© DEC 2025 | IRE Journals | Volume 9 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I6-1712680

Hybrid Ensemble Deep Learning Framework for Blood Cancer Identification

WASEEM SHAREEF K S[1], SYED SHARFUDDIN SHUAIB[2], SYED UMAR[3], SYED MAAZ ATHAR[4], SYEDA MUZAMMIL[5] (ASSISTANT PROFESSOR)
[1, 2, 3, 4]*B.E Student, Department of Computer Science and Engineering, Ghousia College of Engineering, Ramanagara, Karnataka, India*
[5]*Assistant Professor, Department of Computer Science and Engineering, Ghousia College of Engineering, Ramanagara, Karnataka, India*

Abstract- Leukemia is a life-threatening malignancy that originates in the blood-forming tissues and rapidly affects the production and morphology of white blood cells. Conventional diagnosis relies on manual inspection of peripheral blood smear (PBS) images by hematologists, a procedure that is time- consuming, subjective, and difficult to scale in real clinical settings. In this work, we present a Hybrid Ensemble Deep Learning framework that combines a transfer-learning based Convolutional Neural Network (CNN), a Multi-Layer Perceptron (MLP), and an Ensemble Voting Classifier to identify blood cancer from microscopic images. The proposed system integrates robust preprocessing, aggressive data augmentation, feature extraction using MobileNetV2, fully connected decision layers and both hard and soft voting schemes. Experiments on a publicly available Kaggle dataset of leukemic and normal smear images achieve an overall accuracy above 95%, with strong precision and F1-score across malignant classes. A graphical user interface implemented using Python Tkin- ter and a Flask web back-end demonstrate that the model can be deployed for real-time, image-based decision support in hospital environments.

Keywords - Blood Cancer; Leukemia; Deep Learning; Ensemble Clas- sifier; CNN; Medical Imaging; Hybrid Model

I. INTRODUCTION

Leukemia is a hematological cancer that alters the production and maturation of white blood cells in the bone marrow. Abnormal leukocytes proliferate rapidly and crowd out normal cells, leading to anemia, infections and, if untreated, death. Early detection is therefore essential for initiating appropriate therapy and improving survival outcomes.

In routine practice, hematologists visually inspect stained peripheral blood smear (PBS) images to assess the morphology and distribution of cells. Although highly informative, this manual procedure is inherently subjective and depends heavily on the experience of the expert. In addition, visual analysis is labor-intensive and not easily reproducible, which motivates the development of auto- mated methods.

Recent advances in deep learning have transformed medical image analysis. Convolutional Neural Networks (CNNs) automatically learn hierarchical features directly from images and have been successfully applied to a variety of cancer detection tasks. However, a single CNN model can still suffer from overfitting and may not generalize well to new data distributions. Ensemble strategies, where multiple diverse models are combined, have been shown to provide more stable and accurate predictions.

In this paper we propose a hybrid ensemble architecture for blood cancer identification. A transfer-learned Mo- bileNetV2 CNN is used as a feature extractor, followed by an MLP classifier. Predictions from two independently trained mixed neural networks are then fused using an Ensemble Voting Classifier. The system is trained and evaluated on a curated Kaggle dataset of leukemic and normal smear images and integrated into an interactive user interface suitable for clinical settings.

II. BACKGROUND ON BLOOD CELLS AND LEUKEMIA

Leukemia originates in the bone marrow, where pluripotent stem cells differentiate into myeloid and lymphoid lineages. Malignant transformation at any of these stages leads to abnormal blasts entering the bloodstream. These abnormalities can be captured in

© DEC 2025 | IRE Journals | Volume 9 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I6-1712680

peripheral smear images and analyzed using AI-based methods.

Acute lymphoblastic leukemia (ALL) is characterised by the uncontrolled proliferation of immature lymphoblasts. Leukemic cells typically exhibit large nuclei, irregular chromatin patterns, and atypical cytoplasmic features. Differentiating these from normal leukocytes is a key challenge addressed by automated image analysis.

III. RELATED WORK

Early computer-aided detection systems relied on hand- crafted features and classical machine learning. Demir et al. [3] extracted morphological and texture features from leukocytes and used traditional classifiers to distinguish Acute Lymphoblastic Leukemia (ALL) and Multiple Myeloma (MM). Ranjitha et al. [4] performed nucleus seg- mentation using statistical techniques and k-means clustering.

