Tech Science Press

Histogram-Based Decision Support System for Extraction and Classification of Leukemia in Blood Smear Images

Neenavath Veeraiah[1,*], Youseef Alotaibi[2] and Ahmad F. Subahi[3]

[1]Department of Electronics and Communications, DVR & DHS MIC Engineering College, Kanchikacharla, A.P., 521180, India
[2]Department of Computer Science, College of Computer and Information Systems, Umm Al-Qura University, Makkah, 21955, Saudi Arabia
[3]Department of Computer Science, University College of Al Jamoum, Umm Al-Qura University, Makkah, 21421, Saudi Arabia
*Corresponding Author: Neenavath Veeraiah. Email: neenavathveeru@gmail.com
Received: 23 July 2022; Accepted: 08 December 2022

Abstract: An abnormality that develops in white blood cells is called leukemia. The diagnosis of leukemia is made possible by microscopic investigation of the smear in the periphery. Prior training is necessary to complete the morphological examination of the blood smear for leukemia diagnosis. This paper proposes a Histogram Threshold Segmentation Classifier (HTsC) for a decision support system. The proposed HTsC is evaluated based on the color and brightness variation in the dataset of blood smear images. Arithmetic operations are used to crop the nucleus based on automated approximation. White Blood Cell (WBC) segmentation is calculated using the active contour model to determine the contrast between image regions using the color transfer approach. Through entropy-adaptive mask generation, WBCs accurately detect the circularity region for identification of the nucleus. The proposed HTsC addressed the cytoplasm region based on variations in size and shape concerning addition and rotation operations. Variation in WBC imaging characteristics depends on the cytoplasmic and nuclear regions. The computation of the variation between image features in the cytoplasm and nuclei regions of the WBCs is used to classify blood smear images. The classification of the blood smear is performed with conventional machine-learning techniques integrated with the features of the deep-learning regression classifier. The designed HTsC classifier comprises the binary classifier with the classification of the lymphocytes, monocytes, neutrophils, eosinophils, and abnormalities in the WBCs. The proposed HTsC identifies the abnormal activity in the WBC, considering the color and shape features. It exhibits a higher classification accuracy value of 99.6% when combined with the other classifiers. The comparative analysis expressed that the proposed HTsC model exhibits an overall accuracy value of 98%, which is approximately 3%–12% higher than the conventional technique.

Keywords: White blood cells; leukemia; segmentation; histogram; blood smear

1 Introduction

Leukemia occurs in the human organ because of abnormal white blood cells (WBCs), which are referred to as cancer. Leukemia in the blood is evaluated through a microscopic examination of the WBCs. Peripheral Blood Smear (PBS) is used to assess laboratory blood-related disorders [1]. This is accomplished through the identification of microscopic diseases and the diagnosis of treatment options. The examination is based on collecting blood samples and performing a peripheral blood smear evaluation through a microscope [2].

Additionally, the comparison of the blood is based on the evaluation of red blood cells (RBCs), white blood cells (WBCs), platelets, and plasma. The count and morphology of WBCs are observed in laboratories to diagnose leukemia. WBC can be classified into five categories: lymphocytes, monocytes, neutrophils, eosinophils, and basophils. In addition, WBC comprises a dark nucleus and a pale cytoplasm region around the middle. The appearance of these cells on a stained peripheral blood smear is noticed when observed under a microscope [3]. They vary in characteristics such as size, shape, color, and texture. The cells' appearance and deviation from the average count are observed and recorded. Changes in the cytoplasm, nuclei intensity, platelet count and appearance, and RBC size and distribution all contribute to WBC maturation [4].

Examination by microscopic analysis is performed either qualitatively or quantitatively. Through quantitative analysis, blood cells are counted; those are generally carried out using hematology analyzers. The type of blood count can be classified as a complete blood count (CBC) or a differential blood count (DC). The CBC is involved in the computation of the WBC's total counts, platelets, and RBCs, whereas the DC provides each WBC's type count [5]. In addition, hematology analyzers compute indirect parameters such as cell impedance or density. Examining the morphological characteristics, on the other hand, performs qualitative analysis that can be calculated manually. To estimate the abnormal conditions, quantitative and qualitative studies were performed with the blood cell irregularities count in diagnosing the various diseases.

Leukemia originates in stem cells due to mutations. It is classified into four main types based on its growth rate and the type of affected blood cell [6]. The four types of classified leukemia are acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), chronic myeloid leukemia (CML), and chronic lymphocytic leukemia (CLL). Commonly observed morphological changes of WBCs during leukemia are as follows [7]: cytoplasmic amount, which is a reduction in the area of cytoplasm compared to normal WBCs; and the presence of nucleoli, which is a small bubble-like structure within the cell nucleus. The presence of an Auer rod is a clump of granular material that forms elongated needle-like structures seen in the cytoplasm of leukemic WBCs. The vacuole is formulated with a small cytoplasmic cavity bounded by a single membrane containing food, metabolic waste, and water. Intense staining due to the affinity of the cytoplasm for a particular (basophilic) dye. This causes the basophils' deep blue color, coalescent granules' formation, to be caused by improper functioning of the cell coalescent granules [8].

Leukemia diagnosis requires PBS analysis to study the morphological changes in the WBCs. PBS analysis is based on microscopic evaluation of blood smears, considered the gold-standard technique requiring prior training and expertise. Manual evaluation of tiny blood samples is a tedious and time-consuming procedure. The results depend on the skill and experience of the laboratory technician [9]. Pathologists are often overburdened with large sets of such data, which need to be analyzed carefully to arrive at a decision. This results in inter-observer variability, which is hard to reproduce.

The problem arises when a large number of blood samples need to be examined by pathologists. Processing time and skill may limit the result's speed and accuracy. Manual microscopic evaluation is time-consuming and may produce erroneous results. Therefore, it is necessary to construct a cost-effective, robust, and automatic technique to detect leukemia [10]. Through automatic microscopic

evaluation, pathologies increase the blood sample to improve accuracy. Pathologists analyze peripheral blood smears under a microscope to evaluate the shape, size, color, and presence of inclusions in WBCs. Abnormalities in appearance and count of WBCs are reported. These features can be used to design an automated computer-aided diagnosis (CAD) system, which can be used to minimize the pathologists' workload [11]. In addition, automated cell analysis of blood exhibits fast and effective results. It is effectively involved in the effective handling of a massive amount of data. There have been many attempts to develop CAD systems; however, they depend on the availability of a fully automated workflow. Images acquired from a manual setup suffer from illumination and staining variations [12]. In such cases, the automated systems fail to detect leukemia accurately. Developing a robust method that can handle the variations usually present in images acquired using a manual setup is desirable.

This paper concentrated on the segmentation and classification of leukemia in blood images. To perform segmentation and classification, this paper introduces the HTsC scheme. The rest of this work contribution is presented as follows:

i) A method for detecting nuclei resistant to color, shade, and illumination variations common in images obtained through manual setup
ii) A novel adaptive mask generation method for accurate detection of WBCs
iii) A hybrid classifier to classify WBCs with very high accuracy of 99.6%
iv) A novel convolutional neural network (CNN) for classifying white blood cells (WBC-Net) with an average accuracy of 98%.

This paper puts forward a novel decision-support system that is highly accurate and, thus, could be used for screening and diagnosis. Only abnormal cases need to be considered for the pathologist's review. Moreover, this method has been explicitly designed for high robustness to handle common variations that are usual in images acquired from a manual processing workflow. Hence, it offers a practically feasible solution for the automation of image analysis.

The paper structure is defined as follows: Section 2 describes the related works. Section 3 establishes the detection of the nucleus with the proposed HTsC. Simulation results and analysis are explained in Section 4. Finally, the conclusion of the work is described in Section 5.

2 Related Works

Many research groups have proposed various methods for the automated analysis of WBCs. The state-of-the-art techniques are categorized based on WBC detection and classification as follows:

1. WBCs detection
2. Identification of five types of normal WBCs
3. Classification of WBCs (normal and abnormal)

Hegde et al. [13] created a leukocyte nucleus enhancer method to improve the nucleus region in PBS images. Multi-level thresholding was used to detect nuclei, achieving a Dice score of around 0.96. Furthermore, WBC's classification relies on the genetic algorithm to consider nuclei features via k-means clustering. The classification is based on consideration of the texture and shapes, with appropriate classification accuracy values of 81% and 98%, respectively, for the WBC type. In addition, Ghaderzadeh et al. [14] employed arithmetic operations, Otsu's thresholding, and a minimum filter to detect nuclei. Images were obtained from the veterinary clinical pathology database, in which all the photos were stained using Wright's staining method. The reported segmentation accuracy was between 85% and 98%, depending on the type of WBC.

