CrossMark

An Automatic and Robust Decision Support System for Accurate Acute Leukemia Diagnosis from Blood Microscopic Images

Zeinab Moshavash[1] · Habibollah Danyali[1] · Mohammad Sadegh Helfroush[1]

Abstract

This paper proposes an automatic and robust decision support system for accurate acute leukemia diagnosis from blood microscopic images. It is a challenging issue to segment leukocytes under uneven imaging conditions since features of microscopic leukocyte images change in different laboratories. Therefore, this paper introduces an automatic robust method to segment leukocyte from blood microscopic images. The proposed robust segmentation technique was designed based on the fact that if background and erythrocytes could be removed from the blood microscopic image, the remainder area will indicate leukocyte candidate regions. A new set of features based on hematologist visual criteria for the recognition of malignant leukocytes in blood samples comprising shape, color, and LBP-based texture features are extracted. Two new ensemble classifiers are proposed for healthy and malignant leukocytes classification which each of them is highly effective in different levels of analysis. Experimental results demonstrate that the proposed approach effectively segments leukocytes from various types of blood microscopic images. The proposed method performs better than other available methods in terms of robustness and accuracy. The final accuracy rate achieved by the proposed method is 98.10% in cell level. To the best of our knowledge, the image level test for acute lymphoblastic leukemia (ALL) recognition was performed on the proposed system for the first time that achieves the best accuracy rate of 89.81%.

Keywords Leukemia · Blood smear microscopic image · Leukocyte segmentation · Robust segmentation · Cell classification · Ensemble classifier

Introduction

Blood is one of itinerant tissues which plays an important role in health of the body. Erythrocytes, leukocytes, and platelet are three types of blood cells floating in liquid called plasma [1]. Leukocytes have vital roles in the body, hence extracting information about them is valuable for hematologist to diagnosis many diseases. Leukemia is a cancer of blood which starts from bone marrow then spreads into the bloodstream and other vital organs. In the case of leukemia, immature blast

✉ Habibollah Danyali
danyali@sutech.ac.ir

Zeinab Moshavash
z.moshavash@sutech.ac.ir

Mohammad Sadegh Helfroush
ms_helfroush@sutech.ac.ir

[1] Department of Electrical and Electronics Engineering, Shiraz University of Technology, P.O. Box 71555-313, Shiraz, Iran

cells or malignant leukocytes increase in the bone marrow and then enter in to the blood stream causing uncontrolled accumulation of blood cells [2]. It is known that leukemia can be fatal if left untreated. Based on lymphoid or myeloid stem cells become cancerous, leukemia can be divided into myeloid leukemia and lymphoblastic leukemia. There are two subtypes of lymphoblastic and myeloid leukemia: acute and chronic which show how fast leukemia progresses in the body. Therefore, leukemia is generally classified as acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), chronic myeloid leukemia (CML), and chronic lymphoblastic leukemia (CLL) [1]. ALL is a fast-growing cancer that affects children under the age of 5 years old and adults over 50 years of age. The recovery of patient depends on early diagnosing because symptom of ALL is similar to other common diseases such as the flu. Unfortunately, in most cases, the disease is not usually detected in early stages.

Manual microscopic leukocyte analysis is one of the available diagnostic procedures to recognize ALL in which normal and abnormal leukocytes existing in peripheral blood smears are distinguished and counted by skilled operator [3]. This

technique is not only tedious, repetitive time consuming and slow, but also it is affected greatly by human errors.

Due to the increase of computational power, image analysis and pattern recognition techniques have been utilized extensively to assist hematologist in analyzing blood cells. These tools lead to a more accurate and standard analysis in automatic computer-assisted microscopy (CAM) systems [4]. Image processing-based system not only improves accuracy and speed of manual methods but also saves time, manpower, and costs. A captured image from peripheral blood smear under microscope is the only input of these systems for ALL detection. Then blood cells are separated from background through segmentation step and leukocytes are used to extract lymphocytes. Healthy lymphocyte and lymphoblast (malignant lymphocytes) are distinguished by analysis of external and morphological deformation of cytoplasm and nucleus of cells [5].

Many attempts have been made in the past to assist hematologist in analyzing blood smear images for ALL detection [6–10]. Generally, the most important step in automatic blood smear image analysis is leukocyte segmentation. Hence, a huge amount of work has been performed to extract leukocytes from other blood components [11–21]. In [21], a leukocyte segmentation technique is suggested in which leukocyte nuclei are identified by Gram-Schmidt orthogonalization-based method. Then basophils are isolated from other types of leukocytes via extracting some features from nuclei and finally it segments leukocyte consisting of nucleus and cytoplasm using snake algorithm [22]. A nucleus segmentation technique using multilevel Otsu thresholding is presented in [23] which leukocyte nucleus enhancer is introduced to achieve region of interest by enhancing the nucleus area. Ghosh et al. [12] developed a leukocyte segmentation algorithm using hedge-operator-based fuzzy divergence. Meanwhile, few authors claimed to achieve robust segmentation performance under variable staining and uneven imaging conditions [14, 24].

A robust approach for segmentation of leukocyte's nucleus, which performs quite well for noise-affected images, was reported in [14] using new exponential intuitionistic fuzzy divergence-based thresholding method. The threshold value is attained by minimization of intuitionistic fuzzy divergence between the ideally thresholded image and actual image.

Partial analysis of blood microscopic images of leukemic patient have been addressed by multiple authors [7, 11, 14, 25, 26]; few examples of automated ALL detection systems that can extract and analyze leukocytes and discriminate lymphoblast (malignant leukocytes) from healthy leukocytes have been reported in the literature [24, 27, 28].

Putzu et al. [27] presented an automatic system for ALL detection. Leukocytes segmentation was performed by thresholding-based segmentation methods including Zack and Otsu algorithms. Presence of overlapped leukocytes and abnormal component were considered throughout leukocyte extraction process. The total number of 131 features including shape, color, and texture descriptors was utilized to discriminate lymphoblast from healthy leukocytes. According to reported experimental results, the best accuracy of 93.2% was attained by RBF kernel-based SVM under ten-fold cross validation. It is known that the proposed approach by Putzu presents a precise attempt in order to provide a fully automatic procedure to support medical activity for ALL diagnosis which is used as a based reference for next works [24]. The advantage of [27] is that the proposed system has been tested on the standard dataset which is available in public so its results are comparable with other similar systems. Although the presence of overlapped leukocytes and abnormal component were considered through leukocyte extraction process, only sub-images containing individual leucocytes were utilized for classification performance evaluation so effects of adjacent leukocytes and abnormal component are disregarded in final classification results. Moreover, the segmentation procedure is not robust against various types of blood smear images.

Mohapatra et al. [28] improved an automatic ALL diagnosing system based on hematologist visual criteria. A shadow C-means-based segmentation algorithm is used to identify cytoplasm and nucleus of lymphocytes. According to the malignant lymphocytes characteristics, as suggested by the hematologist, 44 features were extracted from segmented lymphocyte sub-images. Using an effective ensemble of classifiers led to a 99% accuracy. The drawback observed in this system is that the numerical results of this paper are not comparable with other related works since they employed their proposed system on their own dataset.

Neoh et al. [24] proposed an intelligent decision support system for ALL detection. To achieve a robust segmentation performance for nucleus and cytoplasm of lymphocytes/lymphoblasts, a novel clustering approach with stimulating discriminant measures (SDM) of both within- and between-cluster scatter variances was developed. An entire set of 80 features comprising shape, texture, and color descriptors of the nucleus and cytoplasm regions was used to distinguish normal and abnormal lymphocytic cells. Multi-layer perceptron, support vector machine (SVM), and Dempster-Shafer ensemble are implemented to classify healthy and malignant lymphocytes in which the Dempster-Shafer ensemble classifier produced the highest accuracy rate of 96.72% in cell level test.

The proposed approaches reported for leukocytes segmentation and ALL detection are greatly susceptible to their dataset, hence they cannot be used for different kinds of images. In fact, the color and intensity of provided images in different laboratories are different because of uneven lighting condition, variable blood smear staining techniques, and using image capturing devices with different calibrations.

Meanwhile, a common drawback observed in the existing ALL detection systems is that they often classify leukocyte sub-images cropped from the whole blood microscopic images. In this case, some important problems in blood microscopic image analysis are not considered, subjects such as presence of overlapped leukocytes and abnormal components, which are required to receive great attention in real application. Accordingly, to improve some of the flaws related to the previous researches, as mentioned above in the present work, an automatic image processing-based system for ALL detection is proposed. In this system, leukocytes are first segmented from blood microscopic image robustly, then a new feature set is extracted from segmented leukocytes and finally an ensemble classifier is applied to discriminate malignant leukocytes (lymphoblast) from healthy ones.

The novelties of the proposed system are as follows:

I) Three assumptions about blood smear images have been defined, which are used as principles of robustness in the proposed leukocyte segmentation method.

