AUTOMATED LEUKEMIA DETECTION FROM BLOOD SAMPLES USING DEEP LEARNING

[1] V.ANIL KUMAR, [2] MR.B. NAGARAJU, [3] A. HEMA VISHNU VARDHAN, [4] S. JAIVANTH, [5] Y. GURU DIKSHITH, [6] K. RANGA SWAMI
[2] Assistant Professor, ECE Dept, RISE Krishna Sai Prakasam Group of Institutions, Valluru,AP
[1,3,4,5]Students, ECE Dept, RISE Krishna Sai Prakasam Group of Institutions, Valluru, AP
[1]vinjamurianil123@gmail.com,[2]nagchinna9@gmail.com,[3]ankireddyvishnu0@gmail.com,[4]Jaivanthsanapathi@gmail.com, [5] yalakapatigurudikshith@gmail.com,[6] kanalarangaswamy5@gmail.com

ABSTRACT:

Leukemia, a dangerous condition that affects white blood cells and bone marrow and can be fatal if left untreated, is included in blood cancer, sometimes referred to as hematopoietic cancer. In this paper, a deep learning-based Convolutional Neural Network (CNN)-based automated leukemia detection system from blood samples is presented. The suggested method uses an end-to-end CNN model that automatically learns discriminative features from microscopic blood smear images after image preprocessing to improve image quality and dataset variety.

The CNN based system carries out feature extraction and classification at the same time, in contrast to traditional machine learning approaches that rely on manually created features and intricate segmentation processes. Using publicly accessible blood picture datasets, such as samples from Kaggle, the model is trained and assessed. It has great generalization capabilities and good classification accuracy.

The outcomes of the experiment verify that the suggested deep learning method is capable of successfully differentiating between leukemic and normal blood samples. In order to support early leukemia detection and assistance in clinical decision-making, this system provides a dependable and scalable computer-aided diagnostic tool.

Index Terms: Leukemia detection, deep learning, Convolutional neural network (CNN), blood smear images, Image preprocessing.

I INTRODUCTION

Leukemia is a severe type of blood cancer that affects the bone marrow and white blood cells (WBCs), interfering with the body's immune system and normal production of blood cells. Early and accurate detection of leukemia is crucial to improving survival rates and lowering mortality. Conventional diagnostic methods mainly rely on skilled pathologists manually examining blood smear samples under a microscope. While this method is effective, it is labour-intensive, time-consuming, and subjective interpretation.

Deep learning methods, especially Convolutional Neural Networks (CNNs), have become potent tools for medical image interpretation due to the quick development of artificial intelligence. CNNs are ideal for assessing microscopic blood smear images because they can automatically learn complicated properties from images. With the goal of offering a quick, precise, and dependable computer-aided diagnostic solution, this research focuses on creating an automated leukemia detection system from blood samples using deep learning.

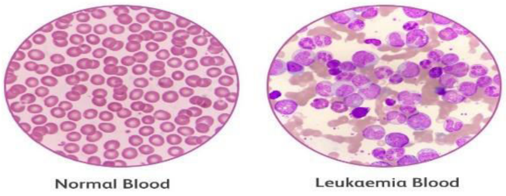


Fig.1Microscopic images for leukemia detection

II LITERATURE SURVEY

Manu Y. M. & M. E. Priyanka (2024) created an automated approach for detecting leukemia by applying conventional machine learning methods to blood smear images. In order to distinguish between leukemic and non-leukemic samples, their study focuses on feature extraction, classifier training, and performance evaluation.

A. Al-Ghraibah (2024) suggested an automated method for detecting leukemia that relies on machine learning and image processing. The system uses efficient segmentation and classification techniques to distinguish between normal and aberrant white blood cells from microscopic blood pictures.

A MATLAB-based leukemia diagnosis system utilizing SVM and KNN algorithms was created by Savithri, Santhosh, Rajkumar, and Sreerambabu (2023). Cell characteristics including shape and texture are extracted for classification following preprocessing and segmentation. When compared to manual analysis, the accuracy of the system is improved.

