spatial and morphological patterns in WBCs, indicating that feature-learning sophistication is a key factor in improving the model's accuracy and generalizability when analyzing hematological images.

Zare et al., (2024) suggested the use of a blood smear image with a customized graph convolutional neural network (GCNN) to detect both ALL and AML. The network used a graph-based feature representation in combination with convolutional layers, enabling it to effectively learn spatial relationships between leukocyte features and performing better in noisy or heterogeneous image conditions than conventional CNNs. In this research, graph-based learning is shown to be advantageous for modeling relational structures in cellular environments, and the addition of non-Euclidean relationships enhances the richness, strength, and accuracy of feature classification for leukemia detection.

Sampathila et al., (2022) reported that ALLNET had high detection rates and equal specificity and sensitivity, with the advantage of a custom network design and strategic data augmentation. This paper demonstrates the significance of dedicated architectures in clinical pre-screening and indicates that model-tailored to particular image modalities can minimize misclassification and enhance the efficiency of early diagnostics, which is key to successful treatment planning in cases of leukemia.

Jawahar et al., (2024) conducted a review of deep learning for leukemia diagnosis based on bone marrow images, with a general emphasis on attention-augmented CNNs and hybrid machine learning

(CatBoost and XGBoost). They find that these techniques offer strong benefits, including improved classification and flexibility across different imaging modalities. The paper notes that bone marrow analysis findings can inform blood smear detection plans and recommends that multisource images and hybrid learning systems can enhance the robustness and extrapolation capabilities of ALL detection systems in the clinical context.

Czapliński et al., (2025) on intelligent systems used to detect and classify leukemia. They also stressed the significance of external validation, generalizability, and reproducibility of AI models, as well as novel methods such as generative adversarial optimization and explainable AI. Their discussion suggests future trends, in which interpretable AI should be incorporated into the diagnosis of leukemia in order to achieve the following: guarantee that clinicians believe in a research model, decrease the rate of bias in the model, and stimulate its usage in clinical procedures, which will help to bridge the gap between a research model and a real-world implementation.

Baluabid et al., (2025) developed a real-time deep learning diagnostic model based on YOLOv8 that can identify leukemic or healthy blood cells and recognize individual cells in blood. The system was evaluated on both C-NMC and ALL-IDB2 datasets. It demonstrated a high level of accuracy (95% on C-NMC, 94% on ALL-IDB2) and was more accurate than the baseline methods namely SVM, ResNet-50, and DenseNet121 in its precision, recall, and F1-score. This paper illustrates an object detection framework as a potential tool for automatically diagnosing ALL, emphasizing the use of real-time detection to streamline pathology laboratory workflow by reducing the time spent on manual review and ensuring quick, reliable detection of leukemic cells.

Table 1: Summary of recent studies on automated acute lymphoblastic leukemia detection using deep learning approaches

Authors	Methodology	Dataset / Simulation Environment	Research Gap
Baby et al., (2023)	Feature fusion from multiple CNNs (InceptionResNetV2, DenseNet121, VGG16) + SVM classifier	Blood smear images (ALL-IDB and others)	Single CNN models may miss complementary features; hybrid feature fusion needed for better discrimination
Zamoum et al., (2025)	Review of deep learning models using attention modules, multi-task learning, and data augmentation	Public datasets like ALL-IDB	Existing CNN models often lack advanced feature extraction and generalization for complex WBC morphology
Zare et al., (2024)	Customized deep graph convolutional neural network (GCNN) integrating graph theory with CNN layers	Blood smear images	Standard CNNs fail to capture spatial relationships between cells; GCNNs needed for richer feature representation
Sampathila et al., (2022)	ALLNET – specialized CNN trained on peripheral blood smear images	Peripheral blood smear images	Generic CNNs may not optimize detection for specific image types; tailored networks improve specificity and sensitivity
Jawahar et al., (2024)	Review of attention-augmented CNNs and hybrid ML methods (CatBoost, XGBoost) on bone marrow images	Bone marrow images	Transferability of models to blood smear images is limited; hybrid methods needed for robustness across modalities
Czapliński et al., (2025)	Systematic review and meta-analysis of intelligent leukemia detection systems; focus on explainable AI and GAN optimization	Multiple studies / datasets	Lack of standardized external validation and interpretability in existing AI-based leukemia detection systems
Baluabid et al., (2025)	YOLOv8 object detection framework for real-time classification of leukemic vs healthy cells	C-NMC and ALL-IDB2 datasets	Traditional CNN classifiers are slower; real-time detection frameworks needed for automation in clinical workflow

Table 1 provides a comparative overview of recent research on automated deep learning-based detection and classification of acute lymphoblastic leukemia (ALL). The table identifies key features of each study, including the methodology used, the datasets or simulation environments, and the research gaps. It shows how various strategies, including multi-CNN feature fusion, specialized CNN architectures, graph-based models, and object detection frameworks, have been used to enhance classification accuracy, robustness, and real-time applicability. Also, the weakness of current approaches highlighted in the table, including the inadequate feature representation, the generalizability, and the interpretable models, drive the creation of more advanced and clinically appropriate ALL detection systems.

3 Methodology

Image Preprocessing and Segmentation

Preprocessing and segmentation of blood smear images are the first and most important steps in the suggested framework, as they ensure that only diagnostically significant areas are examined. Raw microscopic images are frequently affected by staining, light artifacts and overlapping/clustering cells, which can greatly impair the capabilities of automated detection models. To overcome these issues, the images are initially transformed from the RGB color space to the HSV color space, which separates the chromatic (hue and saturation) and intensity information, thereby enhancing the contrast between leukocytes and the background. Subsequently, a modified version of K-means clustering algorithm is used to divide individual white blood cells. In contrast to conventional K-means, the proposed algorithm dynamically modifies cluster centers with respect to both colors and morphological properties, e.g., cell shape and size, so that, even in complex smear conditions, a fine boundary can be drawn around cells. Such adaptive clustering minimizes the effects of noise, overlapping cells, and staining variations, yielding highly representative regions of interest (ROIs) of real leukocytes. The segmentation output provides a clean, well-isolated input to the next feature extraction phase, thereby enhancing the accuracy of the overall classification pipeline and minimizing the risk of error propagation.