II) In contrast to many available leukocyte segmentation methods, the proposed approach for leukocyte segmentation is not susceptible to dataset and it is able to find out leukocytes in different kinds of blood microscopic images.

III) A robust method of leukocyte nucleus segmentation from blood smear images has been presented in which we improve a leukocyte nucleus segmentation stage based on intuitionistic fuzzy divergence, by selection of appropriate color space for fuzzy calculation. In order to demonstrate robust performance of the proposed segmentation algorithms, they have been examined on images of various datasets.

IV) An effective leukocytes separation procedure based on watershed algorithm is suggested.

V) An impressive collection of 53 discriminator features for healthy and malignant leukocyte classification has been introduced in which using local binary pattern (LBP)-based texture descriptors instead of gray-level co-occurrence matrix (GLCM)-based descriptors results in better demarcation between healthy and malignant leukocytes.

VI) Two ensemble classifiers have been developed in this paper to improve classification accuracy of previous ALL detection systems. These classifiers are highly effective in different levels of analysis.

VII) To the best of our knowledge, the image level test for ALL recognition which labels blood microscopic images instead of blood cells is performed on the proposed system for the first time to challenge and evaluate our system more effectively.

The rest of this paper is organized as follow. The "Proposed Method" section describes the proposed ALL detection system in detail including Leukocyte segmentation, feature extraction, and classification. The "Datasets" section gives a brief description of utilized image dataset. The "Experimental Results" section enlists the experimental results of the proposed system obtained in different level of simulation and concludes with some comparison results. Finally, "Conclusion" section concludes the paper.

Proposed Method

Generally, the proposed robust decision support system for accurate acute leukemia diagnosis from blood microscopic images is described in three main sections and is shown in Fig. 1. We aim to have a robust system which means that it will be able to work on different kinds of dataset provided in laboratories with various imaging conditions. As the system is divided into three main sections, it is expected that each section performs robustly.

Leukocyte Segmentation

The blood microscopic image is usually in RGB color space which can be separated into four significant regions including leukocytes nucleus area, leukocytes cytoplasm area, erythrocytes, and background. However, different imaging condition and staining methods would vary the color of these regions, using unchangeable features of blood microscopic image could help to segment these regions under uneven imaging conditions. Therefore, three assumptions are defined in this paper as follows:

1. The darkest purple area shows leukocytes nucleus regions.
2. The lightest area belongs to background.
3. Erythrocytes are nucleus free component.

Unlike the fact that the word assumption may bring to mind some limitation for proposed approach, it causes the method to be applicable and valid for different types of dataset acquired in different staining and imaging conditions. Therefore, these assumptions have been used to design a robust leukocyte segmentation method.

Figure 2 shows the flowchart of the proposed algorithm for robust leukocyte segmentation, which gives a detailed depiction of six steps for extraction of leukocyte regions.

In contrast to other approaches that identify leukocytes directly [11, 21, 24, 27, 28], the proposed method looks for erythrocytes and other components in the images and removes them, so that what remains are leukocytes. Furthermore, adjacent cells separation and removal of abnormal component are considered in the proposed segmentation method that is disregarded in many approaches. The proposed method can be illustrated in terms of the following steps.

Fig. 1 The main sections of the
proposed system

Step 1: Leukocytes nucleus and background extraction

In this step, leukocytes nucleus is identified using first assumption. Leukocyte nucleus extraction plays a key role in the proposed approach since the final leukocytes segmentation result relies greatly on the correct nucleus identification. In other words, every minor errors of this step cause a large error in the final results. Therefore, a correct choice of nucleus segmentation algorithm is very crucial.

Many leukocytes nucleus segmentation algorithms have been persisted in the literature [14, 21, 29]. In [14], a robust technique has been proposed to segment leukocytes nucleus automatically by using a new

Fig. 2 Flow chart of the proposed leukocyte segmentation algorithm

exponential intuitionistic fuzzy divergence (IFD)-based thresholding algorithm. In this method, gray level of image is used which is not an appropriate space for robust segmentation. In order to have a robust segmentation method for various types of blood smear image datasets, in this paper a* component of La*b* color space is used which leukocytes nucleus are more contrasted in. Figure 3 shows segmentation results of IFD-based thresholding using grayscale and a* component of La*b* color space. The method has been tested on four different images from various types of datasets.

It is clear that using a* component of La*b* color space makes IFD-based thresholding method effectively robust against uneven imaging condition. In fact, IFD-based thresholding method is a search process for finding two optimum thresholds (t1, t2) that can divide blood microscopic image into three regions. Therefore, the improved method has been utilized for leukocytes nucleus segmentation step which is described briefly as follows.

1. Original RGB blood microscopic image is converted into La*b* color space. The image obtained from the a* component of this color space is used as the input image later on.

2. As mentioned before, two thresholds are required for blood microscopic image. For each pair of (t1, t2) varying from 0 to 255 image is divided into three distinct regions r1, r2, and r3. Then if $f(r_i)$ shows intensity of any pixel located in the region i ($i = 1, 2, 3$), we have

$$0 \leq f(r_1) < t_1 \leq f(r_2) < t_2 \leq f(r_3) < 255 \qquad (1)$$

3. The membership of every pixel (i, j) is defined as follows:

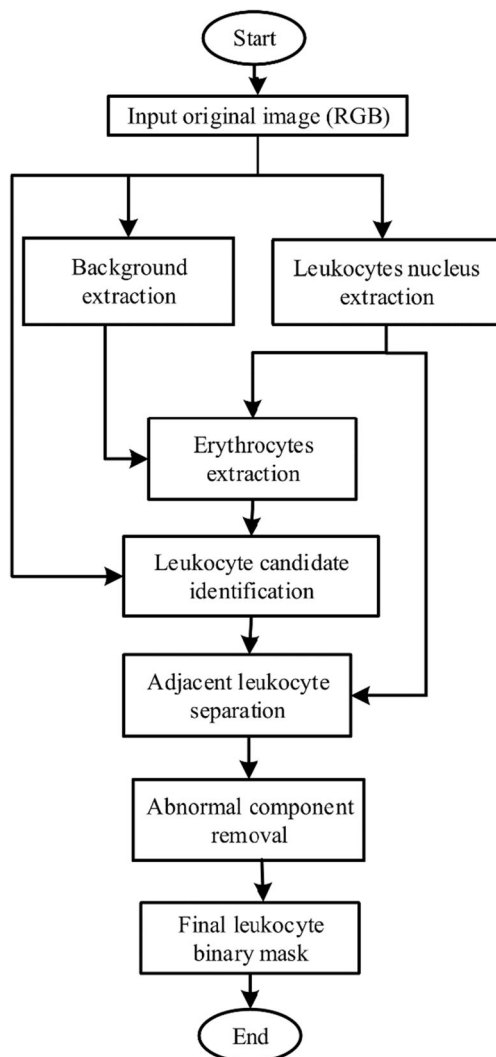$$\mu(i,j) = \exp\left(\frac{f(i,j) - m_k}{2\sigma^2}\right) \qquad (2)$$

where m_k and σ^2 are mean intensity of region k and variance, respectively. The variance is taken as follows:

$$\sigma^2 = f_{max} - f_{min} \qquad (3)$$

4. Non membership value is evaluated as follow:

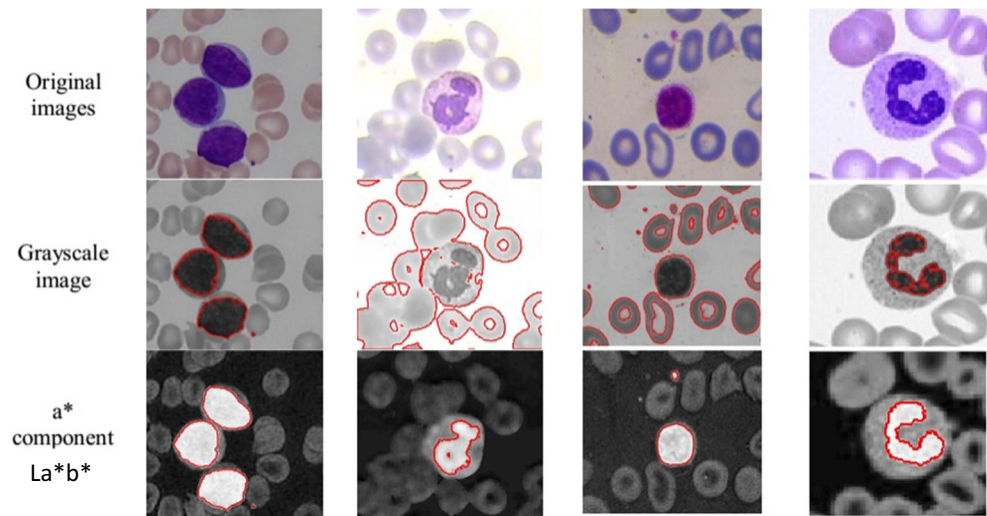$$\nu(i,j) = \frac{1 - \mu(i,j)}{1 + \lambda\mu(i,j)} \; ; \; \lambda > 0 \qquad (4)$$

Which have $\lambda = 2$ (similar to [14]).

