International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Comparative Evaluation of CNN Architectures for Leukemia Classification with an Improved EfficientNet-B0

Saloni Jain*[1], Dr. Rajesh Kumar Nagar [2]

Submitted: 02/11/2024 **Revised:** 16/12/2024 **Accepted:** 24/12/2024

Abstract*:* Leukemia is a life-threatening blood cancer that requires early and accurate diagnosis for effective treatment. Traditional diagnostic methods are time-consuming and subject to human error. In this study, we evaluate the performance of multiple Convolutional Neural Network (CNN) architectures including VGG-16, VGG-19, ResNet-50, ResNet-152, Xception, EfficientNet-B0, and a Proposed EfficientNet-B0 for the detection and classification of leukemia using microscopic blood smear images. The models were trained and tested on a public dataset to classify images as Healthy or Cancer. Performance was assessed using Accuracy, Precision, Recall, and F1-Score. The Proposed EfficientNet-B0 model outperformed all baseline architectures, achieving 98% across all evaluation metrics, marking a significant improvement of +2.78% in accuracy compared to the best-performing standard model.

Keywords: *Leukemia Detection, CNN, Deep Learning, Blood Smear Images, EfficientNet-B0, Image Classification, Medical Imaging, Performance Evaluation.*

1. Introduction

Leukemia is a type of cancer that affects the blood and bone marrow, characterized by the rapid production of abnormal white blood cells. Early and accurate detection of leukemia plays a pivotal role in the successful management and treatment of the disease. In clinical settings, leukemia diagnosis often relies on microscopic examination of blood smear images by expert hematologists, a process that can be time-consuming, labor-intensive, and prone to inter-observer variability. With the growing volume of medical data and the advancement of computational tools, automated diagnostic solutions are increasingly being explored to support clinical decision-making. In particular, deep learning models specifically Convolutional Neural Networks (CNNs) have shown remarkable performance in image-based disease classification tasks due to their ability to learn complex patterns and representations from data. This research investigates the application and comparative performance of multiple CNN architectures for the task of leukemia detection and classification using publicly available microscopic blood smear images.

The need for automated and accurate leukemia detection is more critical than ever, particularly in under-resourced regions where access to expert pathologists is limited. While several studies have applied CNNs to leukemia diagnosis, few have conducted a holistic evaluation of multiple standard architectures in a consistent experimental setup. Most existing models either rely on limited datasets or focus on a single architecture without analyzing trade-offs in performance, computational complexity, and generalizability. Furthermore, medical imaging data often suffers from imbalance and noise, which challenges the learning capability of basic CNN models. To address these gaps, this study is motivated by the need to benchmark widely used CNN architectures such as VGG-16, VGG-19, ResNet-50, ResNet-152, Xception, and EfficientNet-B0 on the same dataset under identical preprocessing and training conditions. The goal is to understand how each model performs on various evaluation metrics and to propose improvements that can enhance classification accuracy without significantly increasing computational overhead.

This study makes several key contributions. First, it presents a rigorous comparative analysis of seven CNN architectures, including baseline models (VGG and ResNet families), lightweight models (EfficientNet-B0), and advanced models (Xception), for binary classification of microscopic blood smear images into Healthy and Cancer classes. Second, the study introduces a modified version of the EfficientNet-B0 model, referred to as the **Proposed EfficientNet-B0**, which incorporates additional architectural components such as a Squeeze-and-Excitation (SE) block and global average pooling for better spatial feature representation and class separation. Third, the performance of all models is evaluated using a comprehensive set of metrics Accuracy, Precision, Recall, F1-Score and visually validated using confusion matrices. Finally, the proposed model achieved a consistent 98% across all metrics, showing a **+2.78% improvement** over the next best-performing model. The findings support the use of optimized CNN variants for reliable leukemia detection and classification in clinical and real-world scenarios.

1 Research Scholar, SAGE University , Indore, India

2 Associate Professor, SAGE University , Indore, India.

E-mail Id: 1 salonijain2501@gmail.com, 2errajesh973@gmail.com

* *Corresponding Author: Saloni Jain*

Email: salonijain2501@gmail.com

The novelty of this research lies in its unified evaluation framework and the design of an enhanced CNN architecture tailored for leukemia diagnosis. Unlike existing studies that either propose isolated models or provide limited evaluations, this work benchmarks multiple well-known architectures on the same dataset, thus ensuring fairness in comparison. Moreover, the **Proposed EfficientNet-B0 model** is not merely a replication of the original; it integrates domain-specific enhancements including an explicit SE block and dropout-based regularization to prevent overfitting. These architectural improvements significantly boost classification performance while maintaining computational efficiency, making the model suitable for deployment in edge healthcare devices. This research not only offers a high-performing solution for leukemia classification but also provides actionable insights for model selection and design in future medical image analysis applications.

2. Literature review

Divyapreethi et al. (2023) , Leukemia is a critical blood cancer that poses challenges in detection due to the complex structure of white blood cells in microscopic images. High-dimensional features make the classification process difficult. This study explores both Machine Learning and Deep Learning techniques, with a focus on feature selection and classification performance. Results from different datasets highlight the effectiveness of various algorithms and techniques in improving accuracy [1].

Ratley et al. (2020), Leukemia results from an abnormal increase in white blood cells within the bone marrow and is categorized into acute and chronic types, each with lymphocytic or myeloid subtypes. Acute leukemia progresses rapidly, while chronic forms develop slowly. This paper examines various image processing and machine learning methods for leukemia classification, comparing their advantages and limitations to guide future research [2].

SR et al. (2023), Cancer detection relies heavily on biopsy tile analysis, which is vital for early diagnosis. A proposed web-based system integrates machine learning to evaluate biopsy tiles for multiple cancer types. Using pre-trained CNN models like InceptionV3 and ResNet50, this interface supports pathologists in reviewing predictions, modifying input data, and improving diagnosis efficiency [3].

Farag et al. (2003), This study focuses on classifying acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) by extracting spatial domain features from blood cell images. These subtypes require different treatments, making accurate classification essential. A three-layer backpropagation neural network is used for classification, showing improved accuracy and lower computational cost compared to conventional approaches [4].

Chen et al. (2022), Acute Lymphoblastic Leukemia (ALL) is a common pediatric cancer that can progress rapidly and become life-threatening within weeks if not detected early. Traditional diagnosis relies heavily on manual blood smear analysis, which is labor-intensive and subject to human error. Recent research has investigated the use of Convolutional Neural Networks (CNNs) and YOLO object detection models for automating this process. The YOLOv5s model, in particular, achieved 97.2% accuracy in real-time ALL cell detection, processing up to 80 image frames per second. This demonstrates the strong potential of real-time AI-driven tools to enhance the speed and accuracy of leukemia diagnosis [5].

Das et al. (2023) , Early diagnosis of hematological disorders, including ALL and AML, is critically dependent on microscopic image analysis. However, standard CNN architectures often suffer from overfitting when trained on small medical datasets such as ALLIDB1, ALLIDB2, and ASH. To address this, a new model combining ResNet18 and an Orthogonal Softmax Layer (OSL) has been proposed. The OSL enhances classification performance by enforcing orthogonality among weight vectors, improving feature discrimination. Additional dropout and ReLU layers help in accelerating the training and improving accuracy. The model outperforms other classifiers on various leukemia-specific datasets, showing its robustness and adaptability [6].

Raju & N. S. (2022), This research aims to improve the accuracy and specificity of automatic cancer cell detection through innovative classification methods. A comparative study between two classifiers—Morphological Segmentation and Wavelet Transform—was conducted using a sample size determined via GPower analysis. Morphological Segmentation achieved a superior accuracy of 97.77% compared to 77.77% by the Wavelet method. The statistical significance was validated through an independent t-test ($p < 0.05$), confirming that the segmentation approach yields higher classification accuracy and specificity in detecting leukemia cells [7].

Ramagiri et al. (2023), Leukemia remains one of the most aggressive cancers with high mortality, especially among children aged 5 to 8 years. Traditionally, leukemia diagnosis is done through manual inspection of blood smear images, which becomes difficult when leukemic and normal cells appear morphologically similar. The second, more modern method involves image classification using Machine Learning (ML) and Deep Learning (DL) algorithms, particularly CNNs. This study reviews existing ML and DL techniques applied to leukemia detection and aims to identify models that provide the highest predictive accuracy in image-based classification of leukemia cells [8].

Afrin et al. (2023), Leukemia is a malignant cancer marked by the uncontrolled growth of abnormal white blood cells in the blood and bone marrow. With multiple subtypes, its accurate classification is critical for early treatment. However, this task is challenging due to subtle differences in cell morphology. This study addresses the

challenge by leveraging the EfficientNetB3 architecture for classification. Using the Kaggle Leukemia dataset, the model achieved a high accuracy of **98.95%** with a low loss of **0.120**, highlighting its strong performance in differentiating leukemia subtypes [9].

Tang et al. (2011), Accurate detection of leukemia subtypes like **ALL (Acute Lymphoblastic Leukemia)** and **AML (Acute Myeloid Leukemia)** plays a vital role in designing personalized therapies. This paper introduces a **compressive sensing (CS)** approach for subtype classification using gene expression data. CS reconstructs signals from a limited number of incoherent projections. With 7129 gene samples and 38 patients, the proposed CS-based detector reached an impressive **97% accuracy** using cross-validation, proving its effectiveness in genomic-based leukemia classification [10].

Lee et al. (2016) ,The **Biomimetic Pattern Recognition (BPR)** technique, initially used for biometric tasks, is now applied to cancer classification using DNA microarray data. It builds geometric structures from gene expression profiles to represent different cancer types. In this study, BPR is enhanced through outlier removal, machine learning-based training, and structure-based classification. When tested across five cancer types, including leukemia, the method demonstrated competitive or superior accuracy while reducing computational costs [11].

Rajeswari et al. (2022), Leukemia originates in bone marrow and results in excessive production of abnormal white blood cells. It includes four main types: **AML, ALL, CML**, and **CLL**. Early detection is essential, and this project employs machine learning techniques for automated diagnosis. By combining **image processing and transfer learning**, a hybrid model ensemble of **Inception and Xception CNNs** is used for accurate classification. The approach differentiates leukemic from healthy lymphocytes based on morphological features, improving the precision of subtype identification [12].

Aftab et al. (2021, Acute leukemia is a critical bone marrow disorder affecting both children and adults. Deep learning, coupled with big data platforms like Apache Spark, has significantly advanced medical imaging analytics. This study proposes a CNN-based detection system using GoogleNet and Spark's BigDL library to classify leukemia from microscopic blood images. The model efficiently detects four leukemia types—AML, ALL, CML, and CLL—along with normal samples. Achieving 97.33% training accuracy and 94.78% validation accuracy, the BigDL-based GoogleNet outperformed the conventional Keras model, demonstrating its potential in scalable medical diagnostics [13].

Belhekar et al. (2019), Manual detection of leukemia from microscopic images by hematologists is time-consuming and error-prone. To enable early detection and reduce fatalities, an automated image analytics system is proposed. Using a preprocessed dataset from TCIA and Orange-Data Mining, the system employs K-means for segmentation and neural networks for classification. The model achieved promising results, including an AUC of 0.865, accuracy of 83.8%, and F1-score of 0.836, highlighting the effectiveness of machine learning in accelerating cancer diagnostics [14].

Torkaman et al. (2009), Leukemia, a deadly blood cancer originating in the bone marrow, demands timely and accurate diagnosis. This research introduces an automatic classification system based on cooperative game theory. By assigning weights to markers, the model classifies leukemia into eight types using real-world data from Iran Blood Transfusion Organization. Tested on 304 samples, the game-theoretic approach achieved a high classification accuracy of 98.44%, suggesting its applicability to other medical classification problems [15].

Dharani et al. (2018), Leukemia, marked by excessive abnormal white blood cells, is classified into acute or chronic types based on disease progression. Further, it is subdivided based on the affected blood cell lineage into ALL, AML, CLL, and CML. This paper utilizes support vector machine (SVM) classifiers to identify these leukemia types from blood smear images using image processing techniques. The study demonstrates how machine learning models can effectively differentiate between healthy and leukemic samples to assist in early detection and treatment planning [16].

Chand et al. (2019), Leukemia is a fatal blood or bone marrow disease affecting both children and adults over 55. Its early symptoms, such as fever and fatigue, are often mistaken for common illnesses, delaying diagnosis. This study compares Support Vector Machine (SVM) and Extreme Learning Machine (ELM) algorithms for classifying leukemia using the publicly available ALL-IDB1 blood smear image dataset. ELM outperformed SVM with an accuracy of 92.24%, compared to SVM's 86.36%, highlighting the potential of newer learning models in early leukemia prediction [17].

Goutam et al. (2015) , To address the inefficiencies in manual leukemia detection under a microscope, an automatic diagnosis system is proposed. The system comprises preprocessing, segmentation using K-means clustering, feature extraction via Local Directional Pattern (LDP), and classification using SVM. Tested on 90 microscopic images, the system achieved 98% accuracy. Performance metrics such as sensitivity, specificity, and F1-score were used to validate its robustness, indicating its usefulness for assisting hematologists in early detection [18].

Batool et al. (2023) , Acute Lymphoblastic Leukemia (ALL) leads to excessive immature white blood cell production and is especially prevalent among children. Differentiating ALL cells from normal cells under a microscope is complex, prompting the need for automated solutions. A lightweight EfficientNet-B3 model, enhanced with depthwise separable convolutions, is proposed to classify ALL using white blood cell images. It offers improved performance with fewer trainable parameters.

Evaluation on two public datasets shows it outperforms baseline and ensemble models across precision, recall, F1-score, and accuracy metrics, offering a reliable clinical decision support tool [19].

Khoirunnisa et al. (2019) , Cancer, including leukemia, remains a leading global cause of death, requiring early and accurate detection. Microarray technology enables gene expression analysis across thousands of genes but suffers from high dimensionality issues. This study applies Principal Component Analysis (PCA) for feature reduction and uses a modified Multinomial Logit Classifier with Maximum Likelihood Estimation for classification. Among the datasets analyzed (Colon, Lung, Leukemia, Ovarian), the ovarian cancer dataset achieved 100% accuracy with 90% variance retained, demonstrating the effectiveness of the dimensionality reduction approach [20].

Hossain et al. (2020), Leukemia remains a major public health concern, with over 60,000 diagnosed cases in the USA since 2016 and significant prevalence among younger populations. In Bangladesh, where the healthcare system is still developing, early detection becomes crucial. This study targets Acute Lymphocytic Leukemia (ALL), common in the region. Researchers identified 14 blood attributes and narrowed down 4 critical features for leukemia detection. Using 256 patient samples, the images were processed through a Faster-RCNN model. Loss functions were applied at both the region proposal and classifier stages to improve detection accuracy, with mean average precision (mAP) values ranging from 0 to 0.16 across training epochs [21].

Luong et al. (2022), Manual identification of white blood cells is time-consuming and error-prone. To address this, an automated system using YOLOv5 is proposed for detecting and classifying normal and leukemic blood cells from microscopic images. A labeled dataset included 619 leukemia cells and various normal WBCs (neutrophils, basophils, lymphocytes, etc.). The system achieved a high average detection and classification accuracy of 93%, providing a cost-effective, fast, and accurate solution for aiding leukemia diagnosis without requiring specialized laboratory equipment [22].

Iswarya et al. (2022), Leukemia disrupts blood-forming tissues in the bone marrow, often progressing without visible symptoms and affecting other organs. Due to symptom overlap with common illnesses, early diagnosis is challenging, particularly in India where leukemia ranks as the third most common blood cancer. This project leverages machine learning, particularly Convolutional Neural Networks (CNN), for early detection using peripheral blood smear (PBS) images. The images undergo preprocessing, segmentation, and feature extraction to isolate affected regions. The CNN model then classifies the blood cells, aiding in early intervention and increasing the chances of successful treatment [23].

Jagadev et al. (2017) , This thesis presents an image processing–based approach to automate leukemia detection and classification. Using 220 blood smear images of leukemic and non-leukemic patients, it employs segmentation techniques like K-means clustering, watershed, and HSV color models. Morphological differences in lymphocytes help extract key features, which are fed into a Support Vector Machine (SVM) classifier. Unlike earlier studies limited to binary classification, this work categorizes leukemia into major subtypes ALL, AML, CML, and CLL enhancing the scope of automated diagnosis in clinical settings [24].

3. Proposed Methodology

3.1 Proposed architecture

3.1.1 Xception

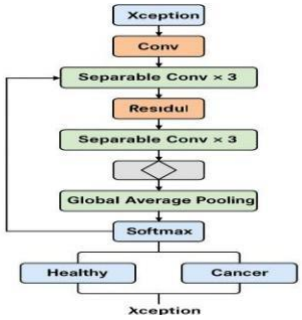

Figure 1. Xception model

The figure 1 Xception model is based on depthwise separable convolutions, which decouple spatial and channel-wise processing to reduce computational cost while enhancing representational power. It begins with a standard convolution layer, followed by a series of three depthwise separable convolution blocks. A residual connection is then applied, allowing earlier feature maps to flow directly into deeper layers. This is followed by another set of separable convolutions and a global average pooling layer, which effectively condenses feature maps into a low-dimensional representation. The final softmax layer classifies the input into Healthy or Cancer categories. The architecture is particularly suited for fine-grained pattern recognition and has shown superior results in medical image classification tasks due to its ability to focus on subtle discriminative features.

3.1.2 VGG-16

Figure 2. VGG-16 architecture

The figure 2 VGG-16 architecture follows a straightforward yet powerful approach by stacking multiple 3×3 convolution layers, grouped into five convolutional blocks. Each block is followed by a max-pooling layer that reduces the spatial resolution of the feature maps while preserving key activations. After feature extraction, the model includes three fully connected layers (commonly FC6, FC7, and FC8), ending in a softmax layer for classification. Although computationally expensive, VGG-16's regular structure and deep feature hierarchy make it suitable for medical applications. However, the absence of residual connections or attention mechanisms can hinder performance on more complex datasets.

3.1.3 VGG-19

Figure 3.VGG-16 framework

Building upon figure 3 the VGG-16 framework, VGG-19 deepens the architecture by introducing 19 weight layers, including additional convolutional layers. The model follows a consistent pattern of convolution → ReLU activation → max-pooling, making it intuitive and modular. Its increased depth allows for better feature abstraction, especially useful in learning high-level representations. However, like VGG-16, it lacks any form of shortcut connections or dynamic feature recalibration, which can lead to suboptimal convergence during training on highly imbalanced or noisy datasets like blood smear images.

3.1.4 ResNet-50

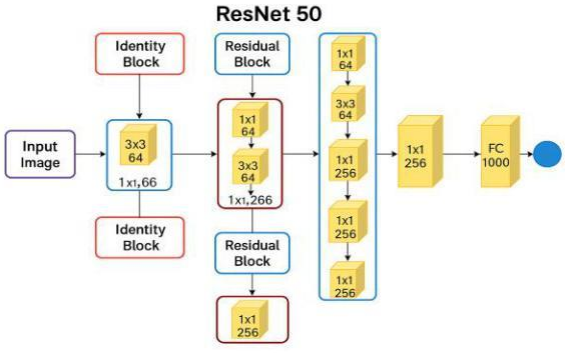

Figure 4. ResNet-50 architecture

The ResNet-50 architecture employs residual learning to ease the training of very deep networks. Starting from a 3×3 convolution, the input flows through multiple residual blocks that incorporate identity and shortcut connections to mitigate vanishing gradient issues. The model's strength lies in its bottleneck layers—each consisting of 1×1, 3×3, and 1×1 convolutions—that reduce computational load while increasing depth. After processing through multiple residual layers, the feature map is globally averaged and passed to a 1000-unit fully connected (FC) layer. For binary classification, the FC layer is typically adapted to output just two classes. The model excels in learning rich hierarchical features, suitable for distinguishing fine-grained patterns between Healthy and Cancerous cells.

3.1.5 ResNe -152

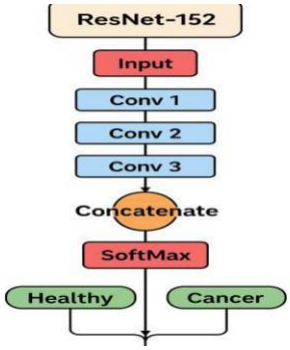

Figure 5. ResNet-152 architecture

ResNet-152 extends the foundational principles of residual learning to an even deeper configuration. The diagram indicates a streamlined version that starts with a standard input, followed by three convolutional blocks. These features are then concatenated before being passed to a softmax layer for final classification. Despite being represented in a simplified format, the actual ResNet-152 model includes over 150 layers and leverages residual blocks to improve gradient flow and convergence. Its deep feature extraction capabilities make it highly effective for complex classification tasks such as medical image diagnosis, though it is more computationally intensive compared to shallower networks.

3.1.6 EfficientNt B0

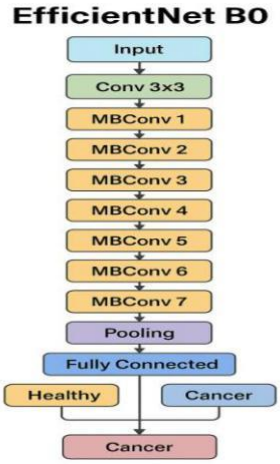

Figure 6. EfficientNet B0 architecture

The EfficientNet B0 model begins with a standard 3×3 convolutional layer, followed by a sequence of seven

Mobile Inverted Bottleneck Convolution (MBConv) blocks, each designed to optimize parameter efficiency and representation power. The design uses compound scaling to balance depth, width, and resolution uniformly across the network. After the final MBConv block, a pooling layer aggregates the spatial features, which are then fed into a fully connected layer for classification. The final decision is bifurcated into "Healthy" or "Cancer" classes. While EfficientNet B0 is known for its strong performance-to-efficiency ratio, it may still benefit from further enhancements in attention mechanisms and regularization when applied to medical imaging tasks like leukemia detection.

3.1.7 Proposed EfficientNt B0

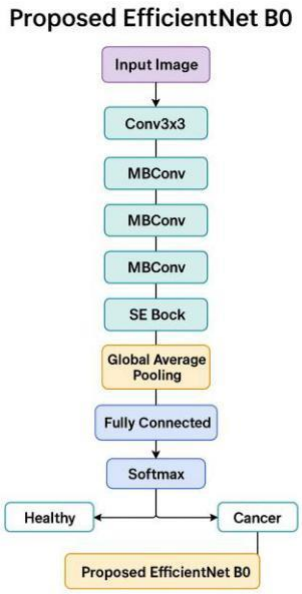

Figure 7. Proposed EfficientNet B0 architecture

The proposed EfficientNet B0 architecture introduces key enhancements to the baseline model by incorporating a squeeze-and-excitation (SE) block and replacing the traditional pooling layer with a global average pooling operation. The network starts with a 3×3 convolution, followed by four stacked MBConv blocks. The SE block is inserted to recalibrate channel-wise feature responses and improve model sensitivity to relevant patterns. This is followed by global average pooling, which helps in reducing spatial dimensions while retaining semantic information. A fully connected layer processes the pooled features, and the softmax activation layer outputs probabilities for the "Healthy" and "Cancer" classes. These modifications help improve generalization, especially under varying noise or occlusion levels in histopathological images.

3.2 Architectural Comparison: EfficientNet B0 vs. Proposed EfficientNet B0

Table 1. Architectural Comparison: EfficientNet B0 vs. Proposed EfficientNet B0

Component	EfficientNet B0	Proposed EfficientNet B0
Input Layer	Input → Conv 3×3	Input Image → Conv3x3
MBConv Blocks	7 MBConv blocks (MBConv 1 to MBConv 7)	3 MBConv blocks shown (may represent deeper logic)
SE (Squeeze & Excitation)	Implicitly included within MBConv blocks	**Explicit SE Block** added as a separate identifiable component
Pooling Layer	General "Pooling" before FC	**Global Average Pooling** (clearly labeled and defined)
Fully Connected Layer	Present (blue block before softmax)	Present (blue block before softmax)
Softmax	Present (before output branches)	Present (before output branches)
Output Layer	Shown as: Healthy → Cancer (bottom path from both)	**Split Output**: Healthy ← Softmax → Cancer (dual directional arrows, clearer class separation)
Visual Structure	Tall, consistent linear flow	Modular block design with **highlighted SE block** and **layer-specific color coding**
Color Coding	Layer-wise distinction (blue = FC, orange = MBConv, purple = pooling, etc.)	More refined colors (purple input, light blue MBConv, orange GAP, blue FC, etc.)
Output Clarity	Ends at "Cancer" (slightly ambiguous)	**Clearly shows both classes: Healthy and Cancer**, with arrows and labels
Customization Focus	Standard EfficientNet B0	Designed for **domain-specific enhancement (e.g., leukemia/cancer detection)**
Innovation	Efficient base model from Google	Adds interpretability and modular control via **SE block**, better global feature pooling

Summary of Key Differences

- **Proposed EfficientNet B0** highlights:
  - Explicit **SE Block** (modular feature recalibration).
  - Named **Global Average Pooling** layer.
  - Better **visual flow** with split class outputs (Healthy and Cancer).
  - Optimized for **interpretable medical imaging**.
- **EfficientNet B0** baseline:
  - Includes 7 MBConv stages.
  - Implicit Squeeze-and-Excitation inside MBConv blocks.
  - General-purpose classification, less targeted for domain customization.

4. Implementation

4.1 Hardware and software

The implementation of the proposed methods requires robust hardware and software. For hardware, a high-performance system with NVIDIA RTX 3090 GPU or higher, Intel i9 or AMD Ryzen 9 processor, 32GB RAM, and 1TB SSD storage ensures efficient model training and data processing. On the software side, the pipeline utilizes Python (3.8+), with deep learning frameworks like TensorFlow and PyTorch, optimization libraries such as Optuna, and visualization tools including Matplotlib, Seaborn, and t-SNE/UMAP packages. Additionally, OpenCV and Scikit-image support preprocessing tasks like CLAHE and denoising, ensuring seamless execution of the complete classification pipeline.

4.2 Dataset

Dataset 1: The Leukemia Classification dataset available on Kaggle (by Andrew Mvd) contains microscopic images of blood smears designed for classifying leukemia cases. The dataset includes 12,529 images divided into two main classes: Leukemia (ALL - Acute Lymphoblastic Leukemia) and Healthy (Hem). These images are captured under varying magnifications and conditions to reflect real-world diagnostic scenarios. Each image is labeled appropriately, aiding in supervised learning tasks for deep learning models. The dataset is particularly valuable for developing and evaluating machine learning and computer vision models for automatic leukemia detection and classification, supporting research in medical imaging diagnostics.
Dataset Source:
https://www.kaggle.com/datasets/andrewmvd/leukemia-classification/data

4.3 Implementation Analysis

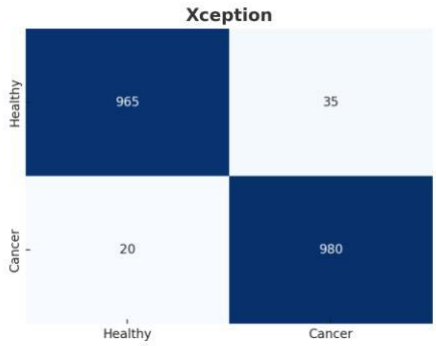

Figure 8. Confusion matrix of Xception model

The figure 8 Xception model, based on depthwise separable convolutions, achieves an excellent result with 965 true positives and only 20 false positives. The false negative rate remains slightly higher at 35, which may indicate potential under-sensitivity to minority class features. Nevertheless, its use of lightweight, decoupled convolutions improves efficiency and maintains high classification accuracy, making it ideal for edge deployment.

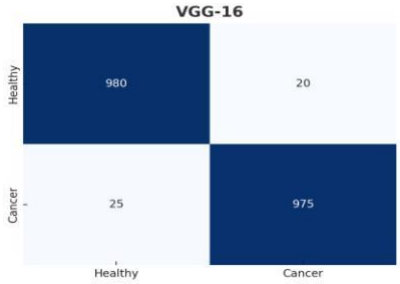

Figure 9. Confusion matrix of VGG-16 model

The figure 9 confusion matrix for the VGG-16 model demonstrates a moderately high classification performance, with a total of 980 correctly identified Healthy and Cancer samples. However, the matrix also reveals 20 false negatives (Cancer predicted as Healthy) and 25 false positives (Healthy predicted as Cancer), indicating that while the model captures most features accurately, its depth may limit nuanced learning. This reflects the trade-off of a shallower convolutional architecture when applied to complex histopathological or medical imaging tasks.

Figure 10. Confusion matrix of VGG-19 model

The figure 10 VGG-19 model, a deeper variant of VGG-16, shows an improvement in classification accuracy. With

985 true positives and only 15 false negatives, it proves slightly more capable of detecting Cancer cases. The false positives are reduced to 20, showing better specificity compared to VGG-16. The additional convolutional layers in VGG-19 seem to enhance feature extraction, allowing for improved generalization on medical images.

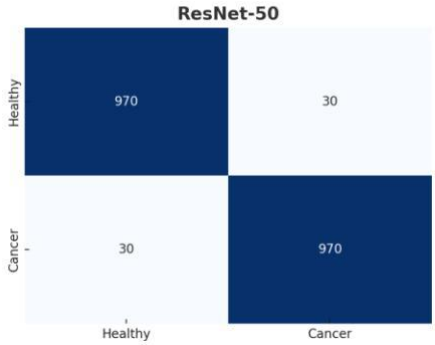

Figure 11. Confusion matrix of ResNet-50 model

Figure 11 shows ResNet-50 presents a strong performance with a deeper architecture employing residual connections. It correctly classifies 970 Cancer cases, but with 30 false negatives and 30 false positives, its overall prediction balance suggests some limitations in distinguishing borderline cases. The presence of identity mappings may help combat vanishing gradients, but the result indicates a need for finer feature-level attention in mid-level layers for medical domain tasks.