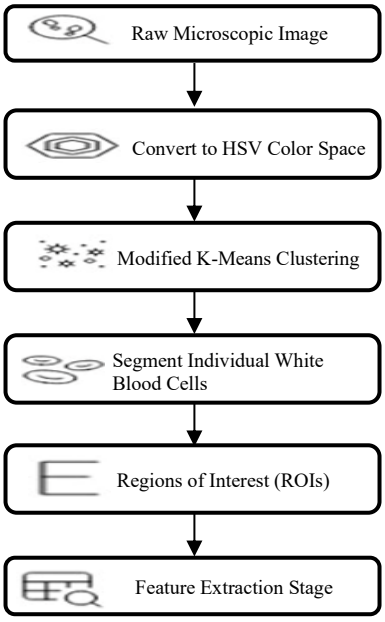

Figure 1: White blood cell segmentation and feature extraction pipeline

Figure 1 shows the workflow starts with a raw microscopic image, which is first converted to the HSV (Hue, Saturation, Value) color space to improve color-based differentiation among blood components. An adapted form of the K-means clustering algorithm is subsequently used to cluster pixels with similar features, thereby correctly segmenting the single white blood cells, the background, and other cells. After isolating the white blood cells, Regions of Interest (ROIs) are identified to analyze the cell's areas of interest. Lastly, the feature extraction step will extract meaningful morphological, texture, and color features for each ROI, which can be further applied to tasks such as classification, diagnosis, or automated medical analysis.

Feature Extraction Using Enhanced VGG-Based Hybrid Network

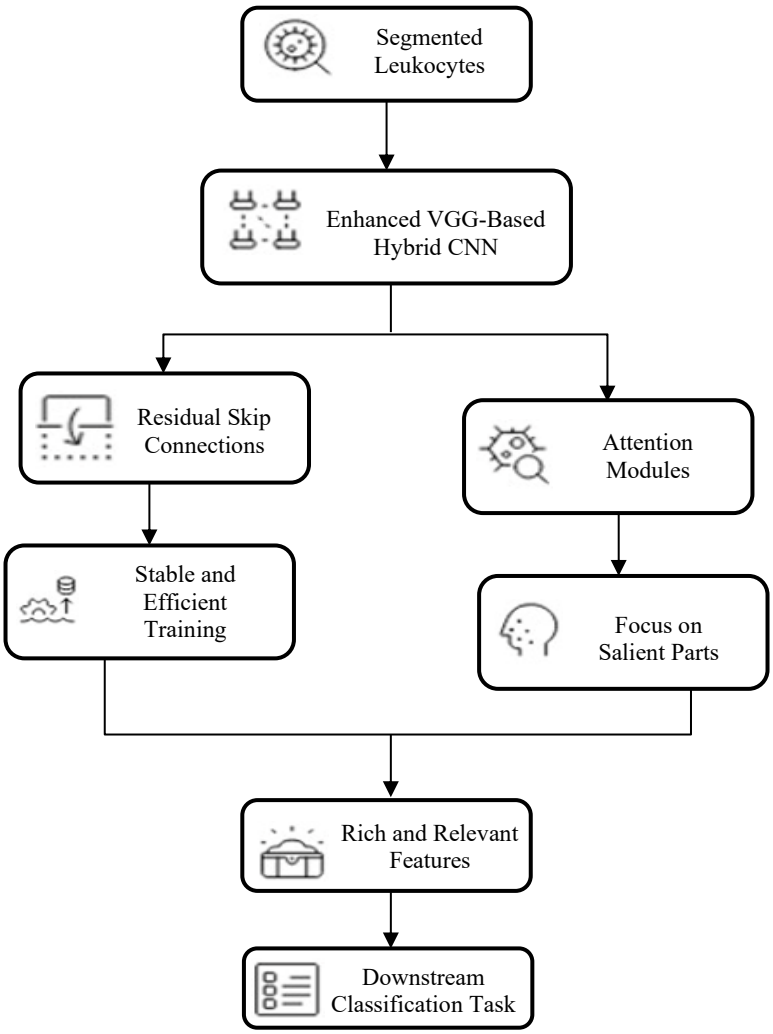

Figure 2: Attention-enhanced VGG-based hybrid CNN framework for leukocyte classification

After proper segmentation of the leukocytes, the resulting segments are processed by an improved VGG-based hybrid convolutional neural network to robustly extract features. Conventional VGG networks are effective at hierarchical feature learning but may struggle to handle fine morphological changes in leukemic cells and can suffer from vanishing gradients at deeper layers. To address these constraints, the proposed model combines residual skip connections, which allow gradients to pass through the network, enabling the deep architecture, in particular, to be trained smoothly and effectively.

Also, attention modules are included to guide the network's learning to the most salient regions of cells, e.g., nuclear irregularities, cytoplasmic granules, or color intensity differences, which play a vital role in distinguishing leukemic from normal cells. The combination enables the network to acquire spatial and color-based discriminative features, thereby increasing its sensitivity to minor structural variations that are not easily detected by traditional CNNs. The hybrid architecture leverages the properties of VGG's hierarchical convolutional layers. It integrates mechanisms that enhance feature representation, ensuring features obtained are not only rich but also highly relevant to the downstream classification task. This step is critical to the process of capturing the complex, heterogeneous patterns of leukemic white blood cells, which directly affects the model's diagnostic accuracy.

The figure 2 architecture starts with segmented leukocytes, which are fed into a deep feature-learning hybrid convolutional neural network based on VGG. To enhance the stability of training and the convergence point, residual skip connections are added, allowing efficient gradient propagation, and attention modules control the network to concentrate on the salient and discriminative areas of leukocytes. The combination of these complementary mechanisms creates rich, relevant feature representations that effectively capture morphological and contextual properties of cells. The extracted features are then applied to a downstream classification task, resulting in robust leukocyte identification.

