Early Diagnosis of Acute Lymphoblastic Leukemia in Bone Marrow Histopathological Images through Deep Learning

Sunitha N V
Assistant Professor
Department of CSE
MITE, Moodabidri

Dr. Pradeep B S
Professor
Department of CSE
MITE, Moodabidri

Abstract: The detection and classification of white blood cell (WBC) cancers, such as leukemia, through the analysis of histopathological images are crucial for enabling early diagnosis and effective treatment planning. Traditional diagnostic methods largely depend on the manual interpretation of these images by pathologists. This manual process is often labor-intensive and susceptible to human error, which can impact the accuracy of diagnoses. To address these challenges, our research utilizes advanced deep learning techniques, specifically convolutional neural networks (CNNs), to improve both the accuracy and efficiency of diagnosing WBC cancers from histopathology images. We have developed and rigorously tested CNN models, including MobileNet, Xception, and InceptionV3, using a comprehensive dataset consisting of bone marrow images. Subsequently, an ensemble model was created to harness the strengths of each individual CNN architecture. This ensemble approach successfully achieved an impressive accuracy rate of 96.5%.

Keywords - White Blood Cell (WBC) cancers, Acute Lymphoblastic Leukemia, MobileNet, Xception, InceptionV3, Histopathology, Convolutional Neural Networks (CNNs), Deep Learning

1. INTRODUCTION

Oncology faces substantial challenges in diagnosing white blood cell (WBC) cancers, including their diverse subtypes, variable clinical presentations, and complex diagnostic processes. The timely and precise identification of these cancers is vital for improving patient outcomes and guiding appropriate treatment strategies. Histopathological examination of blood smears is a cornerstone in diagnosing WBC cancers, offering crucial insights into cell distribution and morphology. However, the manual interpretation of these histological images by pathologists is often laborious and prone to variability, leading to potential diagnostic errors and delays in initiating treatment.

Recent advancements in deep learning, particularly through convolutional neural networks (CNNs), have revolutionized medical image analysis. These sophisticated algorithms automate the processes of feature extraction and classification, noticeably improving the accuracy, efficiency, and consistency of cancer diagnoses. By training on extensive annotated histopathology datasets, deep learning models can identify subtle changes and complex patterns indicative of various cancer subtypes, providing superior precision and reliability compared to traditional diagnostic methods.

Our research focuses on developing a robust deep learning system designed to automatically recognize and classify WBC cancers from histopathological images. The goal is to address the limitations of manual interpretation and deliver a scalable, dependable diagnostic solution through CNNs. We rigorously assess our model's effectiveness and applicability through thorough testing, which includes systematic training and validation across diverse datasets.

Specifically, this research targets the creation of a reliable deep learning system for diagnosing acute lymphoblastic leukemia (ALL) from bone marrow histopathological images. The complexity of these images—characterized by various cell types, features, and abnormalities—makes it particularly challenging to differentiate between benign and malignant cells. The dependence on manual diagnosis by skilled pathologists highlights the pressing need for automated solutions. Our approach harnesses advanced computational techniques, especially deep learning algorithms, to deliver accurate and consistent identification and classification of leukemia cells, thereby addressing the urgent need for more efficient and effective diagnostic tools in oncology.

2. RELATED WORK

In a study presented by Zakir Ullah et al. [1], the necessity for quicker and less invasive methods to detect acute lymphoblastic leukemia (ALL), a common bone marrow form of the illness, is discussed. The pain and delay associated with traditional procedures, like bone marrow and blood examinations, is driving the need for alternatives. A CNN-based approach is proposed to extract improved deep features from medical pictures by combining an attention module (ECA) with the VGG16 architecture. Augmentation techniques improve the amount and caliber of training data. Using the C-NMC dataset, the CNN model distinguishes between normal and malignant cells with 91.1% accuracy. The study identifies the flaws in order to increase the precision of ALL diagnosis.

Prellberg et al. [2] presented a study that addresses the challenge of diagnosing leukemia using blood microscopic images, particularly in settings where flow cytometry is not available. To assist in this diagnostic process, the authors propose an automated classification system leveraging a ResNeXt convolutional neural network enhanced with Squeeze-and-Excitation modules. This method was evaluated in the C-NMC online challenge, where it achieved a weighted F1-score of 88.91% on the test set, demonstrating its effectiveness. The system successfully distinguishes between normal B-lymphoid precursors and malignant B-lymphoblasts. The authors also suggest that future research could benefit from access to raw, unprocessed images, which would enable broader application and refinement of classification techniques beyond the current preprocessing pipeline.