Besides, Khan et al. [15] utilized Markov random fields to segment the nucleus and cytoplasm from bone marrow cell images. The CIE LAB color space representation was utilized to extract color features, and the 2D world decomposition method was used to extract texture features reported for the segmentation of WBCs. Consequently, an overall segmentation accuracy of around 95% was reported. The dataset with both natural and synthetic images was considered in this study. Moreover, Dese et al. [16] presented another WBC segmentation method using PBS and bone marrow images. Marker-controlled watershed segmentation and circle-fit algorithms were used to separate overlapped WBCs, thereby obtaining precision and recall rates of 0.94 and 0.98, respectively, for the segmentation of WBCs. As reported in [16], this study used 31 images.

Grochowski et al. [17] proposed a method for detecting and tracking WBCs in video sequences using a level-set algorithm. They concluded that the performance of the level-set was better compared to the correlation tracking scheme. Abdullah et al. [18] proposed another WBC detection method and WBCs segmentation into the nucleus and cytoplasm using 108 images from the ALL-IDB1 dataset. The WBC detection and background removal are performed with Zack's thresholding these images. The nuclei in the images are detected using Otsu's thresholding and the color spaces CMYK and CIE LAB for detection. The WBCs are grouped by watershed segmentation, providing 92% detection accuracy. Ratley et al. [19] presented an algorithm based on fuzzy cellular neural networks for detecting WBCs. This method included a combination of thresholding, morphological operations, and fuzzy logic. The dataset consisted of 50 microscopic images with illumination and staining variations. A detection accuracy of around 98% was reported.

Gebremeskel et al. [20] achieved WBC detection using the dual-threshold method by considering grayscale images and the H component of HSV color space representations of images to select the threshold values. The dataset consisted of 130 pictures of the ALL-IDB dataset [14]. This paper reported an overall segmentation accuracy of 98%. Rastogi et al. [21] presented a WBCs detection technique using the learning-by-sampling method. A Support Vector Machine (SVM) classifier was trained to learn the color information of WBCs by constructing a color look-up table for WBC extraction. The authors used images of Wright-Giemsa-stained peripheral smears from different hospitals. An overall error rate of 0.156 was reported using 65 photos. Jha et al. [22] proposed another WBC detection and classification method from PBS images. The developed technique was utilized for detecting the nuclei in the WBC's active contour model. The seed points for active contours were generated using the H and S components of HSV color space representation images to detect WBCs and nuclei, respectively. Shape features of nuclei and texture features derived from the gray level entropy matrix and gray level cooccurrence matrix (GLCM) of cytoplasm were extracted for classification using the Naive Bayes classifier. Analysis was carried out using 237 Leishman-stained PBS images consisting of 267 WBCs. The authors reported an overall classification accuracy of around 92%.

Additionally, Umamaheswari et al. [23] constructed a WBC segmentation and classification model. The developed model uses the Gram-Schmidt orthogonalization contour model to segment the nuclei and cytoplasmic regions. The analysis is based on extracting nuclei and cytoplasm from extracted local binary patterns (LBP) and GLCM, as well as shape and texture considerations. Components were selected using the sequential forward selection method. The classification was carried out using two commonly known classifiers: SVM and NN. The authors reported an average segmentation accuracy of 93% and 92% for nuclei and cytoplasm, respectively.

Furthermore, a classification accuracy of 96% was reported using the SVM classifier. Alagu et al. [24] presented an automated method for WBC detection and classification. Fuzzy c-means clustering and morphological operations were used for WBC segmentation, considering the A and B components of CIE LAB color space representation. Shape features of nuclei and color and texture features of cytoplasm

were considered to train the SVM classifier. The reported average classification accuracy was around 95%. The WBCs segmentation is performed with the integration of k-means clustering and expectation maximization based on the color, texture, and shape of the nucleus and cytoplasm. The analysis considers different classifiers such as SVM, K-NN, NN, and Naïve Bayes. The dataset consisted of 115 PBS images stained using the May-Grunwald-Giemsa stain. The reported classification accuracy was around 97% using the NN classifier and 94% using the SVM classifier.

Raphael et al. [25] presented a WBC classification method using PBS images acquired from Wright-stained peripheral blood smears. Threshold-based segmentation was proposed using the R and B components of the original pictures, and a segmentation accuracy of around 93% was reported. For classification, shape features of nuclei and WBCs, as well as color features of cytoplasm, were extracted. A two-step classification method was proposed in which the first step was to classify WBCs into "segmented" and "non-segmented" cells based on the features of the nuclei. In the second step, WBCs were further classified into five types. Linear discriminant analysis was used for the classification of WBCs. Using 1938 sub-images, an overall classification accuracy of around 94% was obtained. Alam et al. [26] proposed a scheme for detecting WBCs and classifying blood images. The analysis is based on the estimation of the morphological operation. The detection is carried out using ellipse-curve fitting. The WBC's classification is performed to identify five types using Naïve Bayes. The classification considers the cytoplasm, shape, and color of the nuclei. The analysis is based on viewing the two datasets with image counts of 555 and 477 sub-images. The developed scheme exhibits an overall classification accuracy of 98%.

Kassim et al. [27] proposed a mobile-cloud-based method for WBC classification using 1030 PBS images. A color-based k-means clustering method was used for the detection of WBCs. The considered features included shape features, statistical properties, and texture features of the WBCs, which were utilized for training an ensemble multi-class SVM. The reported segmentation accuracy was around 96%, and the classification accuracy was about 94%.

Girdhar et al. [28] developed an acute leukemia classification scheme for blood smear images. The WBCs are detected and examined using Markov random fields in conjunction with the k-means algorithm for the CIE LAB color representation space used for WBC imaging. To identify effective classification, the ensemble-based particle swarm optimization technique is used to improve classification performance. The datasets used for the analysis consist of the results of the 633 leukemia cell images and an estimation of the different stain variations in color. The proposed model exhibited an overall accuracy of 97% for adequate classification.

Cheuque et al. [29] developed a decision-support system for estimating microscopic images based on ALL diagnoses and CIELAB representation. The experimental analysis stated that the fuzzy divergence and Zack's thresholding compute the WBC image segmentation. Through the estimation of the watershed segmentation, overlapped cells are evaluated. The experimental analysis involved respect for the watershed segmentation for the overlapped cell computation based on classification features such as shape, color, and LBP. The classification is computed based on the ensemble classifier, which has a classification accuracy value of 97%.

Meenakshi et al. [30] have developed a method for the microscopic evaluation of ALL diagnoses. The proposed scheme employs a stimulating segmentation and detection scheme based on the measurement of WBCs. The segmentation is based on the cytoplasm and nucleus of the WBCs. The proposed method comprises the particle swarm optimization model for computing the discriminant features in the blood smear images. The results expressed that the proposed classifiers, SVM and NN, exhibit discriminant features extracted from the collection of 180 photos in the ALL-IDB2 dataset. The proposed model exhibits a classification accuracy of 95% for the NN classifier. Table 1 provides an overview of the literature for WBC segmentation and classification.

Table 1: Overall summary of the literature

Reference	Published year	Method	Region	Limitation
[13]	2019	Multilevel thresholding	Leukocyte nucleus	Only consider shape and texture
[14]	2021	Otsu's thresholding	Detection of nuclei	Segmentation accuracy alone was evaluated
[15]	2020	Markov random fields	Segmentation of nucleus and cytoplasm	Classification is not performed
[16]	2021	Watershed segmentation	Segmentation of WBCs	Only 31 images are considered
[17]	2019	Level-set algorithm	Tracking of WBCs	Performance is not explained clearly
[18]	2019	Zack's thresholding	Segmentation of WBCs into the nucleus and cytoplasm	Other classes are not considered
[19]	2020	Fuzzy cellular neural networks	Detection of WBCs	Only 50 microscopic images are considered
[20]	2021	Dual-threshold method	WBC detection	The performance is effective for the ALL-IDB dataset alone
[21]	2022	Support vector machine (SVM)	Detection of WBCs	The evaluation is based on 65 images alone
[22]	2019	Active contours method	Detection and classification of WBCs	The classification performance is not accurately measured
[23]	2018	GLCM and local binary patterns (LBP)	Segmentation and classification of WBCs	The evaluation is based on color features; shapes are not considered
[24]	2021	Fuzzy c-means clustering	Detection and classification of WBCs	It uses the conventional technique
[25]	2019	Two-step classification	Classification of WBCs	The performance is complex

3 Detection of Nucleus with Proposed Histogram Threshold Segmentation Classifier (HTsC)

Initially, the proposed HTsC focused on detecting the blood smear's nucleus. The proposed HTsC technique is involved in the computation of critical features that effectively contribute to the WBC's detection. The computation is based on thresholding, filtering, and morphological operations to identify the image's nucleus region [31–35]. Fig. 1 presents the overall process involved in the proposed HTsC for the computation of the nucleus in the blood smear.