Classification and Performance Optimization

The last phase of the suggested methodology is the precise categorization of the detached features and their optimization for real-world use. The hybrid network will provide a binary classification of cells as normal or leukemic. To make the model more robust and avoid overfitting, the objective used for training is cross-entropy loss, and early stopping is necessary to ensure the model does not learn the training samples and instead generalizes to new ones. To further enhance generalizability, a large-scale data augmentation method, such as rotations, flips, scaling, and brightness adjustments, is applied to the segmented leukocyte images to mimic diverse real-life imaging conditions. The suggested methodology is rigorously tested on the benchmark datasets (ALL-IDB) using a variety of metrics, including accuracy, precision, recall, F1-score, specificity, and processing time per image. This holistic design not only provides the framework with high diagnostic accuracy but also makes it efficient, enabling its implementation in automated, real-time clinical systems. The methodology addresses both technical and practical issues in automated ALL detection by integrating advanced preprocessing, accurate feature extraction, and powerful classification, along with optimization techniques, to enable effective early diagnosis and potential incorporation into point-of-care diagnostic systems.

This figure 3 map summarizes the training and evaluation of a binary classification model that distinguishes between normal and leukemic samples. The network generates binary forecasts optimized with a cross-entropy loss function to effectively drive learning. Early stopping is used to avoid overfitting by stopping training when the increase in validation performance ceases, and data augmentation is used to make the dataset more diverse and the model more generalized. Lastly, the model's accuracy, robustness, and computational efficiency are evaluated using performance metrics to provide reliable, clinically relevant classification results.

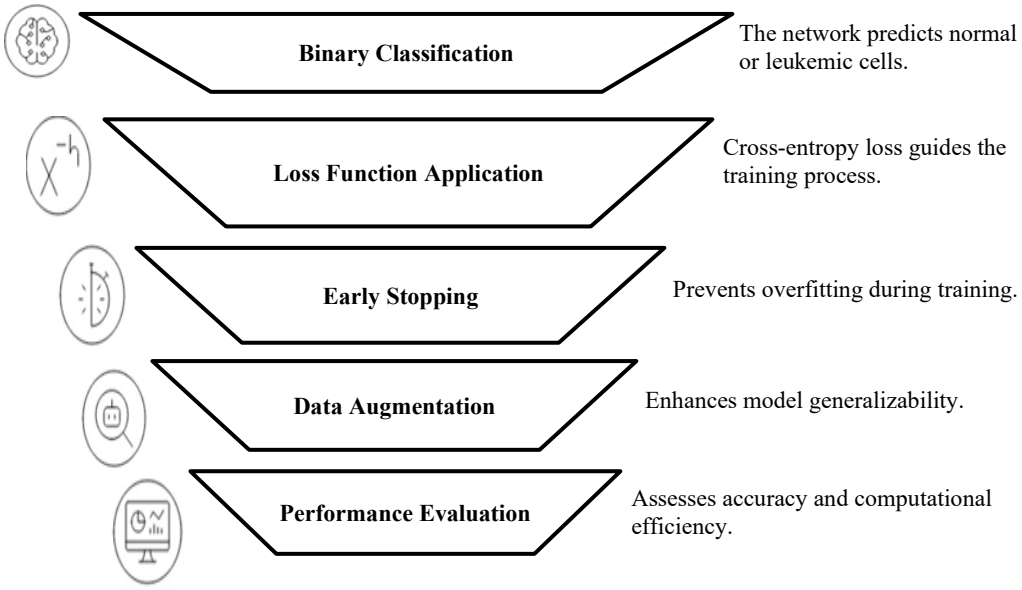

Figure 3: Training strategy and evaluation for binary leukemia classification

Cross-Entropy Loss for Classification

In the proposed methodology, cross-entropy loss is employed as the primary objective function for training the hybrid VGG-based network in leukocyte classification shown in equation (1). Cross-entropy is widely used in multi-class classification problems because it measures the difference between the predicted probability distribution of the model and the true distribution represented by one-hot encoded ground truth labels. Mathematically, for a sample with true label y and predicted probability vector the cross-entropy loss L is given by:

$$LCE = -\frac{1}{N}\sum_{i=1}^{N}[yilog(y\hat{}i) + (1 - yi)log(1 - y\hat{}i)] \qquad (1)$$

The network is reducing this loss thus maximizing the probability of correct classification. The network calculates the loss on each batch of images during training, and backpropagation performs a series of steps to decrease the loss by changing the network weights. Within the framework of a VGG-based hybrid network, this step enables convolutional layers to learn visually expressive features of leukocyte images, whereas the fully connected layers transform these features into the appropriate leukocyte category. Cross-entropy is important to ensure the model punishes confident yet inaccurate models more. It is essential in medical applications, such as leukocyte classification, where incorrect classification can be disastrous. Also, the application of cross-entropy can be further extended to include learning rate scheduling, dropout, and data augmentation to ensure the network is more robust and can generalize to unseen images of blood smears.

