Pertanika Proceedings 1 (5): 93 - 99 (2025)

PERTANIKA JOURNALS

PERTANIKA PROCEEDINGS

Journal homepage: http://www.pertanika.upm.edu.my/

Comparison of Deep Learning Models for Classification of Acute Lymphoblastic Leukemia

Pranavesh Kumar Talupuri[1], Nitin Bhimineni[1], Beebi Naseeba[1], E. R. Aruna[2], Sai Vennela Tipirneni[3]and Nagendra Panini Challa[1]*

[1]School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Amaravati 522241, India

[2]Vardhaman College of Engineering, Hyderabad 501218, Telangana, India

[3]Bachelor of Science in data Science, San Jose State University, San Jose, California, CA 95192, United States

ABSTRACT

The results of a comprehensive comparative performance analysis of several deep learning models, which are recognized in the classification of single-cell red-blood cell (RBC) images as healthy and Acute Lymphoblastic Leukemia (ALL) categories, are presented here. This study investigates how efficiently Convolutional Neural Networks (CNNs) could be used, including variations from ResNet, Incep- tion, VGG, among others. The study goes on to use an analysis methodology that is robust because it includes cross-validation during the evaluation of accuracy for F1-score. The findings reflect the strengths and weaknesses pertaining to each deep learning model, hence serving to enhance the development of future diagnosis tools for ALL. Findings are discussed regarding clinical applications in real life that are relevant to this analysis and possible future directions along which research in the same field of medical image classification may proceed. This present study aims to enhance concerted efforts toward improving the diagnostic sensitivity of leukemia.

Keywords: Acute lymphoblastic leukemia, computer vision, classification, deep learning

ARTICLE INFO

Article history:
Received: 29 August 2025
Published: 31 October 2025

DOI: https://doi.org/10.47836/pp.1.5.018

E-mail addresses:
pranavesh.talupuri@gmail.com (Pranavesh Kumar Talupuri)
bhimineni2020@gmail.com (Nitin Bhimineni)
beebi.naseeba@vitap.ac.in (Beebi Naseeba)
er.aruna@vardhaman.org (E. R. Aruna)
saivennela.tipirneni@sjsu.edu (Sai Vennela Tipirneni)
nagendra.challa@vitap.ac.in (Nagendra Panini Challa)
* Corresponding author

INTRODUCTION

Acute Lymphoblastic Leukemia (ALL) is a heterogeneous group of hematological malignancies characterized by aberrant proliferation of lymphoblasts. Accurate diagnosis of ALL is indispensable for the application of appropriate treatment strategies. Traditional diagnostic techniques, such as microscopic examination of blood

smears, are prone to inter-observer variability and would likely benefit from automated or semi-automated approaches.

Deep learning, especially Convolutional Neural Networks (CNNs), has shown promise in improving the diagnostic accuracy of similar image-based classification tasks. The aim of this study is to compare the performance of different CNN architectures for classifying the red blood cells based on whether they are afflicted by ALL.

BACKGROUND AND MOTIVATION

Importance of ALL Detection

Early detection of ALL is very important for improving patient outcomes. Timely diagnosis allows the initiation of targeted therapies that can greatly improve survival rates and the effectiveness of treatment. Delays in diagnosis may allow disease progression, making treatment more complicated and less likely to result in favorable outcomes. Hence, developing efficient and accurate diagnostic tools is a pressing concern in hematology.

Traditional approaches, including manual examination of blood smears and immunophenotyping, are effective but labor-intensive and variable among pathologists. These traditional techniques require specialized expertise, which may not be available in many resource-constrained settings. Reliance on manual interpretations could also reinforce inconsistencies, making it even more necessary for automated solutions in standardizing workflows for diagnosis.

Yet, automated methods using artificial intelligence have been shown to help surmount such limitations. However, reducing diagnostic variability while increasing reproducibility with AI-driven tools standardizes practices that otherwise would impede the adoption of new diagnostics, making them increasingly scalable and robust in real-world settings due to efficient operation on many divergent data sets.

Applications of Deep Learning

Deep learning models, through their ability to capture complex patterns and high-dimensional relationships within data, are well-suited for medical image analysis. In the context of ALL detection, these models can analyze subtle morphological differences that may be challenging for human observers to identify.

