Multimedia Tools and Applications (2019) 78:17879–17898
https://doi.org/10.1007/s11042-018-7107-x

Development of a robust algorithm for detection of nuclei of white blood cells in peripheral blood smear images

Roopa B. Hegde[1,2] · Keerthana Prasad[1] · Harishchandra Hebbar[1] · Brij Mohan Kumar Singh[3]

Received: 15 March 2018 / Revised: 23 November 2018 / Accepted: 20 December 2018 /
Published online: 15 January 2019

Abstract

Microscopic evaluation of peripheral blood smear analysis is a commonly used laboratory procedure to diagnose various diseases such as anemia, malaria, leukemia, etc. Manual microscopic evaluation is laborious and hence many research groups have attempted to automate smear analysis. Variations in staining procedure and smear preparation introduces color shade variations into peripheral blood smear images. Illumination provided by point source bulb introduces brightness variations across the smear which affects the performance of an automated method. In this paper we present an image processing algorithm for detection of nuclei of white blood cells which is robust to color and brightness variations. In the proposed method we used two different datasets and also five datasets which were derived from original images by introducing brightness variations. We also compared the results of the proposed method with four state-of-the-art methods. The results demonstrate that the proposed method detects nuclei accurately with an average accuracy of 0.99 and Dice coefficient of 0.965.

Keywords Peripheral blood smear images · Nuclei detection · WBC detection · Illumination variations · Computer aided detection

✉ Keerthana Prasad
keerthana.prasad@manipal.edu

Roopa B. Hegde
roopabhegde@gmail.com

Harishchandra Hebbar
harish.hebbar@manipal.edu

Brij Mohan Kumar Singh
brij.singh@manipal.edu

[1] School of Information Sciences, MAHE, Manipal 576104, India

[2] Department of ECE, NAMAMIT, Nitte 574110, India

[3] Department of Pathology, KMC, MAHE, 574106, Manipal, Karnataka 576104, India

1 Introduction

Peripheral Blood Smear (PBS) analysis is a routine work carried out in laboratories to evaluate the health status of a person. The analysis involves collection of blood sample, preparation of peripheral blood smear and evaluation of blood cells under a microscope [19]. Manual microscopic examination is performed to find the cause of a disease to facilitate suitable treatment to the affected individual [18] by counting blood cells and/or studying size, color, shape and texture of cells. Hence it is a laborious procedure, time consuming and results show poor repeatability [1]. Outcome of the analysis depends on instruments, skill and experience of the technician and methods used to analyze the blood sample.

Peripheral blood comprises of three types of cells namely Red Blood Cells (RBCs), White Blood Cells (WBCs) and platelets. Blood cells can be differentiated under a microscope mainly based on their color and size as shown in Fig. 1. Many diseases such as malaria, leukemia, anemia and thrombocytopenia can be diagnosed by studying shape, size, color and distribution of the blood cells [15]. Platelets are the smallest among the three types. WBCs are larger cells and consist of darker nucleus and outer pale cytoplasm. RBCs and platelets do not contain nucleus. WBCs are of five types namely Neutrophil, Eosinophil, Basophil, Lymphocyte and Monocyte. These cells vary in shape, size, color and texture. WBCs are darker compared to the other objects in stained blood smear images. This is due to the presence of nucleus. Thus study of nuclei is an important step towards WBC detection and classification. Many disorders such as leukemia, lymphoma, eosinophilia, neutrophilia and bacterial infections can be diagnosed by identifying WBCs in peripheral blood smear image. Several efforts have been made for automated detection of nuclei using PBS images [36].

The main aim of utilizing image processing techniques in medical field is to imitate human evaluation. Many researchers have tried to automate detection of blood cells from peripheral blood smear images using image processing in order to help hematologists in diagnosing diseases as a decision support system [6, 9, 15]. Cell segmentation is one of the most challenging task in medical image processing due to its complex biological appearance, staining method used and illumination variations in acquired images [20]. Inconsistency in staining procedure leads to various levels of expression of staining [6]. This causes the cells to take up varying color shades. The technique distributes cells in gradually decreasing cell concentration across the smear. Since the cells are unevenly distributed, cells appear dark in the region of high concentration and pale in the region of low concentration. Hence the location in the smear where the image is acquired also contributes for