In the proposed test, the images are cropped based on the region of interest in the image. The nucleus is approximated using an arithmetic operation to avoid the problem associated with the WBC's overlapping automated cropping method and region. With automated nucleus location, original images are cropped, and the approximate process is presented in Fig. 2, as opposed to images that are upgraded with added G

components from the original image. The addition of the G component and contrast enhancement is applied over the addition process in the grayscale images. The approximation in the nucleus region is evaluated with a threshold value of 110 [36–40]. The approximation is computed based on the cropping of the original image's nucleus. The image cropped from the blood smear is presented in Fig. 3, as represented in the bounding box with the nucleus approximation region. The black box denotes the cropped region used for further processing.
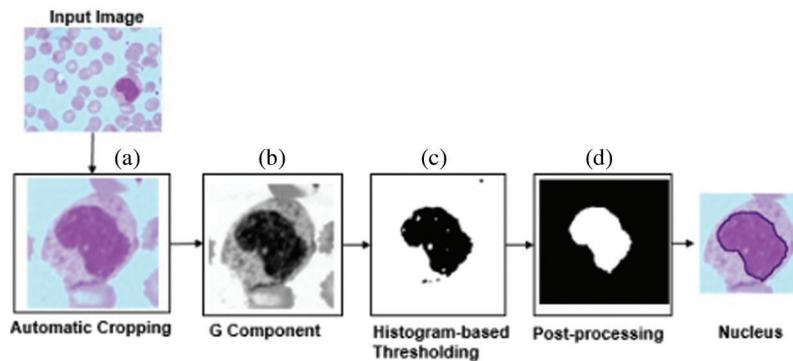
Figure 1: Nucleus detection with HTsC
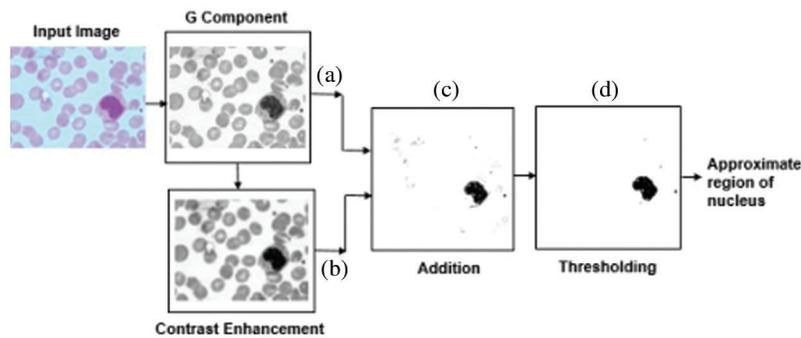
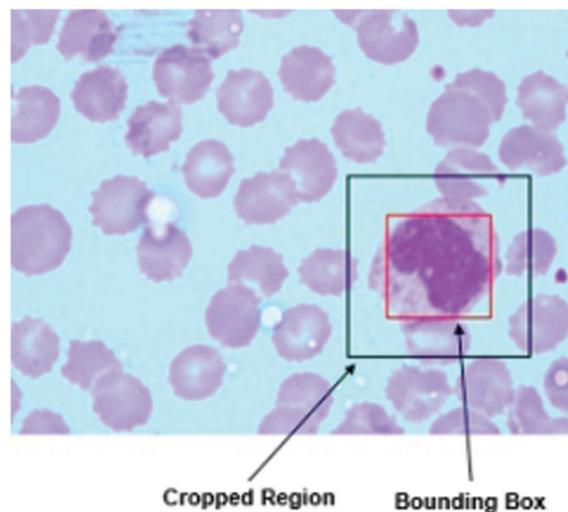**Figure 2:** Approximate nucleus detection in HTsC

Figure 3: Result of automatic cropping

3.1 Histogram-Based Thresholding

The variation in the color combination is evaluated with the histogram-based thresholding method to be used to select the optimal value threshold. The cropped image histogram is obtained with a size bin value of 150 and an estimation of the accurate threshold value, maximum pixel count, and corresponding gray image value for the measured histogram value. The threshold value of the gray level is evaluated based on the enhanced contrast G component to estimate the nucleus region of the image. In the image dataset, the threshold value is estimated to be between 4 and 137. The identified threshold value was 137 for the image shown in the top row (a). This indicates that the image is bright [41–45]. The identified threshold value was 4 for the image shown in the bottom row (b). This indicates that the image is dark. It is clear from the two representative images that the proposed histogram-based thresholding method manages brightness variations. In Fig. 4, the processed histogram value of the image with the proposed HTsC is presented.
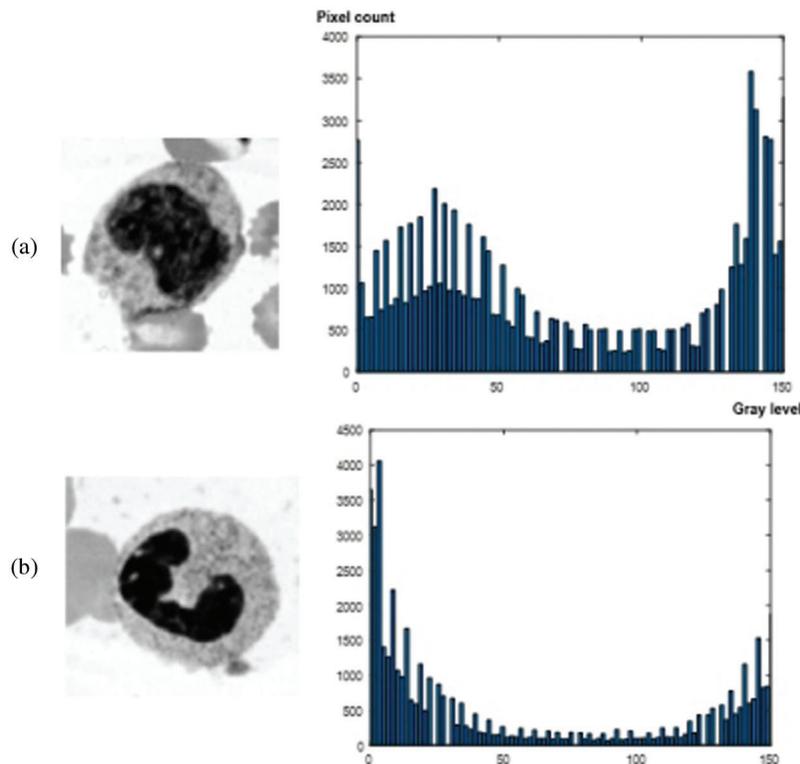

Figure 4: Histogram of two representative images

The original images were subjected to automatic cropping, and the cropped images were used to extract WBCs from the PBS images. The block diagram of the segmentation method is demonstrated in Fig. 5. To minimize color variations in the images, the color transfer method was applied to the cropped images. A grayscale representation of the color-transferred images was considered for further processing. Details of the various methods used are provided in the subsequent sections.
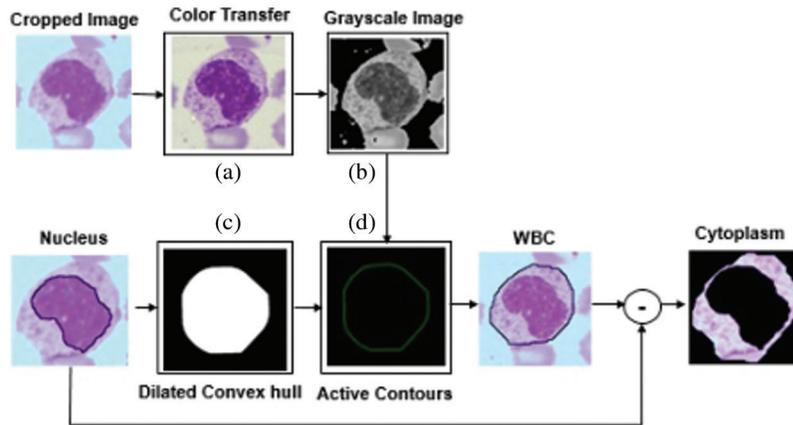
Figure 5: WBC extraction and segmentation method