Fig. 3 Examples of blood images after the leukocyte nucleus identification process by IFD-based thresholding using different color spaces (border highlighted)

Hence, input image is defined as an intuitionistic fuzzy set. Generally, IFD between two images A and B of size M × N is defined as follows:

$$
\text{IFD} = \frac{1}{2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \begin{bmatrix} 4 - \{1 - \mu_A(a_{ij}) + \mu_B(b_{ij})\} e^{\mu_A(a_{ij}) - \mu_B(b_{ij})} \\ - \{1 - \mu_B(b_{ij}) + \mu_A(a_{ij})\} e^{\mu_B(b_{ij}) - \mu_A(a_{ij})} \\ - \{1 - \nu_B(a_{ij}) + \nu_B(b_{ij})\} e^{\nu_A(a_{ij}) - \nu_B(b_{ij})} \\ - \{1 - \nu_B(b_{ij}) + \nu_A(a_{ij})\} e^{\nu_B(b_{ij}) - \nu_A(a_{ij})} \end{bmatrix} \quad (5)
$$

5. Let A be thresholded image by $(t1, t2)$ and B be the ideally thresholded image. Then $\mu_B = 1$ and $\nu_B = 1$.

Hence, Eq. (5) is simplified as follows:

$$
\text{IFD} = \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \begin{bmatrix} 4 - \{2 - \mu_A(i,j)\} e^{\mu_A(i,j)-1} - \mu_A(i,j) e^{1-\mu_A(i,j)} \\ - \{1 - \nu_A(i,j)\} e^{\nu_A(i,j)} - \{1 + \nu_A(i,j)\} e^{-\nu_A(i,j)} \end{bmatrix} \quad (6)
$$

6. IFD is calculated for every possible pair of thresholds $(t1, t2)$. The best thresholds are those which minimize IFD and finding $T = t2$ $(t2 > t1)$ is selected as optimum threshold to extract leukocytes nucleus.

Detailed derivation of IFD formula and thresholding can be found in [14].

After nucleus extraction, it is required to isolate blood cells from the background. Many attempts have been made in the literatures to extract background [27]. While a whitish background of blood microscopic image is easy to extract, a robust method is necessary since background color is affected by noise, staining technique and imaging condition. Robust background extraction is implemented by conversion of RGB image into the CMYK color space, because blood components are more contrasted in M component of CMYK color models. Figure 4 shows visual results for CMYK color transformation.

In fact, Magenta color is presented in all blood components; hence, background can be extracted in this color component easily. Here, the threshold value is achieved by Zack algorithm which is a histogram-based method. Generally, a blood smear image histogram contains peaks and valleys.

Based on assumption 2, the highest peak region is often belong to the background and the threshold value is often located in the valley beside that. Based on Zack algorithm, the accurate threshold value is particularly determined when high and weak peaks of histogram are separated by clear valleys in which M component of CMYK provides a histogram with clear peaks and valleys (see Fig. 5).

The detailed algorithm for background extraction is as follow:

1. Original RGB image is converted into CMYK color model.
2. Local maximums of the M component histogram are determined.
3. The highest and the weakest local maximum are connected by a straight line.
4. The distance between marked line and histogram values is calculated.
5. The threshold value is a gray level where the distance reaches its maximum.

Step 2: Erythrocytes extraction

Based on assumption 3 ("Leukocyte Segmentation" section), nucleus could not be found inside erythrocytes. So in this level, we aim to segment erythrocyte using this characteristic. Leukocyte nucleus and background extraction results in two binary masks which partition image into three regions, each background connected component which has no nucleus inside is labeled as erythrocytes area, but the obtained area does not contain the whole existing erythrocytes.

Step 3: Leukocytes candidate extraction

As mentioned in previous step, only some of erythrocyte areas are extracted so other background connected components contain at least one leukocyte and the adjacent erythrocytes. If
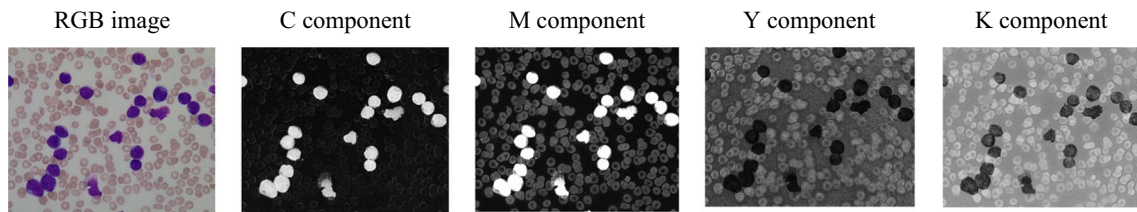
Fig. 4. A visual result for CMYK color transformation

these erythrocytes can be removed efficiently, the remaining area can be signified as leukocytes candidate area.

RGB colors associated with extracted erythrocytes area are utilized to identify all existing erythrocytes. Based on these color vectors, the whole leukocytes consist of cytoplasm and nucleus are extracted as follows:

Let B and I be the background binary mask and original M × N × 3 image respectively, then C can be defined as follows:

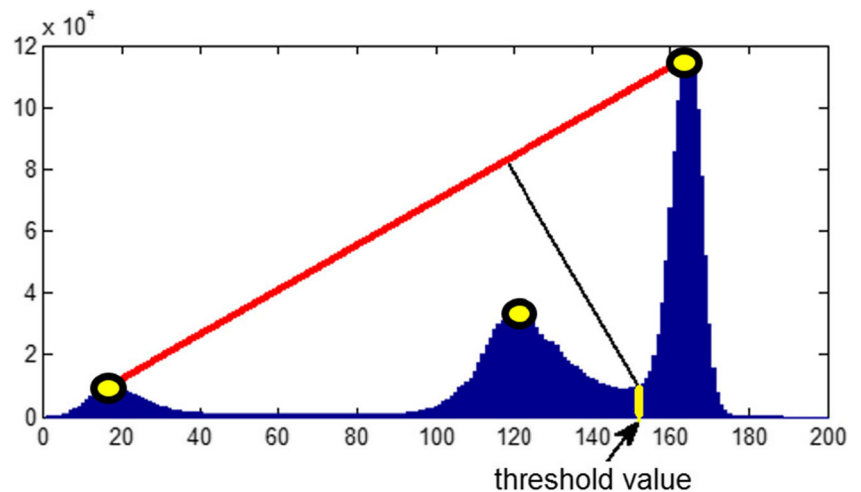$$C(i, j) = \begin{cases} I(i, j) & if \ B(i, j) = 1 \\ 0 & otherwise \end{cases} \quad (7)$$

suppose $V = \{v_1, v_2, v_3, \ldots, v_m\}$ denotes erythrocytes color set in which v_i is known as a 1 × 3 vector, so the following equation is applied to attain leukocyte candidate binary mask W.

$$W(i, j) = \begin{cases} 1 & if \ C(i, j) \notin V \& B(i, j) = 1 \\ 0 & otherwise \end{cases} \quad (8)$$

Step 4: Adjacent leukocytes separation

Adjacent leukocytes usually exist in blood microscopic images during blood smear preparation. Adjacent leukocyte separation is an important issue because it is not possible to extract correct features in the presence of overlapped leukocytes. Researches show that most of leukocyte identification methods are not able to separate overlapped leukocyte. These methods select dataset with adjacent leukocyte free images in order to make it possible, but in the real world scenario, it is necessary to take this level into account.

Watershed algorithm [30] is particularly effective when there are overlapped objects in the image, but it is not usually applied to the intensity image directly since it produces inaccurate segmentation lines. To improve segmentation accuracy, distance transform is applied to overlapped object binary mask before watershed algorithm implementation which results in better segmentation lines. Although distance transform improves watershed performance, when adjacent objects are not with round shape, the algorithm does not perform well. In this case, a marked image is often superimpose to distance image, in order to improve watershed segmentation lines. The marked image contains connected blobs of pixels within each of the objects which shows an overall estimation of center and shape of overlapped objects. The finer the marked image gets, the more accurate the separation lines are produced.

In this section, a new nucleus-based marked image is introduced in order to obtain separation lines which best match to the leukocytes contour. The details of the proposed adjacent Leukocytes separation process are as follows:

1. Roundness measure is used to identify the presence of adjacent leukocytes which is defined as follows:

$$Roundness = \frac{4\pi \times area}{convex \ perimeter} \quad (9)$$

2. Objects having roundness value smaller than the 0.8 (as used in [27]) are labeled as adjacent leucocytes and proceed to the next step.