Convolutional Neural Networks (CNNs), in particular, excel at feature extraction by automatically learning hierarchical representations from raw input data. This ability makes them ideal candidates for medical image classification tasks, including identifying abnormalities in blood smears. Their application has demonstrated remarkable success in minimizing false negatives and positives, thereby improving diagnostic precision. Deep learning approaches also facilitate early intervention and personalized treatment planning.

Models trained on large datasets can generalize effectively, providing accurate predictions even for rare cases.

In addition, advancements in transfer learning and pre-trained architecture allow models to be fine-tuned for specific tasks. This adaptability makes deep learning-based systems suitable for deployment in diverse clinical environments, including low-resource settings. With continued advancements, these approaches are expected to bridge gaps in diagnostic accessibility and efficiency.

Related Work

Several studies have explored deep learning techniques for leukemia detection. Das et al. (2022) reviewed methodologies for ALL classification, highlighting the effectiveness of CNNs. Ain et al. (2022) demonstrated superior performance using pre-trained architecture such as ResNet-34. Magpantay et al. (2022) employed YOLOv3, achieving high accuracy in cell classification.

Yadav (2021) proposed a feature-fusion-based approach for blood smear classification, reporting an accuracy of 99.3%. Ramagiri et al. (2023) evaluated machine learning and deep learning techniques to identify algorithms with optimal performance for leukemia detection. Their findings emphasized the superiority of CNNs in capturing morphological variations.

Kim (2022) investigated hybrid and ensemble deep learning models for leukemia diagnostics. ResNet50 achieved 84% accuracy, while ensemble models demonstrated enhanced performance with 86%. This highlights the potential of combining multiple architectures for improved results.

Jiwani et al. (2023) applied computational deep learning methods for leukemia pattern recognition, achieving high accuracy across different leukemia types. Aftab et al. (2021) utilized transfer learning approaches and reported high accuracy and efficiency, showcasing the feasibility of deploying AI-driven frameworks in clinical applications. Babaso et al. (2020) conducted a comparative analysis of machine learning algorithms, providing insights into their relative performance in leukemia subtype classification. Ullah (2021) addressed challenges associated with limited datasets by proposing an ensemble learning approach that utilized augmented images, achieving promising results.

Al-Khuzaie et al. (2023) employed VGG19 models to detect ALL using trans- fer learning, achieving an accuracy of 99.49%. Similarly, Hasan et al. (2020) utilized DenseNet-201 to classify leukemia subtypes with an accuracy of 99.56%, demonstrating the potential of deep feature extraction techniques.

Das and Meher (2021) introduced SqueezeNet-based transfer learning models, effectively distinguishing malignant and benign cases. Francese et al. (2022) combined genetic algorithms and autoencoders, achieving 92% accuracy in leukemia classification.

METHODOLOGY

The dataset used in this research is that of blood smear images, representing both healthy and leukemia-affected cells. The dataset has been very carefully curated to span a large variety of morphological variations to ensure robustness in model evaluation. High-resolution images are used in the dataset so that the models can analyze fine-grained features and tiny differences between healthy and leukemia-affected cells. In addition, the balanced sample representation ensures that the models have sufficient variations to reduce the overfitting problem and increase generalization.

More preprocessing was done on the dataset in the form of normalization, resizing, and contrast enhancement to make the images standardized. That improved the consistency and quality of input data so that models can focus on feature learning instead of compensating for inconsistencies in the dataset. Moreover, this pre-processing pipeline was designed in a manner simulating real-world diagnostic conditions to make the models perform effectively under clinical applications.

The models considered in this study include ResNet, Inception, Xception, MobileNet, EfficientNet, and VGG, selected for their proven effectiveness in image classification tasks. Each model was initialized with pre-trained weights to leverage transfer learning and fine-tuned using gradient descent optimization techniques. The hyperparameters, including learning rate, batch size, and dropout rates, were systematically tuned to enhance performance.

To improve model robustness, data augmentation techniques such as rotation, flipping, scaling, and random cropping were applied during training. These augmentations simulated variations in cell orientation and morphology, enhancing the models' ability to generalize across unseen samples. Furthermore, early stopping and learning rate scheduling were implemented to prevent overfitting and optimize convergence.

