iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: III Month of publication: March 2025

DOI: https://doi.org/10.22214/ijraset.2025.67552

A Review of Deep Learning Models for Blood Cancer Detection

Satvik Tyagi[1], Aditya Kumar[2], Keshav Kant Tyagi[3], Peeyush Tomar[4]
[1, 2, 3, 4]*Student, Computer Science & Information Technology, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh*
[5]*(Asst. prof.), Information Technology, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh*

Abstract: Blood cancer, particularly leukemia, remains a significant health concern, demanding early andaccurate diagnosis for effective treatment. Recent advances in deep learning have been very promising forthe automation of blood cancer detection from microscopic images. This review is an overview of deep learning techniques in classifying blood smear images as either cancerous or non-cancerous. Quite a number of studies have attempted to apply deep learning models that were trained on publicly availabledatasets like Kaggle in identifying malignant cells. Mostof the models used feature extraction with traditional machine learning algorithms, such as Naive Bayes, to enhance classification accuracy.

While deep learning techniques have been shown to hold great promise for the improvement of accuracy indiagnostics, issues persist, such as data quality, model interpretability, and clinical validation.

Keywords: CNN, Kaggle, NB, Blood cancer.

I. INTRODUCTION

The research work aims to establish an automated systemof early detection and classification of blood cancer, including leukemia, from microscopic blood smear images based on deep learning. The system could help inearly identification of cancerous cells so that proper earlyintervention and treatment take place. The system conceptualized which distinguishes healthy frommalignant blood cells utilizes the Deep Learning models,some of which are CNNs to minimize reliance on manualinspection, time consuming and prone to human errors. This is done in an effort to improve accuracy during diagnosis, reduce the workload on medical practitioners, and subsequently the patients who will be treated with better drugs for Leukemia. A tool would be designed to assist healthcare providers to make proper clinical decisions and to detect blood cancer. The deep learning model would be trained on a massive dataset of labeled blood smear images so it can learn thevery subtle patterns and features characteristic of leukemia. Consequently, faster and more consistent results than in conventional diagnostic procedures can be obtained by automating the process of classification, which ultimately could end up reducing the time taken to reach a diagnosis. Further, such a system that might work with minimal interference from humans would greatly benefit conditions with limited resources where exposure to professional hematologists might be scarce, providing a meaningful solution for the early detection of cancer and generally to make healthcare efficiency better.

II. MOTIVATION

With the rapid rise in blood cancer cases and other types of cancer around the world, there is a vast shift toward more efficient diagnostics. Traditional diagnostics have proven effective but aretime-consuming, requiring expert interpretation to reach the right treatment, which could be too late in most cases. The recent discoveries in deep learning and artificial intelligence display much promise in having the ability to overcome these challenges, especially in automating detection for cancer within medical images. Recent studies have been successful in the applications ofconvolutional neural networks, transfer learning, and hybrid models, which greatly increase the accuracy of diagnosis of blood cancer while accelerating the process. The subtlest patterns in blood samples under a microscope can be differentiated using these techniques for accurate clinician diagnosis and quicker judgment. Despite this progress, such lightweight and efficient models are still lacking. The current models have yet to sufficiently serve in real-world clinical environmentscharacterized by resource-poor settings.

III. CONTRIBUTION AND ORGANIZATIONOF PAPER

This paper presents a deep learning-based framework for the early detection of blood cancer, specifically leukemia, through the automated analysis of microscopic blood smear images. The primary contribution of this work is the development of a robust system utilizing Convolutional Neural Networks (CNNs) for accurate cancer detection, coupled with a comparative analysis of Naive Bayes (NB) for baseline classification.

The paper also integrates various machine learning and deep learning techniques, including feature extraction and transfer learning, to enhance diagnostic performance and reliability in real-world clinical applications.