The dimension of the images is evaluated based on the consideration of image segmentation features with different image pixel values as presented in Eq. (1).

$$F = \{F_1, F_2, \ldots, F_h, \ldots, F_z\} \tag{1}$$

The process of segmentation is performed with histogram computation through the extracted images' pixels as F, to perform the computation of the function utilized for the evaluation. This is presented in Eq. (2).

$$M_\mu = \sum_{i=1}^{q} \sum_{h=1}^{r} d_{ih}^\mu \times E_{ih} \tag{2}$$

where, $E_{ih} = \|e_i - F_h\|$. The computed image pixels are evaluated based on the computed histogram value for the segmentation process as shown in Eq. (3).

$$F_h = \frac{\sum_{i=1}^{u} d_{ih}^\mu . e_i}{\sum_{i=1}^{u} d_{ih}^\mu} \tag{3}$$

The segmentation in the blood smear image for the segmentation process is represented as, $Q_{u,v}^{FCM}$. The proposed HTsC comprises the histogram-based segmentation for the WBCs' identification. The practical component of the image components is designed based on specific selection criteria. The histogram segmentation process is examined using Eq. (4).

$$Q_{u,v} = \begin{cases} Q_{u,v}^A; & if \;\; Q_{u,v}^A == Q_{u,v}^{FCM} \\ M; & if \;\; Q_{u,v}^A \neq Q_{u,v}^{FCM} \end{cases} \tag{4}$$

where, $Q_{u,v}^A$ represents the model's histogram's segmented value output. The extracted features of the image are processed and evaluated based on the segmentation of the features in the images that are considered to have a similar pixel value. The proposed HTsC model's classification accuracy is evaluated by considering the various image features. The analysis is based on consideration of the features represented in Eq. (5).

$$M^A = MI\left(Q_{u,v}^A\right) \tag{5}$$

where, $MI(Q_{u,v}^A)$ denotes the contours of the segmented image, and different windows are represented as W_1, W_2. The pixel size of the image is represented as W_1 is 3×3, and the window size of W_2 is 4×4. The computation of the image features is denoted in Eq. (6).

$$MI(W_1, \ W_2) = E(W_1) + E(W_2) - E(W_1, \ W_2) \tag{6}$$

where, $E(W_1)$ is the window W_1 entropy and $E(W_1, \ W_2)$ is the joint entropy of the image features, and the joint probability of the image is represented by Eqs. (7) and (8).

$$E(W_1) = -\sum_u pw_1(u) \log pw_1(u) \tag{7}$$

$$E(W_1, \ W_2) = -\sum_{u,v} p_{w_1,w_2}(u, \ v) \log pw_1 w_2(u, \ v) \tag{8}$$

where, $pw_1(u)$ is represented as the image conditional probability, and the segmented image features are presented in Eq. (9) as follows:

$$M^{FCM} = MI\left(Q_{u,v}^{FCM}\right) \tag{9}$$

The histogram-based segmentation process in the blood smear images is denoted in Eq. (10).

$$M = \begin{cases} Q_{u,v}^{A}; \ if \ Q_{u,v}^{A} == Q_{u,v}^{FCM} \\ Q_{u,v}^{FCM}; \ else \end{cases} \tag{10}$$

The above Eq. (10) consists of the proposed HTsC model to retain segmentation in the images, represented as F = {F1, F2}.

3.2 Feature Extraction and Classification for Cropped Images

In the proposed HTsC, the cropped image with a single WBC was considered an input image. The color transfer method involves the estimation of the appropriate contrast image based on the image template. In the template image, the method converts the color characteristics of the input images. The color transfer method was applied to the cropped images, which consisted of only one WBC each. The size of the template image is equal to the size of input images with a single WBC [46–52]. The selected template image is shown in Fig. 6. This template image provides good contrast between the WBC and the background. The result of the color transfer method is illustrated in Fig. 6. It can be observed from Fig. 6a that the cytoplasm and the background region in the input image are not well differentiated. The obtained color-transfer image shows that the background and WBC regions can be differentiated, as shown in Fig. 6c. This was achieved through the careful selection of the template image.
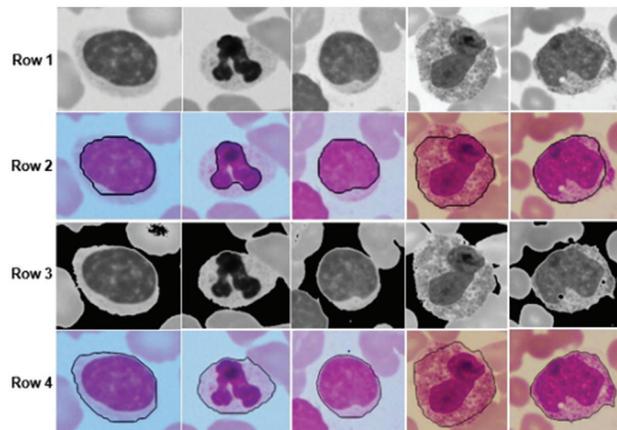

Figure 6: Effect of background removal for active contours method

Row 1: Grayscale images; Row 2: Results of the active contours method on grayscale images; Row 3: Background-removed images; Row 4: Results of the active contours method on background-removed images.

To classify WBCs accurately, the texture features of the nucleus and cytoplasm need to be studied separately. This requires the segmentation of the WBC into the nucleus and cytoplasm. After detecting the nucleus and WBC regions, subtracting the nucleus from the WBC results in the WBC being segmented into the nucleus and cytoplasm. The segmentation results are shown in Fig. 7.
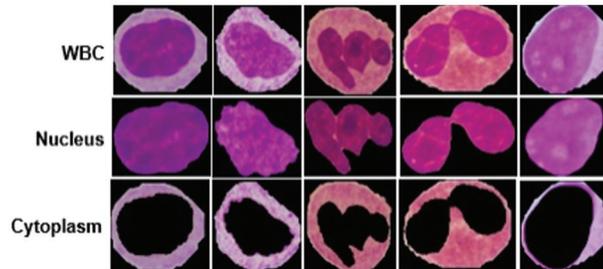

Figure 7: Extracted and classified features in blood smear

3.3 Feature Extraction with Three Classification Methods

In the classification method, a combination of three classifiers, namely an SVM, a NN, and a "binary classifier," was used. The following block diagram of the proposed classification method is shown in Fig. 8. The classification of WBCs was implemented in three steps.
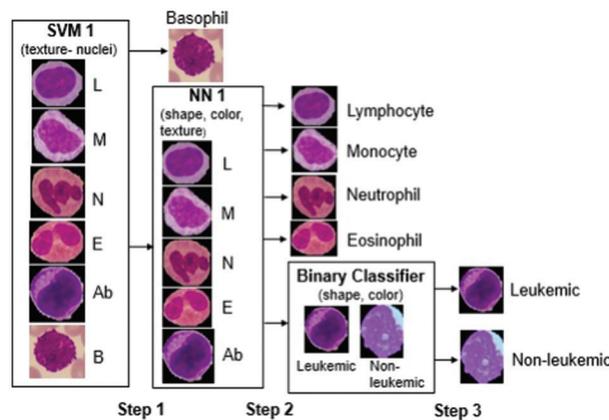

Figure 8: Classification steps of the proposed method

A few sample images of degenerating cells and WBCs considered to be in the "leukemic" class are illustrated in Fig. 9. Row 1: degenerating cells; Row 2: leukemic WBCs. Mean: The blood smear image features are identified with the statistical image information, computed based on Eq. (11).

$$D_2^{k.c} = \eta = \frac{1}{p} \times \sum_{b=1}^{p} Q_{u,v}^{R.p} \tag{11}$$

where, P denotes the image size dimensions ($\exists 1 \times \exists 2$) and F _x0005_, P u, v represents the κth segment of the jth image pixel value.
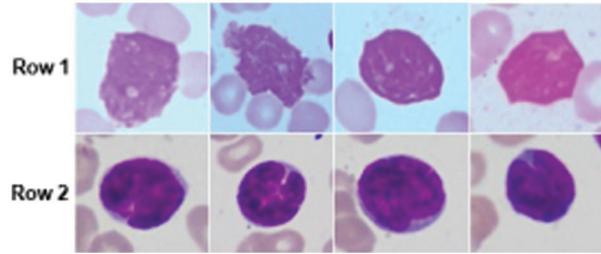
Figure 9: Sample images of the abnormal class

Variance: The statistical image features are computed based on the estimated mean value of the images. The segmented image variance of the blood smear is denoted in Eq. (12).

$$D_3^{k.c} = \frac{\sum_{b=1}^{p} \left(Q_{u,v}^{R.c} - \eta \right)}{P} \tag{12}$$

where, η denoted the segmented image mean.