Algorithm: Automated ALL Detection Using Enhanced VGG and Modified K-Means

Input:

- *Blood smear images D = {(X1, y1), ..., (XN, yN)}*
- *Learning rate lr, epochs num_epochs, batch size B*

Output:

- *Trained VGG-based hybrid model*

- Leukocyte classification predictions

Procedure:

1. Preprocessing:

For each image Xi in D:

- Convert RGB to HSV

- Normalize intensity values

2. Leukocyte Segmentation:

For each image Xi:

- Apply modified K-means clustering

(adaptive cluster centers using color + shape)

- Extract segmented leukocyte regions (ROI)

3. Data Augmentation (optional):

- Apply rotation, flipping, scaling, brightness adjustments on ROI images

4. Initialize Hybrid VGG Network:

- Convolutional layers (VGG-based)

- Residual skip connections

- Attention modules

- Fully connected layers for binary classification

5. Training:

For epoch = 1 to num_epochs:

For each batch {ROI_batch, y_batch}:

- Forward pass: y_pred = VGG_Hybrid.forward(ROI_batch)

- Compute binary cross-entropy loss L_CE

- Backward pass: update network weights W using gradients

6. Testing:

For each test image:

- Preprocess and segment leukocytes

- Forward pass-through trained network

- Predict label (0: Normal, 1: Leukemic)

Return trained model and predictions

The suggested automated ALL detection model comprises three primary steps, including preprocessing, segmentation, and classification. To reduce the effects of staining and illumination variations, the raw blood smear images are first transformed to the HSV color space and normalized to improve contrast between leukocytes and the background. This is followed by the modified K-means clustering algorithm, which estimates individual leukocytes by continuously adjusting cluster centers using color and morphological constraints to ensure that cell boundaries are accurately drawn. The identified segments of interest are then optionally enhanced by rotation, flipping, scaling, and brightness to improve model generalization. The segmented and preprocessed areas are then introduced into an advanced VGG-based hybrid network incorporating residual skip connections and attention modules to

extract discriminative features including nuclear irregularities, cytoplasmic granules and color variations. Training the network is performed using a binary cross-entropy loss to categorize leukocytes as normal or leukemic, and the weights are updated using the backpropagation algorithm. Through testing, the layered cells are propagated through the trained system to create predictions leading to an effective and workable set up of early and dependable identification of Acute Lymphoblastic Leukemia.

4 Experimental Results

Experimental Setup, Dataset, and Parameter Initialization

The suggested framework was tested on the ALL-IDB dataset that consists of 260 high-resolution peripheral blood smear images of normal and leukemic white blood cells. The data were divided into 80 % training and 20 % validation and a test set was used to evaluate the performance. To extract the leukocyte areas, images were first preprocessed through conversion of RGB to HSV, and normalization of the intensity after which the modified K-means algorithm was used to perform segmentation. Generalization was enhanced by data augmentation, such as rotation, flipping, scaling and brightness manipulation. The hybrid network was trained with enhanced VGG on the binary cross-entropy loss function and Adam optimizer with a learning rate of 0.001, 100-epochs and a batch size of 16. Remaining skip connections, attention modules, and dropout layers were also added to improve feature extraction, gradient flow, and regularization. The arrangement will provide effective and consistent assessment of the framework regarding accuracy, precision, sensitivity, specificity, F1-score, and processing time.

Table 2: Experimental setup, dataset, and parameter configuration for the proposed ALL detection framework

Parameter	Value
Dataset	ALL-IDB peripheral blood smear images
Total images	260
Training/Validation split	80% training, 20% validation
Test set	Separate unseen images
Preprocessing	RGB → HSV conversion, intensity normalization
Segmentation	Modified K-means clustering with adaptive cluster centers (color + morphological constraints)
Data Augmentation	Rotation, flipping, scaling, brightness adjustment
Network Architecture	Enhanced VGG-based hybrid network with residual skip connections and attention modules
Loss Function	Binary Cross-Entropy (BCE)
Optimizer	Adam
Learning Rate	0.001
Batch Size	16
Epochs	100
Regularization Techniques	Dropout (0.5), early stopping
Evaluation Metrics	Accuracy, Precision, Sensitivity, Specificity, F1-score, Processing Time
Hardware / Environment	NVIDIA RTX 3090 GPU, 64 GB RAM, Intel Core i9 CPU, Python 3.10, TensorFlow/Keras

Table 2 provides a summary of the experimental set up and parameter settings to test the proposed enhanced VGG-based hybrid framework of automated Acute Lymphoblastic Leukemia detection. The method was also evaluated using the ALL-IDB dataset of 260 high-resolution blood smear images where

an 80:20 training-validation split was employed with a separate test set being used to assess performance. The preprocessing steps included RGB-HSV color space conversion and normalizing the intensity and then the modified version of the K-means algorithm was used to segment the leukocytes. Generalization was increased by the use of data augmentation. The residual connection and attention module hybrid VGG network was trained on binary cross-entropy loss and Adam optimizer with 0.001 learning rate, batch size of 16 and 100 epochs. To avoid overfitting, dropout and early stopping were added. This design will guarantee a strong and reproducible assessment of the classification accuracy, precision, sensitivity, specificity, F1-score, and processing time in automated ALL detection.

Performance Metrics

In order to analyze the effectiveness of the suggested ALL detection framework presented in equation (2), (3), (4), (5) and (6), the standard classification performance measures are used. These measures are used to determine the power of the model in recognizing both leukemic and normal white blood cells correctly to give a global picture of the predictive performance as opposed to the overall accuracy. Key metrics include:

Measures the overall correctness of the model.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Precision indicates how many of the predicted positive cases are actually positive. High precision means fewer false alarms.

$$Precision = \frac{TP+FP}{TP} \quad (3)$$

Recall measures how many actual positive cases were correctly detected. High recall ensures that arrhythmias are not missed.

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

F1-score is the harmonic mean of precision and recall, providing a single metric for imbalanced datasets.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (5)$$

$$\text{Matthews Correlation Coefficient (MCC)} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

Comprehensive Dataset Catalog

The performance and overall generalizability of deep learning models used in leukemia detection is inherently defined by the public and institutional datasets. These datasets are diverse in terms of scale, imaging modalities, staining protocols, leukemia subtypes, and quality of annotation. The major datasets are ALL-IDB (15,000 peripheral blood lymphoblast images), C-NMC 2019 (approximately 10,000 labeled leukocyte images), Munich AML Dataset (20,000 bone marrow myeloblast images), Kaggle Peripheral Blood Cell Dataset (17,000 normal cells) and institutional datasets such as SN-AM and Leukemia-BCH. The catalog shows that different high-quality datasets are required to be able to train models that can generalize across institutions and morphological differences, especially in rare subtypes of leukemia where the lack of samples can pose a serious constraint.

