Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.5, May : 2025

A CNN-DRIVEN DEEP LEARNING APPROACH FOR DETECTING LEUKEMIA IN MICROSCOPIC IMAGES

Dr. S. Sudhakar, M.E., Ph.D., Associate Professor/CSE, SSM Institute of Engineering and Technology, Dindigul, Tamilnadu.
Mrs. M. Backialakshmi M.E., (Ph.D.), Assistant Professor/AI&DS, SSM Institute of Engineering and Technology, Dindigul, Tamilnadu.
Mrs. A. Padmapriya, M.Tech., (Ph.D.), Assistant Professor/CSE, SSM Institute of Engineering and Technology, Dindigul, Tamilnadu.

ABSTRACT
Leukemia is a form of cancer characterized by the abnormal proliferation of blood cells. Traditionally, diagnosing leukemia involves manually inspecting microscopic images—a process that is not only labor-intensive but also prone to human subjectivity. This paper seeks to advance the integration of deep learning with medical imaging by exploring a promising approach in the field of hematologic diagnostics. Given their exceptional performance in image classification tasks, deep learning models—particularly Convolutional Neural Networks (CNNs)—are well-equipped to address the challenges of detecting cancerous cells in microscopic imagery. This research presents the design and assessment of a CNN-based deep learning model trained on a carefully assembled dataset to identify key morphological features of blood cells and accurately detect leukemia.

Keywords—
Leukemia Detection, Deep Learning, CNN, Microscopic Image Analysis

I. Introduction
Leukemia is a serious disorder that arises due to the premature or abnormal formation of white blood cells (WBCs). Commonly known as white blood cell cancer, its exact cause remains unclear. However, a major contributing factor is believed to be cell mutations that result in genetic alterations in a person's DNA. Leukemia is generally categorized into four primary types: Acute Lymphocytic Leukemia (ALL), Acute Myelogenous Leukemia (AML), Chronic Lymphocytic Leukemia (CLL), and Chronic Myelogenous Leukemia (CML). Among these, AML is considered the most severe and is commonly observed in young children and teenagers. Global cancer statistics suggest that leukemia affects males more frequently than females. Early and accurate detection plays a vital role in guiding successful treatment approaches and improving patient prognosis. The proposed methodology involves several stages: preprocessing, segmentation, feature extraction, and classification of blood cells. In the preprocessing stage, microscopic images are enhanced using a low-pass Gaussian filter to smooth out rough edges—similar to enhancing a photo for better clarity. In segmentation, where thresholding techniques are applied to distinguish leukemia cells from the rest of the image, helping the computer focus on the relevant areas. After segmentation, feature extraction identifies the unique characteristics of the blood cells. A CNN is then applied to extract these features using Conv2D layers, which operate like trained observers recognizing key visual elements. Together, these processes build an intelligent system capable of accurately detecting leukemia cells in microscopic images - bringing us closer to a powerful, AI-driven tool for automated medical diagnosis.

II. Related Works
The detection and classification of leukemia using deep learning (DL) techniques has garnered growing attention due to its potential for accuracy, scalability, and automation in medical diagnostics. Mustaqim et al. [3] conducted a systematic literature review analyzing progress in detecting subtypes of Acute Lymphoblastic Leukemia (ALL) from 2018 to 2022. Reviewing 65 research articles, they identified major challenges, including limited publicly available datasets and the high computational

demands of DL models. Complementing this, Likith et al. [9] highlighted key obstacles such as insufficient diversity in training data, the complexity of implementing deep learning frameworks, and the emerging role of transformer architectures in medical image analysis. Kumar et al. [7] emphasized the practical challenges in conventional leukemia diagnosis, which relies heavily on expert evaluation of blood smear slides—an approach that is not only time-consuming but also vulnerable to human error. These foundational works collectively underscore the urgent need for scalable and automated diagnostic tools that can alleviate the burden on healthcare systems while maintaining or improving diagnostic accuracy.

