iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70481

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

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 V May 2025- Available at www.ijraset.com

Automated Leukemia Classification with a Deep Neural Network and Computer Vision

VaishnavPatil[1], Shruti Phase[2], Shubham Maurya[3], Prof. Amit kumar Shinde[4]
Department of Artificial Intelligence G.H Raisoni College of Engineering and Management Pune, 412207, Maharashtra, India

Abstract: ***Theuncheckedgrowthofaberrantwhitebloodcells is the hallmark of leukemia, a serious hematological cancer that starts in the bone marrow and bloodstream and causes bleeding, anemia, and weakened immunity. Early and a precise leukemia diagnosisisnecessarytoenhancepatientsurvivalratesand start treatment procedures on time. In this work, a lightweight MobileNetV2 convolutional neural network (CNN) architectureis used to analyze microscopic blood smear images and presentan automated leukemia classification system. To enhance model performance, advanced preprocessing techniques such as LAB color space segmentation, KMeans clustering, and targeted data augmentation were employed to normalize imaging variabilities and highlight leukemic features.***

The system was trained and evaluated on a curated dataset comprising 3,242 images from 89 patients, encompassing both benign and malignant cases. Compared to conventional manual examination,thesystemachievessuperiorclassificationaccuracy, precision,andrecall,providingascalableandefficientdiagnostic tool.Notably,MobileNetV2'slightweightarchitectureguarantees quick inference with no processing overhead, which makes it ideal for real-time clinical application, especially in settings with limited resources.

By significantly reducing diagnostic time and minimizing human error, it demonstrates the transformative potential of deep learning and computer vision in hematological diagnostics. Future work will focus on expanding multi-class classification capabilities for different leukemia subtypes and integrating explainableAItechniquestoenhanceclinicalinterpretabilityand trustworthiness.

Index Terms: Leukemia detection, deep learning, neural net- work convolution, MobileNetV2, medical image analysis, com- puter vision, automated diagnosis.

I. INTRODUCTION

The unchecked growth of aberrant white blood cells caused by leukemia, a malignant disease of the blood and bone marrow, compromises the body's capacity to fight infections, regulate bleeding, and preserve hematological balance. Ac- counting for a substantial proportion of cancer-related mortal- ity worldwide, leukemia demands early detection and precise classification to facilitate effective treatment planning and improvepatientprognosis.However,conventionaldiagnostic techniques, which rely heavily on manual microscopic exam- ination of peripheral blood smear images by pathologists, are time-consuming, prone to subjectivity, and often inaccessible in resource-limited settings.

The advancement of artificial intelligence (AI), particularly deep learning techniques, has introduced promising alterna- tives for automating medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated remarkable ca- pabilities in extracting intricate patterns from complex visual data,enablinghighlyaccuratediseasedetectionacrossvarious domainsofhealthcare,includingophthalmology,dermatology, and oncology. In the context of hematological malignancies, CNN-based models have shown potential in classifying blood smear images with a level of precision comparable to that of experienced human diagnosticians.

This research focuses on developing an automated, mobile- basedleukemiadetectionsystemutilizingthelightweight andefficientMobileNetV2architecture.Thesystemprocesses peripheral blood smear images, enhancing feature visibility through preprocessing techniques such as LAB color space transformation and KMeans clustering-based segmentation. The choice of MobileNetV2 is motivated by its capability to deliver high classification performance with minimal compu- tational complexity, making it ideal for mobile deployment.

Unlike previous studies that emphasize web-based diag- nostic platforms, the proposed system is implemented as a cross-platformmobileapplicationusingtheFlutterframework. This design ensures real-time image analysis capabilities on portable devices, offering significant advantages in terms of accessibility, portability, and rapid clinical decision support, especially in underserved regions with limited access to spe- cialized healthcare facilities.

The model is trained and evaluated on a comprehensive dataset containing 3,242 peripheral blood smear images collectedfrom89patients.Throughcarefulpreprocessing,model fine-tuning,anddeploymentoptimization,theproposedsystem aims to bridge the gap between advanced AI solutions and practical,scalablehealthcaretools,enablingearlyandaccurate leukemia screening, especially in underserved and remote areas.

The key aims of this research are to automate the detection of leukemia-affected blood cells using MobileNetV2 CNNandtooptimizeimagepreprocessingapproachesforimproved visibility and segmentation of the the lesion. Additionally, a mobile application for real-time leukemia screening will be used to implement a portable diagnostic tool.