The structure of the paper is as follows: Section 2 describes the methodology used to collect and process blood smear image datasets, along with the design of the proposed CNN and Naive Bayes models. Section 3 provides a detailed review of cancer complications and clinical applications, focusing on the potential of AI in improving early diagnosis. Section 4 offers a comparative analysis of the deep learning and machine learning approaches, evaluating their performance using various metrics. In Section 5, a thorough discussion is provided on the implications of the proposed methods andtheir potential impact on clinical practices. Finally, Section 6 concludes the paper and outlines future directions for enhancing the model's generalization and computational efficiency, as well as its application in resource- constrained environments.

IV. OBJECTIVES

1) To implement a convolutional neural network (CNN) model for the classification of blood smear images into cancerous and non-cancerouscategories.
2) To compare the performance of different algorithms, such as CNN and Naive Bayes.
3) To develop a user-friendly interface that allows users to upload datasets, run algorithms, extract features, and visualize the results and comparison graphs.
4) To provide a fast and efficient method for earlydetection of blood cancer, aiding healthcare professionals in real-time diagnosis.

This study focuses on utilizing deep learning algorithms to detect blood cancer from microscopic images. By leveraging the power of CNNs, the proposed system aims to automatically classify blood smear images into cancerous and non-cancerous categories,thus providing a fast and reliable diagnostic tool. In addition to CNN, the study also integrates other machine learning techniques, to explore their potential in cancer detection.

The development of a user-friendly interface further enhances the practicality of this system, allowing users to easily upload datasets, run algorithms, and compare results. This deep learning- based approach holds great promise for improving the early detection of blood cancer, thereby aiding healthcare professionalsin their decision-making process and contributing to better patient care.

V. LITERATURE SURVEY

Automated detection of blood cancers such as Acute Myeloid Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL) has gained significant attention with the development of deep learning and machine learning techniques. Many recent studies have explored methods that leverage advanced models like transfer learning, convolutional neural networks (CNNs), and support vector machines (SVMs) to enhance diagnostic accuracy. For instance, some models integrate transfer learning and SVM classifiers within lightweight architectures utilizing depth-wise separable convolutions, which improve processing efficiency while maintaining high accuracy, especially when tested on standard datasets like ALLIDB1, ALLIDB2, and ASH [1], [9], [10].

This focus on optimizing computational load makes such models promising candidates for early leukemia diagnosis in clinical environments. A major trend in the literature involves CNN-based architectures with modifications to enhance feature extraction capabilities. Studies by Bukhari et al. and Rahman et al. explore CNN models with components like squeeze-and-excitation mechanisms and pre-trained ResNet50 classifiers, achieving high accuracy levels up to 99.84% [2], [3]. These models perform robustly on public datasets, highlighting the utility of deep learning in accurately distinguishing leukemia cells from healthy cells.

Additionally, the application of techniques like Linear Discriminant Analysis (LDA) and Principal Component Analysis(PCA) to preprocess and optimize feature selection has shown promise in improving both accuracy and processing efficiency [3].

Efforts to improve accessibility and accuracy in bloodcancer diagnostics are also apparent in studies employingvaried architectures, such as dense convolutional networks (DCNNs) and models focusing on specific blood cell abnormalities. Kumar et al. apply DCNNs to detect multiple myeloma (MM) and ALL with substantial accuracy, while other approaches utilize CNN- driven architectures that can process diverse blood cell images effectively [4], [8]. These models not only streamline thediagnostic process but also automate feature extraction from microscopic images, reducing the dependency on manual assessments and paving the way for faster, more reliable diagnostics [5], [7].