To address these issues, several studies have proposed DL models designed specifically for image-based leukemia detection. Batool and Byun [1] introduced a lightweight EfficientNet-B3 architecture utilizing depthwise separable convolutions to improve the classification of white blood cell images. Their model aimed to overcome the limitations of traditional DL algorithms, particularly in terms of generalization and convergence speed. Similarly, El-Shalshawi et al. [15] combined image preprocessing with optimized deep neural networks to enhance the detection accuracy of ALL, demonstrating how architectural refinements can substantially improve classification performance. Transfer learning approaches have also been widely adopted. Magpantay et al. [5] utilized a deep convolutional neural network (CNN) with transfer learning and achieved a training accuracy of 97.2% along with a mean average precision (mAP) of 99.8%. Sheet et al. [6] deployed MobileNetV2 in conjunction with CNNs to detect leukemia, reporting 96.58% accuracy, 95.17% sensitivity, and 98.58% specificity. These results validate the effectiveness of transfer learning in reducing training time while maintaining high diagnostic precision. Further supporting this trend, U. Ul Ain et al. [13] developed a CNN-based model to classify leukemia from microscopic images, demonstrating scalability and strong diagnostic capabilities. Taking a step further, Chen et al. [14] implemented a real-time detection model for ALL using CNNs, enabling rapid diagnosis suitable for clinical environments where timely decisions are crucial.

Advanced object detection frameworks have also been applied to leukemia diagnostics. Chatzisofroniou [4] used YOLOv3 to classify ALL and normal cells, achieving 97.2% accuracy and a mAP of 99.8%, indicating the potential of real-time detection in reducing manual diagnostic errors. Malik et al. [8] proposed a BiCNN-CML model combining Faster R-CNN and SSD to detect Chronic Myeloid Leukemia (CML). Their hybrid approach enhances both speed and localization accuracy in complex medical images. Additionally, hybrid feature engineering strategies have led to further improvements in diagnostic performance. Elhassan et al. [12] integrated Fuzzy C-Means (FCM), Particle Swarm Optimization (PSO), and Gray Level Co-occurrence Matrix (GLCM) techniques into a Bi-LSTM and CNN-based classifier, targeting Acute Myeloid Leukemia (AML) detection. Bukhari et al. [11] employed a Squeeze-and-Excitation (SE) framework that enhances channel interdependencies, reporting classification accuracies of 97.57% and 96.41% across primary and secondary datasets, respectively. In parallel, Hossain et al. [2] presented a symptom-based diagnosis model using explainable AI. Leveraging decision tree algorithms on patient data from Bangladeshi hospitals, they achieved an accuracy of 97.45% and an ROC value of 0.783, highlighting the importance of transparency in AI-driven healthcare tools. Finally, Mustaqim et al. [10] introduced a novel DL model for early-stage ALL detection using minimal blood samples. Their findings highlight that effective leukemia identification can be achieved with minimal data through thoughtful model design, reinforcing the significance of early intervention for better patient prognosis.

III. Methodology

Deep learning plays a vital role in automatically identifying key features from microscopic images of blood cells, aiding in the early detection and diagnosis of leukemia. In this study, a deep Convolutional Neural Network (CNN) is trained to recognize and classify five categories of cells: Acute Lymphoblastic Leukemia (ALL), Chronic Lymphocytic Leukemia (CLL), Acute Myeloid Leukemia (AML), Chronic Myeloid Leukemia (CML), and healthy (non-cancerous) cells. The workflow

involves several stages, including preprocessing, image segmentation, and classification. Two different models—one based on statistical analysis and the other utilizing the ResNet34 architecture—are employed for classification tasks. Given that ALL progresses quickly, especially in children, additional models such as a conventional CNN and the YOLO (You Only Look Once) algorithm are also explored for efficient cancer cell detection. Among the various deep learning models tested, DenseNet169 stands out, achieving a remarkable classification accuracy of 97.62%, surpassing the other approaches.

3.1 Scope and Assumptions

Although several techniques—including traditional machine learning and image analysis—are used for leukemia detection, deep learning models excel in extracting meaningful features with higher precision. Detecting all four types of leukemia cells presents challenges, but deep learning, supported by vast datasets, enhances diagnostic accuracy and enables earlier detection. Additionally, these algorithms can predict the risk level of the disease based on patient data, aiding in more effective patient management.

3.2 CNN Algorithm

A Convolutional Neural Network (CNN) is a specialized type of artificial neural network that's especially powerful in deep learning tasks like image recognition and feature extraction. Its growing popularity in image processing stems from its impressive ability to identify and learn meaningful patterns in images. CNNs operate through three key layers: the convolutional layer, pooling layer, and fully connected layer. Among these, the convolutional layer is the workhorse—it handles most of the computation and allows the network to process and learn from vast amounts of data. This structure enables CNNs to automatically detect important features in images, making them a vital tool in modern image analysis.

3.3 Detection of Disease

The detection of disease using microscopic images through deep learning algorithms involves several key processes. Initially, data collection includes classifying all four types of white blood cell cancers. The first phase is meticulous data preprocessing, where raw microscopic images undergo cleaning, resizing, and normalization to ensure consistency and optimal model training. Next, segmentation techniques isolate relevant regions of interest, allowing the model to focus on critical areas for leukemia detection. Following segmentation, feature extraction captures discriminative features from these regions to create a rich representation of underlying patterns. A CNN architecture is then designed and implemented, integrating the preprocessed, segmented, and feature-extracted data for robust learning. The training phase involves fine-tuning the CNN through iterative epochs, optimizing its ability to distinguish leukemia cells from normal cells. Rigorous evaluation and validation on distinct datasets assess the model’s performance across diverse scenarios. This methodology leverages the power of CNN-based deep learning while emphasizing the importance of thorough preprocessing, segmentation, and feature extraction to enhance accuracy and generalization. Ultimately, the proposed approach lays the foundation for an effective, automated leukemia detection system, contributing significantly to advancements in medical image analysis and diagnostic tools.

3.4 Proposed Methodology

The methodologies and tools used to detect four types of leukemia involve preprocessing, segmentation, and feature extraction. The datasets are trained and tested using a Convolutional Neural Network (CNN), which classifies the output into leukemia and normal cells.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.5, May : 2025

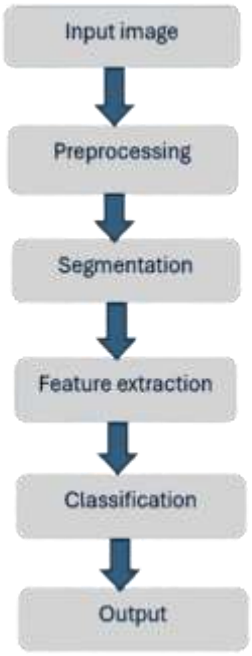

Figure 1: A block diagram of the proposed method.

The process of leukemia detection using a Convolutional Neural Network (CNN) begins with acquiring microscopic images of blood samples. These input images, which may contain both healthy and abnormal white blood cells, serve as the raw data for the system. The first step involves preprocessing these images to enhance their quality and consistency. Common preprocessing tasks include resizing the images, normalizing pixel values, removing noise, and adjusting contrast or brightness levels. These steps ensure that the data fed into the CNN model is clean and standardized. Following this, segmentation techniques are applied to isolate the regions of interest—typically the white blood cells—from the rest of the image. By focusing only on relevant areas, segmentation enhances the accuracy of subsequent analysis and reduces the processing burden on the model, as shown in Figure 1.

After segmentation, the system enters the feature extraction phase, where the CNN automatically learns key features such as shape, texture, and color intensity associated with leukemia cells. These learned features are passed into the classification phase, where the CNN predicts whether the image shows signs of leukemia. Certain models are further trained to recognize specific leukemia subtypes, such as Acute Lymphoblastic Leukemia (ALL) and Acute Myeloid Leukemia (AML). In the final step of the process, the model generates an output that includes the predicted classification along with a confidence score, which reflects how certain the model is about its prediction.

In some implementations, the system also highlights the detected abnormalities within the image. This end-to-end workflow using CNNs not only speeds up the diagnostic process but also minimizes human error, providing a powerful tool for aiding medical professionals in early and accurate leukemia detection.

3.4.1 Image Preprocessing

The image preprocessing phase, a crucial step before further analysis, involves applying a low-pass Gaussian filter to enhance the quality and clarity of microscopic images used in leukemia detection. Initially, raw images are resized to a standardized input size to ensure consistent processing. The resized images then undergo smoothing through the Gaussian filter, which reduces high-frequency noise while preserving essential features. The filter's parameters, including the standard deviation, are carefully tuned to balance noise suppression with feature retention. This preprocessing step is vital for improving the robustness of subsequent segmentation and feature extraction, enabling more accurate identification of leukemia cells. By standardizing and refining the input images, the use of a low-pass

Gaussian filter significantly contributes to the effectiveness of the automated leukemia detection system.

3.4.2 Segmentation

In the Convolutional Neural Network (CNN) architecture for leukemia detection, segmentation is performed using a threshold-based method. Following the preprocessing phase, the grayscale microscopic images undergo segmentation to isolate the relevant regions of interest, particularly the leukemia cells. This method involves applying a carefully selected threshold—determined through histogram analysis or similar techniques—to differentiate the foreground (leukemia cells) from the background. As a result, a binary mask is generated, where pixels with intensities above the threshold are classified as foreground, effectively highlighting the critical areas for further analysis. This threshold-based segmentation is a vital step that directs the CNN's focus toward meaningful features in the images, facilitating accurate feature extraction and classification. Its adaptability allows the model to capture subtle patterns linked to leukemia cells, thereby enhancing the accuracy and reliability of the automated detection system.

3.4.3 Features Extraction

At the core of a Convolutional Neural Network (CNN) lies the Conv2d layer, a pivotal component responsible for detecting intricate patterns and extracting essential features from input images. This layer begins by initializing filters or kernels designed to capture distinct visual characteristics. During the convolution operation, each filter slides systematically over the input image, generating a feature map that highlights relevant patterns. By applying activation functions like ReLU, the model gains the flexibility to understand and represent complex relationships in the input data. Optionally, pooling layers may follow to perform downsampling, which reduces dimensionality and improves computational efficiency. The resulting feature maps, which encapsulate the image's salient features, are then passed to subsequent layers, enabling the network to build hierarchical representations of the input. By stacking multiple Conv2d layers, the CNN can progressively abstract more intricate and meaningful information. During training, the filter weights are optimized using backpropagation, allowing the network to learn task-specific patterns such as identifying leukemia cells in microscopic images. Thus, the Conv2d layer, through its filtering and convolution operations, plays a fundamental role in enabling CNNs to perform effective image analysis and classification.

3.4.4 Classification

In the classification of microscopic images for identifying various types of leukemia, fully connected layers serve as the final and crucial stage of a Convolutional Neural Network (CNN). These layers take the abstract features extracted by the preceding convolutional and pooling layers—such as cell morphology and structural patterns—and flatten them to form a vector that is connected to output neurons, each representing a different leukemia class. In this stage, the fully connected layer takes the features learned by the network and translates them into classification probabilities for various leukemia types, including Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML), Chronic Lymphocytic Leukemia (CLL), Chronic Myeloid Leukemia (CML), and other subtypes.By training the network on a labeled dataset of leukemia images, the fully connected layer learns to distinguish between these categories, effectively acting as a robust classifier that enables accurate diagnosis based on image data.

IV. Results and Discussion

The obtained results from our deep learning model are promising, demonstrating an impressive accuracy of 92% in detecting all four types of leukemia—Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML), Chronic Lymphocytic Leukemia (CLL), and Chronic Myeloid Leukemia (CML)—using microscopic images. This highlights the model's strong ability to differentiate between normal and leukemia cells across various subtypes. A key factor behind this success is the meticulous curation of a diverse and representative dataset used for training.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.5, May : 2025

Figure 2. Grey image and Segmented image

Nonetheless, it is important to recognize ongoing challenges such as inconsistencies in staining techniques, variations in image quality, and the occurrence of overlapping cell types, which may impact the model's performance. As illustrated in Figure 2, preprocessing steps such as conversion to grayscale and segmentation significantly aid in isolating meaningful features from the images, supporting more accurate classification.

Despite these challenges, such as variability in staining techniques, image quality, and overlapping cells that may impact the model's performance, the clinical relevance of the model is evident, offering a promising tool for early and accurate leukemia diagnosis, thereby enabling timely medical intervention. Moving forward, continued collaboration with healthcare professionals, refinement of the dataset, and the integration of interpretability methods will be crucial for enhancing the model's robustness and validating its effectiveness in real-world clinical environments.

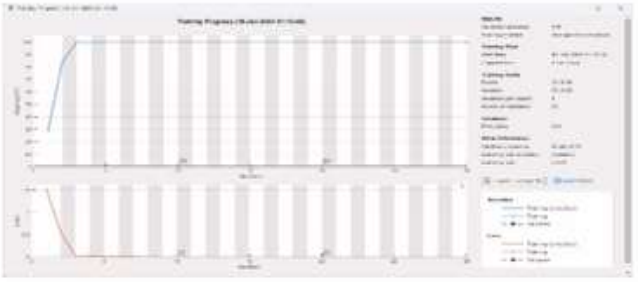

Fig.3. Accuracy graph

The notably short training time of 1 minute and 3 seconds underscores the efficiency of the model's convergence, highlighting its practicality by minimizing computational resource usage and enabling faster deployment. Figure 3 demonstrates the training accuracy and loss curves, indicating stable convergence over time. Additionally, Figure 4 shows the confusion matrix, which visually confirms the model's strong predictive performance across all four leukemia subtypes.

Fig .4 AML is detected

V. Limitations and Future Directions

Our proposed CNN model for leukemia detection using microscopic images shows promising results but faces several limitations. The relatively small dataset may not fully capture biological variability,

risking overfitting due to model complexity. Variations in image quality, labeling inconsistencies, and high computational demands further challenge scalability and performance. To overcome these, future work should focus on expanding the dataset with diverse clinical samples, applying transfer learning, optimizing model parameters, and integrating clinical data to improve generalizability. Collaborations with medical experts are crucial for enhancing interpretability and validating the model on external datasets. Additionally, incorporating explainability techniques will increase transparency and trust in clinical use. Future advancements may include multimodal data integration, real-time diagnosis, automated lesion segmentation, and linking with Electronic Health Records (EHR). The adoption of Explainable AI (XAI), transfer learning, and AI-human collaborative systems will advance personalized medicine and expand CNN applicability in resource-limited settings. Continued research, ethical considerations, and interdisciplinary collaboration are key to realizing these goals.

VI. Conclusion

This study presented the development and evaluation of a CNN model for leukemia detection from microscopic images. Through careful preprocessing, segmentation, and feature extraction, the model demonstrated strong learning capabilities, with accuracy and loss metrics indicating effective training. Testing on unseen data yielded promising results, suggesting potential clinical applicability. Confusion matrix analysis highlighted the model's strengths and areas needing improvement. However, limitations such as dataset size and diversity may affect generalizability, and overfitting risks persist despite mitigation efforts. Future research should emphasize dataset expansion, hyperparameter tuning, exploring advanced architectures, and integrating clinical information. Collaboration with domain experts, application of explainable AI techniques, and external validation are essential to ensure reliability in real-world settings. Overall, while the CNN model shows significant potential in leukemia detection, ongoing refinement is necessary for practical healthcare deployment.

References

[1] Amreen Batool and Yung-Cheol Byun, "Lightweight efficientNetB3 model based on depth wise separable convolutions for enhancing classification of leukemia white blood cell images," IEEE, vol. 11, 2023.
[2] M.A. Hossain et al., "Symptom based explainable artificial intelligence model for leukemia detection," IEEE, vol. 10, 2022.
[3] T. Mustaqim et al., "Deep learning for the detection of ALL subtypes on microscopic images: A systematic literature review," IEEE, vol. 11, 2023.
[4] G. Chatziso froniou, "The Known Beacons Attack," 34th Chaos Communication Congress, 2021. Available: https://labs.com/news/2018/02/01/known-beacons-attack-c3
[5] L. D. C. Magpantay, H. D. Alon, Y. D. Austria, M. P. Melegrito, and G. J. O. Fernando, "Transfer Learning-Based Deep CNN Approach for Classification and Diagnosis of Acute Lymphocytic Leukemia Cells," 2022 International Conference on Decision Aid Sciences and Applications (DASA), 2022.
[6] J. Sheet, C. Ghosh, and B. K. Das, "Deep Learning-Based Transfer Learning for the Detection of Leukemia," 2023 International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC), 2023.
[7] V. S. Kumar, R. S., P. K. Pareek, V. H. C. De Albuquerque, D. Gupta, and A. Khanna, "Detection of Leukemia from Histopathological Image using Deep Learning Techniques," 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), 2022.
[8] V. Malik, R. Mittal, and A. Rana, "BiCNN-CML: Hybrid Deep Learning Approach for Chronic Myeloid Leukemia," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), 2022.

[9] Likith N.G., Uzair Alam, Dr. Jagruthi H., "Leukemia cancer detection using deep learning," https://doi.org/10.56726/IJRMETS38913.

[10] Mustaqim, C. Fatichah, and N. Suciati, "Deep Learning for the Detection of Acute Lymphoblastic Leukemia Subtypes on Microscopic Images: A Systematic Literature Review," IEEE Access, vol. 11, 2023, pp. 16108-16117.

[11] Maryam Bukhari, Sadaf Yasmin, Saima Sammad, Ahmed A. Abd El-Latif, "A Deep Learning Framework for Leukemia Cancer Detection in Microscopic Blood Samples Using Squeeze and Excitation Learning," Mathematical Problems in Engineering, vol. 2022, Article ID 8012274.

[12] T. A. M. Elhassan, M. S. M. Rahim, T. T. Swee, S. Z. M. Hashim, and M. Aljurf, "Feature Extraction of White Blood Cells Using CMYK-Moment Localization and Deep Learning in Acute Myeloid Leukemia Blood Smear Microscopic Images," IEEE Access, vol. 10, 2022, pp. 16577-16591.

[13] Q. Ul Ain, S. Akbar, S. A. Hassan, and Z. Naqvi, "Diagnosis of Leukemia Disease through Deep Learning using Microscopic Images," 2022 2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2), Rawalpindi, Pakistan, 2022.

[14] E. Chen, R. Liao, M. Y. Shalaqinov, and T. H. Zeng, "Real-time Detection of Acute Lymphoblastic Leukemia Cells Using Deep Learning," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, USA, 2022.

[15] M. R. El-Shalshawi, B. E. El-Demersh, and A. Mohamed, "A Deep Learning Approach for Diagnosis and Detection of Acute Lymphoblastic Leukemia," 2023 Intelligent Methods, Systems and Applications (IMSA), Giza, Egypt, 2023.