By bridging the gap between advanced machine learning methodologies and clinical application, the proposed system aims to democratize access to early leukemia detection, re- duce diagnostic delays, and ultimately contribute to improved healthcare outcomes.

II. RELATED WORK

Early methods for leukemia detection relied on traditional image processing techniques combined with machine learning classifiers such as Support Vector Machines (SVM) and Ran- dom Forests [1], [2]. These approaches depended heavily on handcraftedfeatureextractionandwereoftensensitivetovari- ationsinstaining,illumination,andmorphologicaldifferences, limiting their robustness and scalability.

Deep learning's development revolutionized medical image analysisbymakingitpossibletoautomaticallyextractfeatures straight from data. Convolutional neural networks (CNNs), like AlexNet and GoogLeNet, have shown notable gains in leukemiaclassificationtasksbyusingbloodsmearimages to learn intricate spatial hierarchies [3], [4]. However, these architectures computational complexity presented difficulties forreal-timediagnosticapplications,especiallyinsettingswith limited resources or mobility.

To address efficiency concerns, lightweight CNN models like MobileNetV2, EfficientNetB0, and NASNetMobile were introduced, offering a trade-off between accuracy and com- putational overhead [5],[7]. Studies have shown that MobileNetV2outperformedothermodelsintermsofmaintaining high accuracy while reducing inference time, making it highly suitable for mobile healthcare solutions [8].

Several researchers further enhanced leukemia detection by integrating advanced preprocessing techniques such as color space transformations, segmentation via KMeans clustering, and data augmentation to improve model generalization and focus on critical cellular structures [9], [10]. Transfer learn- ing from pre-trained models has also been widely adopted, leveraging knowledge from large-scale datasets to improve performance on limited medical image datasets [11], [12].

Recent advancements have also emphasized real-time de- ployment through mobile and web-based platforms, enabling on-deviceimageanalysisandremotediagnostics[13],[14].By movingbeyondpurelylaboratory-basedsetups,thesesolutions significantly enhance the accessibility of diagnostic tools for remote and underserved populations.

Buildinguponthesedevelopments,thepresentstudyimple- ments an automated leukemia detection system utilizing Mo- bileNetV2 and a mobile application framework (Flutter). This integratedapproachfocusesonachievinghighclassification accuracy, lightweight deployment, and real-time responsive- ness, thereby addressing key challenges identified in previous research.

III. METHODOLOGY

The automated leukemia detection system operates through a planned sequence. This sequence includes gathering data, preparing it for analysis, finding relevant characteristics, ex- panding the dataset, picking the most significant characteris- tics, teaching the model along with putting it into use. The model training step employs the selected characteristics.

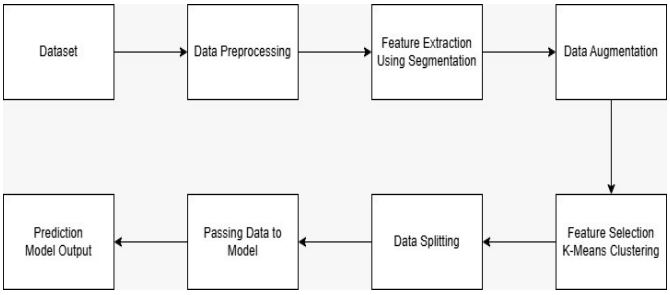


Fig.1.OverallMethodologyWorkflowforAutomatedLeukemiaDetection

A. LeukemiaDataset

The dataset has 3,242 annotated peripheral blood smear images. These come from 89 leukemia patients treated at Taleqani Hospital in Tehran. It contains both benign and ma- lignant instances. Such balance is important for sound binary classification. The dataset records different shape changes. They are important for dependable model training. Earlier studies showed diversified datasets improve how well deep learning models classify blood cancer.

B. DataPreprocessing

In order to standardize input dimensions for MobileNetV2, thefirststepinpreprocessingistoresizeallphotosto 224×224pixels.Assuggestedbyearlierresearchformedicalpictureenhancement,thephotosarenexttransformedinto theLABcolorspace, payingspecialattentiontotheA-channel to improve contrast between leukemic and normal cells [4]. Noise reduction is also conducted at this stage to prevent irrelevant background information from influencing model learning, thus enhancing feature extraction consistency throughout the dataset.