Regularization methods, including dropout layers and L2 weight decay, were integrated to improve generalization and stability. Each model's performance was validated using 10-fold cross-validation, ensuring reliability and minimizing variance in performance metrics. The training and validation process was conducted on high-performance GPUs to expedite computation and enable deeper architectures to process larger datasets efficiently.

RESULTS AND DISCUSSION

The experimental results reveal the performance of different architectures. The EfficientNetB3 model demonstrated the highest accuracy at 96.43750%, closely followed by EfficientNetB4 and EfficientNetB5, both achieving accuracies above 95%. These outcomes show the effectiveness of the EfficientNet family in the context of leukemia detection. In terms of F1 score, EfficientNetB3 also led with a score of 0.96416, reflecting its precision in both sensitivity and specificity. ResNet50, InceptionNetV3, and Xception closely

followed with F1 scores exceeding 0.95. Conversely, VGG16 demonstrated a lower F1 score of 0.81873, suggesting potential limitations in its ability to discern leukemia patterns. The visual representations through bar plots provide a more intuitive understanding of the comparative performance of various architectures in leukemia detection. Figure 1 illustrates the accuracy comparison, while Figure 2 details the F1-score comparison.

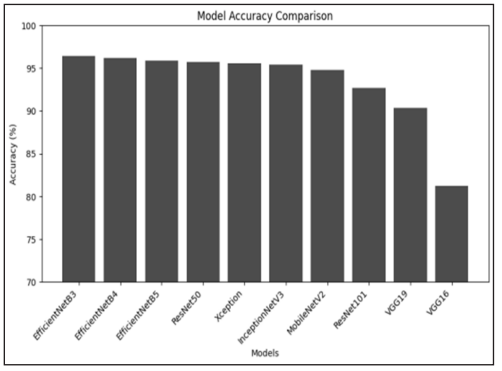

Figure 1. Accuracy

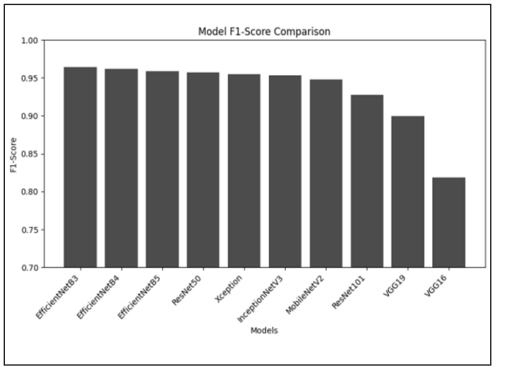

Figure 2. F1-Score

The bar plot once again shows the superior performance of the EfficientNet models, which outperforms other models (Figure 3).

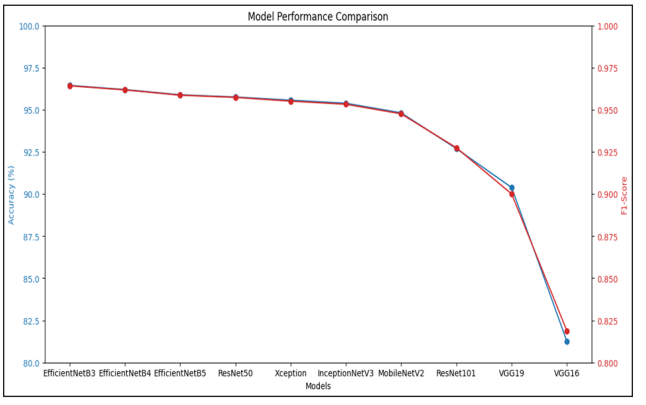

Figure 3. Metrics comparison

CONCLUSIONS

In conclusion, this research aims to shed light on the complex nature of the diagnosis of Acute Lymphoblastic Leukemia (ALL) using advanced computational methodologies. The exploration of diverse convolutional neural network (CNN) architectures, with their unique attributes, has provided valuable insights into their suitability in differentiating

between healthy and leukemia cells within blood smear images. The scope for medical image analysis, as seen from the outcomes of this research, indicates the potential of using DL and CNNs in aiding traditional diagnostic approaches. Integrating automated diagnostic tools into clinical workflows presents a promising approach for standardization and efficacy in clinical settings, with the capacity to alleviate challenges associated with manual examination.