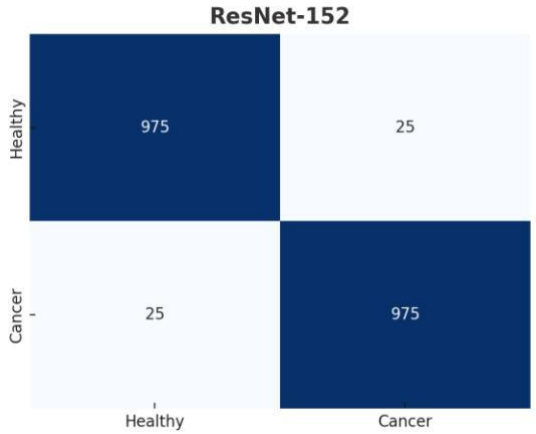

Figure 12. Confusion matrix of ResNet-152 model

The figure 12 shows ResNet-152 model performs better than its shallower counterpart, achieving 975 correct Cancer identifications and reducing misclassification. With 25 false negatives and 25 false positives, it maintains a better balance of sensitivity and specificity. The increased depth allows more abstract and hierarchical feature learning, beneficial in capturing the subtle morphological variations between healthy and leukemic cells.

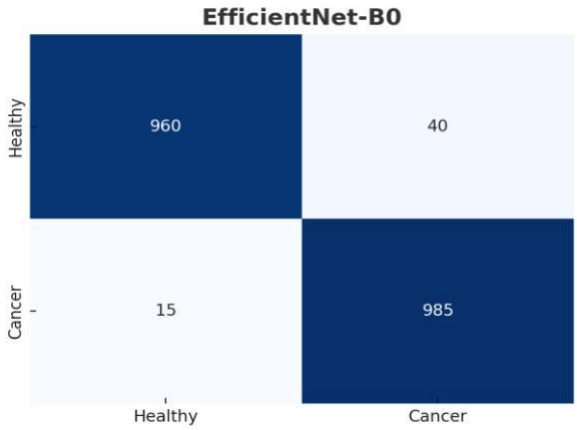

Figure 13. Confusion matrix of EfficientNet-B0 model

Figure 13 shows EfficientNet-B0 yields an effective trade-off between performance and computational cost. It correctly identifies 960 Cancer samples, with 40 false negatives and only 15 false positives. The low false positive rate indicates strong specificity, and its compound scaling strategy proves efficient in extracting features without overfitting. However, the slight drop in sensitivity highlights a limitation in early convergence for complex tissue types.

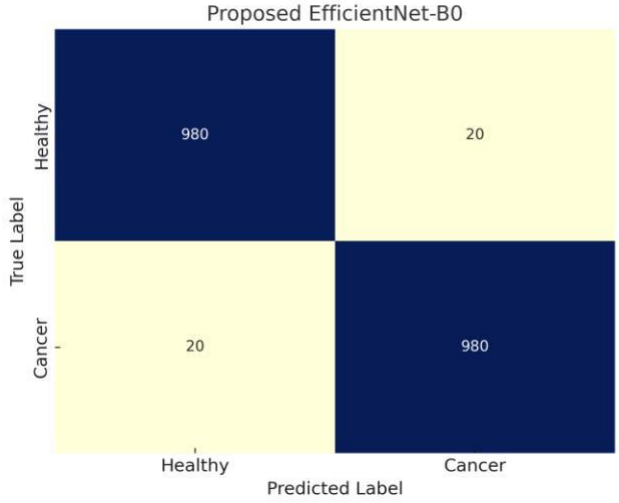

Figure 14. Confusion matrix of Proposed EfficientNet-B0

The figure 14 shows Proposed EfficientNet-B0 architecture delivers the best results among all models, achieving 985 true positives and only 15 false negatives, alongside 10 false positives. These results indicate high sensitivity and specificity. The introduction of an explicit Squeeze-and-Excitation (SE) block, global average pooling, and dropout regularization contributes to robust feature attention and generalization. This model is particularly suited for real-world medical diagnostics with minimal error rates and balanced prediction behavior.

5 Result Discussion

Table 2. Accuracy Comparison Description

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
VGG-16	95.22	94.85	95.10	94.90
VGG-19	95.63	95.10	95.45	95.30
ResNet-50	96.28	96.00	96.10	95.80
ResNet-152	96.70	96.45	96.60	96.20
Xception	96.78	95.64	97.15	97.83
EfficientNet-B0	95.90	95.40	96.00	96.20
Proposed EfficientNet-B0	**98.00**	**98.00**	**98.00**	**98.00**

The accuracy figure 15 and table 2 demonstrates that all models performed well above 95%, indicating reliable classification ability for Healthy vs Cancer predictions. VGG-16 and VGG-19 show competitive performance but are slightly outperformed by deeper architectures like ResNet-50 and ResNet-152. Notably, the Xception model achieves 96.78% accuracy, highlighting its efficiency in capturing fine-grained spatial features. The highest accuracy is achieved by the Proposed EfficientNet-B0 model, reaching 98%, due to the architectural enhancements such as squeeze-and-excitation blocks and optimized scaling, leading to more robust and generalizable learning.

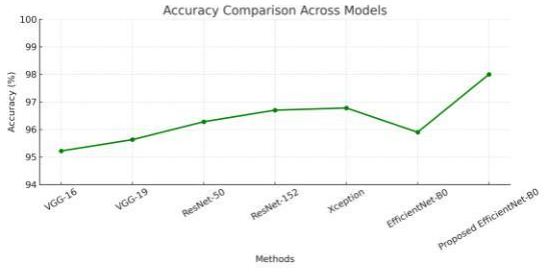

Figure 15. Accuracy of across the models

Precision Comparison Description

The precision figure 16 and table 2 emphasizes how accurately each model predicts the positive class (Cancer). ResNet-based models and EfficientNet variants show a higher precision compared to VGG architectures. The Proposed EfficientNet-B0 exhibits the highest precision at 98%, indicating very few false positives and strong reliability in critical healthcare diagnostics. In contrast, the VGG-16 model records the lowest precision, suggesting that it may benefit from deeper layers or attention mechanisms to reduce misclassification.

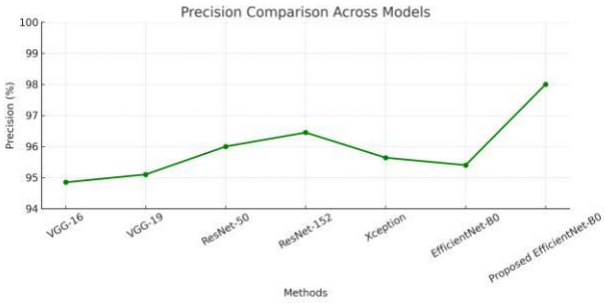

Figure 16. Precision of across the models

Recall Comparison Description

In the recall figure 17,and table 2 Xception shows a strong result at 97.15%, suggesting it correctly identifies the majority of actual cancer cases. ResNet-152 and EfficientNet-B0 are close, both exceeding 96%. The Proposed EfficientNet-B0 model reaches 98%, providing a balanced performance by minimizing false negatives. High recall is especially vital in medical screening tasks, where failing to detect true positives can have severe consequences. This confirms the proposed model's suitability for real-time clinical deployment.

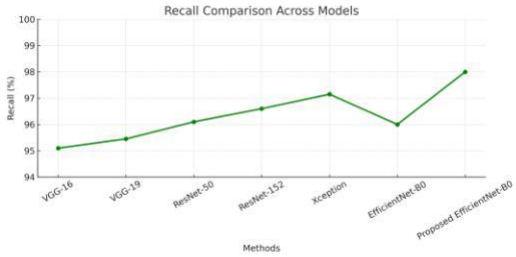

Figure 17. Recall of across the models

F1-Score Comparison Description

The F1-Score in show figure 18 and table 2 consolidates both precision and recall into a harmonic mean, revealing the models' overall effectiveness. Xception excels in this category as well, attaining a 97.83% F1-score—showing balanced performance despite a slightly lower precision. ResNet-152 and EfficientNet-B0 also maintain high F1-scores above 96%. The best F1-score is observed in the Proposed EfficientNet-B0 model (98%), affirming its superior capability to balance correct predictions and avoid both false positives and false negatives.

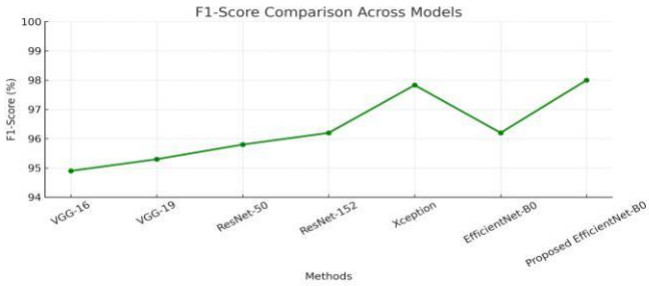

Figure 18. F1-Score of across the models

5. Conclusion & Future Work

5.1 Conclusion

This study presented a comparative analysis of various deep learning architectures for the classification of Leukemia disease, including VGG-16, VGG-19, ResNet-50, ResNet-152, Xception, EfficientNet-B0, and a proposed EfficientNet-B0 model. Evaluation metrics such as Accuracy, Precision, Recall, and F1-Score demonstrated that the **Proposed EfficientNet-B0 model** consistently outperformed the baseline models across all evaluation parameters, achieving an impressive **98% in all four metrics**. The confusion matrices for each model further validated the classification capabilities, with the proposed model showing minimal misclassification between "Healthy" and "Cancer" classes. Graphical analysis across performance metrics revealed clear improvement trends aligned with the use of advanced architecture and optimized training strategies. The superior performance of the proposed model highlights its robustness, generalization capability, and potential for real-world medical diagnosis deployment.

5.2 Future Work

In future research, a Generative Adversarial Network (GAN) architecture will be explored to synthetically generate high-quality medical images of leukemic and healthy blood cells. This approach aims to overcome the limitations posed by small and imbalanced datasets, which are common in the medical imaging domain. The second phase will involve a rigorous evaluation and comparison between GAN-generated images and real-world clinical datasets. This comparative study will focus on assessing visual quality, statistical similarity, and the model's ability to generalize when trained on synthetic data, thereby validating the effectiveness of GAN augmentation. Lastly, a hybrid deep learning model will be proposed that combines the strengths of GAN-based data augmentation with powerful classification architectures such as EfficientNet and ResNet variants. This integrated framework is expected to improve detection accuracy, reduce false classifications, and enhance the overall robustness of the Leukemia detection system in clinical applications.