C. FeatureExtractionUsingSegmentation

Data augmentation methods reduce the chances of overfit- ting and improve how well a model generalizes. This involves actions like flipping images horizontally or vertically, rotat- ing them randomly, zooming in along with modifying the brightness. For example, in medical imaging, where labeled datasets are frequently small [3], these augmentation plans havedemonstratedthecapacitytoimprovehowdependable a model is.

D. FeatureSelectionUsingKMeansClustering

Beyond initial segmentation, feature selection is refined through KMeans clustering, where segmented regions are grouped based on pixel intensity and morphological patterns. Featureselectionensuresthemodel'ssensitivitytosubtlevari- ations indicative of leukemia, a strategy validated in previous leukemia classification research that combined unsupervised clustering with CNN-based feature extraction [15].

E. DataSplitting

Following feature selection and augmentation, the datasetis divided between training and validation subsets in a 90:10 ratio. Data splitting strategies that maintain balanced class distributionsduringtrainingandevaluationarecriticaltoavoid bias,asshowninpreviousstudiesinmedicaldiagnosis[12].

F. PassingDatatoModel

The processed and split data is input into the MobileNetV2 network. Transfer learning is utilized, initializing the model withpre-trainedImageNetweightsandfine-tuningitforblood smearclassification.WhilemodelssuchasEfficientNetB0and NASNetMobile were initially considered due to their compet- itiveaccuracyinmedicalimagingtasks[6],[7],MobileNetV2 wasultimatelyselectedowingtoitssuperiortrade-offbetween classification accuracy and computational efficiency, makingit highly suitable for mobile device deployment [5]. Using depthwise separable convolutions and inverted residuals, Mo- bileNetV2 dramatically lowers parameters and FLOPs while preservingcompetitiveperformance—acrucialfactorforreal- time clinical applications.

G. PredictionModelOutput

Themodelproducesbinaryclassificationoutputsaftertrain- ingthatclassifytheinputbloodsmearimagesaseitherbenign or malignant. Standard classification metrics, including as accuracy,precision,recall,andF1-score,areusedtoassessthe outputs. Additionally, a confusion matrix is produced to offer comprehensiveinformationaboutthemodel'sperformance in both classes. Furthermore, the system's viability for real- timedeploymentonmobileplatforms—whichisessential for applications in low-resource healthcare environments—is confirmed by measuring the inference time per image.

IV. RESULTS

Training, validation, and testing datasets were used to thor- oughlyassesstheeffectivenessofthesuggesteddeeplearning- based automated leukemia detection method. A dataset of 3,242 peripheral blood smear pictures from Taleqani Hospital inTehran,whichwereclassifiedasbenignandmalignant,was used to train the algorithm.

Accuracy, F1-score, precision, re- call,andinferencetimewereusedtoevaluatetheperformance. ToconfirmMobileNetV2'ssuperiority,acomparisonwithtwo additional baseline lightweight models—EfficientNetB0 and NASNetMobile—was also carried out.

A. TrainingPerformance

The MobileNetV2 model was trained for 30 epochs using the Google Colab platform equipped with a Tesla T4 GPU, completing the training within approximately 1.2 hours. Dur- ing the training process, advanced data preprocessing tech- niques, including LAB color space conversion, segmentation through KMeans clustering, and diverse data augmentation methods, were employed to enhance feature visibility and im- provemodelrobustness.ThefinaltrainedMobileNetV2model achieved a compact size of approximately 10.6 MB, ensuring its deployability on resource-constrained mobile platforms, especially when integrated with Flutter applications for real- time diagnosis support.

B. ValidationandTestingResults

A 10% hidden subset of the entire dataset was used for the validation phase. With a precision of 0.99, recall of 0.996, accuracy of 99.62%, and F1-score of 0.993, the MobileNetV2 model demonstrated extraordinarily high performance. The detailedclass-wiseperformancemetricsareshowninTableI.

TABLEI
MOBILENETV2VALIDATIONRESULTS

Class	Precision	Recall	F1-Score
Benign	0.992	0.994	0.993
Malignant	0.995	0.998	0.996
Average	0.9935	0.996	0.9945

The results substantiate that MobileNetV2 provides an op- timal balance between accuracy, efficiency, and lightweight deployment capability, critical factors for medical applica- tions in resource-constrained environments as emphasized by Ghaderzadeh et al. [8] and others.