Fig. 5 Histogram based threshold selection by Zack algorithm

3. Distance transform is applied to adjacent leukocytes binary mask and the regional maximal of the image is computed as a binary mask for further analysis.

4. Regional maxima is replaced with predefined markers, each of which represents an object.

5. Using markers, number of objects presented in each leukocytes nucleus region is counted then if the region contains at most one object it replaces with its convex hull otherwise it remains unchanged.

6. The new binary mask obtained from previous level is utilized to construct marker image using distance transform and finding regional maxima.

7. Distance transform is applied to leukocytes binary mask and new marker image is superimpose to the resulting image.

8. Finally, watershed algorithm is applied to separate adjacent leukocytes.

Figure 6 illustrates an example of the final separation results of seven adjacent leukocytes.

Step 5:   Abnormal component removal

The presence of abnormal component in blood microscopic image degrades the accuracy of image processing-based analysis. In leukemia detection, it requires the attention of researchers since abnormal component are usually extracted through leukocyte segmentation step and labeled as leukocyte that causes classification errors. Although the removal of all abnormal components is a complicated process, solidity measure is practically effective when abnormal components are with irregular shape in which ratio of the area of an object to the area of a convex hull of the object is calculate as solidity value.

$$Solidity = \frac{area}{convex\ area} \qquad (10)$$

It can be observed from experimental results that a threshold value equals to 0.93 can be used to signify abnormal component from leukocytes. Using leukocyte binary mask, connected components having solidity value greater than 0.93 are labeled as leukocyte.

Feature Extraction

The percentage of lymphoblast present in the peripheral blood or bone marrow samples is an essential criteria during ALL diagnosis. Discriminating lymphoblast from lymphocytes is a difficult task because they have some similar visual characteristics, so here we have used the current visual criteria which is followed by hematologists [31] to differentiate lymphoblast from lymphocyte. In fact, nuclear and cytoplasmic changes are two basic characteristics for distinguishing normal and blast lymphocytes. Nuclear changes include variation in shape, size, boundary, and chromatin pattern which are used not only for ALL detection but also for other types of
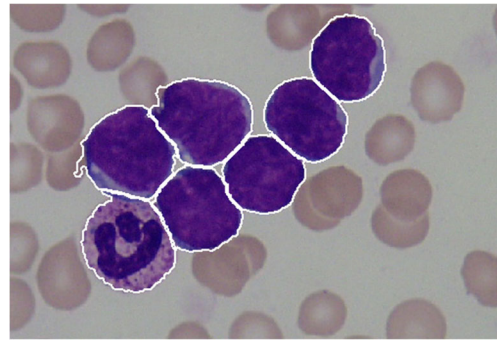

Fig. 6 An example of overlapped leukocytes separation

leukemia and some other cancer types. Cytoplasmic changes consist of variation in amount of cytoplasm, boundary contour, and chromatin pattern that are useful for ALL detection. In addition, the ratio between the area of the cytoplasm and the nucleus is utilized to indicate the maturity of a cell.

In image processing-based system, the quantitative features are required for nuclear and cytoplasmic change investigation so a total of 53 features consisting of 15 shape, 32 texture, and 6 color descriptors are computed from extracted nucleus and cytoplasm regions. Shape descriptors such as area, perimeter, convex area, convex perimeter, major axis, and minor axis are extracted from binary version of nucleus and cytoplasm image directly which are used to calculate other shape descriptors.

The 15 shape features are nucleus perimeter, nucleus area, cytoplasm area, nucleus to cytoplasm ratio, nucleus shape measure (roundness, elongation, compactness, and shape factor), nucleus boundary (Hausdroff dimension (HD)), nucleus, and cell contour signature (variance, skewness, and kurtosis of all the distances between nucleus or cell centroid and contour pixels).

Although shape features are considered as essential discriminating factors, they are susceptible to segmentation error therefore texture parameters of leukocytes are used which are extracted from gray scale image. The gray-level co-occurrence matrix (GLCM) is widely used for texture analysis which is applied to gray level of an image. To extract GLCM-based texture features, statistical measures such as energy, contrast, correlation, and homogeneity are computed from the 2D gray-level co-occurrence matrix (each component of matrix indicates probability of two pixels having particular gray levels at particular spatial relationships). These features are calculated for angles of 0, 45, 90, and 135. Many histopatological cell detection methods employ local binary pattern (LBP)-based texture features [32–34] which are robust against illumination changes. Therefore to have a robust feature set, LBP-based features have been used in this work in which LBP operator is applied on gray scale image before calculating GLCM features. In LBP method, each image pixel is compared with its neighborhood pixels having particular spatial distance in order to produce a binary pattern, whereas if the pixel intensity is

higher than that of the neighborhood sample the binary digit of pattern is set to 1 otherwise is set to zero.

As texture features are extracted from gray scale image, they are affected by background intensity causing inaccurate texture features. To solve this problem, all pixels intensity are shifted by 1 and background pixels are set to zero using background binary mask. Then, the first row and column of co-occurrence matrix associated with the new image have to be removed to attain improved co-occurrence matrix in which influence of background texture is eliminated. Texture features are extracted from both nucleus and cytoplasm regions. Thus the total number of texture features is 32.

As binary mask and gray level image do not contain color information of the leukocyte, it is required to extract some features from color channels of the RGB image. Mean color intensity in red, green, and blue color channels is computed as color feature to represent color changes. These features are measured for nucleus and cytoplasm regions.

Generally, the final feature vector for each leukocyte consists of 53 descriptors with dissimilar range of value in which lower-magnitude data are masked by higher-magnitude data. To overcome this problem, feature normalization is applied to standardize the dataset. For n feature vectors with i features, each input feature x_i^n can be normalized as follows:

$$x_{T_i}^n = \frac{x_i^n - \overline{x}_i}{\sigma_i} \tag{11}$$

where $\overline{x}_i$ and σ_i are mean and standard deviation values of ith feature along all feature vectors respectively.

Classification

In this phase, the goal is to discriminate lymphoblast from healthy leukocytes. In fact, classifiers are implemented to divide feature vectors into various classes using feature similarity. Although many researchers utilize a single classifier for healthy and malignant leukocytes [27, 29], they are not efficient in complex pattern recognition problem with complex decision boundary to be learned. Thus, ensemble classifiers are employed instead of single classifiers in order to provide more stable predictions for noisy data. An ensemble classifier is formed by combination of multiple individual classifiers to have multiple views of the same classification problem which is effectively robust in terms of classification result. Two ensemble classifiers are introduced in this paper which are beneficial for different levels of applications.
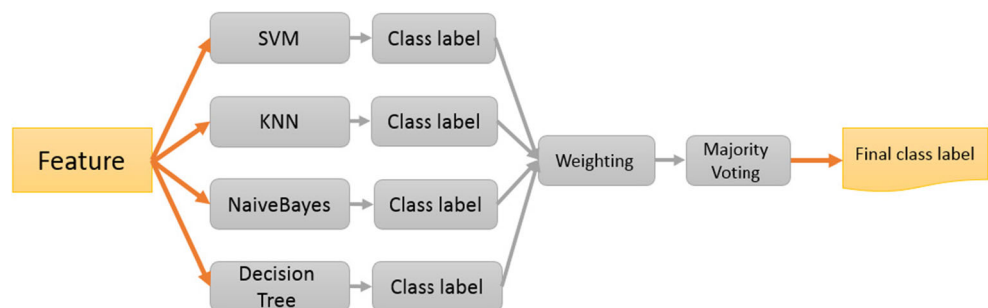
Ensemble Classifier 1

Support vector machine (SVM), K-nearest neighbor (KNN), naive Bayes (NB), and decision tree (DT) are the most common classifiers that are used for leukocyte classification [27]. A collection of these classifiers are employed to construct the ensemble classifier in which each of classifiers is trained with the same training set individually, then the classifiers are tested with test feature vectors, combination of class label introduced by the individual classifiers is selected as the final class label. Here, the majority voting principle is performed to combine class labels in which the final class label equals to the class label chosen by most of the classifiers. Weighted majority voting is based on the idea that not all classifiers should have the same amount of influence over the final decision.