Standard Deviation: The standard deviation is computed from the image pixel variation for the segmented images, and accurate estimation is based on Eq. (13).

$$D_4^{k.R} = \sqrt{\frac{\sum_{b=1}^{p} \left(Q_{u,v}^{R.P} - \eta \right)^2}{P - 1}} \tag{13}$$

The algorithm for the proposed HTsC for segmenting leukemia in the blood smear is presented in Algorithm 1.

Algorithm 1: Proposed HTsC for leukemia segmentation

Begin

Select the image element defined as C_o–0; where o = 1

Evaluate every image pixel as $c \in C_o$, where, $\mu k(o, c) > 0 - Q$

where $Q = \varphi$

do

        compute and eliminate the entropy value of image c from entropy set as presented in Q

    $f_{val} \leftarrow \max_{d \in c_0}[\min(f_0(F), \mu_k(c, F))]$

if $f_{val} > f_o (c)$ then

$f_o (c) \leftarrow f_{val}$

evaluate image histogram of the image

$\mu_k(c, e) > 0;\ f_{val} > f_0(e) f_{val} > f_0(e)$ and $\mu_k(c, e) > f_0(e)$

end if

end while

end

Compute the image histogram for the values as C_o–0; where o = 1

(Continued)

Algorithm 1 (continued)

      Evaluate $o - Q$

while $Q = \phi$

if $f_o(c)$ is maximal computed histogram value of the image Q

 for every image histogram $\mu_k(c, e) > 0$

do

$f_{val} \leftarrow \min(f_0(c), \mu_k(c, e))$

if

$f_{val} > f_0(e)$ for $f_0(e) \leftarrow f_{val}$

Compute the histogram value of the image based on the entropy e

End if

End for

End while

End

  The proposed HTsC model includes image segmentation techniques for blood smear images. As presented in Algorithm 1, the images in the dataset are computed based on the estimation of the entropy value in the images. Upon the computed entropy values, the images within the entropy region are calculated, and segmentation is performed. Based on the added segmentation value of the images, the features are estimated, and the classification of the images is performed using Algorithm 1.

4 Results and Discussion

4.1 Data Collection

  The blood smear image dataset was collected from Leishman-stained slides collected from an Olympus CX51 microscope with a resolution of 1600 × 1200 in the JPEG format [14]. The data were collected from 1159 images in PBS with a lymphocyte count of 170, a monocyte count of 109, a neutrophil count of 297, an eosinophil value of 154, a basophil value of 81, and an abnormal WBC value of 604. The multiple WBCs comprise the different images that process the 1418 images of WBC and the 1159 PBS images. In WBCs, the abnormal activity is computed based on the reactive lymphocyte cells that are degenerated, leukemic, and myelocytes that are defined as blasts. To assess the variation in the microscopic images, the abnormal cells in WBCs were compared to leukemic WBCs with developed HTsC. In Fig. 10, the dataset image for different intensities, brightnesses, and color shades comprises the platelets, WBCs, and RBCs.
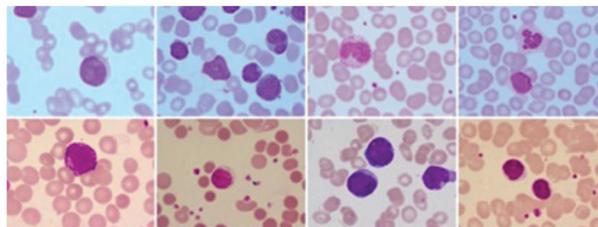

Figure 10: Sample images of the dataset

Adding or removing a constant amount alters the brightness of the photos. Pixel intensity is increased or decreased evenly throughout the picture by adding or removing an endless amount. The 'C' value was changed in 10-step increments from −20 to +20. This resulted in either darker or brighter photos than they would have been otherwise. There are many examples of pictures with consistent brightness fluctuations shown in Fig. 11. A uniform change in brightness can be observed in the images. The estimated image features of the image for the computation are presented in Eq. (14).

$$I_{Uniform}(i, j) = I_{Original}(i, j) + C \tag{14}$$

where $I_{Uniform}$ is the resultant image with uniform brightness after adding or subtracting the value C and $I_{Original}$ is the original image.
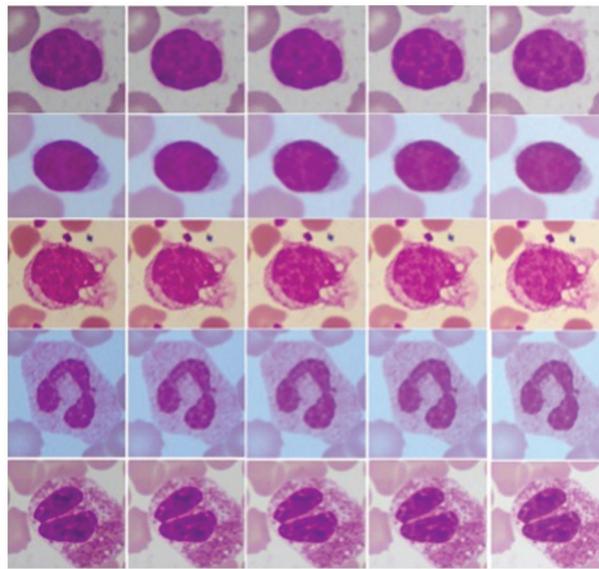

Figure 11: Uniform brightness variations (a) C = −20 (b) C = −10 (c) original images (d) C = +10 (e) C = +20

The brightness non-uniformity is developed through the scaling profile S; those are multiplied by every row in the original image defined in Eq. (15).

$$I_{non-Uniform} = I_{Original}(i, j) * S(Column) \tag{15}$$

In Eq. (15), S(column) represented the linearly spaced scaling profile; $I_{non-Uniform}$ is denoted as the brightness variation in the non-uniform resultant images, and the original image is denoted as $I_{Original}$. The variation in the scaling profile is computed in the horizontal direction with the generator value of S(column) as defined as follows in Eq. (16):

$$S(Column) = start + (index - 1) * \frac{(Stop - Start)}{(number\ of\ Columns - 1)} \tag{16}$$

The scaling profile for the linearly spaced image computed through the developed algorithm lies between the values 0.75 and 1.25. In shape, all rows are multiplied by the variation in the non-uniform brightness image demonstrated in Fig. 12.
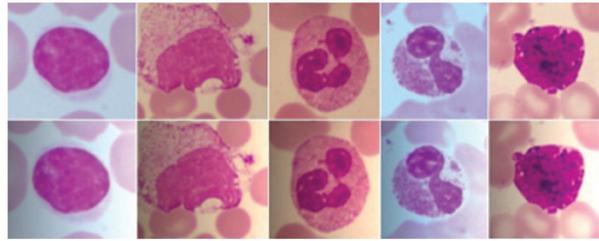
Figure 12: Non-uniform brightness variations

Row 1: Original images; Row 2: Brightness-varied images Dataset 1 corresponds to the original images; Dataset 2 and Dataset 3 correspond to brightness-varied images, and Dataset 4 is an online dataset named ALL-IDB2 [14]. This paper used Dataset 1 for the design of the proposed method, whereas Dataset 2 to Dataset 4 were used for the performance evaluation of the nuclei and WBC detection methods. Furthermore, this paper used Dataset 1 for WBCs classification using a traditional machine learning approach. Due to data imbalance, Dataset 4 was excluded from the traditional machine learning approach. This dataset consists of lymphocytes, lymphoblasts, and very few other WBC types. As it is known that the deep learning approach requires a large dataset, all the datasets used in WBC's classification are suitable for the deep learning approach.

4.2 Performance Metrics

The evaluation of the proposed HTsC is based on the metrics considered, such as true positive rate (TPR), true negative rate (TNR), and accuracy.