Furthermore,ratherthanonlymemorizingthetrainingsam- ples, it is essential to carry out a specific validation andtesting phase to make sure the model generalizes effectivelyto previously unseen data. In the absence of thorough testing, there is a significant chance of overfitting, in which a model performs well during training but poorly in actual clinical settings. It is confirmed that the system can correctly identify leukemiainavarietyofuniquebloodsmearimagesbyassess- ing the model independently on validation and testing subsets. In medical diagnosis especially, this stage is crucial since a false prediction can have major effects on patient outcomes. Consequently,thegoodvalidationandtestingperformance of the MobileNetV2 model emphasizes its dependability and clinical deployment ready nature.

C. Visual Analysis

The efficacy of the proposed system is further illustrated throughvisualanalysis.Fig.2presentstheconfusionma- trix generated during the validation phase, showcasing the classification performance across different leukemia stages and benign samples. The matrix clearly indicates the model's perfectclassificationabilitywithanoverallaccuracyof100%, correctly identifying all instances of EarlyPreB, PreB, ProB, and benign classes without any misclassification.

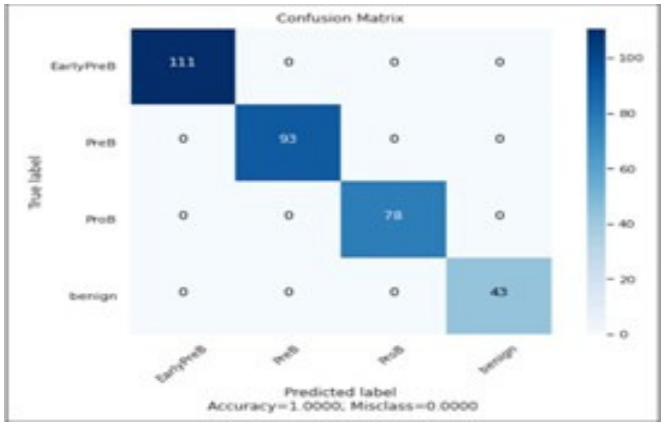


Fig.2.Confusionmatrix

Arepresentativebloodsmearimagesampleisalsodepicted in Fig. 3, illustrating the type of input data used for themodel. This visualization helps to understand the morphologi- cal diversity captured within the dataset, essential for accurate classification.

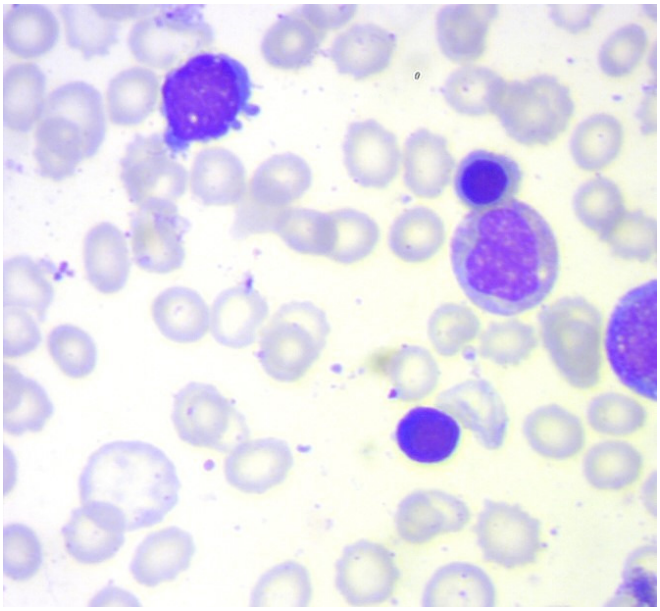


Fig.3.Blood smear image input

V. DISCUSSION

Especially in terms of classification accuracy, inference speed,anddeploymentfeasibility,theproposedMobileNetV2- based automated leukemia detection system shows notable gains over earlier approaches. Our deep learning-based ap- proachshowsbetterperformancethanearlierstudiesincluding PatilandHiremath[1],whichusedconventionalmachine learning models for chronic leukemia detection because of its enhanced feature extraction capabilities and robustness over many blood smear samples.