From experimental results presented in [27], SVM has showed good results for lymphoblast discrimination compared with other classifiers so the weight associated with SVM is 2 to increase the influence of SVM in the final decision. Block diagram of proposed ensemble classifier 1 is depicted in Fig. 7. This classifier is practical when evaluations are in cell levels which is discussed in next section.
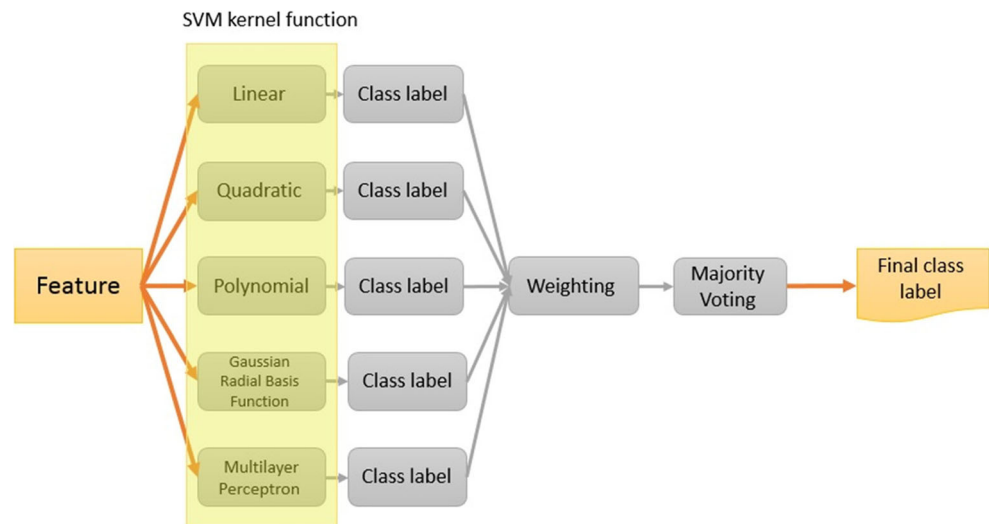
Ensemble Classifier 2

In [27], SVM with the most common kernel has been investigated. However, Gaussian radial basis kernel-based SVM achieved the best accuracy in ALL recognition; our ensemble classifier 2 has utilized five SVM kernel function: linear, quadratic, polynomial, multi-layer perceptron, and Gaussian radial basis, to achieve more robust performance specially when the whole blood microscopic image consists of overlapped leukocytes, abnormal shape and noise, is considered and classified. In this case, a more complex model is required to be employed, which is not possible by an individual kernel, hence combination of the SVM classifiers with various types

Fig. 7 Ensemble classifier 1 for lymphoblast discrimination in cell level

Fig. 8 Ensemble classifier 2 for lymphoblast discrimination in image level

of kernels that generate a different decision boundary may provide better classification results and reduce the total error. Second proposed ensemble classifier is shown in Fig. 8.

Datasets

In this paper, the proposed segmentation method has been applied on images from three datasets which are provided under uneven imaging and staining condition with different qualities in order to evaluate robustness of the method. Generally, the aim of proposed system is to identify lymphoblast in blood microscopic images in order to detect ALL. In fact, many of ALL detection systems are tested with their own datasets which are not public available datasets. ALL-IDB is a standard, public, and free available blood microscopic image dataset taken from healthy individuals and leukemic patients which has been proposed by Ruggero Donida Labati et al. [5]. We have used this dataset in order to evaluate our system and fairly compare with some available systems using ALL-IDB dataset. The ALL-IDB dataset includes two distinct versions microscopic images (ALL-IDB1 and ALL-IDB2). The ALL-IDB1 is composed of 108 images which can be used both for evaluating classification systems, as well as testing segmentation capability of algorithms. The ALL-IDB2 includes 260 blood microscopic images. This image set can be used for testing the performances of classification systems in ALL detection. This image set is a collection of cropped area which indicates normal and blast cells belong to the ALL-IDB1 dataset. Authors usually use this dataset in order to evaluate their systems.

Dataset 1 Samples of ALL-IDB were collected at M. Tettamanti Research Center–Monza, Italy. The images were acquired using optical laboratory microscope at different magnifications, ranging from 300 to 500. ALL-IDB dataset images were captured by Canon PowerShot G5 camera and saved in jpg format with 24-bit color depth. Table 1 shows a short description of ALL-IDB dataset

Dataset 2 The samples were collected at Hematology–Oncology and BMT Research Center of Imam Khomeini hospital in Tehran, Iran. The dataset including 400 samples were stained by Gismo-Right technique. They were attained by a light microscope (Microscope-Axioskope 40) using an achromatic lens with a magnification of 100. The images were taken by a digital camera (Sony Model No. SSC-DC50AP) and saved in BMP format in which the resolution is 720×576 pixels. The dataset includes manual leukocyte segmentation which has been performed by an expert.

Dataset 3 The samples were collected at Special Hematology Department of the General Hospital "Dr. Juan Bruno Zayas Alfonso" from Santiago de Cuba. The microscope slides were stained by Giemsa technique. The images were acquired by

Table 1 Characteristic of ALL-IDB dataset

Image acquisition setup		
Camera	Canon PowerShot G5	
Magnification of the microscope	300 to 500	
Image format	JPG	
Color depth	24 bit	
	ALL-IDB1	ALL-IDB2
Images	109	260
Resolution	2592×1944	257×257
Elements	39,000	260
Candidate lymphoblast	510	130
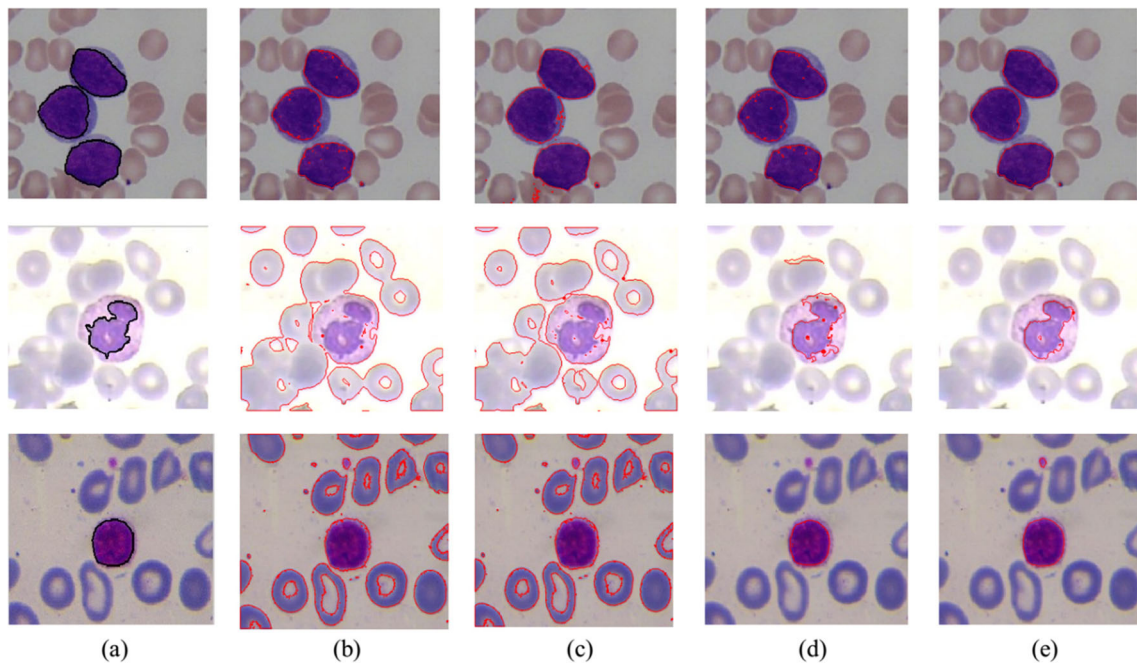
Fig. 9 Leukocyte nucleus segmentation result under uneven imaging condition **a** ground truth segmentation result obtained by expert, **b** IFD-g result, **c** Otsu's result, **d** GSO result, and **e** result of the proposed method

Leika microscope with ×100 augmented lens and a Kodak EasyShare V803 camera. Each image has been saved in jpg format with size of 3264 × 2448 pixels and depth of 24 bits per pixel.

Experimental Results

All the algorithms were implemented with MATLAB on a Windows 7 operating system with an 8 GHz Intel Core i7 CPU and 2 GB memory.

Nucleus Segmentation Results

The proposed nucleus segmentation method based on intuitionistic fuzzy divergence using a* component of La*b* color space has been compared with three competitor nucleus
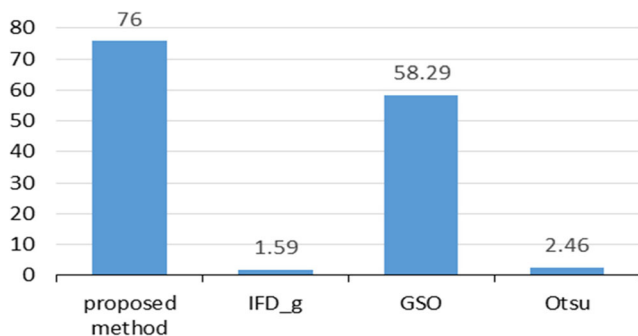
Fig. 10 Average similarity measure of four nucleus segmentation methods

segmentation algorithms including Otsu method [35], nucleus segmentation using Gram-Schmidt orthogonalization (GSO) [21], and intuitionistic fuzzy divergence based for gray level thresholding (IFD-g) [14]. Figure 9 illustrates three examples of segmental results using images taken from various datasets without any extra processing to ensure the robust performance of our method.

