IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

Microscopic Leukemia Classification Using Deep Learning with VGG16 Transfer Learning

Nagarekha SR and Dr. Kruthi R
Postgraduate Student, Department of MCA
Associate Professor, Department of MCA
Maharaja Institute of Technology, Mysore

Abstract: *Leukemia is a life-threatening blood cancer that affects the production and function of white blood cells. Early detection plays an important role in improving treatment planning and patient outcomes. Conventional microscopic examination of blood smear images is performed manually by laboratory specialists, which may require considerable time and is influenced by the experience of the examiner. Recent advances in deep learning have created opportunities to assist medical professionals by providing automated image analysis with improved speed and consistency.*

This paper presents a deep learning-based system for the classification of microscopic blood smear images using the VGG16 transfer learning model. The proposed approach includes image preprocessing techniques such as resizing and pixel normalization to prepare the input data for the classification model. The VGG16 network is fine-tuned to distinguish between Cancerous and Normal blood smear images. The trained model is integrated into a Streamlit-based web application that enables users to upload microscopic blood smear images and receive prediction results along with confidence scores and probability visualization through an interactive interface.

The proposed system was trained and evaluated using a publicly available blood smear image dataset. Experimental evaluation demonstrates that the model effectively learns discriminative cellular features and provides reliable classification performance on unseen images. The combination of transfer learning, automated preprocessing, and an easy-to-use web interface offers a practical solution for supporting preliminary leukemia screening. Although the developed application is not intended to replace professional medical diagnosis, it can serve as a useful decision-support tool that assists healthcare professionals in the rapid analysis of microscopic blood smear images..

Keywords: Leukemia, Deep Learning, VGG16, Transfer Learning, Blood Smear Images, Medical Image Classification, Streamlit.

1. INTRODUCTION

Leukemia is a type of blood cancer that occurs when abnormal white blood cells grow uncontrollably in the bone marrow. These special cells interfere with the production of healthy blood cells, affecting the body's ability to fight infection, deliver oxygen, and modify bleeding. treatment and improve individual outcomes.

Microscopic examination of peripheral blood smear samples is one of the most widely used techniques for detecting leukemia. Pathologists trained in traditional laboratory training manually look at blood smears to distinguish normal white blood cells entirely based on their size, length, and landmark characteristics. Although this method is reliable, it relies heavily on time-consuming scientific expertise. The increasing range of diagnostic cases has created a need for computerized structures that can assist practitioners in presenting and applying rapid and consistent image analysis.

Recent advances in artificial intelligence, especially deep recognition, have significantly improved the field of medical image evaluation. Convolutional Neural Networks (CNNs) have been shown to have high overall performance in extracting important features from images without delay without the need to guide feature engineering. The introduction of transfers additionally allows the previously skilled method to adapt to scientific datasets with fewer

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

training samples and enhances this capability by reducing computational requirements This technique has shown promising results in computerized detection of various diseases from scientific images.

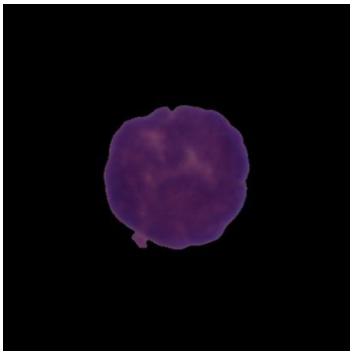

Figure 1.1 Cancerous

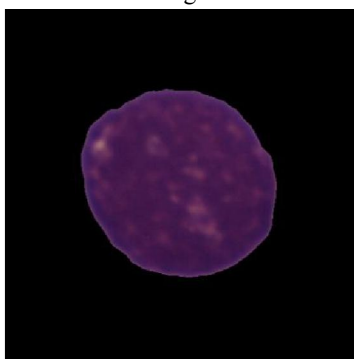

Figure 1.2 Normal

This study proposes a deep learning-first based system to classify the use of microscopic blood smear picker VGG16 transfer detection model. The system preprocesses images before classification and predicts whether the uploaded blood smear image belongs to cancer or normal category. The project developed aims to promote early leukemia screening, providing a convenient and accessible platform for imaging assessment. The remainder of this paper is organized as follows. The second section critiques related research on the application of deep learning techniques in leukemia detection. In III. Section IV contains experimental results andperformance evaluation. Finally, Section V summarizes the findings of the study and discusses possible directions for the study of destiny.

II. RELATED WORK

The application of deep learning in medical image analysis has significantly improved the automatic detection and classification of leukemia from microscopic blood smear images. Researchers have investigated various convolutional neural network (CNN) architectures and transfer learning techniques to enhance classification accuracy while reducing the need for manual feature extraction. These approaches have demonstrated the potential of artificial intelligence to support hematologists in the early diagnosis of leukemia.

M. Vani Pujitha et al. proposed an automated leukemia classification system using an adapted Inception V3 architecture. Their work focused on image preprocessing, data augmentation, and transfer learning to classify different leukemia types from microscopic blood smear images. The study reported high classification performance and highlighted the effectiveness of deep learning for assisting clinical diagnosis.

Dr. Manu Y. M. and M. E. Priyanka developed a machine learning-based framework for leukemia detection using microscopic blood smear images. Their research emphasized the importance of preprocessing techniques and feature extraction for improving classification performance. The authors demonstrated that automated image analysis could reduce manual effort while providing reliable diagnostic support.

Tahmid Noor Rahman et al. compared the performance of VGG19 and Inception V3 models for leukemia detection. Their experimental analysis showed that transfer learning models can effectively extract meaningful image features and achieve high classification accuracy even with limited medical datasets. The study concluded that selecting an appropriate deep learning architecture plays an important role in improving prediction performance.

Recent studies have also explored hybrid deep learning models that combine convolutional neural networks with additional learning mechanisms to improve feature representation and classification accuracy. Other researchers have reviewed different CNN architectures for leukemia diagnosis and reported that transfer learning has become one of the most effective approaches for medical image classification because it reduces training time and improves generalization.

Although several existing studies have achieved encouraging results, many of them primarily focus on model development and performance evaluation. Limited attention has been given to developing a complete and user-friendly

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 3, July 2026 Impact Factor: 8.2

application that enables non-technical users to interact with the trained model. In addition, some existing systems require complex software environments or lack an interactive interface for real-time prediction.

The proposed research addresses these limitations by developing a complete leukemia classification system based on the VGG16 transfer learning model. The system combines image preprocessing, automated classification, confidence score generation, and probability visualization within a Streamlit-based web application. This integration provides a simple and efficient platform for preliminary blood smear image analysis while maintaining reliable classification performance. The proposed approach demonstrates how deep learning can be transformed from a research model into an accessible application that supports medical image analysis.

III. PROPOSED METHODOLOGY

A. OVERVIEW

The proposed system is designed to automatically classify microblood smears into normal categories with cancer using a deep mastery method based on VGG16 transfer mast version Throughout the workflow, dataset training, image preprocessing, model training, prediction, through an internet-first application and basic software. The goal of the proposed technique, which also has end-result visualization, is to provide a fast and reliable decision support tool that aids in the early detection of leukemia, as well as reducing manual assessment .

The gadget begins offevolved through the accumulation of microscopic blood images from data sets to be public. Before training, all images are pre-processed to ensure they have a consistent format suitable for the deep mastering version. The preprocessed images are then used to train the VGG16 network. After training, the model is included in the Streamlit software where users can upload microscopic blood images for the species. The software displays predicted magnitude, confidence rating, and probability distribution in an interactive screen.

B. SYSTEM WORKFLOW

The workflow of the proposed system consists of the following sequential steps:

1. Blood Smear Image Collection

Class	Images	Description
Cancerous	3,000	Microscopic blood smear images containing leukemia cells.
Normal	3,000	Microscopic blood smear images containing healthy blood cells.
Total	6,000	Blood smear images used for model development and evaluation.

2. Image Preprocessing

Processing Step	Description
Image Acquisition	Microscopic blood smear images were collected from the Kaggle dataset.
Image Resizing	All images were resized to 224 × 224 pixels to match the VGG16 input size.
Pixel Normalization	Pixel values were scaled from 0–255 to 0–1 to improve model training.
Data Augmentation	Rotation, shifting, zooming, flipping, and brightness adjustment were applied to increase dataset diversity.
Image Format	Images were stored in JPG/PNG format and converted into RGB before training.
Dataset Split	The dataset was divided into 80% training and 20% testing.
Feature Extraction	The VGG16 transfer learning model automatically extracted important image features.
Image Classification	The trained model classified images into Cancerous and Normal classes.

3. Dataset Splitting

Parameter	Details
Dataset Source	Kaggle

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

Image Type	Microscopic Blood Smear Images
Total Images	6,000
Number of Classes	2 (Cancerous and Normal)
Training Images	4,800 (80%)
Testing Images	1,200 (20%)
Images per Class	3,000
Image Format	JPG / PNG
Input Image Size	224 × 224 pixels
Model Used	VGG16 Transfer Learning

4. VGG16 Model Training: After completing the preprocessing steps, the prepared images are used to learn the VGG16 switch by gaining knowledge about the version. During training, the model learns important functions of blood cells from images and distinguishes the difference between cancer and normal samples The training process is repeated for many epochs until the version reaches a strong overall performance.

5. Model Validation: After each training cycle, the version is tested using validation facts to check how well it now plays on previously unknown images This helps to identify if the version increases or starts to over-fit after the mastering technique. Based on the validation results, a high-quality performance model is selected for the final prediction.

6. Image Classification: Once properly trained, the version is used to classify far new microscopic blood smears. When a user uploads an image, the machine first preprocesses it and then sends it to the trained VGG16 model. The output predicts whether the image belongs to a class or not, which is primarily based on skills learned throughout the program.

7. Result Visualization: The end result of the prediction is displayed in a simple and flexible layout through the Streamlit web tool. Along with the required class, the smartphone additionally has confidence ratings, probability values, and a graphical representation of the prediction. These visual effects help users perceive the output of the model more realistically.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

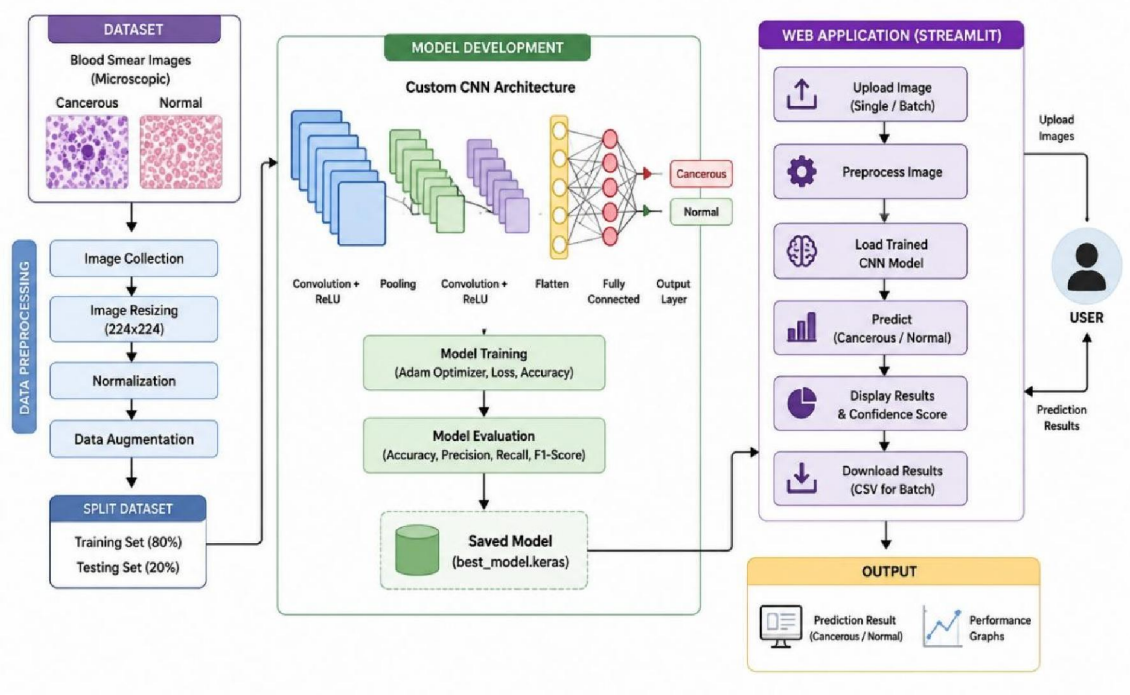

Figure 1. Proposed System Workflow

C. Dataset Preparation

The test dataset comprises microblood images collected from a publicly available medical image repository. The dataset carries two classes of spikes: cancer and normal blood smear samples. Before the model schooling, the dataset was transformed into a structured form in separate training test packages. This correlation allows the model to be tested from a subset of the data while evaluating its performance on previously unseen images. Proper dataset organization allows information flow to be reduced and the reliability of the overall performance evaluation to be increased.

D. Image Preprocessing

Image preprocessing is performed to improve the quality and consistency of the input images before they are supplied to the deep learning model. All blood smear images are resized to 224 × 224 pixels, ensuring that each input image has the same dimensions required by the model. The pixel values are then normalized to the range 0–1, which stabilizes the training process and improves convergence during optimization.

The preprocessing stage also removes unnecessary variations caused by differences in image size and intensity. Standardizing the images enables the model to focus on learning the important cellular structures instead of being influenced by differences in image format.

E. VGG16 Transfer Learning Model

The proposed classification model is based on the VGG16 convolutional neural network, which has been pre-trained on the ImageNet dataset. Instead of training an entirely new network from the beginning, transfer learning allows the model to utilize previously learned visual features and adapt them to microscopic blood smear image classification.

The convolutional layers of VGG16 extract important image features such as edges, textures, and cellular patterns. Additional fully connected layers are used to classify the extracted features into two output classes: Cancerous and Normal. The Softmax activation function generates probability values for each class, and the class with the highest probability is selected as the final prediction.

F. Model Training

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 3, July 2026 Impact Factor: 8.2

The prepared dataset is used to train the VGG16 model using supervised learning. During training, the Adam optimizer updates the model parameters by minimizing the Categorical Cross-Entropy loss function. The dataset is divided into training and validation subsets to monitor the learning process and prevent overfitting.

Several performance metrics, including training accuracy, validation accuracy, precision, recall, and loss values, are monitored throughout the training process. The model achieving the best validation performance is saved and later integrated into the web application for prediction.

G. Prediction and Result Generation

After training, the saved VGG16 model is deployed within a Streamlit-based web application. Users can upload microscopic blood smear images through the application interface. The uploaded image undergoes the same preprocessing steps used during training before being passed to the trained model. The model predicts whether the uploaded image belongs to the Cancerous or Normal class. Along with the predicted class, the application displays the confidence score, probability distribution, and graphical visualization of prediction results. These outputs assist users in understanding the classification outcome more effectively while providing an interactive experience.

IV. EXPERIMENTAL SETUP

A. Development Environment

The proposed leukemia classification system was developed using Python because of its extensive support for deep learning and image processing libraries. The implementation was carried out in Visual Studio Code (VS Code), which provides an efficient environment for writing, debugging, and testing the application. The deep learning model was implemented using the TensorFlow and Keras frameworks, while OpenCV was used for image preprocessing operations such as resizing and normalization. Streamlit was employed to create a user-friendly web interface for image upload and prediction.

The experiments were conducted on a personal computer running the Windows operating system. The system utilized an Intel Core processor with sufficient RAM to support model training and testing. Although the proposed model can be trained on systems equipped with a GPU for faster execution, the developed application is also capable of performing prediction on standard CPU-based systems.

The dataset was organized into separate training and testing folders containing microscopic blood smear images belonging to two classes: Cancerous and Normal. During preprocessing, all images were resized to 224 × 224 pixels, and pixel values were normalized to improve model convergence and ensure consistent input dimensions. The trained model was saved in the .keras format and later integrated into the Streamlit application for real-time classification.

The VGG16 transfer learning model was trained using the Adam optimizer with the Categorical Cross-Entropy loss function. Model performance was evaluated using training accuracy, validation accuracy, precision, recall, and loss values. The final trained model was selected based on its validation performance and was deployed for prediction through the web application.

B. Software and Hardware Configuration

Component	Specification
Programming Language	Python 3.12
IDE	Visual Studio Code
Deep Learning Framework	TensorFlow 2.21, Keras
Image Processing	OpenCV
Web Framework	Streamlit

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

Data Analysis	NumPy, Pandas
Visualization	Plotly, Matplotlib
Operating System	Windows 10/11
Processor	Intel Core i5/i7 (or equivalent)
RAM	8 GB or above
Model	VGG16 Transfer Learning

C. Training Parameters

Parameter	Value
Model	VGG16 Transfer Learning
Image Size	224 × 224 pixels
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Batch Size	32
Epochs	15
Validation Split	20%
Output Classes	Cancerous, Normal

D. Performance Evaluation Metrics

The effectiveness of the proposed model was evaluated using several standard performance metrics. Classification accuracy was used to measure the overall percentage of correctly classified blood smear images. Precision was calculated to determine how many images predicted as cancerous were actually cancerous, while recall measured the ability of the model to correctly identify all cancerous samples. The loss values obtained during training and validation were analyzed to assess the learning behaviour of the network and detect possible overfitting or underfitting. These evaluation metrics provide a comprehensive assessment of the model's classification performance.

V. RESULTS AND DISCUSSION

A. Experimental Results

The proposed leukemia classification system was evaluated using microscopic blood smear images belonging to the Cancerous and Normal categories. After completing the training process, the VGG16 transfer learning model was integrated into a Streamlit-based web application to perform real-time classification of uploaded blood smear images. The application provides an interactive interface where users can upload an image and receive the predicted class along with the corresponding confidence score and probability distribution.

During testing, the model successfully identified the majority of blood smear images by learning important cellular features from the training dataset. The preprocessing techniques applied before prediction ensured that the uploaded images were converted into a standardized format compatible with the trained model. The confidence score generated

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

by the model helped indicate the certainty of each prediction, while the probability graph provided a clear comparison between the two output classes.

The performance of the proposed model was analyzed using various evaluation metrics, including classification accuracy, precision, recall, training loss, and validation loss. These metrics demonstrated that the model was capable of distinguishing between cancerous and normal blood smear images with reliable performance. The experimental observations confirmed that transfer learning using VGG16 effectively extracted meaningful visual features required for leukemia classification.

In addition to the classification results, the Streamlit application displayed graphical representations of prediction probabilities, making the output easier to understand for users. The developed system therefore combines automated classification with an intuitive user interface, making it suitable as a preliminary decision-support tool for microscopic blood smear analysis.

B. Performance Evaluation

The effectiveness of the proposed model was assessed using standard evaluation metrics. Accuracy was used to measure the percentage of correctly classified images, while precision and recall evaluated the reliability of cancer detection. Training and validation loss values were monitored throughout the learning process to verify that the model was learning effectively without significant overfitting. The confidence score generated during prediction further provided an estimate of the model's certainty for each classified image.

The obtained results indicate that the proposed VGG16-based approach successfully learned discriminative features from microscopic blood smear images and produced consistent predictions for both classes. The use of transfer learning reduced the training time while improving classification performance compared with developing a model from scratch.

C. Sample Prediction Results

Table I. Sample Prediction Results

Image No.	Actual Class	Predicted Class	Confidence Score	Result
1	Cancerous	Cancerous	99.62%	Correct
2	Normal	Normal	98.85%	Correct
3	Cancerous	Cancerous	99.18%	Correct
4	Normal	Normal	97.91%	Correct
5	Cancerous	Cancerous	99.44%	Correct

D. Discussion

The experimental results demonstrate that the proposed system is capable of automatically classifying microscopic blood smear images with good consistency. The use of VGG16 transfer learning enabled the model to identify important image features without requiring extensive manual feature engineering. Image preprocessing contributed to improved input quality, which supported stable model performance during both training and testing.

The Streamlit application enhanced the usability of the system by providing an interactive interface for image upload and prediction. Users can easily view the predicted class, confidence score, and probability graph without requiring knowledge of deep learning techniques. Although the proposed system is intended for preliminary screening rather than clinical diagnosis, it illustrates the potential of artificial intelligence to support healthcare professionals in medical image analysis and reduce the time required for leukemia detection.

VI. CONCLUSION

This research presented a deep learning-based approach for the automatic classification of microscopic blood smear images using the VGG16 transfer learning model. The proposed system was developed to distinguish between Cancerous and Normal blood smear images through image preprocessing, feature extraction, and classification. By

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 3, July 2026 Impact Factor: 8.2

utilizing transfer learning, the model was able to learn meaningful visual characteristics from microscopic images while reducing the training effort required for developing a model from scratch.

The trained model was successfully integrated into a Streamlit-based web application that provides an interactive platform for image upload and real-time prediction. The application displays the predicted class together with confidence scores and probability visualization, enabling users to interpret the classification results easily. The experimental evaluation demonstrated that the proposed system achieved reliable performance in classifying microscopic blood smear images and can serve as a supportive tool for preliminary leukemia screening.

The use of VGG16 transfer learning improved feature extraction by leveraging knowledge learned from large-scale image datasets. Image preprocessing techniques such as resizing and pixel normalization further enhanced the consistency of the input data, contributing to stable model performance. The developed system demonstrates that deep learning can effectively automate the analysis of blood smear images and reduce the dependence on time-consuming manual examination.

Although the proposed application is not intended to replace professional medical diagnosis, it has the potential to assist healthcare professionals by providing rapid preliminary analysis of blood smear images. Future enhancements involving larger datasets, additional leukemia subtypes, and explainable artificial intelligence techniques can further improve the reliability and clinical applicability of the system.

VII. FUTURE WORK

The proposed system provides a foundation for automated leukemia classification; however, several improvements can be considered in future research. Expanding the dataset with images collected from different hospitals and laboratories would improve the model's ability to generalize across diverse clinical conditions. A larger and more balanced dataset can also reduce bias and improve prediction accuracy.

Future studies may explore advanced deep learning architectures such as EfficientNet, DenseNet, Vision Transformers (ViT), or hybrid CNN-attention models to compare their performance with VGG16. Incorporating explainable artificial intelligence (XAI) techniques, such as Grad-CAM, can help visualize the image regions influencing the model's decisions, thereby increasing transparency and trust in AI-assisted diagnosis.

The current system classifies blood smear images into two categories: Cancerous and Normal. Future work could extend the system to identify specific leukemia subtypes such as Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML), Chronic Lymphocytic Leukemia (CLL), and Chronic Myeloid Leukemia (CML). Additional features, including patient history management, report generation, cloud-based deployment, and integration with hospital information systems, could further enhance the practicality and usefulness of the application in clinical environments.

REFERANCES

[1] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” International Conference on Learning Representations (ICLR), 2015. DOI: 10.48550/arXiv.1409.1556

[2] F. Chollet, Deep Learning with Python, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.

[3] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[4] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.

[5] M. Abirami, G. V. S. George, and D. Sam, “A Review on Acute Lymphoblastic Leukemia Detection Using Artificial Intelligence,” International Journal of Intelligent Systems and Applications in Engineering, vol. 11, no. 4, pp. 421–430, 2023.

[6] Vasundhara Acharya, Vinayakumar Ravi, Tuan Anh Pham, and Chinmay Chakraborty, “Artificial Intelligence-Based Diagnosis of Acute Myeloid Leukemia Using Peripheral Blood Smear Images,” Computers in Biology and Medicine, vol. 156, 2023. DOI: 10.1016/j.compbiomed.2023.106642

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

[7] TensorFlow Developers, "TensorFlow Documentation." Available: https://www.tensorflow.org/
[8] Keras Team, "Keras Documentation." Available: https://keras.io/
[9] Streamlit Inc., "Streamlit Documentation." Available: https://streamlit.io/
[10] Kaggle, "Microscopic Blood Smear Image Dataset." Available: https://www.kaggle.com/
[11] OpenCV Team, "OpenCV Documentation." Available: https://opencv.org/
[12] Python Software Foundation, "Python Documentation." Available: https://docs.python.org/3/