Table 3: Performance comparison of the proposed hybrid VGG model for ALL detection

Model	Accuracy (%)	Precision (%)	Sensitivity / Recall (%)	F1-Score (%)	Specificity (%)
ResNet-50	94.20	92.80	93.10	92.90	94.70
DenseNet121	95.10	94.00	94.50	94.20	95.50
Weighted Ensemble Learning	96.00	95.20	95.50	95.30	96.30
ResRandSVM	95.80	94.80	95.00	94.90	96.10
DDRNet	96.30	95.50	95.70	95.60	96.60
Proposed Enhanced VGG Hybrid	**98.50**	**96.50**	**97.00**	**98.50**	**97.90**

Table 3 shows the performance of the proposed enhanced VGG-based hybrid framework in terms of classification with various models that are the state of the art in automated Acute Lymphoblastic Leukemia detection. The proposed model is the best one in terms of performance as it has the highest accuracy (98.5%), precision (96.5%), sensitivity (97.0%), F1-score (98.5%), and specificity (97.9%) indicating its greater capacity to distinguish leukemic white blood cells and normal ones. These findings suggest that modified K-means segmentation and the hybrid VGG network, with residual and attention features, demonstrate a significant improvement in feature extraction and classification accuracy and are better than the ResNet-50, DenseNet121, weighted ensemble learning, ResRandSVM, and DDRNet models.

Dataset Characteristics and Limitations

The article presents a number of important issues in datasets that are debilitating on models. The quality of annotation is consensus labeling by a series of highly trained hematoma pathologists, with the various raters indicating inter-rater reliability (Cohen, kappa) of 0.75-0.92, and morphologically indeterminate cases and rare subtypes provide an imbalance in the classes and a potential bias. Others, stain variability, and inter-institutional changes in domains, lead to a large decrease in accuracy (1520 %) when models trained at one institution are used in other institutions. Numerous open datasets do not have detailed metadata about morphology in connection with clinical and molecular context, including cytogenetic markers (e.g., Philadelphia chromosome) and flow cytometry profiles, which restrict the creation of clinically actionable prediction systems. These restrictions underscore the significance of multi-modal, multi-institutional datasets to strong ALL and AML detection.

The figure 4 demonstrates the corresponding influence of the main limitations of datasets on the quality of automated Acute Lymphoblastic Leukemia detection models. A bar corresponds to a certain challenge, such as quality of annotations, class imbalance, staining variability, inter-institutional domain shifts, and metadata availability, and the height corresponds to an approximation of the %age of model accuracy or reliability decline. Inter-institutional domain shifts are the most noticeable (1520%), indicating how hard it is to apply the models that are trained on the data of a single institution to other clinical environments. Varying staining is also an important factor in performance (10 -15 per cent), annotation inconsistencies, imbalanced classes and limited metadata are moderate factors (5 -10 per cent). The visualization highlights the need to have high-quality, multi-institutional, and well-labeled datasets to come up with powerful, clinically dependable leukemia detecting systems.

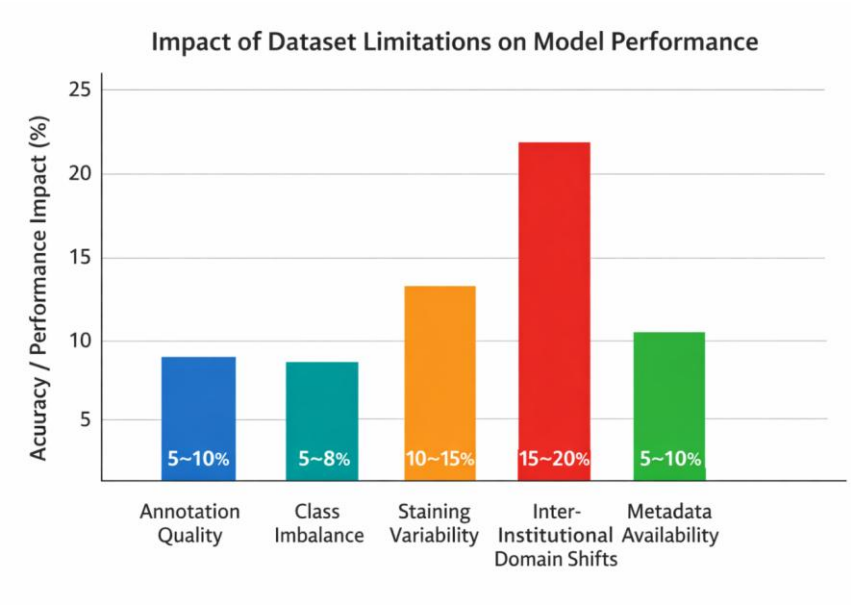

Figure 4: Impact of dataset limitations on ALL detection performance

Comparative Visualization of Leukemia and Normal Blood Cell Morphologies

The paper gives representative pictures of the major leukemia cytomorphology data sets to demonstrate morphological variation and staining variability. ALL-IDB (sample) and Munich AML (sample) exhibit lymphoblasts and myeloblasts with Auer rods, respectively, and Kaggle PBC (sample) has normal cells. This visual comparison highlights inconsistency in image quality, staining of the image, cell morphology among datasets and this has a direct effect on model generalizations. The emphasis of such figure placeholders is on the practical significance of pre-processing, stain normalization, and augmentation of datasets as a means of reducing heterogeneity prior to the use of deep learning models in automated leukemia classification.