Accuracy: The measured accuracy for the proposed HTsC for the closeness classification classifier is computed in Eq. (17).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{17}$$

TPR: It computes the proposed HTsC-identified samples, denoted in Eq. (18).

$$TRP = \frac{TP}{TN + FP} \tag{18}$$

TNR: It presents the proposed HTsC values; those are samples that were negatively rejected for the computed value as presented in Eq. (19).

$$TNR = \frac{TN}{TN + FP} \tag{19}$$

The dice similarity coefficient (DSC) or dice coefficients used to evaluate the region's performance have segmented and ground truth values. The value computed for the DSC is presented in Eq. (20).

$$DSC = \frac{2TP}{(FP + 2TP + FN)} \tag{20}$$

where TP signifies true positive, TN refers to true negative, FN states the false negative, and FP refers to false positive.

4.3 Simulation Results

The proposed HTsC is involved in segmenting and classifying leukemia in the blood smear images. In the classification process based on the extracted texture features, the nuclei are removed to compute the blood smear. The performance of the proposed HTsC model is simulated and examined in the simulation software Python. Basophil detection is used to evaluate the developed training and testing for the confusion matrix. The measured parameters stated that the testing accuracy value was 99.6% and the training accuracy value was 100%—the classification with the proposed HTsC involved in the classification, as shown in Table 2.

Table 2: Training and testing data for HTsC

	Training		Testing	
	Basophil	Others	Basophil	Others
Basophil	65	0	15	0
Others	0	1070	1	267

In the second step, the classification of WBCs into lymphocytes (1), monocytes (2), neutrophils (3), eosinophils (4), and abnormal WBCs (5) was considered using the "NN 1" classifier. The confusion matrix, which gives the classification results for each type of WBC, is demonstrated in Fig. 13. It shows that matrix (a) represents the training performance of the classifier, and matrices (b) and (c) illustrate the performance of validation and testing, respectively. The values indicated along the diagonal of the matrix represent the number of each type of WBC that is correctly classified, and the other entries in the matrix represent the number of WBCs that are misclassified.
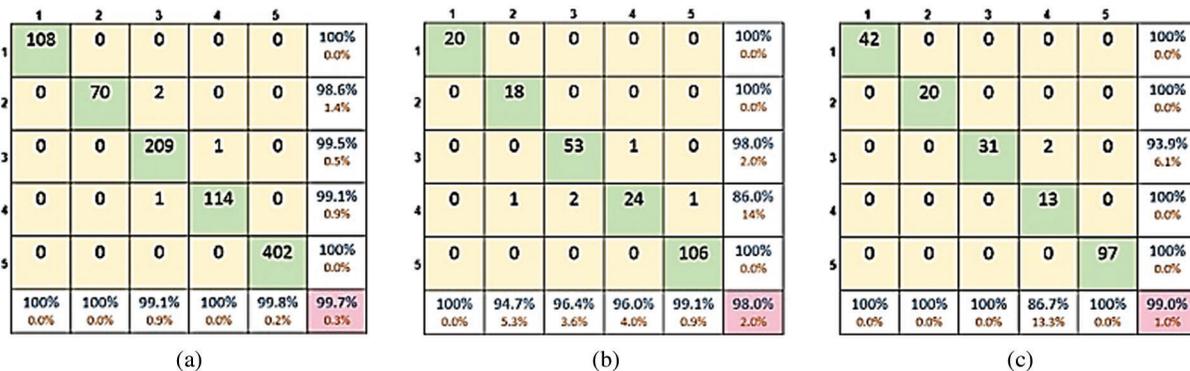

Figure 13: Confusion matrices of "NN 1" classifier (a) training performance (b) validation performance (c) testing performance

The accuracy for the computation is observed at 99.4%, and the computed accuracy value for the abnormal WBCs is measured at 99.9%. After estimating the BC's classification, the standard and abnormal activity are computed to detect the abnormality. The classification of normal WBCs into their types helps count the number of each type of WBC. This can diagnose count-related diseases such as neutropenia, eosinophilia, basophilia, and more. The classified results for the proposed HTsC are based on the different modules in the blood smear and are presented in Table 3.

Table 3: Performance metrics for HTsC

	Accuracy		Sensitivity		Specificity	
	Mean %	S.D	Mean %	S.D	Mean %	S.D
Lymphocyte	99.8	0.18	100	0	99.4	0.34
Monocyte	98.9	1.11	99.1	0.79	98.3	1.95
Neutrophil	98	0.33	98.8	0.64	99.4	0.54
Eosinophil	98.7	0.71	98.7	1.64	98.7	1.32
Abnormal	98.8	0.13	99.3	0.21	98.8	0.18

The abnormal class consists of 124 degenerated cells, out of which 122 were correctly identified using shape features. The "leukemic" class consists of 483 cells, out of which 12 WBCs were identified as degenerating by considering only shape features, as mentioned in step 7. This misclassification rate was corrected by considering the mean value of the component, as mentioned in step 8. The confusion matrix of the "binary classifier" is shown in Fig. 13. An accuracy of 99.6% was obtained for detecting leukemic WBCs. Furthermore, it can be seen from the figure that all the leukemic WBCs were correctly identified. The performance metrics considered for the different image parameters are shown in Table 4.

Table 4: Performance metrics for different images

Images	Dice score	Accuracy	Precision	Recall	Simulation time (s)
1	0.98	0.99	0.98	0.99	1.12
2	0.98	0.98	0.98	0.98	0.98
3	0.96	0.99	0.99	0.98	0.87
4	0.99	0.99	0.98	0.99	1.46
5	0.99	0.98	0.98	0.99	1.34

Table 4 shows that the proposed HTsC model achieves the dice score value range of 0.96 to 0.99. The accuracy, precision, and recall value of the proposed HTsC are measured at 0.98 to 0.99, which is significantly higher for processing. The study of active contours without an adaptive mask was carried out to demonstrate adaptive mask generation. In this case, the detected nucleus was considered a mask and dilated by selecting the structuring element as a disk of size 50. The disk size was experimentally found to cover the largest WBC in the dataset. The simulation time for the proposed HTsC model is presented for the varying number of images. According to the simulation analysis, the proposed model has a simulation time of at least 0.87 s. It can be observed that the adaptive mask generation approach offers more accurate segmentation. The mask is initially generated without dilation. This method failed in many cases with round nuclei (lymphocytes). Because the nuclear borders are darker than the cytoplasm, a non-dilated mask detects nuclei rather than the cytoplasm. The precision and recall rates above 0.95 indicate that over-segmentation and under-segmentation cases were fewer, indicating that the proposed method's results matched well with the ground truth. The performance metrics considered for the analysis of the proposed HTsC are presented in Fig. 14.

To extract the WBCs accurately, the adaptive mask generation method was used to address the size and shape variations. The dataset consists of images with varying color shades, and hence the color transfer method was used to address this issue. The results indicate that the proposed method can accurately