Deep learning approaches, especially CNN-based architectures, have demonstrated superior performance. Shafique and Tehsin [5] used pretrained CNN models for ALL subtype classification with high accuracy. Hallek et al. [6] studied biomarker-driven diagnosis for hematological malignancies.

Jayachitra and Umarkathaf [2] introduced a hybrid ensemble deep learning approach combining mixed neural networks with an Ensemble Vote Classifier. Their work achieved accuracy above 95%. Our approach builds upon this foundation while enhancing robustness with an im- proved MobileNetV2 feature extractor and a refined voting strategy.

IV. DATASET

The dataset used in this study is sourced from a public Kaggle repository containing 3242 microscopic smear images. Each image is labeled as:

- Normal
- Malignant Early Pre-B
- Malignant Pre-B
- Malignant Pro-B

A 70:15:15 split is applied for training, validation, and testing. Images vary in staining, illumination, size, and noise, necessitating preprocessing and augmentation.

V. PROPOSED METHODOLOGY

The workflow consists of preprocessing, deep feature extraction, classification, and ensemble fusion.

5.1 Preprocessing

Images are resized to 224 × 224, normalized, and aug- mented using:

- Rescaling
- Shearing
- Zooming
- Horizontal/Vertical Flipping

5.2 Mixed Neural Network (MNN)

Two MNN models are trained independently with:

- MobileNetV2 feature extractor
- Global Average Pooling
- Dense ReLU layers
- Softmax classifier

5.3 Ensemble Voting Classifier

For m classifiers with outputs $H_j(x)$:

$$\hat{i} = \text{mode}\{H_1(x), H_2(x), \ldots, H_m(x)\}$$

Soft voting:

$$\hat{i} = \arg\max_i \sum_{j=1}^{m} w_j p_{ij}$$

Equal weights $w_j = 1/m$ are used.

VI. EXPERIMENTAL SETUP

Training is performed on Google Colab using an NVIDIA Tesla K80 GPU. The Adam optimizer, categor- ical cross-entropy loss, early stopping, and learning rate scheduling are implemented. Class weights address data imbalance.

VII. RESULTS AND DISCUSSION

Both MNN models converge steadily with validation accuracy stabilizing above 95%. The ensemble exhibits:

- 95–96% accuracy on test data
- High precision for malignant classes
- Balanced F1 scores across all four classes

The ensemble significantly improves robustness and generalization compared to individual models.

VIII. CONCLUSION

This paper presents a Hybrid Ensemble Deep Learning framework for automated leukemia detection. The inte- gration of transfer learning, deep feature extraction, and ensemble classification results in accuracy exceeding 95%. The model's lightweight architecture and GUI deployment demonstrate practical clinical applicability.

Future work includes:

- Adding more leukemia subtypes
- Incorporating explainable AI techniques
- Multi-center validation

IX. ACKNOWLEDGMENT

The authors express gratitude to Mrs. Syeda Muzammil, Assistant Professor and Project Guide, Department of CSE, Ghousia College of Engineering, for her continuous support and encouragement.

REFERENCES

[1] S. Muzammil, Assistant Professor & Project Guide, Department of CSE, Ghousia College of Engineer- ing, 2024.
[2] J. Jayachitra and N. Umarkathaf, "Blood cancer iden- tification using hybrid ensemble deep learning tech- nique," in Proc. 2nd Int. Conf. on Electronics and Renewable Systems (ICEARS), 2023.
[3] K. Kessenbrock, V. Plaks, and Z. Werb, "Matrix met- alloproteinases: Regulators of the tumor microenvi- ronment," Cell, vol. 141, no. 1, pp. 52–67, 2010.
[4] A. Rehman, N. Abbas, T. Saba, S. I. U. Rahman, Z. Mehmood, and H. Kolivand, "Classification of acute lymphoblastic leukemia using deep learning," Microscopy Research and Technique, vol. 81, no. 11, pp. 1310–1317, 2018.
[5] S. Shafique and S. Tehsin, "Acute lymphoblas- tic leukemia detection and classification using pre- trained CNN," Technology in Cancer Research and Treatment, vol. 17, 2018.
[6] M. Hallek, P. L. Bergsagel, and K. C. Anderson, "Multiple myeloma: Increasing evidence for a mul- tistep transformation process," Blood, vol. 91, no. 1, pp. 3–21, 1998.