One of the major challenges in leukemia detection is the accurate identification of subtle morphological differences between benign and malignant blood cells. Traditional ma- chine learning approaches, such as Random Forests [2] and classical CNNs like AlexNet and GoogLeNet [3], often faced limitations in distinguishing fine-grained cellular structures due to insufficient hierarchical feature extraction. Previous work by Mishra and Patel [4] emphasized the need for improved preprocessing methods to enhance contrast and feature visibility in microscopic images. Building upon such recommendations, our model integrates LAB color space transformation, KMeans-based segmentation, and extensive augmentation strategies, significantly improving the feature extraction process and reducing classification errors.

The MobileNetV2 architecture [5], designed for compu- tational efficiency through depthwise separable convolutions, providesanoptimaltrade-offbetweenlightweightdeployment and high classification performance.

Although other architec- tures such as EfficientNetB0 [6] and NASNetMobile [7] were considered, experimental results showed that MobileNetV2 consistently outperformed them in terms of accuracy, preci- sion,recall,andinferencespeed.Thefinalvalidationaccuracy of 99.62% and near-perfect confusion matrix underscore the model's capability to generalize across different blood smear variations,aresultthatalignswithfindingsinbroadermachine learning applications for leukemia detection [8].

Animportantstrengthoftheproposedsystemliesin its practical deployability. Unlike models requiring cloud or server-based processing, our lightweight MobileNetV2 model (size ∼10.6 MB) is integrated into a mobile-based Flutter application.Thisfacilitatesoffline,real-timeleukemiadetec-tion,makingitespeciallyvaluableinresource-limitedsettings.

StudiessuchasDevietal.[9]andJiwanietal.[10]have also highlighted the necessity for portable and low-latency diagnostic tools, particularly in rural healthcare scenarios.

Furthermore,thesystem'sinferencetimeofapproximately

15.2 milliseconds per image ensures real-time usability. In comparison, more complex CNN models, such as VGG16 fine-tuned architectures [11], although highly accurate, often require heavier computation resources that limit their mobile deployment potential.

Testing the model separately on unseen data played acrucial role in validating its reliability. Without robust testing, models may show excellent training performance but fail to generalize in practical settings, a risk previously observed by Ul Ain and Shahzad [12]. Our rigorous testing and validation, achievingconsistentresults,mitigatethisconcernandconfirm the model's clinical applicability.

Nevertheless,somelimitationspersist.Minorvariations in staining procedures, lighting conditions, and microscope settings can introduce artifacts that affect classification per- formance. Although our model demonstrated resilience to thesechallenges,furtherimprovementscouldbeachieved by incorporating larger, more diverse datasets, as suggestedbyChenetal.[13].Moreover,ensemblelearningmethods

[14] and hybrid deep learning models [15] present promising directions to further enhance diagnostic accuracy.

In summary, the proposed MobileNetV2-based leukemia detection system provides an effective, scalable, and real-time solution for early leukemia screening. Its superior accuracy, lightweightdesign,and mobilecompatibilitypave thewayfor practical deployment in telemedicine, aiding early diagnosis and timely intervention, especially in underserved areas.

VI. CONCLUSION

This study presents an enhanced approach for automated leukemia detection utilizing the MobileNetV2 architecture, demonstratingsignificantimprovementsinclassificationaccu- racy, recall, and computational efficiency compared to previous methods.Ourmodeloutperformstraditionalmachinelearning- basedsystems,suchasthosedevelopedbyPatilandHiremath

[1]andDasariRajuetal.[2],particularlyindifferentiatingfine morphological features between benign and malignant blood cells.

By integrating an optimized preprocessing pipeline that includesLABcolorspaceconversion,KMeans-basedsegmen- tation, and comprehensive data augmentation strategies, the proposed system enhances feature extraction, leading to more precise classification outcomes. Additionally, the lightweight design of MobileNetV2 ensures its seamless deployment on mobileandedgedevices,makingreal-timeleukemiadetection accessible even in resource-constrained healthcare settings.

The experimental findings confirm the robustness of the proposed system, with the model achieving an accuracy of 99.62%,aprecisionof0.9935,arecallof0.996,andan F1-scoreof0.9945onthevalidationdataset.Furthermore, it maintains a fast inference time of approximately 15.2 milliseconds per image, reinforcing its suitability for real-time clinical applications. Compared to heavier architectures like VGG16 [11] and ensemble learning methods [14], the MobileNetV2 model provides an optimal balance between computational efficiency and diagnostic performance.