detect the entire WBC region. An overall Dice score of 0.95 was obtained. The results of WBC segmentation directly affect the classification results. This is because the utilized classifiers are trained with the features extracted from WBCs. Accurate segmentation of WBCs leads to correct classification by considering the relevant feature set. Parts of the nucleus and cytoplasm need to be studied individually to identify the types of WBCs appropriately. Therefore, in the proposed method, segmentation of the WBC into the middle and cytoplasm was considered before the feature extraction step.
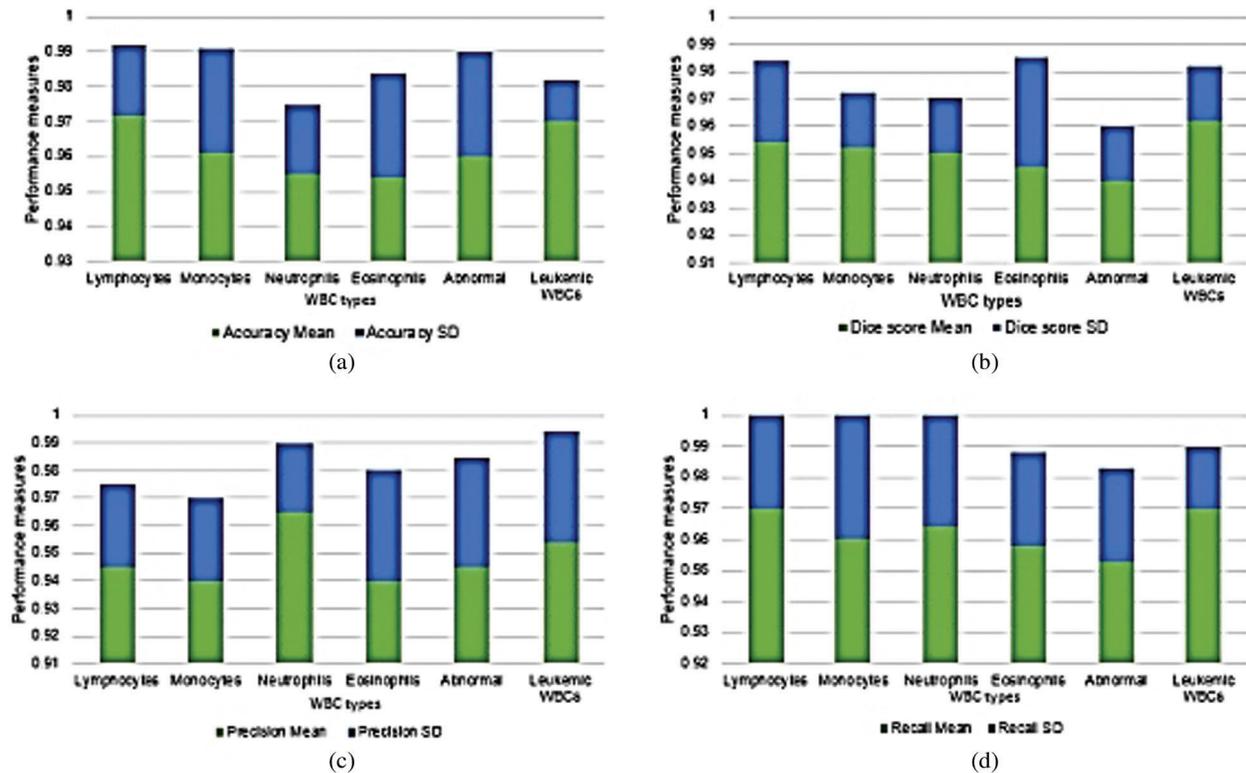

Figure 14: Performance measures of the WBC detection method for dataset 1 (a) mean and SD of accuracy (b) mean and SD of the dice score (c) mean and SD of precision rate (d) mean and SD of recall rate

The proposed HTsC model's performance is compared to that of the existing classifier. The comparative analysis of the proposed HTsC with the current classifier model is presented in Table 5.

Table 5: Comparative analysis

Method	Overall accuracy (%)
Otsu's method [14]	92
Color space decomposition [15]	91.3
Otsu's thresholding [18]	74
Proposed HTsC	98

The proposed HTsC model's performance is compared to the conventional technique, and it is stated that the proposed HTsC model has a higher overall accuracy value of 96%. The comparative analysis of the

proposed HTsC with the existing Otsu color space and Otsu thresholding demonstrated that the proposed HTsC model exhibits higher performance than the current techniques. The performance of the proposed HTsC is approximately 3%–13% higher than that of the conventional approach. Finally, the proposed HTsC method shows better accuracy than the existing methods.

5 Conclusion

An HTsC model for leukemia classification in blood smear images was presented in this paper. The goal of this study was to detect leukemia from peripheral blood smear images accurately. The images consisted of WBCs, RBCs, platelets, and staining artifacts. The presence of the inner nucleus and outer cytoplasm region distinguishes WBCs. Five types of WBCs were categorized in this work: lymphocytes, monocytes, neutrophils, and basophils. The shape, size, color, and texture features vary depending on the type of WBC. These features need to be studied individually for nuclei and regions of cytoplasm for accurate classification. The classification was done by considering two approaches: traditional machine learning and deep learning. To detect WBCs accurately, the active contours method, which is robust to shape and size variations, was used. The analysis stated that he proposes that the HTsC model exhibits higher overall accuracy than the conventional technique. The proposed HTsC model displays the limitation of its higher complexity for processing the different particles in the blood smear. The captured images need to be more apparent for an accurate diagnosis and higher accuracy. In the future, to reduce the complexity, the proposed model can use the pre-defined datasets for processing to increase the classification performance, such as through the ALEXNet architecture.

Conflicts of Interest: The authors declare that have no conflicts of interest to report regarding the present study.

References

[1] R. B. Hegde, K. Prasad, H. Hebbar and B. M. K. Singh, "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 382–392, 2019.

[2] S. Rajagopal, T. Thanarajan, Y. Alotaibi and S. Alghamdi, "Brain tumor: hybrid feature extraction based on unet and 3DCNN," *Computer Systems Science and Engineering*, vol. 45, no. 2, pp. 2093–2109, 2023.

[3] K. K. Anilkumar, V. J. Manoj and T. M. Sagi, "A survey on image segmentation of blood and bone marrow smear images with emphasis to automated detection of leukemia," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, pp. 1406–1420, 2020.

[4] R. V. Tali, S. Borra and M. Mahmud, "Detection and classification of leukocytes in blood smear images: State of the art and challenges," *International Journal of Ambient Computing and Intelligence (IJACI)*, vol. 12, no. 2, pp. 111–139, 2021.

[5] S. Rajaraman, S. Jaeger and S. K. Antani, "Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images, *PeerJ*, vol. 7, pp. e6977, 2019.

[6] P. A. Pattanaik, M. Mittal and M. Z. Khan, "Unsupervised deep learning cad scheme for the detection of malaria in blood smear microscopic images," *IEEE Access*, vol. 8, pp. 94936–94946, 2020.

[7]   R. B. Hegde, K. Prasad, H. Hebbar and B. M. K. Singh, "Development of a robust algorithm for detection of nuclei and classification of white blood cells in peripheral blood smear images," *Journal of Medical Systems*, vol. 42, no. 6, pp. 1–8, 2018.

[8]   S. Mishra, B. Majhi and P. K. Sa, "Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection," *Biomedical Signal Processing and Control*, vol. 47, pp. 303–311, 2019.

[9]   A. Genovese, M. S. Hosseini, V. Piuri, K. N. Plataniotis and F. Scotti, "Acute lymphoblastic leukemia detection based on adaptive unsharpening and deep learning," in *ICASSP 2021–2021 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, pp. 1205–1209, 2021.

[10] A. Genovese, M. Hosseini, V. Piuri, K. N. Plataniotis and F. Scotti, "Histopathological transfer learning for acute lymphoblastic leukemia detection," in *2021 IEEE Int. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Hong Kong, China, pp. 1–6, 2021.

[11] J. Amin, M. Sharif, M. Almas Anjum, A. Siddiqa, S. Kadry *et al.*, "3D semantic deep learning networks for leukemia detection," *Computers, Materials & Continua*, vol. 69, no. 1, pp. 785–799, 2021.

[12] S. Tavakoli, A. Ghaffari, Z. M. Kouzehkanan and R. Hosseini, "New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images," *Scientific Reports*, vol. 11, no. 1, pp. 1–13, 2021.

[13] R. B. Hegde, K. Prasad, H. Hebbar and B. M. K. Singh, "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 382–392, 2019.

[14] M. Ghaderzadeh, F. Asadi, A. Hosseini, D. Bashash, H. Abolghasemi *et al.*, "Machine learning in detection and classification of leukemia using smear blood images: A systematic review," *Scientific Programming*, vol. 2021, pp. 1–14, 2021.

[15] S. Khan, M. Sajjad, T. Hussain, A. Ullah and A. S. Imran, "A review on traditional machine learning and deep learning models for wbcs classification in blood smear images," *IEEE Access*, vol. 9, pp. 10657–10673, 2020.

[16] K. Dese, H. Raj, G. Ayana, T. Yemane, W. Adissu *et al.*, "Accurate machine-learning-based classification of leukemia from blood smear images," *Clinical Lymphoma Myeloma and Leukemia*, vol. 21, no. 11, pp. e903–e914, 2021.

[17] M. Grochowski, M. Wąsowicz, A. Mikołajczyk, M. Ficek, M. Kulka *et al.*, "Machine learning system for automated blood smear analysis," *Metrology and Measurement Systems*, vol. 26, no. 1, pp. 81–93, 2019.

[18] E. L. E. N. Abdullah and M. K. Turan, "Classifying white blood cells using machine learning algorithms," *International Journal of Engineering Research and Development*, vol. 11, no. 1, pp. 141–152, 2019.

[19] A. Ratley, J. Minj and P. Patre, "Leukemia disease detection and classification using machine learning approaches: A review," in *2020 First Int. Conf. on Power, Control and Computing Technologies (ICPC2T)*, Raipur, India, pp. 161–165, 2020.