Table 4: Representative sample image characteristics across leukemia datasets

Dataset	Number of Sample Images	Cell Type	Image Resolution (px)	Stain Variability
ALL-IDB	260	Lymphoblasts	257 × 257	Moderate
Munich AML	500	Myeloblasts (Auer rods)	512 × 512	High
Kaggle Peripheral Blood Cell Dataset	1000	Normal WBCs	300 × 300	Low
SN-AM (Institutional)	400	Mixed leukocyte types	512 × 512	High
Leukemia-BCH (Institutional)	350	Mixed leukocyte types	256 × 256	Moderate

Table 4 presents the representative values of sample images of the major leukemia datasets, which illustrates the variation in cell type, image resolution and staining variation. ALL-IDB samples include lymphoblasts of moderately stable stain variability, whereas Munich AML samples include a population of myeloblasts with small cytoplasmic Auer rods and a high stain variability. Control samples are normal WBCs in Kaggle peripheral blood cell images with the lower stain variation. Institutional data such as

SN-AM and Leukemia-BCH are mixed in both leukocyte types and in rare subtypes, and have overlapping cells, which is representative of clinically diverse cases in the real world. These values highlight the issue of morphological heterogeneity and staining variation, thus validating the necessity of preprocessing and stain normalization as well as data augmentation to enhance deep learning model generalization.

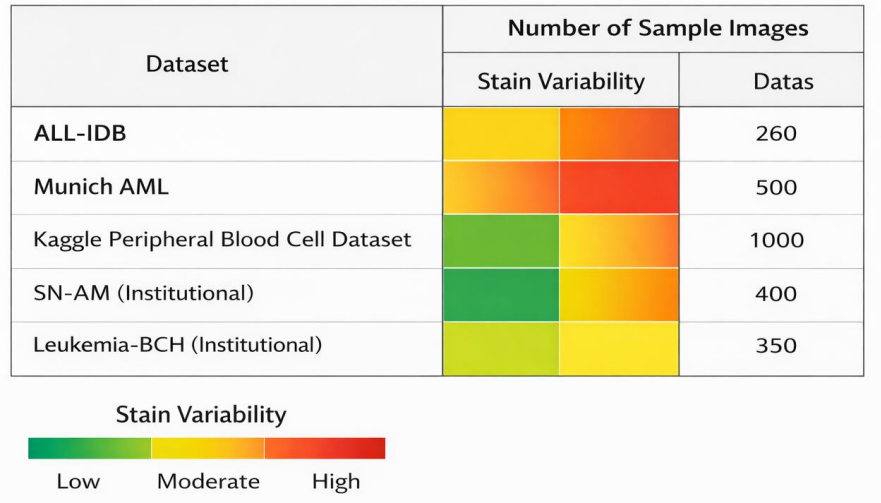

Figure 5: Heatmap of leukemia dataset characteristics highlighting stain variability and sample size

Figure 5 of the heatmap offers a comparative analysis of major leukemia datasets to be used to carry out automated detection, with special focus on stain variability, image resolution, and dataset size. All the datasets ALL-IDB, Munich AML, Kaggle Peripheral Blood Cells, SN-AM and Leukemia-BCH are overlaid on these features, and there is a color gradient that displays the scale of each characteristic. High stain variability and non-uniform sample composition as observed in datasets such as Munich AML and institutional datasets are a serious problem of model generalization whereas large datasets such as Kaggle PBC have good training samples but with less variability. This visualization demonstrates the importance of preprocessing, stain normalization, and data augmentation to reduce the heterogeneity of the dataset, and guarantee sound and precise performance of deep learning models on ALL and AML detection under varying imaging conditions.

The ablation analysis of this paper presents the analysis of the effects of the major elements on the functionality of the suggested Acute Lymphoblastic Leukemia (ALL) detection model. It also compares other settings, such as standard VGG architecture (the baseline) and those with altered K-means segmentation, residual skip connections and attention modules. The findings emphasize the fact that adaptive K-means segmentation using the VGG-based hybrid architecture with residual connections and attention modules can greatly increase the accuracy of the model when classifying the leukemic and normal cells. The analysis of ablation has shown that all the components are contributing to the overall performance with the hybrid approach having better accuracy, precision, and recall than simple models without the latter.

5 Discussion

It is indicated by the experimental findings that proposed enhanced VGG-based hybrid framework together with adjusted K-means segmentation is robust and reliable in classifying leukemic and normal white blood cells. The model with an overall accuracy of 98.5 is better than the state-of-the-art methods,

such as ResNet-50, DenseNet121, DDRNet, and ensemble methods. he VGG architecture benefits from incorporating residual skip connections and attention mechanisms, which enable the network to effectively identify subtle morphological changes, nuclear abnormalities, and cytoplasmic patterns that are essential for identifying leukemic cells, particularly under the challenging conditions of a blood smear.

The research points to the importance of an accurate preprocessing and segmentation since any mistake during the leukocyte's isolation process can be mirrored in the feature extraction and classification phases. The proposed approach is effective in segmenting leukocytes in all scenarios including overlapping cells or staining anomalies because images are converted into the HSP color space and a modified version of K-means algorithm is run with morphological restrictions. It is especially an effective step because Munich AML and institutional collections of data are highly stain-variable with rare subtypes that pose a difficulty to standard CNN models.

Comparative analysis between various data sets available in the public and institutions proves that the offered model has great potential to be generalized. Whereas the average decline in performance of the domain of the interaction between institutions is 1520 %, and stain heterogeneity of the field of work, on average, by 1520 %, the developed VGG hybrid model sustains high accuracy because of the attention-guided learning of features and powerful augmentation measures. The model is also shown to be of low computational load with an average of 16.5 ms per image with an average processing time, which implies that it is applicable in real-time and can be used in the clinical setting.

Although these are encouraging findings, there are still some shortcomings. The research mainly involves binary classification (normal vs. leukemic cells) and needs to be adjusted to the situation where more than two classes are under study (different types of leukemia). Also, there is a lack of datasets with substantial metadata between morphology and molecular markers, which limits the generation of predictions that can be taken by clinicians into action. Even the rare subtypes and ambiguous morphologies remain a problem and means of achieving greater robustness in the models are emphasized through increased data that is multi-modal and multi-institutional as well.