In 2023, Hedieh Sajedi and Muhammad Zolfaghari gave a thorough overview of automated acute leukemia identification and categorization. Their analysis identifies research gaps and future objectives by examining segmentation strategies, deep learning models, traditional machine learning, and performance comparisons.

In 2020, Bodzas, Kodytek, and Zidek presented an automated leukemia detection method that draws inspiration from human vision. To identify leukemia from blood smear images, the method combines traditional image processing with machine learning classifiers like ANN and SVM.

Pallavi, Sumita Mishra, and Sachin Kumar Using k-means clustering for picture segmentation and simple machine learning classifiers, Asthana & Pragya (2018) published one of the earliest automated leukemia diagnosis investigations. Before the advent of deep learning, their work proved that automated leukemia detection was feasible

III EXISTING MECHANISM

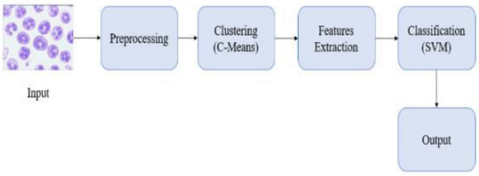


Fig.2 Existing Mechanism for Leukemia Detection Using Machine Learning.

a. Data Acquisition:

The current approach makes use of microscopic images of blood smears obtained using conventional optical microscopy. Red blood cells, white blood cells, platelets, background noise, and staining variances are all present in these pictures. Image quality is impacted by variations in illumination, resolution, and sample preparation, all of which have a direct impact on the precision of later processing and classification steps.

b. Preprocessing:

To increase computational analysis and the quality of blood smear images, image preparation techniques are used at this stage. Grayscale conversion and histogram equalization are used to improve contrast after noise and image blurring are reduced using the median and Wiener filtering techniques. After that, Otsu's thresholding is used to eliminate tiny artifacts, guaranteeing that cellular features are clearly preserved for precise segmentation and additional analysis.

c. Clustering (C-Means):

C-Means clustering is employed in the current system to segment the pre-processed image. Pixels are assigned to several clusters with different levels of membership using this soft clustering technique, which is advantageous for medical pictures where cell borders could be unclear. White blood cells can be successfully separated from other components using C-Means, although its effectiveness is heavily reliant on initial cluster centres and parameter adjustment.

d.Features Extractions & Classification:

To acquire quantitative representations of cellular properties, the segmented white blood cell areas are subjected to feature extraction.

The structural distinctions between normal and aberrant cells are described by handcrafted aspects pertaining to morphology, texture, and intensity.

These characteristics, however, necessitate specialized knowledge and might not adequately represent intricate clinical variations.

A Support Vector Machine (SVM) is used to classify the features that were extracted. Because of its excellent generalization ability and efficacy with less training data, SVM is frequently utilized in current leukemia diagnosis systems.

e. Output:

The current system's final output shows the blood sample's classification outcome, such as normal or leukemic.

IV PROPOSED MECHANISM

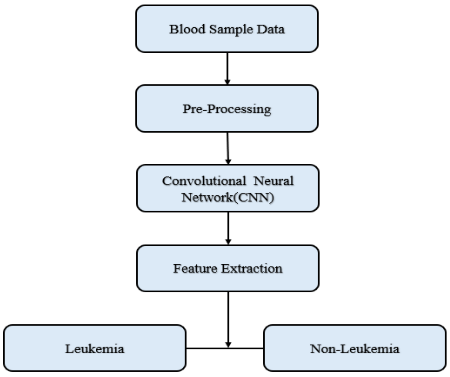


Fig.3 Proposed Mechanism for Automated Leukemia Detection Using Deep Learning.

a. Blood Smear Image Input:

The device uses optical microscopy to obtain tiny images of blood smears. Red blood cells, white blood cells, platelets, staining variations, and background noise are all present in these pictures. The main source of information for identifying anomalies associated to leukemia is the input image.

b. Preprocessing Techniques:

To improve image quality and standardize the input for deep learning analysis, preprocessing is used. In order to comply with CNN input specifications, this step involves noise reduction, contrast improvement, illumination normalization, and image scaling. Preprocessing guarantees constant cellular structure depiction and enhances image quality.