Kumar Deepika et al. [3] using photographs of patients with multiple myeloma (MM) and acute lymphoblastic leukemia (B-ALL), proposed a methodology for image classification using

convolution, max-pooling, and fully connected layers. A Chi-square test was used for feature selection in order to find pertinent features for categorization. Through the use of precision-recall curve analysis, the model was evaluated, demonstrating how well Random Forests categorize cancer cell images. When used to identify cancer from bone marrow images, the deep convolutional neural network (DCNN) model produced results with a high accuracy rate of 97.2%. With their backgrounds in computer science and deep learning, authors Sweta Mittal and Suresh Chandra Satapathy made a substantial contribution to this important development in image processing-based cancer detection.

Manna Ankur et al. [4] documented ensemble technique that combines the predictions made by different base learners using an advanced fusion mechanism. This approach seeks to improve cervical cancer screening accuracy by using the class with the lowest fused score as the final prognosis. The ensemble model outperforms the individual base learners in terms of performance, using the confidence scores of each learner to rank predictions and produce accurate classifications. Although preprocessing measures could be required to guarantee accurate picture analysis, the ensemble model demonstrates robustness on a variety of datasets. By comparing the suggested ensemble technique with well-known fusion schemes and CNN models, the paper highlights how effective it is in improving classification accuracy and developing the field.

Chen et al. [5] devdeloped an AI model for the classification of acute lymphoblastic leukemia (ALL) in microscopic pictures using a Resnet101 ensemble. 1,800 early test photos and over 10,000 training images were taken from the CNMC dataset. The model outperformed individual models, achieving an 85.11% accuracy and an 88.94 F1-score. A three-level L9(34) OA was used to optimize algorithm hyperparameters such as Optimizer, MiniBatchSize, MaxEpochs, and InitialLearnRate. The precision of the model is essential for accurately detecting ALL, which helps with effective screening of blood microscopic pictures and may enhance diagnostic procedures.

Shahiq et al. [6] proposed the use of a pretrained deep convolutional neural network (AlexNet) for the automated identification and classification of acute lymphoblastic leukemia (ALL) and its subtypes. In the literature, comparative analysis performs better than other techniques with a high degree of accuracy and doesn't require microscopic picture segmentation. The use of microscopic blood image analysis in computer-aided leukemia diagnosis procedures has been shown to be more accurate and efficient than manual alternatives, which are also known to be costly and time-consuming. The article also discusses data augmentation and the gathering of additional microscopic blood photographs from Google that have been validated by a specialist oncologist in order to enhance the dataset. Future efforts to further improve the diagnosis system will incorporate deep learning with larger image datasets.

C.Mondal et al. [7] explores the automated use of deep Convolutional Neural Networks (CNNs) for the detection of acute lymphoblastic leukemia (ALL) from images of microscopic cells. By utilizing transfer learning from pretrained models like VGG-16 and Xception, a

weighted ensemble of various deep CNNs can enhance the performance of the ALL cell classifier. It discusses how input resolutions, image pre-processing techniques, and crucial preprocessing steps affect training different CNN architectures to ensure improved ALL prognosis. Apart from assessing machine learning (ML) techniques for ALL analysis and investigating the applicability of deep learning methods for image-based ALL prognoses, a comparative study with previous studies on ALL detection and classification is conducted.

Ying Liu et al. [8] focused on applying deep learning algorithms to analyze acute lymphoblastic leukemia (ALL) cells in microscopic images. To address problems such minor visual changes and a lack of training data, the authors propose an improved bagging ensemble learning technique employing augmented images. They employ a range of preprocessing techniques, including cell segmentation and stain normalization, along with an initial test set consisting of thirty patients. A modified bagging ensemble training technique is used to address the imbalanced dataset. The results indicate possible efficacy, with weighted F1-scores of 0.84 and 0.88 for the preliminary and final test sets, respectively. The paper's conclusion highlights the need for additional training samples, blood cell detection inclusion, and other criteria in order to verify DL-based algorithms in an actual clinical context.