[20] K. D. Gebremeskel, T. C. Kwa, K. H. Raj, G. A. Zewdie, T. Y. Shenkute *et al.*, "Automatic early detection and classification of leukemia from microscopic blood image," *Abyssinia Journal of Engineering and Computing*, vol. 1, no. 1, pp. 1–10, 2021.

[21] P. Rastogi, K. Khanna and V. Singh, "LeuFeatx: Deep learning–based feature extractor for the diagnosis of acute leukemia from microscopic images of peripheral blood smear," *Computers in Biology and Medicine*, vol. 142, pp. 105236, 2022.

[22] K. K. Jha and H. S. Dutta, "Mutual information based hybrid model and deep learning for acute lymphocytic leukemia detection in single cell blood smear images," *Computer Methods and Programs in Biomedicine*, vol. 179, pp. 104987, 2019.

[23] D. Umamaheswari and S. Geetha, "Review on image segmentation techniques incorporated with machine learning in the scrutinization of leukemic microscopic stained blood smear images," in *Int. Conf. on ISMAC in Computational Vision and Bio-Engineering*, Cham, Switzerland, pp. 1773–179, 2019.

[24] S. Alagu and K. B. Bagan, "Computer assisted classification framework for detection of acute myeloid leukemia in peripheral blood smear images," in *Innovations in Computational Intelligence and Computer Vision*, vol. 1189, Singapore: Springer, pp. 403–410, 2021.

[25] R. T. Raphael and K. R. Joy, "Segmentation and classification techniques of leukemia using image processing: An overview," in *2019 Int. Conf. on Intelligent Sustainable Systems (ICISS)*, Palladam, India, pp. 378–384, 2019.

[26] M. M. Alam and M. T. Islam, "Machine learning approach of automatic identification and counting of blood cells," *Healthcare Technology Letters*, vol. 6, no. 4, pp. 103–108, 2019.

[27] Y. M. Kassim, K. Palaniappan, F. Yang, M. Poostchi, N. Palaniappan *et al.,* "Clustering-based dual deep learning architecture for detecting red blood cells in malaria diagnostic smears," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1735–1746, 2020.

[28] A. Girdhar, H. Kapur and V. Kumar, "Classification of white blood cell using convolution neural network," *Biomedical Signal Processing and Control*, vol. 71, pp. 103156, 2022.

[29] C. Cheuque, M. Querales, R. León, R. Salas and R. Torres, "An efficient multi-level convolutional neural network approach for white blood cells classification," *Diagnostics*, vol. 12, no. 2, pp. 248, 2022.

[30] A. Meenakshi, J. A. Ruth, V. R. Kanagavalli and R. Uma, "Automatic classification of white blood cells using deep features based convolutional neural network," *Multimedia Tools and Applications*, vol. 81, pp. 30121–30142, 2022.

[31] Y. Alotaibi, "A new meta-heuristics data clustering algorithm based on tabu search and adaptive search memory," *Symmetry*, vol. 14, no. 3, pp. 623, 2022.

[32] N. Subramani, P. Mohan, Y. Alotaibi, S. Alghamdi and O. I. Khalaf, "An efficient metaheuristic-based clustering with routing protocol for underwater wireless sensor networks," *Sensors*, vol. 22, no. 2, pp. 415, 2022.

[33] Y. Alotaibi and A. Subahi, "New goal-oriented requirements extraction framework for e-health services: A case study of diagnostic testing during the COVID-19 outbreak," *Business Process Management Journal*, vol. 28, no. 1, pp. 273–292, 2021.

[34] Y. Alotaibi, "A new database intrusion detection approach based on hybrid meta-heuristics," *CMC-Computers Materials & Continua*, vol. 66, no. 2, pp. 1879–1895, 2021.

[35] S. Rajendran, O. I. Khalaf, Y. Alotaibi and S. Alghamdi, "MapReduce-based big data classification model using feature subset selection and hyperparameter tuned deep belief network," *Scientific Reports*, vol. 11, no. 1, pp. 1–10, 2021.

[36] R. Rout, P. Parida, Y. Alotaibi, S. Alghamdi and O. I. Khalaf, "Skin lesion extraction using multiscale morphological local variance reconstruction-based watershed transform and fast fuzzy C-means clustering," *Symmetry*, vol. 13, no. 11, pp. 2085, 2021.

[37] A. Alsufyani, Y. Alotaibi, A. Almagrabi, S. Alghamdi and N. Alsufyani, "Optimized intelligent data management framework for a cyber-physical system for computational applications," *Complex & Intelligent Systems*, pp. 1–13, 2021.

[38] J. Jayapradha, M. Prakash, Y. Alotaibi, O. I. Khalaf and S. A. Alghamdi, "Heap bucketization anonymity—An efficient privacy-preserving data publishing model for multiple sensitive attributes," *IEEE Access*, vol. 10, pp. 28773–28791, 2022.

[39] S. S. Rawat, S. Alghamdi, G., Kumar, Y. Alotaibi, O. I. Khalaf *et al.,* "Infrared small target detection based on partial sum minimization and total variation," *Mathematics*, vol. 10, no. 4, pp. 671, 2022.

[40] P. Mohan, N. Subramani, Y. Alotaibi, S. Alghamdi, O. I. Khalaf *et al.,* "Improved metaheuristics-based clustering with multihop routing protocol for underwater wireless sensor networks," *Sensors*, vol. 22, no. 4, pp. 1618, 2022.

[41] Y. Alotaibi, M. Malik, H. Khan, A. Batool, S. Islam *et al.,* "Suggestion mining from opinionated text of big social media data," *CMC-Computers, Materials & Continua*, vol. 68, no. 3, pp. 3323–3338, 2021.

[42] D. Anuradha, N. Subramani, O. I. Khalaf, Y. Alotaibi, S. Alghamdi *et al.,* "Chaotic search and rescue optimization based multi-hop data transmission protocol for underwater wireless sensor networks," *Sensors*, vol. 22, no. 8, pp. 2867, 2022.

[43] S. Bharany, S. Sharma, S. Badotra, O. I. Khalaf, Y. Alotaibi *et al.,* "Energy-efficient clustering scheme for flying Ad-hoc networks using an optimized LEACH protocol," *Energies*, vol. 14, no. 19, pp. 6016–6016, 2021.

[44] G. Li, F. Liu, A. Sharma, O. I. Khalaf, Y. Alotaibi *et al.,* "Research on the natural language recognition method based on cluster analysis using neural network," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–13, 2021.

[45] Y. Alotaibi, "Automated business process modelling for analyzing sustainable system requirements engineering," in *2020 6th Int. Conf. on Information Management (ICIM) IEEE*, London, UK, pp. 157–161, 2020.

[46] H. H. Khan, M. N. Malik, R. Zafar, F. A. Goni, A. G. Chofreh *et al.,* "Challenges for sustainable smart city development: A conceptual framework," *Sustainable Development*, vol. 28, no. 5, pp. 1507–1518, 2020.

[47] Y. Alotaibi, "A new secured E-government efficiency model for sustainable services provision," *Journal of Information Security and Cybercrimes Research*, vol. 3, no. 1, pp. 75–96, 2020.

[48] M. Abdel-Fattah, O. Al-marhbi, M. Almatrafi, M. Babaseel, M. Alasmari *et al.,* "Sero-prevalence of hepatitis B virus infections among blood banking donors in Makkah City, Saudi Arabia: An institutional-based cross-sectional study," *Journal of Umm Al-Qura University for Medical Sciences*, vol. 6, no. 2, pp. 4–7, 2020.

[49] V. Mani, P. Manickam, Y. Alotaibi, S. Alghamdi and O. I. Khalaf, "Hyperledger healthchain: Patient-centric IPFS-based storage of health records," *Electronics*, vol. 10, no. 23, pp. 3003–3003, 2021.

[50] K. Lakshmanna, N. Subramani, Y. Alotaibi, S. Alghamdi, O. I. Khalaf *et al.,* "Improved metaheuristic-driven energy-aware cluster-based routing scheme for IoT-assisted wireless sensor networks," *Sustainability*, vol. 14, no. 13, pp. 7712, 2022.

[51] S. S. Rawat, S. Singh, Y. Alotaibi, S. Alghamdi and G. Kumar, "Infrared target-background separation based on weighted nuclear norm minimization and robust principal component analysis," *Mathematics*, vol. 10, no. 16, pp. 2829, 2022.

[52] Y. Alotaibi and F. Liu, "A novel secure business process modeling approach and its impact on business performance," *Information Sciences*, vol. 277, pp. 375–395, 2014.