- Image Filtering: To standardize inputs for the CNN model. Before the CNN processes microscopic blood smear images, image filtering is used to reduce noise and enhance image quality.
- Segmentation: To concentrate the investigation on cellular features related to leukemia, white blood cell areas are highlighted or isolated using thresholding or clustering approaches.
- Feature Extraction: Without the need for human feature design, Convolutional Neural Networks automatically extract discriminative features from the images, such as cell texture, form, and morphological patterns.

C . Convolutional Neural Networks (CNN):

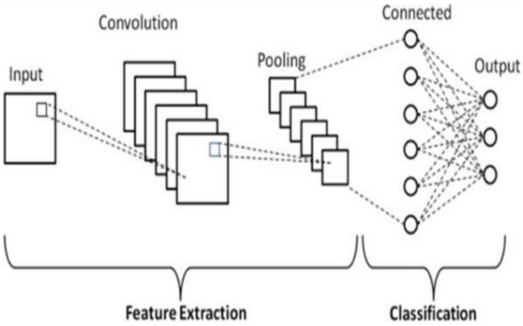


Fig.4 Convolutional Neural Networks architecture

A computer-aided diagnostic method called automated leukemia detection utilizing deep learning examines microscopic blood smear images to accurately detect leukemia. The method uses Convolutional Neural Networks (CNNs) to automatically learn discriminative features from blood cell pictures, decreasing human error and reducing reliance on hematologists' manual scrutiny. This method encourages quicker decision-making, earlier diagnosis, and better treatment results.

- The input layer: The technique starts with digital microscope photographs of microscopic blood smears. To enhance image quality and guarantee consistency for deep learning analysis, preprocessing techniques like scaling, normalization, noise reduction, and contrast enhancement are used.
- Convolution Layer: In order to automatically extract low-level and high-level information including cell boundaries and shapes, nucleus size and texture, cytoplasm patterns, and abnormal morphological structures, the convolution layers apply a number of filters (kernels) to the input image.
- Pooling Layer: Pooling layers preserve crucial information while reducing the spatial dimensions of feature maps. In order to preserve the most important characteristics associated with leukemic disorders, max-pooling is frequently utilized.
- Fully Connected (Dense) Layer: The network learns intricate correlations between characteristics once the retrieved features are flattened and sent to fully connected layers.
- Output Layer (Classification): The final output layer classifies blood samples into normal and leukemia groups using activation functions. The model generates probability ratings that show how confidently leukemia is detected.

Deep learning and CNN architecture-based automated leukemia detection offers a dependable, effective, and precise diagnostic solution. The method greatly improves early leukemia diagnosis and helps medical personnel provide prompt and accurate treatment by fusing cutting-edge image processing with potent feature learning capabilities.

V SOFTWARE DISCRIPTON

A deep learning-based framework for examining images of microscopic blood samples is used to create the automated leukaemia diagnosis system in MATLAB. Preprocessing techniques such as noise reduction, contrast enhancement, normalisation, and scaling are used to improve the quality of digital blood smear photographs that are obtained in common formats. Preprocessing increases the visibility of white blood cell structures, which are essential for leukaemia detection, and guarantees uniformity among samples. White blood cells are separated and aberrant leukemic areas are highlighted using image segmentation techniques such as thresholding, clustering, and morphological processes.

In order to eliminate the requirement for manual feature extraction, deep learning models—specifically, Convolutional Neural Networks (CNNs)—are used to automatically learn discriminative features from the segmented blood cell pictures.

Based on patterns learned about changes in cell shape, texture, and intensity, the trained network divides samples into leukemic and non-leukemic groups. Metrics including accuracy, precision, recall, and F1-score are used to assess system performance and verify diagnostic reliability. The method is appropriate for automated and repeatable leukaemia diagnosis in clinical and research contexts because it makes use of MATLAB's Image Processing Toolbox and Deep Learning Toolbox, which facilitate effective model training, visualisation, and analysis.