The current research has established a solid foundation for the automated use of deep learning-based methods for the identification of leukemia and other associated cancers. Even though these techniques have produced encouraging results, there is still room for improvement in terms of diagnostic accuracy through model development and access to a variety of datasets. In order to overcome the shortcomings of existing approaches, the suggested work aims to improve the early detection of Acute Lymphoblastic Leukemia in bone marrow histopathology pictures by utilizing a hybrid strategy that combines Xception, MobileNet, and Inception V3 models.

3. PROPOSED METHOD

The procedure began with the laborious curation of a diverse dataset that included histology images of white blood cell cancers. Subsequently, preprocessing techniques like resizing, splitting, balancing and augmentation are employed to enhance image quality and standardize characteristics, hence facilitating efficient analysis. The next step involves the development and optimization of the three distinct convolutional neural network (CNN) models: MobileNet, Xception, and InceptionV3. These models use transfer learning to initialize pre-trained weights, accelerating training and enhancing performance. A careful watch was maintained on the models' progress during training, and any necessary adjustments to the hyperparameters were performed to optimize performance.

The post-training performance of every CNN model is then meticulously evaluated using established metrics like accuracy, precision, recall, and F1 score. In order to assess generalization abilities and identify issues like as overfitting or underfitting, the dataset was split into separate sets for testing, validation, and training. Comprehensive comparative studies were also conducted to evaluate the benefits and drawbacks of each CNN design,

accounting for factors such as interpretability, classification accuracy, and computing efficiency. This investigation informed our choice of the model that would be most useful for real-world use in clinical settings. By employing this rigorous approach, our aim as shown in Fig 3.1 was to develop a dependable and effective system for identifying white blood cell cancer, which would ultimately improve patient outcomes in the oncology profession.

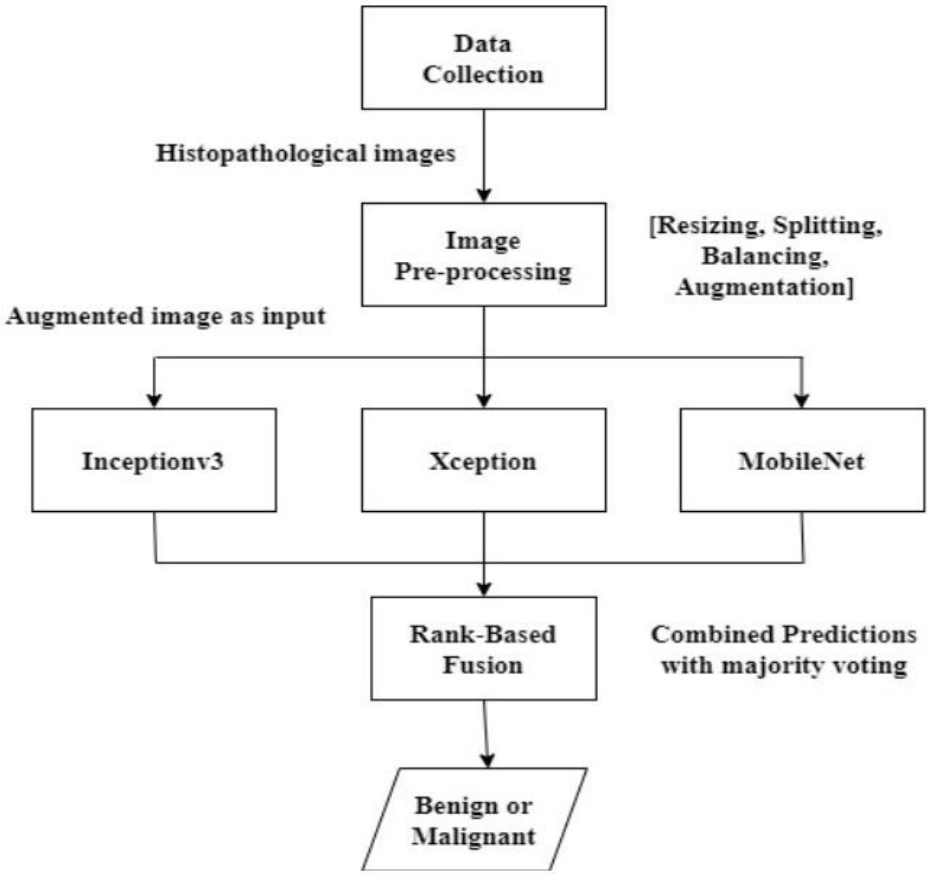
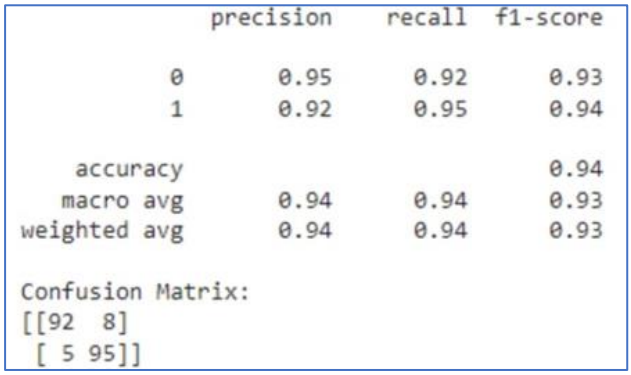


Fig 3.1: Architecture/Design

Our white blood cell cancer detection system features an architecture designed to seamlessly integrate computational components with user interaction. The system includes a web-based interface for easy classification result verification and image upload related to histopathology. On the server side, deep learning techniques perform image processing and categorization tasks, distinguishing between benign and malignant anomalies. Additionally, the system incorporates robust data management modules for efficient preprocessing, retrieval, and storage of histopathology images. This architecture has been meticulously crafted to meet the demands of real healthcare environments, delivering accurate, practical, and scalable solutions for white blood cell cancer detection.

4. RESULT AND DISCUSSION

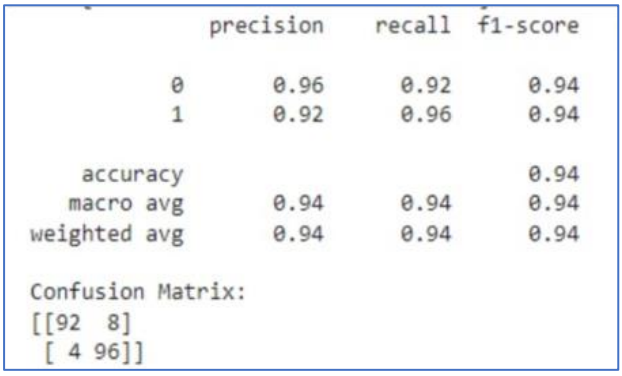
In this study, three deep learning models—MobileNet, Inception V3, and Xception—are employed to classify bone marrow histopathological images into benign and malignant categories. The dataset used is sourced from Kaggle. The models are evaluated based on accuracy, precision, recall, F1 score, and confusion matrices.



	precision	recall	f1-score
0	0.95	0.92	0.93
1	0.92	0.95	0.94
accuracy			0.94
macro avg	0.94	0.94	0.93
weighted avg	0.94	0.94	0.93

Confusion Matrix:
[[92 8]
[5 95]]

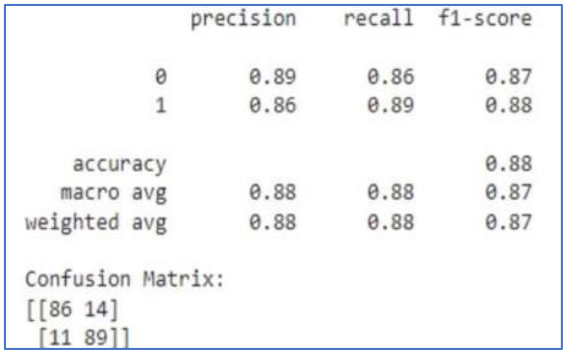
(a) Mobilenet



	precision	recall	f1-score
0	0.96	0.92	0.94
1	0.92	0.96	0.94
accuracy			0.94
macro avg	0.94	0.94	0.94
weighted avg	0.94	0.94	0.94

Confusion Matrix:
[[92 8]
[4 96]]

(b) Xception



	precision	recall	f1-score
0	0.89	0.86	0.87
1	0.86	0.89	0.88
accuracy			0.88
macro avg	0.88	0.88	0.87
weighted avg	0.88	0.88	0.87

Confusion Matrix:
[[86 14]
[11 89]]

(c) Inception v3

Fig 5.1: Classification performance of three models

The results obtained from the individual models, MobileNet, Inception V3, and Xception, as shown in Fig 5.1, demonstrate their effectiveness in accurately classifying bone marrow histopathological images as benign or malignant. Mobilenet and Xception demonstrate very similar performance metrics with both accuracy and F1-scores around 0.94, which indicates high efficiency in classifying the medical images accurately. Their precision, recall, and F1-scores are closely aligned, suggesting that these models are reliable for balanced datasets where the proportion of positive to negative classes is similar. Mobilenet shows a high precision, which means it has a lower rate of false positives. This could be particularly useful in scenarios where falsely identifying a condition as present could lead to unnecessary treatments. Xception provides the best balance between recall and precision, making it potentially the most robust choice for medical image analysis where both missing a condition (low recall) and incorrectly identifying a condition (low precision) have serious implications. Inception v3, shows a less performance in all metrics compared to the other two models but it is still efficient. The lower scores across the board might make it less suitable for critical applications but could be useful in preliminary screening phases where speed is more critical than absolute accuracy.

Additionally, in the proposed work a majority voting technique is used to find final output, resulting in the accuracy of 96.5% showing enhanced ability to correctly classify benign and malignant cases, thereby reducing the likelihood of diagnostic errors.

5. CONCLUSION AND FUTURE WORK

This study demonstrates that deep learning, particularly through ensemble methods, can significantly improve the accuracy and reliability of early detection of Acute Lymphoblastic Leukemia. The results indicate that such methods could potentially be integrated into clinical practice, providing support for pathologists and contributing to more accurate and timely diagnoses. Future work could explore the application of these models to larger and more diverse datasets, as well as the integration of additional features, such as patient demographics or genetic data, to further improve model performance.

REFERENCES

[1] Zakir Ullah, M., Zheng, Y., Song, J., Aslam, S., Xu, C., Kiazolu, G. D., & Wang, L. (2021). An attention-based convolutional neural network for acute lymphoblastic leukemia classification. Applied Sciences, 11(22), 10662. https://doi.org/10.3390/app112210662

[2] Prellberg, J., & Kramer, O. (2019). Acute lymphoblastic leukemia classification from microscopic images using convolutional neural networks. Lecture Notes in Bioengineering, 53–61. https://doi.org/10.1007/978-981-15-0798-4_6

[3] Kumar, D., Jain, N., Khurana, A., Mittal, S., Satapathy, S. C., Senkerik, R., & Hemanth, J. D. (2020). Automatic detection of white blood cancer from bone marrow microscopic images using convolutional neural networks. IEEE Access, 8, 142521–142531. https://doi.org/10.1109/access.2020.3012292

[4] Manna, A., Kundu, R., Kaplun, D., Sinitca, A., & Sarkar, R. (2021). A fuzzy rank-based ensemble of CNN models for classification of cervical cytology. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-93783-8

[5] Chen, Y.-M., Chou, F.-I., Ho, W.-H., & Tsai, J.-T. (2021). Classifying microscopic images as acute lymphoblastic leukemia by Resnet Ensemble Model and taguchi method. BMC Bioinformatics, 22(S5). https://doi.org/10.1186/s12859-022-04558-5

[6] Shafique, S., & Tehsin, S. (2018). Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks. Technology in Cancer Research & Treatment, 17, 153303381880278. https://doi.org/10.1177/1533033818802789

[7] Mondal, C., Hasan, Md. K., Ahmad, M., Awal, Md. A., Jawad, Md. T., Dutta, A., Islam, Md. R., & Moni, M. A. (2021). Ensemble of Convolutional Neural Networks to diagnose acute lymphoblastic leukemia from microscopic images. Informatics in Medicine Unlocked, 27, 100794. https://doi.org/10.1016/j.imu.2021.100794

[8] Liu, Y., & Long, F. (2019). Acute Lymphoblastic Leukemia Cells Image Analysis with Deep Bagging Ensemble Learning. https://doi.org/10.1101/580852