Images of first, second, and third lines are taken from Dataset 1, Dataset 2, and Dataset 3 respectively. It can be clearly observed that the proposed technique for robust nucleus segmentation is much satisfactory intuitively and quite better than the other algorithms.

In order to have a quantitative evaluation, the proposed nucleus segmentation algorithm has been compared with three mentioned nucleus segmentation algorithms in term of similarity measure using the same dataset (Dataset 2). In this case, we have a fair comparison. The quantitative measures for calculating the similarity evaluation of the nucleus segmentation approaches are described as follows:

$$Similarity = 100 \times \frac{A_{algorithm} \cap A_{expert}}{\max\left(A_{algorithm}, A_{expert}\right)} \quad (12)$$

where $A_{algorithm}$ and A_{expert} are the area of segmented region by algorithm and expert respectively.

Dataset 2 is a challenging dataset in the field of robustness testing, since the images used in this dataset include a wide range of colors and intensity variations with blur regions.

As seen in Fig. 10, the proposed nucleus segmentation algorithm has a significantly better segmentation performance in terms of similarity measure when compared to other

Fig. 11 **a** An original blood microscopic image, **b** leukocyte nucleus, **c** background, **d** three intensity regions of image using nucleus and background binary mask, **e** erythrocytes area, **f** connected component contained leukocyte, **g** leukocyte candidate, and **h** final leukocyte segmentation results
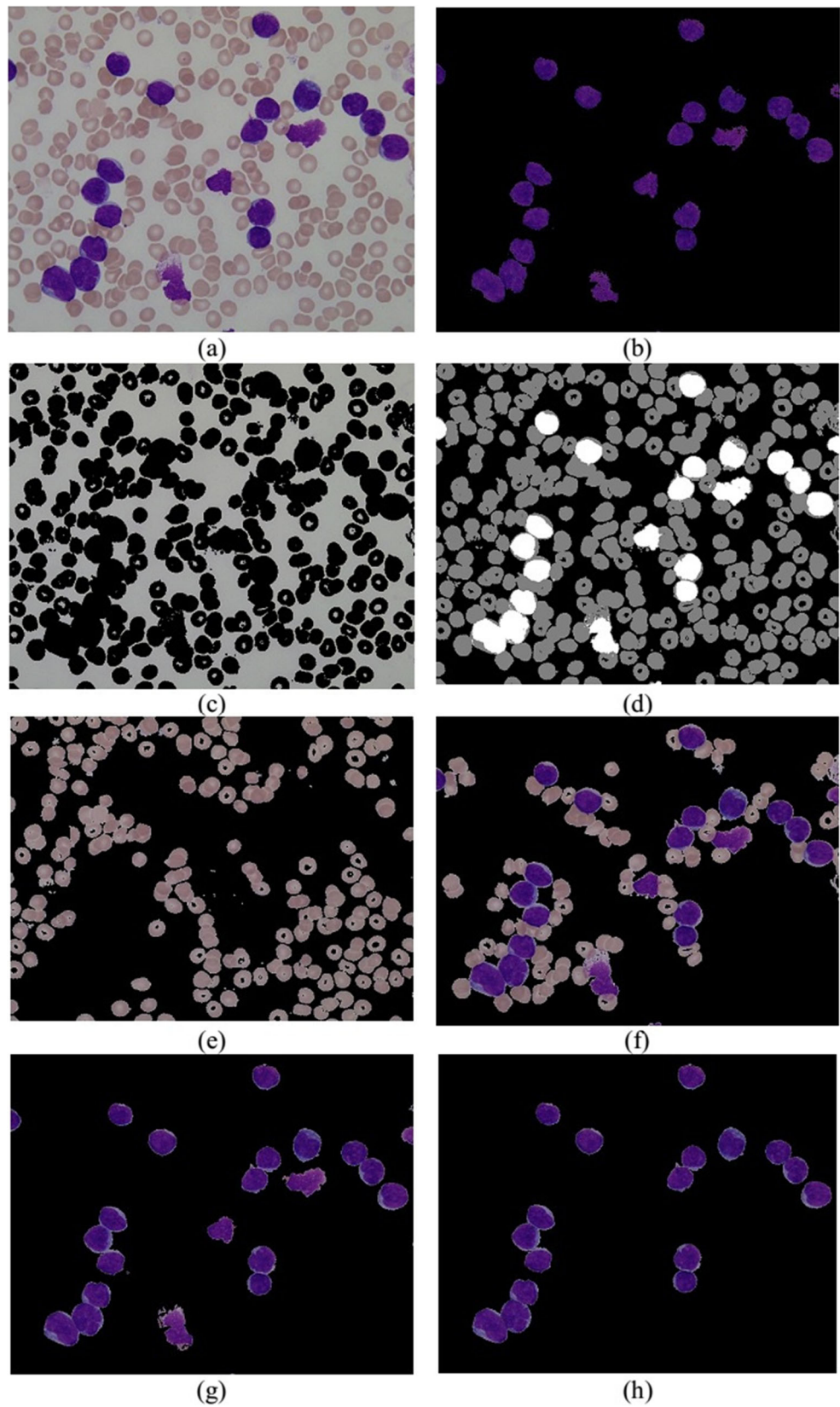
algorithms. Moreover, numerical results show that Otsu and IFD-g are not applicable algorithms for Dataset 2 which prove that many existing algorithms are susceptible to the dataset and cannot be performed as a robust technique.
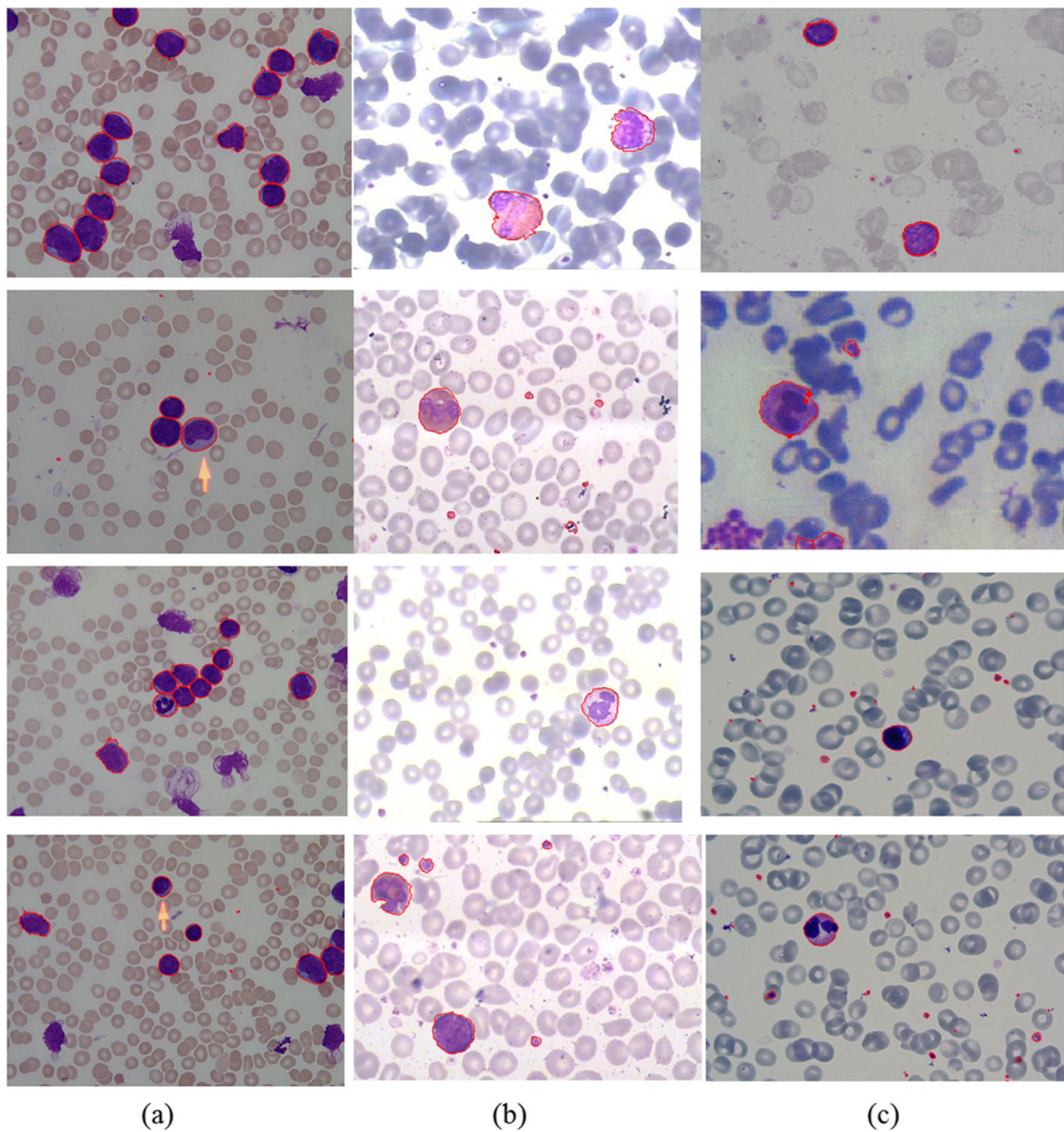
Fig. 12 Examples of the proposed leukocyte segmentation results. **a** Samples of Dataset 1, **b** samples of Dataset 2, and **c** samples of Dataset 3

Leukocyte Segmentation Results

Figure 11 indicates results of each segmentation step superimposed over an example of blood microscopic image.

Figure 12 illustrates the final results of our proposed leukocyte identification algorithm applied when using images acquired from different datasets.

It is shown that the proposed algorithm can find leukocytes in blood microscopic image under uneven condition. Also it has separated adjacent leukocytes and removed abnormal component accurately.

ALL Detection

There are two points of view during blood microscopic images analysis for ALL detection. Presence of at least one lymphoblast in a blood microscopic image is a considerable sign for hematologists in ALL diagnosis. In this case, the number

Table 2 Experimental results using texture features

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F measure (%)
GLCM texture	88.10 ± 0.05	87.92 ± 0.12	88.18 ± 0.05	87.22 ± 0.09	87.16 ± 0.10
LBP-based texture	92.86 ± 0.03	93.40 ± 0.05	91.77 ± 0.04	91.38 ± 0.07	92.07 ± 0.05

Table 3 Comparison of the proposed system for ALL detection with other existing systems in the case of the same dataset and classifier

	Number of features	Accuracy (%)
Putzu [27]	131	93.20
Neoh [24]	80	96.67
Proposed method	53	97.62

of lymphoblasts presented in the image is not considered. Treatment of ALL depends on the extent of the disease, hence each leukocytes presented in image needs to be characterized as a healthy or malignant in order to estimate how the disease is spread in the blood and bone marrow. Therefore, there are two benchmarks in ALL classification as follows [5]:

> Image test: the test is positive when whole blood microscopic image contains at least one lymphoblast.
> Cell test: the test is positive when the cell is malignant (lymphoblast).

Performance of all automated disease recognition systems is evaluated by comparison of hematologist opinion and classifier decision which is represented in confusion matrix:

$$\text{confusion matrix} = \begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \tag{13}$$

In each level of benchmark, confusion matrix component is defined as follows:

TP   the number of cells/images correctly identified as positive by classifier
FP   the number of cells/images incorrectly identified as positive by classifier
FN   the number of cells/images incorrectly identified as negative by classifier
TN   the number of cells/images correctly identified as negative by classifier

According to these definitions, some parameters are defined for performance evaluation as follows:

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \tag{14}$$

$$\text{Sensivity} = \frac{TP}{TP + FN} \tag{15}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{16}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{17}$$

$$F-\text{measure} = \frac{2 \times \text{Precision} \times \text{Sensivity}}{\text{Precision} + \text{Sensivity}} \tag{18}$$

As the size of training and testing sets decreases, classification results vary when different compositions of feature vectors are selected as training and testing sets. In this case, cross validation techniques are applied in order to evaluate classifier performance. K-fold cross validation is a method in which feature vectors are partitioned into K equally sized subsets then, one set is used as testing set and remaining K-1 sets are applied as a training set. This process is performed for K times to apply all feature vectors as testing set. Here, we have employed ten-fold cross validation to evaluate our system performance.

Using the same SVM classifier, it is shown that the proposed feature set results in higher classification performance, also using LBP-based texture features improves discrimination power of feature set. Table 2 provides the results obtained with different texture features using the same SVM classifier. It confirms that using LBP-based texture features improves discrimination power of feature set compared to the GLCM-based texture features. Standard deviation of the results for ten-fold cross validation is considered as the permissible error which is presented in this table.

The obtained accuracy measure and the corresponding numbers of extracted features of three available ALL detection systems are shown in Table 3 in which all the systems used ALL-IDB dataset with the same SVM classifier and ten-fold cross validation. It can be observed from this table, although the proposed feature set contains less features, it provides more accuracy. It is concluded that the proposed leukocyte segmentation method beside the optimum selection of features cause more accurate feature vectors for classification of normal and abnormal leukocytes. In the previous results, we did not implement the proposed ensemble classifiers which confirms that our segmentation and feature extraction process perform better than other systems.

Table 4 Experimental results for ALL detection using different classifiers in cell level test

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F measure (%)
SVM	97.62 ± 0.77	97.57 ± 0.99	97.26 ± 0.68	97.17 ± 0.56	97.27 ± 0.71
KNN	94.76 ± 0.03	92.36 ± 0.04	97.09 ± 0.04	97.22 ± 0.08	94.50 ± 0.06
Naive Bayes	84.76 ± 0.04	86.91 ± 0.09	82.17 ± 0.05	83.05 ± 0.07	84.43 ± 0.07
Decision Tree	87.14 ± 0.04	85.97 ± 0.02	86.00 ± 0.05	88.03 ± 0.08	86.11 ± 0.06
Proposed classifier 1	98.10 ± 0.02	97.57 ± 0.06	98.09 ± 0.04	98.17 ± 0.07	97.80 ± 0.05
proposed classifier 2	96.67 ± 0.02	95.27 ± 0.08	98.09 ± 0.05	98.06 ± 0.09	96.51 ± 0.06

Table 5 Comparison of the proposed system with other existing systems

	Dataset	classifier	Validation	Number of features	Accuracy (%)
Putzu [27]	ALL-IDB	SVM	10-fold	135	93.20
Neoh [24]	ALL-IDB	Ensemble	Bootstrap	80	96.72
Mohapatra [28]	Proprietary	Ensemble	5-fold	32	96.88
Proposed system	ALL-IDB	Ensemble 1	10-fold	53	98.10

Table 4 depicts the classification results obtained from the different classifiers acceptable for most leukemia classification tasks. Standard deviation of the results for ten-fold cross validation is considered as the permissible error which is presented in this table.

The proposed ensemble classifier 1 demonstrated the best classification outcome to identify lymphoblast among other healthy leukocytes.

The best accuracy rate achieved by four ALL detection systems have been presented in Table 5. However, we cannot compare these systems fairly; it is shown that the proposed system achieves better classification result.

However, the proposed ensemble classifier 2 does not perform as well as the proposed ensemble classifier 1 in cell test; it will present a better performance for image level test by using various types of kernels that generate a different decision boundary when compared with some of the existing classifiers. To the best of our knowledge, this is the first ALL detection systems evaluated by image test which is applied on whole blood microscopic images. In fact, the reported ALL classification performance in the literature is calculated only in cell level test which segmentation algorithm and feature extraction process are usually performed on sub-images containing individual leucocytes. Using these leukocytes sub-images not only decreases segmentation errors but also increases classification accuracy. In the real-world scenario, the whole blood microscopic image containing adjacent leukocytes, abnormal component, and noise has to be analyzed so the goal of this paper is to implement image test on blood microscopic images. This is a great challenge which leads to lower accuracy in image and cell classification; in this case, 260 sub-images belong to the ALL-IDB2 were used as training set and 108 images of ALL-IDB1 were considered as test

set. Table 6 shows performance results obtained by image test evaluation using different classifiers.

As seen in Table 6, classification by image test requires an effective classification model which is not provided by some regular classifiers. It is observed that the proposed ensemble classifier 2 provides the best accuracy of 89.81 which proves the robust performance of this classifier in comparison with other standard classifiers. Moreover, the best sensitivity of 100 is achieved by the proposed system which confirms that no leukemic patient is considered as healthy.

Limitation of the Proposed Method

In fact, the proposed system may not work properly on some types of blood microscopic images witch are not provided well. As an example, the proposed algorithm does not perform correctly on images which cells are overlapped and clumped (Fig. 13). Actually, this type of images are not acceptable in laboratories for manual analysis so this limitation of our systems is not a real drawback.