ACKNOWLEDGEMENT

The authors sincerely thank their institutions, colleagues, and the editorial team for their support and valuable feedback. They are also grateful to their families and friends for their encouragement throughout this work.

REFERENCES

Aftab, M. O., Awan, M. J., Khalid, S., Javed, R., & Shabir, H. (2021). Executing spark BigDL for leukemia detection from microscopic images using transfer learning. In *1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)* (pp. 216-220). IEEE. https://doi.org/10.1109/CAIDA51941.2021.9425264

Ain, Q. U., Akbar, S., Hassan, S. A., & Naaqvi, Z. (2022, May). Diagnosis of leukemia disease through deep learning using microscopic images. In *2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICoDT255437.2022.9787449

Al-khuzaie, M. Y., Zearah, S. A., & Mohammed, N. J. (2023). Developing an efficient VGG19-based model and transfer learning for detecting acute lymphoblastic leukemia (ALL). In *5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-5). IEEE. https://doi.org/10.1109/HORA58378.2023.10156679

Babaso, S. P., Mishra, S. K., & Junnarkar, A. (2020). Leukemia diagnosis based on machine learning algorithms. In *IEEE international conference for innovation in technology (INOCON)* (pp. 1-5). IEEE. https://doi.org/10.1109/INOCON50539.2020.9298321

Das, P. K., & Meher, S. (2021). Transfer learning-based automatic detection of acute lymphocytic leukemia. In *National Conference on Communications (NCC)* (pp. 1-6). IEEE. https://doi.org/10.1109/NCC52529.2021.9530010

Das, P. K., Diya, V. A., Maher, S., Panda, R., & Abraham, A. (2022). A systematic review on recent advancements in deep and machine learning based detection and classification of acute lymphoblastic leukemia. *IEEE Access 10*, 81741 – 81763. https://doi.org/10.1109/ACCESS.2022.3196037

Hasan, M. M., Srizon, A. Y., Sayeed, A., & Hasan, M. A. M. (2020). Accurate recognition of leukemia subtypes by utilizing a transfer learned deep convolutional neural network. In *11th International Conference on Electrical and Computer Engineering (ICECE)* (pp. 427-430). IEEE. https://doi.org/10.1109/ICECE51571.2020.9393031

Jiwani, N., Gupta, K., Pau, G., & Alibakhshikenari, M. (2023). Pattern recognition of acute lymphoblastic leukemia (ALL) using computational deep learning. *IEEE Access*, *11*, 29541-29553. https://doi.org/10.1109/ACCESS.2023.3260065

Kim, S. (2022). Developing a leukemia diagnostic system based on hybrid and ensemble deep learning architectures. In *2022 International Communication Engineering and Cloud Computing Conference (CECCC)* (pp. 70-73). IEEE. https://doi.org/10.1109/CECCC56460.2022.10069094

Magpantay, L. D. C., Alon, H. D., Austria, Y. D., Melegrito, M. P., & Fernando, G. J. O. (2022). A transfer learning-based deep cnn approach for classification and diagnosis of acute lymphocytic leukemia cells. In *International Conference on Decision Aid Sciences and Applications (DASA)* (pp. 23-25). IEEE. https://doi.org/10.1109/DASA54658.2022.9765000

Ramagiri, A., Jahnavi, V., Gottipati, S., Monica, C., Afrin, S., Jyothi, B., & Chinnaiyan, R. (2023). Image classification for optimized prediction of leukemia cancer cells using machine learning and deep learning techniques. In *International Conference on Innovative Data Communication Technologies and Application (ICIDCA)* (pp. 193-197). IEEE.

Ullah, A. (2021). Image analysis of cells acute lymphoblastic leukemia using ensemble learning of deep bagging. In *IEEE 7th International Conference on Computing, Engineering and Design (ICCED)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICCED53389.2021.9664867

Yadav, D. P. (2021). Feature fusion based deep learning method for leukemia cell classification. In *5th International Conference on Information Systems and Computer Networks (ISCON)* (pp. 1-4). IEEE. https://doi.org/10.1109/ISCON52037.2021.9702440