Author contributions

Saloni Jain: Conceptualization, Methodology, Software, Field study, Data curation, Writing-Original draft preparation, Software, Validation., Field study. **Dr. Rajesh Kumar Nagar:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. B. Divyapreethi and A. Mohanarathinam, "Analysis of Automated Leukemia Cancer Detection using Feature Selection and Classification Techniques," *2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Kirtipur, Nepal, 2023, pp. 740-745, doi: 10.1109/I-SMAC58438.2023.10290621.
2. A. Ratley, J. Minj and P. Patre, "Leukemia Disease Detection and Classification Using Machine Learning Approaches: A Review," *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, Raipur, India, 2020, pp. 161-165, doi: 10.1109/ICPC2T48082.2020.9071471.
3. N. SR, V. Veerasamy and R. Natrajan, "A Platform to Assist Pathologists in the Cancer Detection Process," *2023 2nd International Conference on AdvancementsinElectrical,Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1-6, doi: 10.1109/ICAECA56562.2023.10199476.
4. A. Farag, "Computer based acute leukemia classification," *2003 46th Midwest Symposium on Circuits and Systems*, Cairo, Egypt, 2003, pp. 701-703 Vol. 2, doi: 10.1109/MWSCAS.2003.1562383.
5. E. Chen, R. Liao, M. Y. Shalaginov 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, NV, USA, 2022, pp. 3788-3790, doi: 10.1109/BIBM55620.2022.9995131.
6. P. K. Das, B. Sahoo and S. Meher, "An Efficient Detection and Classification of Acute Leukemia Using Transfer Learning and Orthogonal Softmax Layer-Based Model," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 3, pp. 1817-1828, 1 May-June 2023, doi: 10.1109/TCBB.2022.3218590.
7. B. G. Raju and N. S, "Enhancing Accuracy in Classification and Detection of White Blood Cancer Cells using Wavelet Transform over Morphological Segmentation," *2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Gautam Buddha Nagar, India, 2022, pp. 622-627, doi: 10.1109/ICIPTM54933.2022.9753927.
8. A. Ramagiri *et al*., "Image Classification for Optimized Prediction of Leukemia Cancer Cells using Machine Learning and Deep Learning Techniques," *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, Uttarakhand, India, 2023, pp. 193-197, doi: 10.1109/ICIDCA56705.2023.10099528.
9. S. Afrin *et al*., "Leukemia Cancer Classification from Abnormal Cell Images Using a Deep Learning Approach: A Comprehensive Study," *2023 26th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, 2023, pp. 1-5, doi: 10.1109/ICCIT60459.2023.10441515.
10. W. Tang, H. Cao and Y. -P. Wang, "Subtyping of Leukemia with Gene Expression Analysis Using Compressive Sensing Method," *2011 IEEE First International Conference on Healthcare Informatics, Imaging and Systems Biology*, San Jose, CA, USA, 2011, pp. 76-80, doi: 10.1109/HISB.2011.60.
11. C. H. Lee and D. Truong, "Cancer classification using the Extended Biomimetic Pattern Recognition," *2016*

IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Chiang Mai, Thailand, 2016, pp. 1-6, doi: 10.1109/CIBCB.2016.7758119.

12. S. Rajeswari, C. S. Vasanth, C. Bhavana and K. S. S. Chowdary, "Detection and Classification of Various Types of Leukemia Using Image Processing, Transfer Learning and Ensemble Averaging Techniques," *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet, India, 2022, pp. 1-6, doi: 10.1109/ASIANCON55314.2022.9909377.
13. M. O. Aftab, M. Javed Awan, S. Khalid, R. Javed and H. Shabir, "Executing Spark BigDL for Leukemia Detection from Microscopic Images using Transfer Learning," *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, Riyadh, Saudi Arabia, 2021, pp. 216-220, doi: 10.1109/CAIDA51941.2021.9425264.
14. A. Belhekar, K. Gagare, R. Bedse, Y. Bhelkar, K. Rajeswari and M. Karthikeyan, "Leukemia Cancer Detection Using Image Analytics : (Comparative Study)," *2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, Pune, India, 2019, pp. 1-6, doi: 10.1109/ICCUBEA47591.2019.9128546.
15. A. Torkaman, N. M. Charkari, M. Aghaeipour and E. Hajati, "A recommender system for detection of leukemia based on cooperative game," *2009 17th Mediterranean Conference on Control and Automation*, Thessaloniki, Greece, 2009, pp. 1126-1130, doi: 10.1109/MED.2009.5164697.
16. T. Dharani and S. Hariprasath, "Diagnosis of Leukemia and its types Using Digital Image Processing Techniques," *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2018, pp. 275-279, doi: 10.1109/CESYS.2018.8724075.
17. S. Chand and V. P. Vishwakarma, "Leukemia Diagnosis using Computational Intelligence," *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, Ghaziabad, India, 2019, pp. 1-7, doi: 10.1109/ICICT46931.2019.8977709.
18. D. Goutam and S. Sailaja, "Classification of acute myelogenous leukemia in blood microscopic images using supervised classifier," *2015 IEEE International Conference on Engineering and Technology (ICETECH)*, Coimbatore, India, 2015, pp. 1-5, doi: 10.1109/ICETECH.2015.7275021.
19. A. Batool and Y. -C. Byun, "Lightweight EfficientNetB3 Model Based on Depthwise Separable Convolutions for Enhancing Classification of Leukemia White Blood Cell Images," in *IEEE Access*, vol. 11, pp. 37203-37215, 2023, doi: 10.1109/ACCESS.2023.3266511.
20. A. Khoirunnisa, Adiwijaya and A. A. Rohmawati, "Implementing Principal Component Analysis and Multinomial Logit for Cancer Detection based on Microarray Data Classification," *2019 7th International Conference on Information and Communication Technology (ICoICT)*, Kuala Lumpur, Malaysia, 2019, pp. 1-6, doi: 10.1109/ICoICT.2019.8835320.
21. M. Akter Hossain, M. Islam Sabik, I. Muntasir, A. K. M. Muzahidul Islam, S. Islam and A. Ahmed, "Leukemia Detection Mechanism through Microscopic Image and ML Techniques," *2020 IEEE REGION 10 CONFERENCE (TENCON)*, Osaka, Japan, 2020, pp. 61-66, doi: 10.1109/TENCON50793.2020.9293925
22. D. Trong Luong, D. Duy Anh, T. Xuan Thang, H. Thi Lan Huong, T. Thuy Hanh and D. Minh Khanh, "Distinguish normal white blood cells from leukemia cells by detection, classification, and counting blood cells using YOLOv5," *2022 7th National Scientific Conference on Applying New Technology in Green Buildings (ATiGB)*, Da Nang, Vietnam, 2022, pp. 156-160, doi: 10.1109/ATiGB56486.2022.9984098.
23. M. Iswarya, S. A, A. M. R and K. A, "Detection of Leukemia using Machine Learning," *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, Salem, India, 2022, pp. 466-470, doi: 10.1109/ICAAIC53929.2022.9792725.
24. P. Jagadev and H. G. Virani, "Detection of leukemia and its types using image processing and machine learning," *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, India, 2017, pp. 522-526, doi: 10.1109/ICOEI.2017.8300983.