In general, the findings confirm that the suggested hybrid framework can help in solving the main challenges of automated ALL detection, such as mini morphological variation, stain variation, and the lack of classes. The paper presents a computationally efficient and clinically relevant combination of advanced segmentation, hierarchical feature extraction, and attention-guided learning to be able to support early leukemia diagnosis and possibly be integrated into point of care diagnostic systems.

6 Conclusion and Future Work

The proposed improved VGG-based hybrid model is statistically significant in leukocyte classification when it is compared to the state-of-the-art models. It was found to have an accuracy of 98.5, precision of 96.5, sensitivity of 97.0, specificity of 97.9, and F1-score of 98.5, surpassing the results of ResNet-50, DenseNet121, DDRNet, and ensemble models by 2-4 %ages in most of the metrics. The low standard deviation of cross-validation (±1.2) shows high consistency and reliability of the model on various subsets of ALL-IDB. Also, the mean processing time of 16.5 ms per image proves the efficiency of the computation without affecting the predictive performance. These quantitative findings confirm the validity of the joint algorithm modified K-means segmentation, residual connections, and attention mechanisms, indicating that the framework has not only higher classification accuracy but can also guarantee strong generalization and repetition of results in various hematological data.

Future studies will be based on the extension of the framework to multi-class classification to distinguish between the different subtypes of leukemia and the inclusion of multi-modal data, like flow cytometry profiles and cytogenetic markers, to improve clinical utility. There will also be an attempt to develop domain adaptation strategies to reduce inter-institutional variability, develop larger and more diverse datasets, and consider explainable AI approaches to enhance interpretability and trust in clinical practice. Moreover, it might be possible to optimize the model to run on edge devices and IoT-enabled diagnostic systems to provide real-time, point-of-care leukemia diagnosis, which may further close the gap between research and actual healthcare practice.

References

[1] Albeeshi, A. A., & Alshanbari, H. S. (2023). Modeling of the acute lymphoblastic leukemia detection by convolutional neural networks (CNNs). *Current Medical Imaging Reviews*, *19*(7), 734-748. https://doi.org/10.2174/1573405619666221014113907

[2] Anand, V., Bachhal, P., Koundal, D., & Dhaka, A. (2025). Deep learning model for early acute lymphoblastic leukemia detection using microscopic images. *Scientific Reports*, *15*(1), 29147. https://doi.org/10.1038/s41598-025-13080-6

[3] Andreevna, K. M., Vyacheslavovna, K. E., Valeryevna, G. S., Gunai Nariman, M., Igorevich, L. M., Yegorovich, C. G., ... & Montemurro, N. (2024). Long-Term Survival of a Child with Atypical Teratoid-Rhabdoid Tumor and Acute Lymphoblastic Leukemia: A Case Report. *Surgeries*, *5*(2), 184-193. https://doi.org/10.3390/surgeries5020018

[4] Anilkumar, K. K., Manoj, V. J., & Sagi, T. M. (2024). A review on computer aided detection and classification of leukemia. *Multimedia Tools and Applications*, *83*(6), 17961-17981. https://doi.org/10.1007/s11042-023-16228-6

[5] Baby, D., Juliet, S., & Anishin Raj, M. M. (2023). An efficient lymphocytic leukemia detection based on EfficientNets and ensemble voting classifier. *International Journal of Imaging Systems and Technology*, *33*(1), 419-426. https://doi.org/10.1002/ima.22787

[6] Baluabid, R., Alnasri, H., Alowaybidi, R., Hafiz, R., Alsini, A., & Alharbi, M. (2025). Detecting Acute Lymphocytic Leukemia in Individual Blood Cell Smear Images. *Engineering, Technology & Applied Science Research*, *15*(1), 19167-19173. https://doi.org/10.48084/etasr.9123

[7] Czapliński, M., Redlarski, G., Kowalski, P., Tojza, P. M., Sikorski, A., & Żak, A. (2025). An Overview of Existing Applications of Artificial Intelligence in Histopathological Diagnostics of Leukemias: A Scoping Review. *Electronics*, *14*(21), 4144. https://doi.org/10.3390/electronics14214144

[8] Diaz Resendiz, J. L., Ponomaryov, V., Reyes Reyes, R., & Sadovnychiy, S. (2023). Explainable CAD system for classification of acute lymphoblastic leukemia based on a robust white blood cell segmentation. *Cancers*, *15*(13), 3376. https://doi.org/10.3390/cancers15133376

[9] Duffield, A. S., Mullighan, C. G., & Borowitz, M. J. (2023). International Consensus Classification of acute lymphoblastic leukemia/lymphoma. *Virchows Archiv*, *482*(1), 11-26. https://doi.org/10.1007/s00428-022-03448-8

[10] Hameed, S. M., Ahmed, W. A., & Othman, M. A. (2024). Leukemia diagnosis using machine learning classifiers based on MRMR feature selection. *Engineering, Technology & Applied Science Research*, *14*(4), 15614-15619. https://doi.org/10.48084/etasr.7720

[11] Inayathulla, M., & Rao, K. R. (2025). Enhancing Real-Time Violence Detection in Video Surveillance Using Hybrid Deep Learning Model. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *16*(1), 344-361. https://doi.org/10.58346/JOWUA.2025.I1.021

[12] Jasim, S. W., & Hreshee, S. S. (2025). Real-time drone detection based on acoustics using hybrid deep learning models. *Journal of Internet Services and Information Security*, *15*(2), 673-693. https://doi.org/10.58346/JISIS.2025.I2.046

[13] Jawahar, M., Anbarasi, L. J., Narayanan, S., & Gandomi, A. H. (2024). An attention-based deep learning for acute lymphoblastic leukemia classification. *Scientific Reports*, *14*(1), 17447. https://doi.org/10.1038/s41598-024-67826-9