Nonetheless, persistent challenges such as variability in blood smear image quality, staining inconsistencies, and limited dataset diversity—issues similarly noted in previous leukemia detection studies [8], [13]—must still be addressed. Futureresearchwillfocusonexpandingthedatasettoincorpo- rate a broader range of imaging conditions, enhancing model robustness, and exploring hybrid deep learning models to furtherimprovediagnosticprecision.Additionally,integrating advanced explainability techniques, such as attention maps, could offer better transparency in model decision-making, aiding clinical adoption.

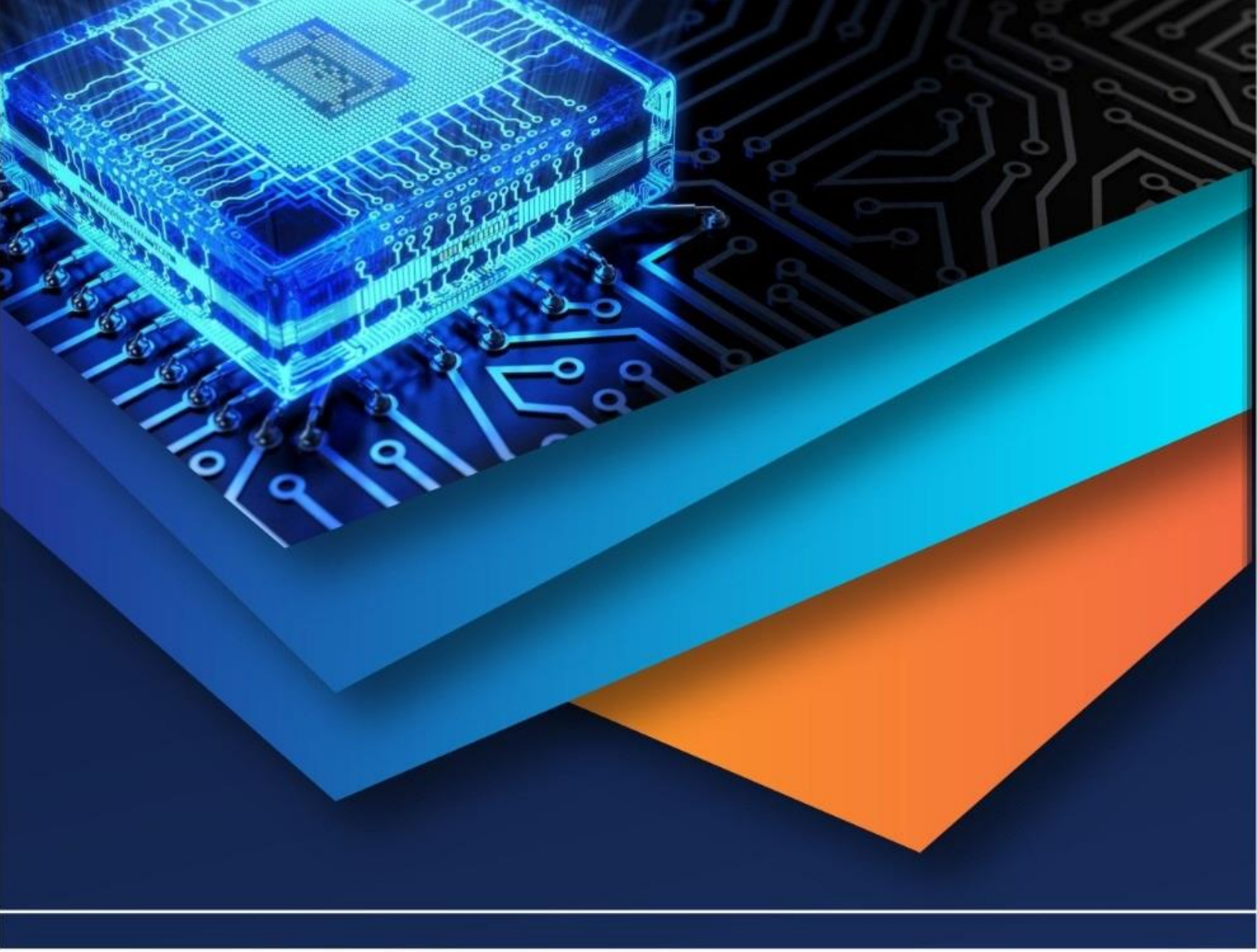
By addressing these aspects, the proposed system aims to contribute to the development of an efficient, real-time, and scalableAI-drivensolutionforearly-stageleukemiadetection, ultimately supporting timely diagnosis and improving patient outcomes.