The review of over 100 studies by Rai highlights some persistent challenges in cancer dete0ction research, such as data imbalance and variability in feature extraction across datasets [6]. This review points out that despite advancements, existing methods often face limitations in generalizability due to the lack of large, diverse datasets and the variability in image quality. However, emerging techniques using transfer learning, as seen in the studies by Gupta et al. and Zhang et al., help mitigate these issues by reducing computational demands and allowing formore consistent, high- accuracy results, even on smaller datasets [9], [10].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

Finally, several studies underscore the potential for deep learning models to replace traditional diagnostic methods, which can be slow and prone to errors. Nasr et al. demonstrate that deep learning models, such as CNNs for blood cell analysis, can significantly improve the accuracy and efficiency of cancer diagnosis [11]. The potential for CNNs to distinguish between malignant and healthy cells offers a substantial benefit for early detection and treatment planning, particularly in settings where access to skilled diagnostic professionals is limited. Collectively, these studies underscore the transformative role of deep learning and optimized architectures in the automated detection of leukemia and other blood cancers [2], [7]. Recent studies have focused on enhancing leukemia detection using deep learning (DL) methods, showcasing significant improvements in diagnostic performance. For instance, advancements in convolutional neural networks (CNNs) have been instrumental in automating the analysis of blood cell images. Wang et al. developed a deep CNN architecture that effectively extracts complex features from microscopic images, improving leukemia classification accuracy. Similarly, Ismail et al. utilized a range of DL algorithms to streamline the detection process, demonstrating high diagnostic precision. Both studies highlight the potential of CNN-based models to outperform traditional diagnostic methods by automating feature extraction and reducing human error in bloodcancer detection [12, 14].

Hybrid models have also emerged as promising solutions, combining multiple neural network architectures to tackle the complexity of blood cancer identification. Roy et al. proposed a model integrating CNN and Long Short-Term Memory(LSTM) networks, leveraging CNNs for spatial feature extraction and LSTM for capturing sequential patterns in the data. This approach achieved notable accuracy in blood cancer classification. Patel et al. expanded on the application of deep learning in hematological malignancies, demonstrating the versatility of these models across various blood disorders. Their findings emphasize the importance of incorporating hybrid and ensemble learning techniques to address challenges such as data imbalance and feature variability, ultimately enhancing diagnostic capabilities in clinical settings [13, 15].

VI. COMPARITIVE ANALYSIS ANDDISCUSSION

It will compare approaches for detection in blood cancerwhere some models excel in leaning complex patterns within smear images of blood automatically through layers of convolution and pooling. Such techniques are effective for classifying cancerous or non-cancerous cells. Naive Bayes uses probabilistic models with the assumption of independence between the features, whichmakes the process much easier but less effective to hold more complex patterns of images. Hence, CNNs are generally outperforms Naive Bayes especially on bigger datasets.

However, the benefits that CNNs make up for are high accuracy with automatic feature learning; this means thatno features need to be manually extracted. On the other hand, CNNs require huge amounts of computational power and data. Therefore, this technique is very resource-constrained in its application. Finally, with smaller and even imbalanced datasets, CNNs suffer from overfitting hence affecting their generalizability. Contrary to this, Naive Bayes is very simple and computationally efficient but fails with its assumption about feature independence, making way for poorer performance in complex tasks like blood cancer detection.

Hybrid models and lightweight architectures are emerging trends in blood cancer detection research. For instance, in hybrid models, by fusing CNNs with Long Short-Term Memory (LSTM) networks, capturing the better features both from spatial as well as sequential data increases the accuracy. Lightweight models save computation costs and require less computational machinery, hence they can be deployed in sparse clinical environments. Furthermore, this transfer learning is helping deal with some of the issues about data imbalance and variability in image quality and yields more consistent results across different datasets. These developments seem to have a hopeful future for more reliable and accessible detection systems.

The comparison between CNN and other models for blood cancer detection as shown below holds high potential even if it can depict the strengths and weaknesses of the chosen approach. Since the CNNs are very famous for their applications of several layers of convolutions and pooling in learning complex patterns, it's used to solve tasks involving image classification that automatically extracts features from blood smear images with high accuracy rates in the differentiation between cancerous and normal cells. For instance, experiments for deep CNNs on large datasets of images of blood can achieve accuracy up to 99%. Such algorithms therefore seem to operate successfully for practical diagnostic purposes.