VI ANALYSIS OF PERFORMANCE AND RESULTS

The performance of the deep learning-based automatic leukaemia detection system is summarised visually in the sample classification results. The model's projected class and the associated ground truth diagnosis are both labelled on each microscopic blood sample image. All twelve test samples, including leukemic and non-leukemic cases, show appropriate blood sample classification because the projected outcomes match the actual labels. These findings demonstrate how well the suggested deep learning method can automatically detect leukaemia from blood smear pictures. The model's dependability and potential for application in automated blood sample analysis and clinical decision support systems are demonstrated by its capacity to accurately differentiate aberrant leukemic cells from normal blood cells.

a. Sample Classification Results:

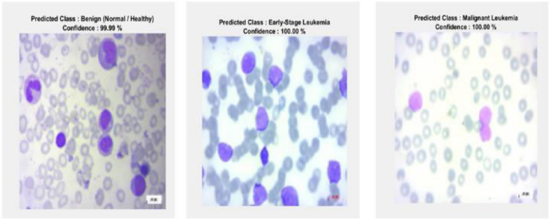


Fig.5 Classified Output Results.

The deep learning-based leukaemia detection model's dependability is demonstrated by the steady agreement between expected and actual results, which shows that it has successfully learned to differentiate leukemic cell patterns from normal blood cell shape.

These findings highlight the suggested system's potential for automated leukaemia identification in clinical and laboratory settings, where accurate and prompt blood sample analysis is essential. Additionally, the model's performance is intuitively validated by the visual depiction of categorisation results, which boosts researchers' and medical professionals' confidence in its diagnostic potential.

b. Confusion Matrix:

Blood samples are divided into three categories in the confusion matrix: benign, early-stage leukaemia, and malignant leukaemia. The deep learning model's projected class labels are shown in the columns, while the rows represent the actual class labels. High prediction accuracy across all three classes is indicated by the strong intensity along the diagonal cells, which indicate that the majority of samples are accurately sorted into their respective categories.

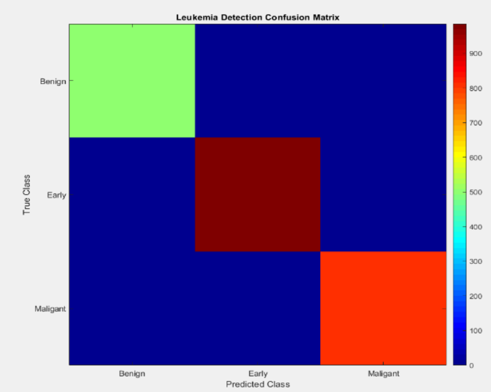
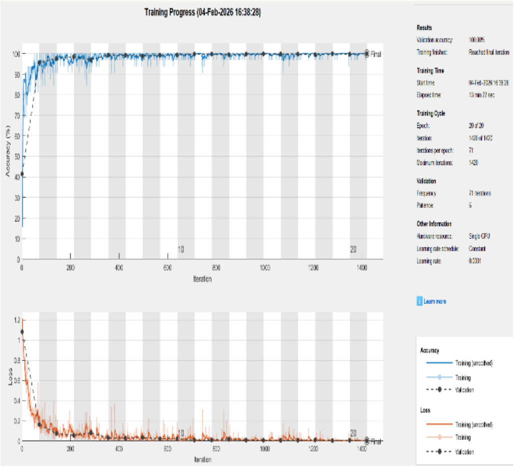


Fig.6 Confusion matrix for leukemia datasets

This confusion matrix shows how well the suggested model learns discriminative leukemic cell properties and reduces misclassification. The system's dependability for automated leukaemia identification from blood samples is demonstrated by the strong alignment between predicted and true labels, which supports its prospective use in clinical and laboratory diagnostic applications.

c. Accuracy and Loss Curve of Training:


Training Progress (04-Feb-2026 16:38:28)

Results
Validation accuracy: 100.00%
Training finished: Reached final iteration

Training Time
Start time: 04-Feb-2026 16:38:28
Elapsed time: 13 min 22 sec

Training Cycle
Epoch: 20 of 20
Iteration: 1420 of 1420
Iterations per epoch: 71
Maximum iterations: 1420

Validation
Frequency: 71 iterations
Patience: 5

Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001

Fig.7 Accuracy and Loss Curve

The training progress graph depicts the deep learning model's learning behaviour after 20 epochs and 1420 iterations. The accuracy curve rises quickly during the first iterations, followed by steady convergence near 100% for both the training and validation stages. This shows that the model quickly learnt significant features from blood sample photos and performed consistently throughout the training period. The validation accuracy of 100% indicates strong generalisation within the dataset.

The loss curve shows a dramatic decline at the start of training, progressively approaching near-zero values as the iterations pass. Both training and validation losses are low and well aligned, showing minimal overfitting. The utilisation of a consistent learning rate and early stopping patience helped to ensure stable convergence. Overall, the training results validate the efficacy and reliability of the proposed deep learning model for automated leukaemia identification from blood samples.

d. Table of Performance Measures:

Leukemia Detection Accuracy = 100.00%

Class-wise Performance:

labels	precision	recall	F1
'Benign'	1	1	1
'Early'	1	1	1
'Malignant'	1	1	1

Leukemia CNN model saved successfully

VII CONCLUSION

This article describes an autonomous method that uses deep learning techniques applied to photos of microscopic blood samples to diagnose leukaemia. To identify aberrant white blood cells and categorise them as leukemic or normal, the system analyses blood smear images. Deep learning eliminates the need for manual feature design by allowing the model to learn key features straight from the photos, increasing detection accuracy. Because it makes deep learning model implementation, training, and visualisation simple, MATLAB is employed as the development platform.

The findings of the trial demonstrate that the suggested method can consistently and accurately identify leukaemia. By lowering manual labour and promoting early diagnosis, this system can help physicians. However, the size and quality of the dataset determine how well the model performs, and different staining and imaging settings can have an impact on the outcomes. Future research can further boost dependability and improve the system's usefulness in actual clinical applications by utilising more and more varied datasets and refining the deep learning model.

REFERENCES

1. SiddhikaArunachalam(2020) , “Applications of machine learning and image processing techniques in the detection of leukemia,” Int. J. Sci. Res. Comput. Sci. Eng., vol. 8, no. 2, pp. 77–82, Apr. 2020.
2. V. V. Nagamani, Sirigiri Lakshman Sai, Puram Pavan, Chilukuru Venkata Sai (2024), “Detection of leukemia disease using image processing and machine learning,” Int. J. Res. Appl. Sci. Eng. Technol., vol. 12, no. IV, pp. – , Apr. 2024, doi:10.22214/ijraset.2024.60399.
3. Sachin Paswan & Yogesh Kumar Rathore (2017) “Detection and classification of blood cancer from microscopic cell images using SVM, KNN and NN classifier,” Int. J. Adv. Res. Ideas Innov. Technol., vol. 3, no. 6, pp. 315–318, Nov. 2017.
4. Arunachalam.S(2020).“Applications of machine learning and image processing techniques in the detection of leukemia,” International Journal of Scientific Research in Computer Science and Engineering, vol. 8, no. 2, pp. 77–82, Apr. 2020.
5. Bodzas, K., Kodytek, P., & Zidek, J. (2020). “Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception,” Front. Bioeng. Biotechnol., vol. 8, art. no. 1005, Aug. 2020, doi:10.3389/fbioe.2020.01005.

6. Peetala, S. & Kumari, B. L. (2022). “Detection of leukemia and its staging using image processing and machine learning techniques,” Int. J. Signal Process., vol. 7, pp. –, 2022.
7. Preeti Jagdev (2017) “Automated detection and classification of leukemia using image processing and machine learning,” Int. J. Adv. Res. Sci. Eng., vol. 6, no. 4, pp. –, Apr. 2017.
8. H. Ramoser, V. Laurain, H. Bischof & R. Ecker (2005) “Leukocyte segmentation and classification in blood-smear images,” in Proc. IEEE-EMBS, 2005, pp. 3371–3374.
9. Srimathi & Yamuna (2018) “Study of cancer detection techniques using various image processing algorithms,” Int. J. Eng. Dev. Res., vol. 6, no. 3, pp. –, 2018.
10. Hazra, Kumar & Tripathy (2017) “Automatic leukemia detection using image processing technique,” Int. J. Latest Technol. Eng. Manage. Appl. Sci., vol. VI, no. IV, pp. 42, Apr. 2017.