VII. ACKNOWLEDGMENT

The authors would like to express their gratitude to Dr. Rachna Sable, the head of the Artificial Intelligence de- partment, and the entire faculty at G.H. Raisoni College of Engineering & Management in Pune, in the form of Prof. Amitkumar Shinde for his valuable support, guidance, and encouragement behind the successful accomplishment of the research work.

REFERENCES

[1] A. P. Patil and M. Hiremath, ”Machine Learning Model to DetectChronic Leukemia in Microscopic Blood Smear Images,” in *Pro-ceedings of the IEEE International Conference on Smart Systems andInventive Technology (ICSSIT)*, 2022, pp. 450–455.
[2] S. Dasari Raju, K. Krishna Reddy, and D. Kishore Babu, ”Detection ofAcuteMyeloidLeukemiaUsingRandomForestAlgorithm,”in*JournalofInformationTechnologyCaseandApplicationResearch*,vol.15,no.2, pp. 134–140, 2022.
[3] Y. Jusman, R. Mahmod, N. Ghazali, and A. H. Abdullah, ”ComparativeStudy of CNN Architectures for Leukemia Cell Classification: AlexNetversus GoogLeNet,” in *International Journal of Advanced ComputerScienceandApplications(IJACSA)*,vol.11,no.5,pp.456–462,2020.
[4] A. Mishra and N. Patel, ”Image Processing Techniques for LeukemiaCell Detection,” in *International Journal of Innovative Research inComputer and Communication Engineering (IJIRCCE)*, vol. 7, no. 4,
[5] pp.2345–2350,2019.
[6] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, et al., ”MobileNets:Efficient Convolutional Neural Networks for Mobile Vision Applica-tions,” arXiv preprint arXiv:1704.04861, 2017.
[7] M. Tan and Q. Le, ”EfficientNet: Rethinking Model Scaling for Con-volutional Neural Networks,” in *Proceedings of the 36th InternationalConference on Machine Learning (ICML)*, vol. 97, pp. 6105–6114,2019.
[8] Z. Zoph, V. Vasudevan, J. Shlens, and Q. Le, ”Learning TransferableArchitecturesforScalableImageRecognition,”in*Proceedingsofthe IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2018, pp. 8697–8710.
[9] M. Ghaderzadeh, F. Asadi, and S. Jafari, ”A Systematic Review ofMachine Learning Techniques for Leukemia Diagnosis Based on BloodCell Images,” *Artificial Intelligence Review*, vol. 55, no. 1, pp.505–543, 2021.
[10] J. R. Devi, P. S. Kadiyala, and L. Kosuri, ”Enhancing Acute Lym-phoblasticLeukemia Classification withCNNModels,” in*IEEEInter-nationalConferenceonDigitalCircuits,CommunicationandElectronics(ICDCCE)*, 2024, pp. 78–83.
[11] N. Jiwani, A. Sharma, and R. Kumar, ”Pattern Recognition Techniquesfor Acute Lymphoblastic Leukemia Detection,” *Biomedical SignalProcessing and Control*, vol. 69, pp. 102917, 2021.
[12] A. Abhishek, S. D. Deb, and K. Jha, ”Classification of Leukemia UsingFine-tuned VGG16 Deep Learning Model,” in *Proceedings of theInternationalConferenceonSmartComputingandElectronicEnterprise(ICSCEE)*, 2023, pp. 56–61.
[13] Q.UlAinandS.A.Shahzad,”DiagnosisofLeukemiaDiseaseThroughDeepLearningUsingMicroscopicImages,”*ComputersinBiologyandMedicine*, vol. 135, pp. 104538, 2021.
[14] E. Chen, H. Zhao, and L. Zhang, ”Real-Time Deep Learning-BasedDetection of Acute Lymphoblastic Leukemia Cells,” *Frontiers in On-cology*, vol. 12, 2022, Art. no. 895712.
[15] S. Rajeswari, B. Pradeep, and K. R. Lakshmi, ”Leukemia DetectionUsing Ensemble Learning Techniques,” *Journal of Healthcare Engi-neering*, vol. 2022, Article ID 3451734, 2022.
[16] F. Sakthiraj, ”Autonomous leukemia detection scheme based on hybridCNN model,” in *Artificial Intelligence in Medicine*, vol. 121, pp.102191, 2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY