Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 10, 2025

FEATURE-BASED MACHINE LEARNING FRAMEWORK FOR THE DETECTION AND CLASSIFICATION OF ACUTE LYMPHOBLASTIC LEUKEMIA

[1]Maryam Bilal, *[2]Fahima Tahir, [3]Wajeeha Malik, [4]Rabia Javed, [5]Huma Tauseef, [6]Khansa Aatif

[1]Department of Computer Science, Lahore College for Women University, Lahore 54000, Pakistan
*[*2]Department of Computer Science, Lahore College for Women University, Lahore 54000, Pakistan*
[3]Department of Computer Science, Lahore College for Women University, Lahore 54000, Pakistan
[4]Department of Computer Science, Lahore College for Women University, Lahore 54000, Pakistan
[5]Department of Computer Science, Lahore College for Women University, Lahore 54000, Pakistan
[6]Department of Computer Science, Lahore College for Women University, Lahore 54000, Pakistan

DOI: https://doi.org/10.5281/zenodo.17403387

Keywords
Microscopic Blood Smear Images, Spatial Features, Acute Lymphoblastic Leukemia, Color Features, Segmentation, Machine Learning, Prediction

Article History
Received on 18 Sept,2025
Accepted on 16 Oct 2025
Published on 17 Oct 2025

Corresponding Author: * Fahima Tahir

Abstract

Early detection of Acute Lymphoblastic Leukemia (ALL) is essential for effective diagnosis and timely treatment. This study presents an automated framework for leukemia classification using microscopic blood smear images combined with machine learning techniques. Spatial features, including shape and texture, along with color-based descriptors, were extracted from microscopic images and analyzed both individually and in fused combinations. Several classifiers, such as k-Nearest Neighbor, Support Vector Machine, and Decision Tree, were employed to evaluate model performance. The results revealed that individual feature sets provided moderate accuracy, while the fusion of all three feature types substantially improved classification performance. Among the tested models, the Support Vector Machine achieved the highest accuracy of 100% using the combined feature set, reflecting its strong generalization capability. The proposed approach offers a cost-effective, accurate, and scalable diagnostic solution, highlighting its potential for integration into clinical workflows and automated leukemia screening in hematological analysis.

1. INTRODUCTION

Cancer is a life-threatening disease caused by the uncontrolled proliferation of ab-normal cells that invade and destroy healthy tissues, thereby disrupting normal physio-logical functions [1]. Among various cancer types, blood cancer is one of the most critical, as it directly impairs the production and functionality of blood cells. The three major forms of blood cancer are leukemia, lymphoma, and myeloma. Leukemia is characterized by the abnormal overproduction of white blood cells, myeloma affects plasma cells in the bone marrow, and lymphoma originates in the lymphatic system, primarily targeting lymphocytes. Each type interferes with normal hematopoietic processes and leads to severe health complications [2].

Leukemia begins in the bone marrow, the soft, spongy tissue inside bones where blood cell formation occurs. In this disease, the bone marrow produces excessive numbers of abnormal white blood cells (WBCs), which hinder the generation of healthy blood cells [3], [4]. The bone marrow contains a complex structure composed of blood vessels, hematopoietic stem cells, fat cells, and fibrous tissue [5]. Blood stem cells differentiate into lymphoid or myeloid stem cells—

lymphoid stem cells generate lymphocytes (a type of WBC), while myeloid stem cells produce red blood cells, platelets, granulocytes, and monocytes. Leukemia develops when stem cells begin producing large numbers of immature white blood cells, known as blast cells, which can multiply un-controllably—where even a single blast cell may give rise to billions of others [6].

Leukemia is classified based on the rate of disease progression and the type of stem cell affected. Abnormal lymphoid stem cell production leads to lymphocytic leukemia, whereas irregular myeloid stem cell production results in myelogenous leukemia [6][7]. Rapidly progressing leukemia is termed acute leukemia, while slow-developing forms are known as chronic leukemia. The four primary types of leukemia include chronic lymphocytic leukemia (CLL), chronic myelogenous leukemia (CML), acute myelogenous leukemia (AML), and acute lymphocytic leukemia (ALL).

Acute lymphoblastic leukemia (ALL) originates from immature white blood cells, known as lymphoblasts, within the bone marrow. This aggressive hematologic malignancy is characterized by the rapid and uncontrolled proliferation of abnormal lymphocytes, which replace normal hematopoietic cells and impair healthy blood cell formation. As a result, patients commonly present with symptoms such as fatigue, frequent infections, easy bruising or bleeding, and bone or joint pain. If not diagnosed and treated promptly, ALL can progress rapidly and become life-threatening within a few months. Although it can affect individuals across all age groups, ALL predominantly occurs in children, making it the most common pediatric leukemia [8].

According to the French-American-British (FAB) classification system, acute lymphoblastic leukemia is divided into three subtypes—ALL-L1, ALL-L2, and ALL-L3—based on the morphological characteristics of the leukemia cells, including their size, shape, and nuclear appearance. In ALL-L1, the white blood cells are typically small, with homogeneous nuclear chromatin and a regular, round nucleus that occasionally exhibits a slight indentation or cleft. [9]. Figure 1a illustrates blood cells exhibiting the typical morphology of the ALL-L1 subtype.

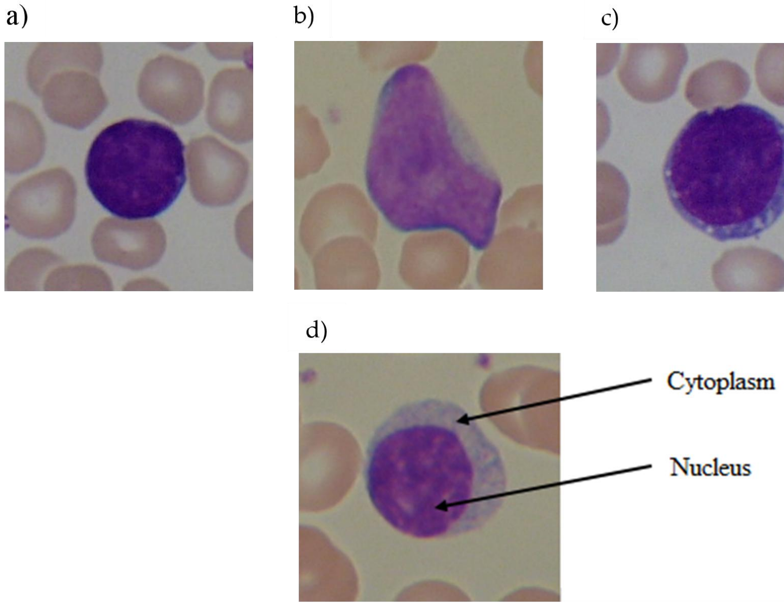

FIGURE 1: MICROSCOPIC BLOOD SMEAR IMAGES OF A) TYPE 1 LEUKEMIA (ALL-L1) B) TYPE 2 LEUKEMIA (ALL-L2) C) TYPE 3 LEUKEMIA (ALL-L3) D) HEALTHY CELL

ALL-L2: WBC's belonging to ALL-2 are large with heterogeneous variable nuclear chromatin and the shape of nuclear is irregular. The amount of cytoplasm is variable, and one or more nucleoli are present in the cell [9]. Figure 1b illustrates the blood cells having ALL-L2.

ALL-L3: WBC's belonging to ALL-L3 are large with homogeneous fine stippled chromatin. The nucleoli of the cell is regular and prominent and cytoplasm is abundant, most distinguishing feature in ALL-L3 is cytoplasmic vacuolation [9]. Figure 1c illustrates the blood cells having ALL-L3 and Figure 1d shows a microscopic blood smear image (MBSI) of a healthy individual.

Various methods have been proposed in the literature for the classification, seg-mentation, and feature extraction of microscopic images. Among these, segmentation plays a crucial role, as its accuracy has a direct impact on the overall performance of the system. The authors introduced a methodology for blast MBSI classification by creating two classes based on local pixel information derived from statistical features [10]. Similarly, in another research authors developed an approach that integrates both col-or-based and morphological segmentation techniques to improve accuracy and robust-ness.[11]. In another research proposed in which they used MBSI and those MBSI were processed by CIELAB color conversion algorithm for accurate segmentation. For Nuclei segmentation they used the k-mean algorithm and then used Local Directional Pattern. Several bin values were calculated for the construction of histogram. Support vector machine (SVM) classifier was applied on normalized values of histogram to detect normal and abnormal cells [12]. To segment the MBSI Amin et al., used k-mean algorithm. Segmented images were used to generate different statistical and geometric at-tributes, and multi-SVM classifier was used to classify healthy cell, ALL-L1, ALL-L2, and ALL-L3 [13].

The segmentation algorithm proposed to detect acute leukemia. Their segmentation algorithm was based on Otsu's threshold method. After that they extract three features which are perimeter, area and circularity, k-Nearest Neighbor (k-NN) classifier used to classify the normal and cancerous lymphocytes [14]. Red, green blue (RGB) to cyan, magenta, yellow, black (CMYK) color model conversion was used in research to identify WBC. Zack algorithm was used for segmentation. In feature extraction phase attributes such as color, shape and texture were extracted and SVM classifier used to classify leucocytes [15]. An approach proposed in which they used Otsu's method for segmentation. NN classifier was used to classify WBC [16]. Deep belief network approach was used for segmentation in research which was proposed by Duggal et al., in which they apply Lab color space using k-mean algorithm. After that they determine the connected components of MBSI. If any of connected components are found, then break this otherwise it was correctly segmented [17]. HSI color model is a technique used for segmentation of MBSI, this technique was used by Singhal et al., in their re-search. They extracted shape features and correlation-based feature selection method was used for feature selection then Naive Bayesian, SVM, k-NN, Radial Basis Function Neural Network and multilayer perceptron (MLP) classifier were applied on selected features [18]. Amin et al., worked on images and proposed a method for ALL and its cell subtypes detection in which fuzzy c-means clustering algorithm was used for segmentation of cell nucleus. First and second-order geometric statistical features were extracted. SVM classifier was used with parameters and different kernels in which it classifies the cells into either cancerous (L1, L2, L3) or noncancerous [19]. To Detect the cancer cells a method proposed in which RGB color image converted into gray scale image. Marker controlled watershed segmentation was applied on enhanced image [20].

In another research marker-based watershed segmentation technique was applied on MBSI. Texture features were extracted from the cytoplasm and nucleus. Probabilistic principal component analysis (PPCA) was applied to get most relevant attributes. Random forest classifier was used for classification [21]. Moreover, the proposed methodology in which they reduce noise MBSI then applied Lab color space for segmentation using k-mean algorithm. Texture and shape-based characteristics were extracted and identify shape of the cell by using k-mean [22]. The authors used marker-controlled watershed algorithm for segmentation, they extracted color, texture and geometrical features. SVM classifier was used to sort the images into healthy and blast cell [23].

Convolutional Neural Networks (CNN) applied on original dataset images for feature extraction. After feature extraction they performed feature reduction using Gain ratio then classified the images using SVM classifier [24]. K-mean algorithm was used for segmentation. Attributes such as color, texture and shape were obtained from segmented images. cluster of differentiation (CD) Marker based technique was used for classification [25]. Research which was proposed deal with the counting of the WBC's, in which they segmented images using Otsu thresholding then Circle Hough Transform was applied to count WBCs [26].

Another research that was again performed the prediction of ALL using blood smear images. They have used transfer learning algorithms i.e. Inception-V3, VGG-16 and ResNet50. To minimize the error rate, they have applied data augmentation and preprocessing. According to their research ResNet50 outer performed [27]. Sampathila et al. predicted leukemia using CNN. They have used Nvidia Tesla P100 GPU method from Google Collaboratory for the training of a custom ALLNET model and achieved 95.54% accuracy [28]. Rangini et al. presented the prediction of ALL using textural features and machine learning algorithms. They applied GLCM and First Order Histograms as textural features along with ANN, NB and logistic regression as classifiers. Results show that ANN along with GLCM outer performs [29].

Most previous research focuses on classifying healthy WBCs and those affected by (ALL), but it often does not address the different types of ALL. There is a need to predict ALL and its subtypes more accurately. Our research aims to address this gap by accurately predicting ALL and its subtypes (L1, L2, and L3) using blood smear images with maximum accuracy.

The detailed description of dataset used and proposed methodology is explained in Material and Methods section, after that Results and Discussion section explains the complete set of experiments performed in this study.

2. MATERIALS AND METHODS

2.1. Data Set Description

This research used a dataset, which was gathered from ALL-IDB, a publicly available website. Each image in the dataset has been categorized by expert oncologists [6]. In this research 86 images are used for classification purpose from which 43 images are healthy, 15 are from ALL-L1, 19 are from ALL-L2 and 09 are from ALL-L3.

2.2. Proposed Methodology

The proposed methodology comprises four main stages: preprocessing, segmentation, feature extraction, and classification. Figure 2 presents an overview of the work-flow. Microscopic Blood Smear Images (MBSI) are employed in this study to detect Acute Lymphoblastic Leukemia (ALL) and its subtypes—ALL-L1, ALL-L2, and ALL-L3. The objective of this research is to support hematologists in accurately identifying ALL and its specific variants. A detailed representation of the complete methodology is illustrated in Figure 3.

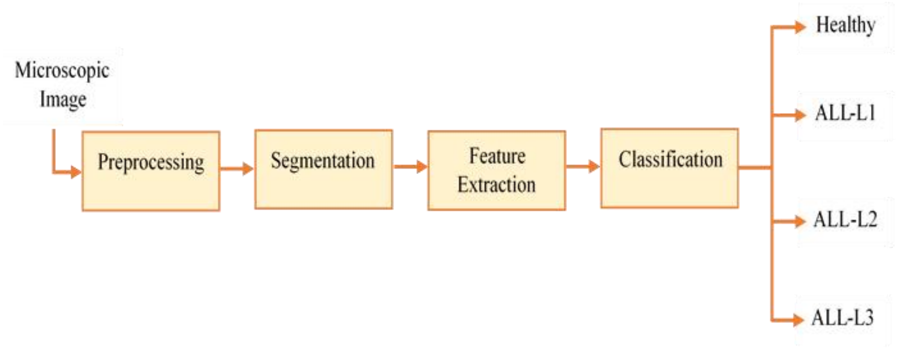

FIGURE 2: BRIEF FLOW OF THE PROPOSED METHODOLOGY FOR THE PREDICTION OF LEUKEMIA

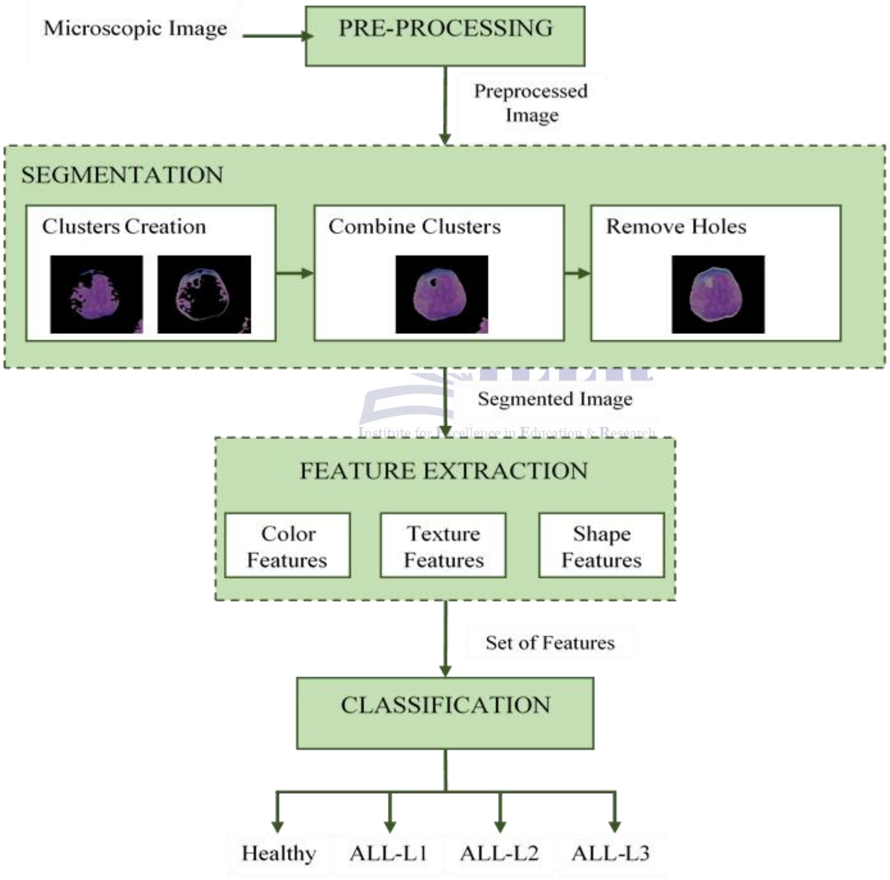

FIGURE 3: DETAILED FLOW OF THE PROPOSED METHODOLOGY FOR THE PREDICTION OF LEUKEMIA

2.2.1 Preprocessing

The early stage of the planned methodology involves preprocessing the MBSI. This phase is crucial as it prepares the images for subsequent stages of analysis. In our approach, scaling was performed using Bulk Image Resizer [30].

2.2.2 Segmentation

Segmentation is the second critical phase of the proposed research, as it significantly impacts the overall accuracy of the system. In the proposed methodology MBSI was segmented on MATLAB R2017a [31].

Segmentation of the research has three basic steps which are:

- Cluster Creation
- Combine nucleus and cytoplasm cluster
- Remove holes and extra regions

2.2.2.1 Cluster Creation

In cluster creation step, we put the same color pixels in one cluster. By using k-mean algorithm on MATLAB make four different colors clusters, first cluster for red blood cells, second cluster for the Nucleus of the WBC, third cluster for cytoplasm of the WBC and last one for background of the image [22]. Figure 4a) illustrates the cluster of the Nucleus of WBC and Figure 4b) illustrates the cluster of the cytoplasm of WBC.

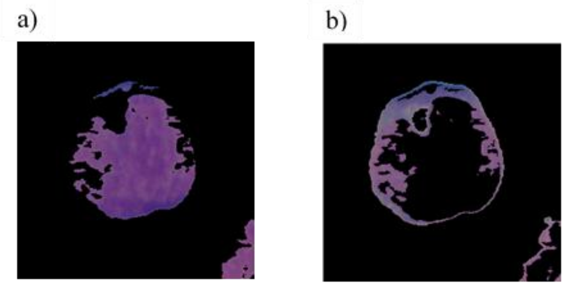

FIGURE 4: REPRESENTATION OF NUCLEUS AND CYTOPLASM CLUSTERS OF WBC. 4A) NUCLEUS 4B) CYTOPLASM

2.2.2.2 Combine Nucleus and Cytoplasm Cluster

After creation of cluster, we combine the nucleus and cytoplasm cluster to make the complete WBC. Nucleus of WBC and cytoplasm of WBC are combined by using MATLAB function [31]. Figure 5 illustrates after combining the nucleus and cytoplasm.

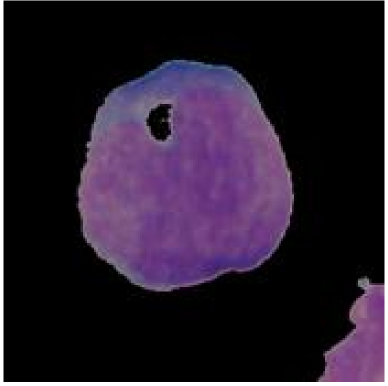

FIGURE 5: COMBINING THE NUCLEUS AND CYTOPLASM CLUSTER

2.2.2.3 Remove Holes and Extra Regions

After combining the nucleus and cytoplasm there are some holes in the image that need to be filled so those holes are filled. After filling the holes there are some extra regions in the image that need to be removed so those regions are removed. Holes and Extra regions are removed by using MATLAB functions [31]. Figure 6 illustrates after re-moving holes and extra region.

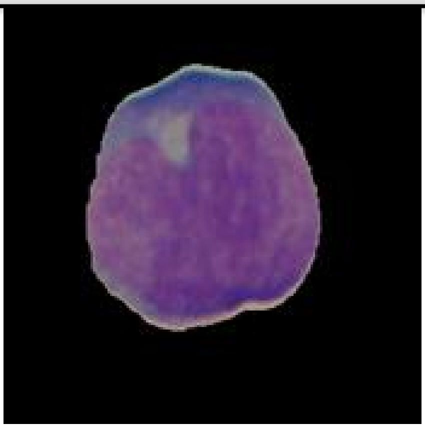

FIGURE 6: REMOVING HOLES AND EXTRA REGIONS

2.2.3 Feature Extraction

The third phase of proposed methodology is featuring extraction. Feature extraction phase converts the segmented image into meaningful values with attributes and parameters. These values are very important to get the maximum accuracy in classification phase. In our proposed methodology three types of features are extracted which are:

- Texture features
- Shape features
- Color features

2.2.3.1 Texture Features

Texture features are a type of spatial feature highlighting the texture of the surface. In our proposed methodology 278 textural features are extracted which provide variety of statistical parameters such as co-occurrence matrix, histogram-based features, run-length matrix, autoregressive model, absolute gradient, and wavelet [32].

Absolute Gradient: Gradient based features are gradient variance, gradient mean, gradient kurtosis and gradient skewness [32].

Histogram Based Features: Histogram based features are variance, mean, skewness and kurtosis [32].

Autoregressive Model: The local interactions were assumed between pixel intensity and sum of the intensities of the neighboring pixels using autoregressive (AR) model [32].

Co-Occurrence Matrix: The Co-Occurrence Matrix (COM) is a second-order histogram that provides a detailed statistical description of the spatial relationships between neighboring pixels within a region of interest. By examining how often pairs of pixel values occur together, the matrix estimates the joint probability distribution of these pixel pairs. COM features include several metrics that describe different aspects of pixel value distributions and their spatial arrangements. COM features include entropy, contrast, correlation, angular second moment, sum average, sum entropy, sum of squares, sum variance, inverse difference moments, difference variance and difference entropy [32].

Run length matrix-based (RLM) Features: These features are calculated by the number of pixels having the same grey level in a run. RLM features are long run emphasis moments, short run emphasis moments, run length nonuniformity, gray level nonuniformity and fraction of images in runs [32].

Wavelet Parameters: The discrete wavelet transforms (DWT) is a linear transformation which transforms numerically different vectors of the same length as input data vectors. It separates data into different frequency components. DWT is computed with low and high-pas filters. MaZda is used to measure these energy features at eight scales using four bands of frequency (LL, LH, HL and HH) [32].

2.2.3.2 Shape Features

Shape features define the region-based features of the image also belong to the category of spatial features. In the proposed methodology area, perimeter, centroid, equivalent diameter, solidity, roundness and bounding box were calculated by using MATLAB function region props. Area is calculated by counting actual number of pixels in the region and the output is a scalar number which is on in the binary image. Perimeter is the distance around the boundary. It is calculated by using binary image and the output is the scalar number. Centroid provides center of region of interest (ROI) image and result is the two scalar numbers one for x-axis and second for y-axis. Equivalent Diameter provides the diameter of the ROI image, and the output is a scalar number. Solidity provides the proportion of the pixels in the convex hull, and the output is a scale number. Bounding Box is the smallest rectangle which provides four scalar numbers which are from the smallest rectangle [33]. Roundness provides the information about the ROI image is how round.

2.2.3.3 Color Features

Color features define the ratio of the color in the images. In our proposed method-ology color moments are calculated by using MATLAB R2017a. Color moments which are calculated in our proposed research are mean, standard deviation, skewness and kurtosis. These are calculated for red, green and blue color separately so 12 color moments were calculated.

2.2.4 Classification

Classification is the last phase of our proposed methodology. This phase is crucial as it serves as the foundational element of the research. During this stage, the images are classified into healthy categories i.e. ALL-L1, ALL-L2, ALL-L3 based on features extracked. For classification, three types of classifiers are used in planned research, which are SVM, k-NN and Decision tree to get the best result. These classifiers are applied in MATLAB where 60% of dataset is used for training and 40% of dataset is used for testing.

3. RESULTS AND DISCUSSION

In this proposed research 86 MBSI are used. These images are separated into two groups, which are Training and Testing. 60% of datasets are used for training purpose and 40% is used for testing purposes. The results of the proposed methodology are based on three types of experiments.

3.1 Experiment 1

In this experiment, a total of 86 microscopic blood smear images (MBSI) were utilized, comprising 43 healthy and 43 cancerous samples. The cancerous images included samples from ALL-L1, ALL-L2, and ALL-L3 categories. The classification was initially performed on individual feature sets, namely shape, color, and texture features. Subsequently, pairwise combinations of feature sets were evaluated — shape + color, texture + shape, and texture + color — to analyze their combined discriminative capability. Finally, all three feature sets (shape, color, and texture) were fused into a comprehensive feature representation, upon which classifiers were applied to achieve the maximum possible accuracy. This Experiment is performed on MATLAB R2017a [31]. The results are shown in Table 1.

TABLE 1: PREDICTION RESULTS OF HEALTHY VS BLAST WBCS

Features	No. of classes	K-NN	Decision Tree	SVM
Shape	2	79.4%	70.6%	79.4%
Color	2	73.6%	70.6%	82.4%
Texture	2	82.4%	76.5%	82.4%

Shape and color	2	79.4%	70.6%	73.5%
Texture and shape	2	85.3%	76.5%	91.2%
Texture and color	2	85.3%	91.2%	91.2%
Texture, Color and Shape	2	85.3%	76.5%	91.2%

When we apply classifiers on shape features SVM and k-NN both provide the maximum accuracy which is 79.4 percent. After this we apply classifiers on color features SVM provide maximum accuracy which is 82.4 percent then we apply classifiers on texture features k-NN and SVM both provide maximum accuracy which is 82.4 per-cent. After this we applied classifiers on the pair of two sets of features first pair is shape and color features, from this k-NN provide maximum accuracy that is 79.4 per-cent. When applying on texture and shape, SVM provide maximum accuracy which is 91.2 percent. When classifiers applied on texture and color features SVM and decision tree both provide maximum accuracy 91.2 percent. In the last we apply these classifiers on all features SVM provide maximum accuracy 91.2 percent. Figure 7 illustrates the graph of the experiment 1.

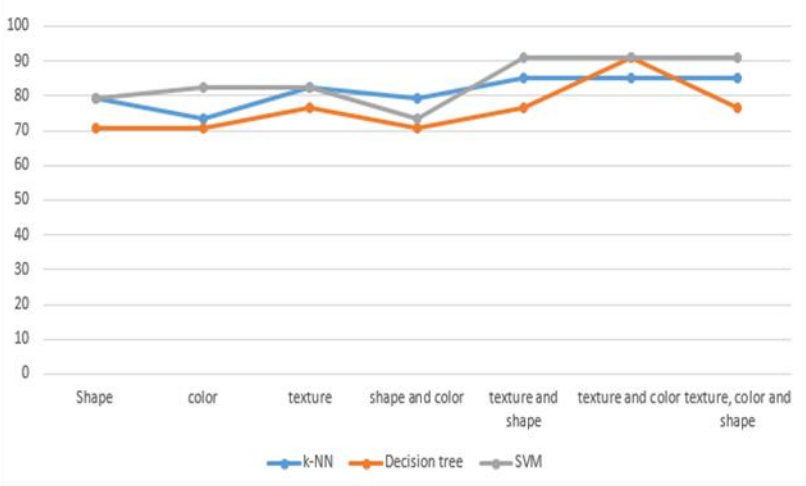

FIGURE 7: GRAPH REPRESENTING ACCURACY OF HEALTHY VS BLAST WBCS

3.2 Experiment 2

In this experiment we take 62 MBSI, 19 images are heathy, 15 images were ALL-L1, 19 Images are ALL-L2 and 9 images are ALL-L3. We take texture, Shape and color features separately and apply k-NN, SVM and Decision Tree classifier. All features are combined and classifiers applied on it. This Experiment was performed on MATLAB R2017a [31]. The results are in shown in Table 2.

TABLE 3: PREDICTION RESULTS OF DIFFERENT FEATURES ON HEALTHY VS ALL-L1VS ALL-L2 VS ALL-L3

Features	**No. of classes**	**SVM**	**Decision Tree**	**k-NN**
Shape	4	50%	33.33%	41.6%

Color	4	50%	37.50%	37.5%
Texture	4	62%	54%	50%
Texture, Color and Shape	4	100%	91.67%	66.7%

Among individual feature sets, the SVM classifier achieved the highest accuracies of 50% each for both shape and color-based features, while texture features yielded a moderately higher accuracy of 62%. However, when shape, color, and texture features were combined into a unified representation, the SVM demonstrated a significant improvement, achieving 100% classification accuracy. These results confirm that feature fusion substantially enhances model performance by capturing complementary discriminative information. The comparative results are illustrated in Figure 8.

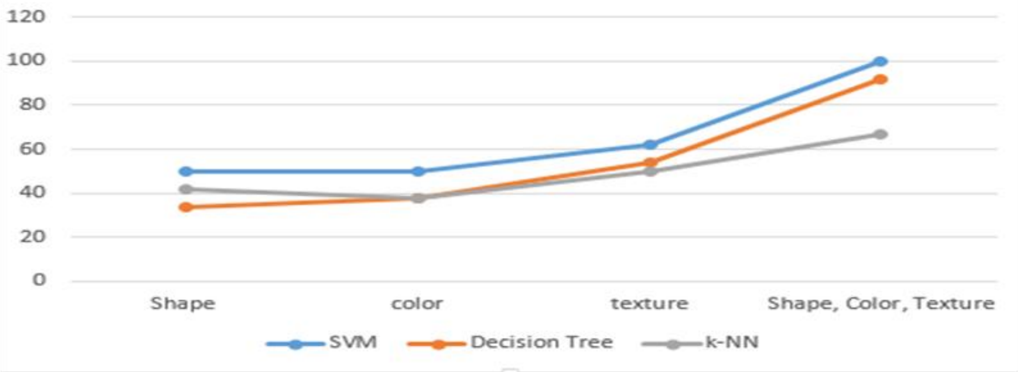

FIGURE 8: GRAPH REPRESENTING ACCURACY OF DIFFERENT FEATURES ON HEALTHY VS ALL-L1VS ALL-L2 VS ALL-L3

Experiment 3

In this experiment, a dataset comprising 62 MBSI was utilized, including 19 healthy samples, 15 ALL-L1, 19 ALL-L2, and 9 ALL-L3 images. The classification was conducted in multiple stages: initially, each leukemia subtype was classified against healthy samples (i.e., ALL-L1 vs. Healthy, ALL-L2 vs. Healthy, and ALL-L3 vs. Healthy). Subsequently, pairwise combinations of leukemia subtypes were analyzed alongside healthy samples (ALL-L1 vs. ALL-L2 vs. Healthy, ALL-L1 vs. ALL-L3 vs. Healthy, and ALL-L2 vs. ALL-L3 vs. Healthy). Finally, a comprehensive multi-class classification was per-formed by combining all leukemia subtypes with healthy samples (ALL-L1, ALL-L2, ALL-L3, and Healthy). All experiments were executed using MATLAB R2017a, and the corresponding results are summarized in Table 3.

TABLE 4: PREDICTION RESULTS OF HEALTHY VS DIFFERENT TYPES OF ALL

Dataset	No. of classes	Decision Tree	K-NN	SVM
L1 vs Healthy	2	75.0%	58.3%	91.7%
L2 vs Healthy	2	92.9%	100%	100%
L3 vs Healthy	2	83.3%	83.3%	83.3%
L1 vs L2 vs Healthy	3	83.3%	83.3%	94.4%
L1 vs L3 vs Healthy	3	77.8%	77.8%	88.8%
L2 vs L3 vs Healthy	3	22.2%	44.4%	44.4%
L1 vs L2 vs L3 vs Healthy	4	91.7%	66.7%	100%

For the ALL-L1 vs. Healthy classification, the SVM achieved the highest accuracy of 91.6%. In the ALL-L2 vs. Healthy experiment, both k-NN and SVM yielded a perfect accuracy of 100%, while for ALL-L3 vs. Healthy, the k-NN, Decision Tree, and SVM classifiers produced identical accuracies of 83.33%. When classifying combined leukemia subtypes, the SVM achieved 94.4% accuracy for ALL-L1 vs. ALL-L2 vs. Healthy, and 88.88% for ALL-L1 vs. ALL-L3 vs. Healthy. For the ALL-L2 vs. ALL-L3 vs. Healthy configuration, both SVM and k-NN attained a maximum accuracy of 44.44%. Finally, in the comprehensive multi-class classification (ALL-L1, ALL-L2, ALL-L3, and Healthy), the SVM again demonstrated superior performance with 100% accuracy. The graphical representation of these results is shown in Figure 9.

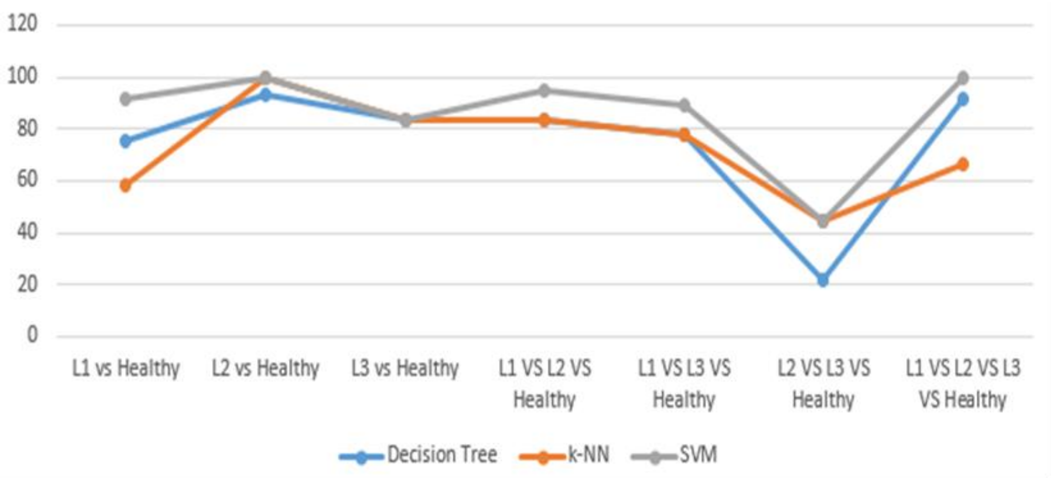

FIGURE 9: GRAPH REPRESENTING ACCURACY OF HEALTHY VS DIFFERENT TYPES OF ALL

A brief comparison of the proposed research with existing research is shown in Table 4. Most of the research studies existing in literature are based on the binary classification i.e. between healthy and blast cells. Some studies exist for classification between some types of ALL and AML. The results of the proposed research effectively show that fusion of textural and color-based features can effectively be used for correct diagnosis of all types of ALL.

TABLE 5: COMPARISON OF EXISTING RESEARCH WORKS FOR THE DIAGNOSES OF ALL

Authors, years	No of Classes	Accuracy
[29]	Healthy or blast cells	99.17%
[28], 2022	Healthy or blast cells	95.54%
[27], 2022	Healthy or blast cells	100%
[34], 2018	Healthy, ALL (L1, L2), AML (M1, M2)	99.67%
Proposed Research	Healthy, L1, L2, L3	100%

5. CONCLUSIONS

This study demonstrates the effectiveness of machine learning classifiers for the automated detection and classification of Acute Lymphoblastic Leukemia using microscopic blood smear images. Through systematic experimentation on individual and combined feature sets (shape, color, and texture) the SVM classifier consistently out-performed other models, achieving up to 100% accuracy when all features were fused. Comparative analyses across binary and multi-class configurations (ALL subtypes vs. healthy samples) further confirmed the robustness of the proposed framework. These results highlight that the integration of diverse feature representations significantly enhances discriminative performance, offering a reliable, cost-effective, and scalable computational approach for leukemia diagnosis that could complement conventional laboratory diagnostic methods.

Funding: This research received no external funding

Data Availability Statement: All the experiments of the proposed research a carried on a dataset, which was gathered from ALL-IDB, a publicly available website [6].

Conflicts of Interest: Authors have no conflict of interests.

REFERENCES

[1] "American Association for Cancer Research (AACR)." Accessed: Oct. 13, 2025. [Online]. Available: https://www.aacr.org/

[2] "American Society of Hematology." Accessed: Jul. 25, 2021. [Online]. Available: http://www.hematology.org/Patients/Cancers/

[3] A. Chennamadhavuni, V. Iyengar, S. K. R. Mukkamalla, and A. Shimanovsky, "Leukemia," StatPearls, Jan. 2023, Accessed: Oct. 13, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK560490/

[4] "Leukemia: Causes, treatment, and early signs." Accessed: Oct. 13, 2025. [Online]. Available: https://www.medicalnewstoday.com/articles/142595

[5] G. Onofrio and G. Zini, "Analysis of Bone Marrow Aspiration Fluid Using Automated Blood Cell Counters," Clinics in Laboratory Medicine, pp. 1–18, 2015, doi: 10.1016/j.cll.2014.10.001.

[6] W. Is and A. Lymphocytic, "About Acute Lymphocytic Leukemia (ALL)".

[7] W. Is and A. Myeloid, "About Acute Myeloid Leukemia (AML) What Is Acute Myeloid Leukemia ?".

[8] "About Acute Lymphocytic Leukemia | American Cancer Society." Accessed: Oct. 13, 2025. [Online]. Available: https://www.cancer.org/cancer/types/acute-lymphocytic-leukemia/about.html

[9] G. Abdul-hamid, "Classification of Acute Leukemia," University of Aden/Hematology Unit Yemen, 1999, doi: 10.1016/S0740-2570(03)00031-5.

[10] S. S. Al-jaboriy, N. N. A. Sjarif, S. Chuprat, and W. M. Abduallah, "Acute lymphoblastic leukemia segmentation using local pixel information," Pattern Recognition Letters, vol. 125, pp. 85–90, 2019, doi: 10.1016/j.patrec.2019.03.024.

[11] H. N. Lim, M. Y. Mashor, N. Z. Supardi, and R. Hassan, "Color and morphological based techniques on white blood cells segmentation," Proceedings - 2015 2nd International Conference on Biomedical Engineering, ICoBE 2015, no. March, pp. 30–31, 2015, doi: 10.1109/ICoBE.2015.7235915.

[12] H. Kodaira, T. Tanaka, and J. Takeuchi, "[Classification of acute lymphoblastic leukemia]," Nihon Rinsho, vol. 70 Suppl 2, pp. 438–441, 2012.

[13] and M. G. O. Morteza Moradi Amin, Saeed Kermani, Ardeshir Talebi, "Recognition of Acute Lymphoblastic Leukemia Cells in Microscopic Images Using K-Means Clustering and Support Vector Machine Classifier," Journal of Medical Signals and Sensors, vol. 5, no. jan-Mar, pp. 49–58, 2015.

[14] M. M. D. Joshi, P. A. H. Karode, and P. S. R. Suralkar, "White Blood Cells Segmentation and Classification to Detect Acute Leukemia," International Journal of Emerging trends &

Technology in Computer Science, vol. 2, no. 3, pp. 147–151, 2013.

[15] L. Putzu, G. Caocci, and C. Di, "Artificial Intelligence in Medicine Leucocyte classification for leukaemia detection using image processing techniques," Artificial Intelligence In Medicine, vol. 62, no. 3, pp. 179–191, 2014, doi: 10.1016/j.artmed.2014.09.002.

[16] S. Nazlibilek, D. Karacor, T. Ercan, M. Husnu, and O. Kalender, "Automatic segmentation , counting , size determination and classification of white blood cells," MEASUREMENT, vol. 55, pp. 58–65, 2014, doi: 10.1016/j.measurement.2014.04.008.

[17] R. Duggal, A. Gupta, R. Gupta, M. Wadhwa, and C. Ahuja, "Overlapping cell nuclei segmentation in microscopic images using deep belief networks," Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing - ICVGIP '16, pp. 1–8, 2016, doi: 10.1145/3009977.3010043.

[18] V. Singhal, "Correlation based Feature Selection for Diagnosis of Acute Lymphoblastic Leukemia," pp. 5–9, 2015.

[19] M. Moradiamin, A. Memari, N. Samadzadehaghdam, S. Kermani, and A. Talebi, "Computer aided detection and classi fi cation of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis," no. April, pp. 1–9, 2016, doi: 10.1002/jemt.22718.

[20] K. Raghul, A. S. Raj, and P. U. Ilavarasi, "Acute Lymphocytic Leukemia Detection by Image Processing Using Matlab," Middle-East Journal of Scientific Research, vol. 24, pp. 263–267, 2016, doi: 10.5829/idosi.mejsr.2016.24.S1.54.

[21] S. Mishra, B. Majhi, P. K. Sa, and L. Sharma, "Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection," Biomedical Signal Processing and Control, vol. 33, no. March, pp. 272–280, 2017, doi: 10.1016/j.bspc.2016.11.021.

[22] S. Mohapatra, "Automated Leukemia Detection in Blood Microscopic," pp. 184–187.

[23] S. C. Neoh and S. Todryk, "Intelligent leukaemia diagnosis with bare- bones PSO based feature optimization," Applied Soft Computing Journal, vol. 56, no. April, pp. 405–419, 2017, doi: 10.1016/j.asoc.2017.03.024.

[24] L. H. S. Vogado, R. M. S. Veras, F. H. D. Araujo, R. R. V. Silva, and K. R. T. Aires, "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification," Engineering Applications of Artificial Intelligence, vol. 72, no. October 2017, pp. 415–422, 2018, doi: 10.1016/j.engappai.2018.04.024.

[25] J. Laosai and K. Chamnongthai, "Classification of acute leukemia using medical-knowledge-based morphology and CD marker," Biomedical Signal Processing and Control, vol. 44, pp. 127–137, 2018, doi: 10.1016/j.bspc.2018.01.020.

[26] S. N. Mohd Safuan, M. R. Md Tomari, and W. N. Wan Zakaria, "White blood cell (WBC) counting analysis in blood smear images using various color segmentation methods," Measurement: Journal of the International Measurement Confederation, vol. 116, no. November 2017, pp. 543–555, 2018, doi: 10.1016/j.measurement.2017.11.002.

[27] A. Saeed, F. Rehman, A. Ashfaq, H. Sharif, R. Zeeshan, and A. Zahid, "Acute Lymphoblastic Leukemia Prediction from Microscopic Blood Smear Images Using Advanced Machine Learning," 2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), 2022, doi: 10.1109/MACS56771.2022.10022378.

[28] N. Sampathila et al., "Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images," Healthcare 2022, Vol. 10, Page 1812, vol. 10, no.

10, p. 1812, Sep. 2022, doi: 10.3390/HEALTHCARE10101812.

[29] M. Rangini, S. Pundir, M. E. M. Soudagar, D. Vekariya, H. Patil, and S. Rajkumar, "Texture-based Feature Extraction and Machine Learning Model for the Detection of Acute Lymphoblastic Leukemia," Proceedings - 2024 5th International Conference on Mobile Computing and Sustainable Informatics, ICMCSI 2024, pp. 505–511, 2024, doi: 10.1109/ICMCSI61536.2024.00079.

[30] "Bulk Image Resizer." Accessed: Dec. 05, 2017. [Online]. Available: https://bulkresizephotos.com/

[31] MathWorks, Learning MatLab 7 Release 14. 2005.

[32] M. Strzelecki and P. Szczypinski, "MaZda User ' s Manual MaZda User ' s Manual MaZda User ' s Manual MaZda User ' s Manual," 1998.

[33] "MATLAB Region Properties." Accessed: Jul. 25, 2021. [Online]. Available: https://de.mathworks.com/help/images/ref/regionprops.html

[34] J. Laosai and K. Chamnongthai, "Classification of acute leukemia using medical-knowledge-based morphology and CD marker," Biomed Signal Process Control, vol. 44, pp. 127–137, Jul. 2018, doi: 10.1016/J.BSPC.2018.01.020.