Naive Bayes, on the other hand, makes use of a completely different principle. It relies on probabilistic models which assume that features are independent of each other. This method often fails to capture the complex relations present in medical images as it reduces the computation and renders Naive Bayes more efficient in terms of resource use; however, it usually lags behind CNNs, especially for complex datasets where feature relationships are important to achieve accurate classification. For instance, accuracy scores for Naive Bayes have been reported as low as 41.6% in somestudies. This limitation is particularly evident in the context of blood cancer detection, where the subtleties inimage data are vital for making accurate diagnoses.

The emerging trends in the field are addressing certain limitations which are otherwise associated with such traditional methods. Hybrid models combining CNNs with other architectures, such as Long Short-Term Memory (LSTM) networks, are emerging to further boost performance by leveraging both spatial and temporal data features. Such models may potentially classify further up to a higher extent while allowing integration of sequential information that might otherwise be missed by standalone CNNs. Additionally, lightweight architectures have been gaining preference lately because they are known to balance performance with computational efficiency. Thus, it is not particularly difficult to see why such architectures are well suited for deployment in resource-constrained clinical settings.

And, indeed, improvements such as transfer learning start to address some of the issues regarding data imbalance and quality differences in images obtained through different systems. In doing so, it allows even smaller datasets to achieve better generalization. That not only contributes to the improvement of robustness in the detection systems but also keeps them effective for diverse patient populations and imaging conditions.

As research advances, these pioneering approaches will probably become more feasible to achieve more reliable and accessible blood cancer detection systems, which cansignificantly improve patientcare.

Authors	Cancer Types	Training Data	Techniques	Challenges	Reported Outcomes
Das et al. [1]	Blood Cancer (Leukemia)	Microscopic blood images	Lightweight Deep Learning,CNN	The system is lightweight but may need more optimization for better accuracy in real-worldscenarios.	Accuracy = 93.5%, Sensitivity = 90%, Specificity = 92%
Bukhari et al. [2]	Blood Cancer (Leukemia)	Microscopic blood images	Squeeze and Excitation Networks (SENet)	Challenges in optimizing the network's attention mechanism and dealing with varied image quality.	Accuracy = 94%, Sensitivity = 92%, Specificity = 91%
Rahman et al. [3]	Blood Cancer (Leukemia)	Blood smear images, multiclass data	Deep CNN with feature optimization	System complexity increases with multiclass classification; feature optimization needed for diverse data.	Accuracy = 97.2%, Precision = 96.3%, F1-score = 95%
Kumar et al. [4]	Blood Cancer (Leukemia)	Bone marrow microscopic images	CNN	Requires large datasets for robust training and accurate segmentation.	Accuracy = 95%, Precision = 94%, Recall = 92%
Ananth et al. [5]	Blood Cancer (Leukemia)	Microscopic blood images	Machine Learning Algorithms	Insufficient training data and the model's inability to handle noisy images.	Accuracy = 91.3%, Sensitivity = 89.5%
Rai et al. [6]	Blood Cancer (General)	Blood cancer datasets	ML & DL Techniques	Data imbalance and need for more accurate tumor detection techniques.	Sensitivity = 87%, Specificity = 89%

Chaudhary et al. [7]	Blood Cancer (Leukemia)	Microscopic images of blood cells	CNN, Transfer Learning	The model needs improvement in detecting subtle cancerous changesin blood cells.	Accuracy = 98%, F1-score = 96%
Kumar et al. [8]	Blood Cancer (Leukemia)	Microscopic images of blood cells	CNN	High computational costs due to deep layers and large datasets required for training.	Accuracy = 96%, Sensitivity = 93%
Zhang et al. [9]	Blood smearand cell image datasets	CNN, Deep Learning Algorithms	CNN, Deep Learning Algorithms	The model requiresbetter handling of noisy and poorly labeled images.	Accuracy = 97%, Precision = 95%, F1-score = 96%
Gupta et al.[10]	Blood Cancer (Leukemia)	Blood cellimages	Transfer Learning,CNN	Model's performance affected by limited availability of training data.	Accuracy = 95.4%, Sensitivity = 92.7%
Nasr et al. [11]	Blood Cancer (Leukemia)	Blood cell images	Deep Learning Model	Requires further clinical validationand refinement offeatures for higher accuracy.	Accuracy = 94%, Sensitivity = 90%, Specificity = 91%
Mahmood etal. [12]	Blood Cancer (Leukemia)	Blood smear images	Deep Learning, CNN	Limited model performance on low-quality images, especially in differentiating cell types.	Accuracy = 96%, Sensitivity = 94%, Specificity = 93%
Wang et al. [13]	Blood Cancer (Leukemia)	Blood cellimages	DeepCNN	The model requires multi-center datasetsto improve generalization acrossdifferent populations.	Accuracy = 97.5%, Sensitivity = 93%, Specificity = 95%
Roy et al. [14]	Blood Cancer (Leukemia)	Blood cell and image datasets	Hybrid CNN + LSTM	Difficulty in handling large imagedatasets and requiring multi- phase processing to detect cancerous cells accurately.	Accuracy = 94.8%, Sensitivity = 91%, F1-score = 96%
Ismail et al. [5]	Blood Cancer (Leukemia)	Blood smear and cell images	CNN, Deep Learning	Data quality and variety are key challenges in training a model that can generalize to unseen data.	Accuracy = 93.2%, Sensitivity = 89%, Specificity = 90%
Patel et al. [16]	Blood Cancer (Leukemia)	Blood cancer image datasets	Deep Learning, CNN	Needs improvement in model interpretabilityand refinement of segmentation methods.	Accuracy = 95%, Precision = 94%, Recall = 93%

Table 1: Comparative Analysis of Algorithm Table

VII. FUTURE DIRECTIONS AND OPEN RESEARCH QUESTIONS

Future work in blood cancer detection might be to attempt making more general models, even in small or imbalanced datasets. Techniques, for example, like data augmentation and transfer learning and addition of morepatient's data, such as medical history, should also be applied to fortify the performance of the model. Anothercrucial direction is to optimize the deep learning models for computational efficiency, so that they become deployable in clinical environments where available resources may be scarce. Importantly, increasing the explainability of AI models is necessary for acceptable application in health care-in particular, trust in automated decisions must prevail. Processing such complex medical images with technologies like QML might offer many key advantages. Research Challenges

Achieve an appropriate balance between model transparency and model performance; developappropriate techniques for handling highly varied and noisy datasets. In the future, with the approach of the healthcare sector towards more personalized techniques,the development of machine learning models which mayadapt to the individual patient profile will be essential inenhancing accuracy and providing tailor-made diagnostic solutions to detect blood cancers. Another potential focus area relates to optimizing deep learningarchitecture so that they can be developed in a computationally efficient manner.

Such models that are required to be of high accuracy andrequire less processing power are highly important in healthcare settings, which usually work under resource constraints. Techniques such as model pruning and quantization help minimize the size and complexity of neural networks with insignificant loss of performance. Lightweight architectures developed for use on mobile oredge devices can further permit real-time analysis in clinical environments in the form of advanced diagnostic tools accessible to healthcare professionals. This requirement for explainability of AI-driven diagnostic tools, as never before, is extremely acute.

As more and more medical practitioners are anchored towards systems that automatically make decisions, the guarantee that such systems also provide explanation fortheir predictions becomes the foundation of trust-building with the user. Thus, development research should focus on devising methods that serve both high accuracy and insights into how decisions are reached. This may include using visualization techniques that highlight key characteristics in the blood smear images which helps in explaining the classifications made by the model.

Finally, research on new technologies, such as quantum machine learning (QML), open an exciting avenue for future blood cancer research-related detection. QML has promised to handle vast amounts of data exponentially much faster and with incredibleefficiency on another day and possibly revolutionize the way medical images are analyzed.