Conclusion

An automatic image processing-based ALL detection system was proposed. The proposed system consisted of three sections, leukocytes segmentation, feature extraction, and classification. Three assumptions which are common between different types of blood microscopic images were defined to be used to construct a robust segmentation algorithm. Experimental results demonstrate that the proposed approach robustly extracts leukocytes under uneven lighting and

Table 6 Experimental results for ALL detection using different classifiers in image level test

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F measure (%)
SVM	72.22	100.00	49.15	62.03	76.56
KNN	72.22	100.00	49.15	62.03	76.56
Naive Bayes	62.96	100.00	32.20	55.06	71.01
Decision Tree	59.26	100.00	25.42	52.69	69.01
Proposed classifier 1	75.00	100.00	54.24	64.47	78.40
proposed classifier 2	89.81	100.00	81.36	81.67	89.91
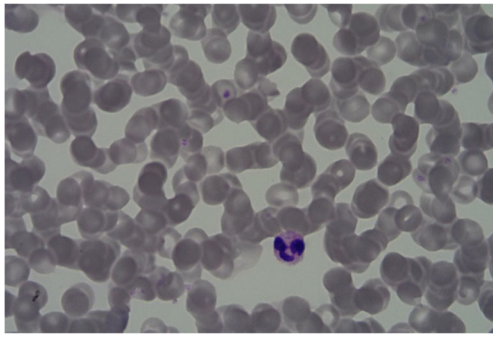
Fig. 13 An examples of inappropriate image for proposed system

imaging conditions. Furthermore, an algorithm for nucleus segmentation was improved which has been compared with some other nucleus segmentation techniques in which our method outperforms the others in terms of segmentation accuracy and robustness.

In feature extraction section, a new set of 53 features was introduced to achieve all the information required to perform ALL classification. Results showed that using LBP-based features improved classification accuracy compared with GLCM features.

Two ensemble classifiers have been developed in this paper which each of them is highly effective in different levels of analysis. The image level test for ALL recognition was performed on the proposed system for the first time to challenge and evaluate our system more accurately. The ensemble classifier 2 based on SVM kernel functions performed the best in this level of evaluation compared with other benchmark classifiers.

The detected lymphoblast in the proposed system can be classified as L1, L2, and L3 according to the French–American–British classification [36]. Classification of candidate lymphoblasts into their subtypes is an important task because it ensures physicians that the patient achieves the correct treatment. However, there are some leukemia classification systems in the literature [26, 37], as future work an algorithm can be designed to classify extracted lymphoblasts and outperform other available ALL classification systems.

References

1. Hall J: Guyton and hall textbook of medical physiology. Elsevier Health Sciences, 2010
2. Inaba H, Greaves M, Mullighan C: Acute lymphoblastic leukaemia. The Lancet 381(9881):1943–1955, 2013
3. G. Voigt and S. Swist, Hematology techniques and concepts for veterinary technicians, John Wiley & Sons., 2011.
4. Münzenmayer C, Schlarb T, Steckhan D, Haßlmeyer E, Bergen T, Aschenbrenner S, Wittenberg T, Weigand C, Zerfaß T: HemaCAM–A computer assisted microscopy system for hematology. In: Microelectronic systems. Berlin Heidelberg: Springer, 2011, pp. 233–242
5. R. D. Labati, V. Piuri and F. Scotti, "All-IDB: The acute lymphoblastic leukemia image dataset for image processing," in *Proc. IEEE ICIP*, Brussels, Belgium, 2011.
6. M. Habibzadeh, A. Krzyzak, T. Fevens and A. Sadr, "Counting of RBCs and WBCs in noisy normal blood smear microscopic images," in *Proc. SPIE7963*, 2011.
7. Mohammed EA, Mohamed MM, Far BH, Naugler C: Peripheral blood smear image analysis: a comprehensive review. Journal of pathology informatics 5, 2014
8. Wermser D, Haussmann G, Liedtke C-E: Segmentation of blood smears by hierarchical thresholding. Computer Vision, Graphics, and Image Processing 25(2):151–168, 1984
9. Osuna V, Cuevas E, Sossa H: Segmentation of blood cell images using evolutionary methods. AISC, Springer 175:299–311, 2013
10. Ong S-H, Lim J-H, Foong K, Liu J, Racoceanu D, Chong A, Tan K: Automatic area classification in peripheral blood smears. Biomedical Engineering, IEEE Transactions on 57(8):1982–1990, 2010
11. Ghosh M, Chakraborty C, Konar A, Ray AK: Development of hedge operator based fuzzy divergence measure and its application in segmentation of chronic myelogenous leukocytes from microscopic image of peripheral blood smear. Micron 57:41–55, 2014
12. Ghosh M, Das D, Chakraborty C, Ray AK: Automated leukocyte recognition using fuzzy divergence. Micron 41(7):840–846, 2010
13. M. Hamghalam, M. Motameni and A. E. Kelishomi, Leukocyte segmentation in giemsa-stained image of peripheral blood smears based on active contour, in *Proc. Signal Progcessing Systems*, singapore, 2009.
14. Jati A, Singh G, Mukherjee R, Ghosh M, Konar A, Chakraborty C, Nagar AK: Automatic leukocyte nucleus segmentation by intuitionistic fuzzy divergence based thresholding. Micron 58:55–65, 2014
15. K. Jiang, Q. -M. Liao and S. -Y. Dai, A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering, in *Proc. Machine Learning and Cybernetics*, 2003.
16. Ko BC, Gim J-W, Nam J-Y: Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake. Micron 42(7):695–705, 2011
17. Li K, Lu Z, Liu W, Yin J: Cytoplasm and nucleus segmentation in cervical smear images using radiating GVF snake. Pattern Recognition 45(4):1255–1264, 2012
18. Pan C, Park DS, Yoon S, Yang JC: Leukocyte image segmentation using simulated visual attention. Expert Systems with Applications 39(8):7479–7494, 2012
19. Saraswat M, Arya K: Automated microscopic image analysis for leukocytes identification: a survey. Micron 65:20–33, 2014
20. F. Zamani and R. Safabakhsh, An unsupervised GVF snake approach for white blood cell segmentation based on nucleus, in *Proc. Signal Processing*, Beijing, 2006.
21. Rezatofighi SH, Soltanian-zadeh H: Automatic recognition of five types of white blood cells in peripheral blood. Computerized Medical Imaging and Graphics 35(4):333–343, 2011
22. Kass M, Witkin A, Terzopoulos D: Snakes: active contour models. International Journal of Computer Vision 1(4):321–331, 1988
23. Huang D-C, Hung K-D, Chan Y-K: A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images. Journal of Systems and Software 85(9):2104–2118, 2012
24. Neoh S, Srisukkham W, Zhang L, Todryk S, Greystoke B, Lim C, Hossain M, Aslam N: An intelligent decision support system for leukaemia diagnosis using microscopic blood images. Scientific reports 5, 2015
25. Escalante H, Montes-y-Gómez M, González J, Gómez-Gil P, Altamirano L, Reyes C, Reta C, Rosales A: Acute leukemia classification by ensemble particle swarm model selection. Artificial intelligence in medicine 55(3):163–175, 2012

26. Fatichah C, Tangel M, Yan F, Betancourt JP, Widyanto MR, Dong F, Hirota K: Fuzzy feature representation for white blood cell differential counting in acute leukemia diagnosis. International Journal of Control, Automation and Systems 13(3):1–11, 2015

27. Putzu L, Caocci G, Di Ruberto C: Leucocyte classification for leukaemia detection using image processing techniques. Artificial Intelligence in Medicine 62(3):179–191, 2014

28. S. Mohapatra, D. Patra and S. Satpathy, Automated leukemia detection in blood microscopic images using statistical texture analysis, in *Proc. Communication, Computing & Security*, 2011.

29. Agaian S, Madhukar M, Chronopoulos AT: Automated screening system for acute myelogenousl eukemia detection in blood microscopic images. IEEE Systems Journal 8(3):995–1004, 2014

30. Meyer F: Topographic distance and watershed lines. Signal processing 38(1):113–125, 1994

31. Tejinder S: Atlas and text of hematology. New Delhi: Avichal Publishing Company, 2010

32. Rezaee K, Haddadnia J, Tashk A: Optimized clinical segmentation of retinal blood vessels by using combination of adaptive filtering, fuzzy entropy and skeletonization. Applied Soft Computing 52: 937–951, 2017

33. Tashk A, Helfroush MS, Danyali H, Akbarzadeh-Jahromi M: Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features. Applied Mathematical Modelling 39(20):6165–6182, 2015

34. A.Tashk, MS. Helfroush, H. Danyali: A computer-aided system for automatic mitosis detection from breast cancer histological slide images based on stiffness matrix and feature fusion. Current Bioinformatics 10(4):476–493, 2015

35. Otsu N: A threshold selection method from gray-level histograms. IEEE Trans. Systems, Man and Cybernetics 9(1):62–66, 1979

36. Bennett J, Catovsky D, Daniel M, Flandrin G, Galton D, Gralnick H, Sultan C: Proposals for the classification of the acute Leukaemias French-American-British (FAB) co-operative group. British journal of haematology 33(4):451–458, 1976

37. O. Sarrafzadeh, H. Rabbani, A. M. Dehnavi and A. Talebi, Detecting different sub-types of acute myelogenous leukemia using dictionary learning and sparse representation, in *Proc. IEEE ICIP*, Quebec City, 2015.