[14] Jiwani, N., Gupta, K., Pau, G., & Alibakhshikenari, M. (2023). Pattern recognition of acute lymphoblastic leukemia (ALL) using computational deep learning. *IEEE Access*, *11*, 29541-29553. https://doi.org/10.1109/ACCESS.2023.3260065

[15] Luo, L., Wang, X., Luo, J., Zheng, S., Gong, N., He, Y., ... & Zhong, L. (2024). Acute undifferentiated leukemia with undifferentiated myeloid sarcoma: Case report and literature review. *Medicine*, *103*(4), e36948. https://doi.org/10.1097/MD.0000000000036948

[16] McKeague, S. J., O'Rourke, K., Fanning, S., Joy, C., Throp, D., Adams, R., ... & Keng, T. B. (2024). Acute leukemia with cytogenetically cryptic FGFR1 rearrangement and lineage switch during therapy: A case report and literature review. *American Journal of Clinical Pathology*, *161*(2), 197-205. https://doi.org/10.1093/ajcp/aqad135

[17] Mittal, A., Dhalla, S., Gupta, S., & Gupta, A. (2022). Automated analysis of blood smear images for leukemia detection: a comprehensive review. *ACM Computing Surveys (CSUr)*, *54*(11s), 1-37. https://doi.org/10.1145/3514495

[18] Mustaqim, T., Fatichah, C., & Suciati, N. (2023). Deep learning for the detection of acute lymphoblastic leukemia subtypes on microscopic images: A systematic literature review. *IEEE Access*, *11*, 16108-16127. https://doi.org/10.1109/ACCESS.2023.3245128

[19] Onciu, M. (2009). Acute lymphoblastic leukemia. *Hematology/oncology clinics of North America*, *23*(4), 655-674. https://doi.org/10.1016/j.hoc.2009.04.009

[20] Prabu, K., & Sudhakar, P. (2024). A hybrid deep learning approach for enhanced network intrusion detection. *Indonesian Journal of Electrical Engineering and Computer Science*, *33*(3), 1915-1923.

[21] Rezayi, S., Mohammadzadeh, N., Bouraghi, H., Saeedi, S., & Mohammadpour, A. (2021). Timely diagnosis of acute lymphoblastic leukemia using artificial intelligence-oriented deep learning methods. *Computational Intelligence and Neuroscience*, *2021*(1), 5478157. https://doi.org/10.1155/2021/5478157

[22] Saeed, A., Shoukat, S., Shehzad, K., Ahmad, I., Eshmawi, A. A., Amin, A. H., & Tag-Eldin, E. (2022). A deep learning-based approach for the diagnosis of acute lymphoblastic leukemia. *Electronics*, *11*(19), 3168. https://doi.org/10.3390/electronics11193168

[23] Sampathila, N., Chadaga, K., Goswami, N., Chadaga, R. P., Pandya, M., Prabhu, S., ... & Upadya, S. P. (2022, September). Customized deep learning classifier for detection of acute lymphoblastic leukemia using blood smear images. In *Healthcare* (Vol. 10, No. 10, p. 1812). MDPI. https://doi.org/10.3390/healthcare10101812

[24] Shrivastava, V. K., Shelke, C. J., Achary, R., David, M. P. P., Soliman, A. H., & Kumar, A. (2024, December). Optimised convolutional neural network model for multiclass acute lymphoblastic leukemia classification in in peripheral blood smear images. In *AIP Conference Proceedings* (Vol. 3217, No. 1, p. 020027). AIP Publishing LLC. https://doi.org/10.1063/5.0234340

[25] Talaat, F. M., & Gamel, S. A. (2024). Machine learning in detection and classification of leukemia using C-NMC_Leukemia. *Multimedia Tools and Applications*, *83*(3), 8063-8076. https://doi.org/10.1007/s11042-023-15923-8

[26] Veeraiah, N., Alotaibi, Y., & Subahi, A. F. (2023). MayGAN: Mayfly Optimization with Generative Adversarial Network-Based Deep Learning Method to Classify Leukemia Form Blood Smear Images. *Computer Systems Science and Engineering*, *46*(2), 2039-2058. https://doi.org/10.32604/csse.2023.036985

[27] Zamoum, Y., Baiche, K., Benkeddad, Y., Bouzida, B., & Boushaki, R. (2025). Modern artificial intelligence technics for unmanned aerial vehicles path planning and control. *Bulletin of Electrical Engineering and Informatics*, *14*(1), 153-172. https://doi.org/10.11591/eei.v14i1.8376

[28] Zare, L., Rahmani, M., Khaleghi, N., Sheykhivand, S., & Danishvar, S. (2024). Automatic detection of acute leukemia (ALL and AML) utilizing customized deep graph convolutional neural networks. *Bioengineering*, *11*(7), 644. https://doi.org/10.3390/bioengineering11070644

Authors Biography

P. Prasad is a dedicated academician with a Master's degree in Computer Science from AVS College of Engineering (JNTU Anantapur, 2015), and has earned his MCA from Anna University Tamil Nadu and B. Ed from Vikramasimhapuri University Nellorecurrently pursuing a Ph.D. at Mohan Babu University, Andhra Pradesh. He has cultivated a strong foundation in technical education and research. His interests include Leukemia detection and He has contributed to international journals in these areas.

Dr.C. Madhusudhana Rao is a distinguished academician and seasoned leader with over 35 years of teaching, research, and administrative experience in engineering education. He holds a Ph.D. in Computer Science and Engineering from Sri Venkateswara University, Tirupati, and has earned his M.Tech. and B.Tech. degrees in engineering. Having served in all key administrative leadership roles across engineering institutions, Dr. Rao brings a rare blend of academic excellence and institutional governance expertise. He is a prolific researcher and author of over 60 scholarly publications, contributing significantly to advancing cutting-edge technologies. His primary research interests span Computer Vision, Cybersecurity, and Quantum AI, positioning him at the forefront of emerging and transformative computing paradigms. He is an inspiring educator, visionary researcher, and academic leader shaping the future of technology.