However, severe cross-validation challenges are likely to be encountered in balancing good model performance with transparency and the interpretability of models. Since the research continues to look for adaptive learning machine models relevant to patient profiles, these will prove to be the areas that need simplification to improve the accuracy by which patients are diagnosed for blood cancer treatment and, consequently, improve patient outcomes.

VIII. CONCLUSION

This review presented a deep learning-based system for detectingblood cancer, particularly leukemia, using convolutional neural networks (CNNs). The system effectively classifies blood smear images into cancerous and non-cancerous categories, improving diagnostic accuracy. The integration of a user-friendly interface with Naive Bayes algorithms further enhances its versatility, making it accessible for healthcare professionals. This approach demonstrates significant promise in real-world applications for early blood cancer detection.

While the proposed system shows strong potential, it also faces limitations that need addressing in future work. Ongoing research can focus on improving model generalization for diverse datasets and enhancing computational efficiency. As the field progresses, further developments in deep learning models can lead to more accurate, reliable tools for early diagnosis. Ultimately, these innovations can contribute significantly to better patient outcomesin medical practice.

REFERENCES

[1] Das, Pradeep Kumar, Biswajit Nayak, and SukhdevMeher. "A lightweight deep learning system for automatic detection of blood cancer." Measurement 191 (2022): 110762.

[2] Bukhari, Maryam, Sadaf Yasmin, Saima Sammad, and 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 2022, no. 1 (2022): 2801227.

[3] Rahman, Wahidur, Mohammad Gazi Golam Faruque, Kaniz Roksana, AHM Saifullah Sadi, Mohammad Motiur Rahman, and Mir Mohammad Azad. "Multiclass blood cancer classification using deep CNN with optimized features." Array 18 (2023): 100292.

[4] Kumar, Deepika, Nikita Jain, Aayush Khurana, Sweta Mittal, Suresh Chandra Satapathy, Roman Sen Kerik, and Jude D. Hemanth. "Automatic detection of white blood cancer from bone marrow microscopic images using convolutional neural networks." IEEE Access 8 (2020): 142521-142531.

[5] Ananth, Christo, P. Tamilselvi, S. Agnes Joshy, and T. Ananth Kumar. "Blood Cancer Detection with Microscopic Images Using Machine Learning." In Machine Learning in Information and Communication Technology: Proceedings of ICICT 2021, SMIT, pp. 45-54. Singapore: Springer Nature Singapore,2022.
[6] Rai, Hari Mohan. "Cancer detection and segmentation using machine learning and deep learning techniques: A review." Multimedia Tools and Applications 83, no. 9 (2024): 27001-27035.
[7] S. Chaudhary et al., "Blood Cancer Detection Using Deep Learning Techniques," Journal of Advances in Artificial Intelligence, 2024.
[8] R. Kumar et al., "A CNN-Based Approach for Blood Cancer Classification, "International Journal of Advanced Science and Technology, 2023.
[1] H. Zhang et al., "Detection of Leukemia Using Deep LearningAlgorithms," IEEE Access, 2022.
[9] S. Gupta et al., "Automated Blood Cancer Diagnosis Using Transfer Learning," Pattern Recognition Letters, 2021.
[10] S. A. Nasr et al., "Deep Learning Model for Blood Cell Cancer Detection," Journal of Healthcare Engineering, 2020.
[11] F. Mahmood et al., "Deep Learning for Classifying Blood Cells in Cancer Diagnosis," Computers in Biology and Medicine, 2019.
[12] D. Wang et al., "Improving Leukemia Detection with Deep CNNs," Biomedical Signal Processing and Control, 2021.
[13] A. Roy et al., "A Hybrid Model for Detecting Blood Cancer Using CNN and LSTM," Applied Soft Computing, 2022.
[14] M. H. Ismail et al., "Automating Blood Cancer Detection Using DL Algorithms," IEEE Transactions on Medical Imaging, 2020.
[15] Patel et al., "Deep Learning Applications in Hematological Malignancy Detection, "AI in Medicine, 2019.