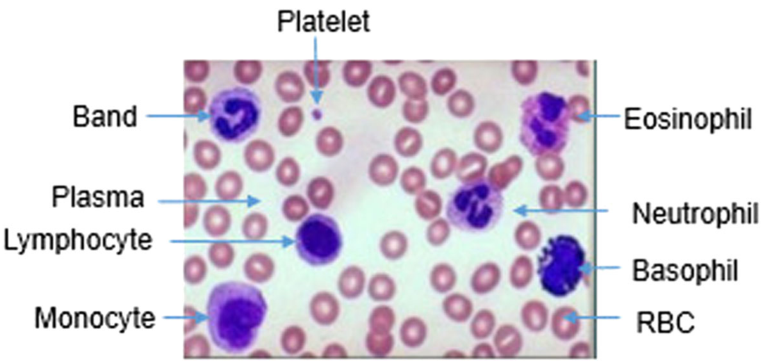


Fig. 1 Peripheral blood smear image [19]

color shade variations of cells. The slide under the microscope is illuminated by a bulb which results in brightness variations across the smear. Illumination is brighter at the center and decreases towards the periphery at the given field of view. Hence a method for detection of WBCs must be robust enough to manage all these variations. However most of the methods developed for automation assume images taken under specific conditions such as uniform illumination [2, 3, 8, 16, 22, 23, 27, 33, 34], carefully acquired images with no overlapping cells [1, 23, 27, 35] and considering images of specific disorders [2, 11, 16, 24, 34, 35]. Der-Chen et al. [8] proposed a method for nuclei segmentation based on Otsu's method using G component of RGB color model and S component of HSV color model. Genetic algorithm based k-means clustering approach was used for classifying the five types of WBCs. They reported the accuracy of segmentation in the range of 81–98% depending on the type of WBC. Lin et al. [16] proposed a method for segmentation of WBCs using GVF snake with integrated squared error criteria estimation. They used 58 images in their study. They reported the accuracy of segmentation of nuclei around 94%. Mazin et al. [27] proposed method for detection of WBCs using image arithmetic and thresholding. Bakht et al. [3] proposed color based segmentation of WBCs using image quantization and morphological operations. They used 50 normal blood smear images in their study. The reported average accuracy varied between 94 and 100%. Morteza et al. [2] employed k-means clustering and SVM for recognition of ALL, thereby obtaining average accuracy around 97%. Seyed et al. [33] proposed image processing algorithms to recognize five types of WBCs. GramSchmidt orthogonalization with snake algorithm was used for nuclei and cytoplasm segmentation. Artificial Neural Network (ANN) and Sup- port Vector Machine (SVM) classifier results were compared. The dataset consisted of 400 samples. An accuracy of 93% was reported for segmentation of nuclei. Mathur et al. [23] proposed a method for segmentation of WBCs into nuclei and cytoplasm using active contours without edges. Also, a method to count number of lobes of nucleus was presented using region splitting algo- rithm for classification of WBCs. A dataset considted of 237 normal images was considered in their study. Overall accuracy around 92% was reported using Naive Bayesian classifier. Jaroonrut et al. [10] proposed a method for detection and classification of WBCs. Thresholding, morphological operation and ellipse curve fitting were used for detection of WBCs and Naive Bayes classifier was used for detection of WBC types. Shape and color features of nuclei and cytoplasm were used to train the classifier. Two datasets consisting of 555 images and 477 cropped images respectively were considered in the proposed method. The segmentation and classifier accuracy around 98% was reported. Detection of WBCs using dual-threshold method was proposed in [14]. A combination of RGB and HSV color space representations was used to detect WBCs. The dataset consisted of 130 images taken from ALL-IDB dataset. The study re-ported an overall segmentation accuracy of 97%. Nurhanis et al. [22] and Rong Chu et al. [34] also presented a method of nuclei segmentation using active contours with an accuracy of 92%. Adnan Khashman [1] proposed a method for identification of blood cells using global pattern averaging and neural net- work, which resulted in correct identification rate of 99.1%. Farnoosh et al. [35] proposed a method for detection of nucleus and cytoplasm using GVF snake algorithm and Zack's thresholding respectively. The study reported the use of 20 images and reported average accuracy of 92 and 78% for detection nuclei and cytoplasm respectively. Minal et al. [11] and Nagabhushana et al. [24] proposed methods using image arithmetic and thresholding on grayscale images for segmentation of WBCs to detect acute leukemia. The reported accuracy of methods was between 88 and 93%. Arithmetic operations and Otsu's thresholding method were also used for detection of nuclei [20, 31]. The authors reported an average accuracy of nuclei detection between 75 and 98% depending on the cell type. Congcong et al. [5] employed a combination of color space decomposition and k-means clustering for detection of nucleus and cytoplasm of WBCs. Various color space representations namely RGB, HSI and

CMYK were utilized for segmentation. The study reported the use of 300 PBS images with segmentation accuracy of 95.7 and 91.3% for nucleus and cytoplasm respectively. Nemane et al. [25] proposed a method for segmentation of WBCs using image arithmetic and thresholding on gray scale images. They reported the accuracy around 99%. Ramin et al. [30] proposed a method for extraction of nucleoli in nuclei of WBCs using k-means algorithm to detect nuclei and curvelet transform to detect nucleoli. The dataset consisted of 100 images of lymphoid and lymphoblast cases. The authors reported nucleoli detection specificity of 80.2% with sensitivity of 84.3%. Hayan et al. [21] proposed segmentation of lymphoblast cells using thresholding and morphological operations. The authors reported the use of 180 images with at least one lymphoblast and sub-images having single cell to segment nucleus and cytoplasm. The reported accuracy of segmentation varied between 90 and 95%. Yiping et al. [37] proposed a method for extraction of WBCs based on S component of HSI color space and B component of RGB color model using AND, XOR and morphological operations. They used 10 representative images in their study. Overall segmentation accuracy around 93% was reported. Leyza Baldo et al. [7] proposed a method for segmentation of WBCs into nucleus and cytoplasm using morphological operations, watershed segmentation and scale-space analysis. The dataset consisted of 100 images of normal cases. Rawat et al. [32] proposed a method for detection of acute lymphoblastic leukemia using hybrid classifiers. Features such as shape, texture and color features of nuclei and cytoplasm were used to train the classifier. Thresholding and morphological operations were used for detection of nuclei. The authors used ALL-IDB2 dataset in their study. Neoh et al. [26] also used ALL-IDB2 dataset for detection of leukemia. Clustering algorithm based on stimulating discriminant measure was employed for detection of nuclei, which resulted in similarity measure around 0.86. The authors reported the use of 180 sub-images in their study. It can be observed that none of the state of the art methods have considered possibilities of variations in color shades and brightness.

In this paper we present a method for detection of nuclei which is robust to color and brightness variations. We used images from ALL-IDB2 dataset [13] and also obtained images from two microscopes with different resolutions. We introduced uniform and non-uniform brightness variations into the images to demonstrate the robustness of the proposed method. The details of data collection, methodology and results are provided in following sections.

2 Materials and methods

This section is divided into data collection, development of a robust method to detect nuclei and introducing brightness variations into the images. We provide the details of data collection in sub-section 2.1. The details of the proposed method for detection of nuclei is given in sub-section 2.2. The details of uniform and non-uniform brightness variations which are introduced to demonstrate the robustness of the proposed method are given in sub-section 2.3.

2.1 Data collection

We used a total of 160 images from ALL-IDB2 dataset available at https://homes.di.unimi.it/scotti/all which we will be referring to as *dataset-1*. More details are available in [13]. This dataset contains both normal and abnormal WBCs. The images in the dataset are cropped images of size 257X257X3. A few sample images of ALL-IDB2 dataset are shown in Fig. 2. Illumination variation can be observed between image (a), image (b) and image (d).

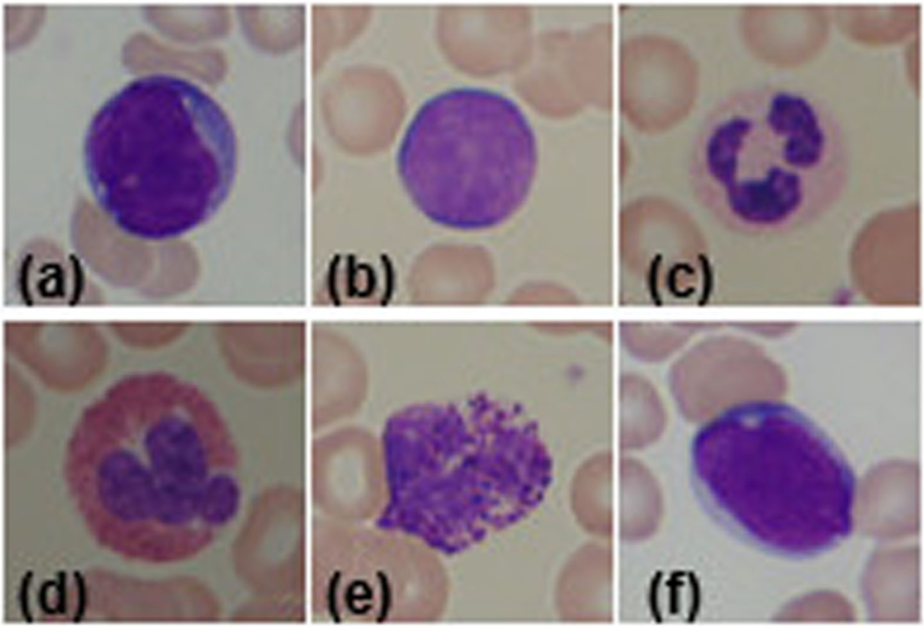


Fig. 2 Sample images of ALL-IDB2 dataset

To demonstrate the robustness of the proposed method, we also considered 160 images of Leishman stained peripheral blood smears which were acquired at 100X magnification from two different laboratories in KMC hospital, Manipal, India. Lab-1 setup uses OLYMPUS BX51 microscope with 2560X1920 resolution. We obtained 80 images from lab-1 setup. Lab-2 setup uses OLYMPUS CX31 microscope with 1600X1200 resolution. We obtained 80 images from lab-2 setup. Figure 3 shows the sample images of the dataset, it can be observed that color of blood cells and brightness vary significantly across these images. Images in upper row were acquired from lab-1 setup and images in bottom row correspond to lab-2 setup. The images contained both normal and abnormal RBCs, WBCs and platelets. We will be referring to these 160 images as *dataset-2*.

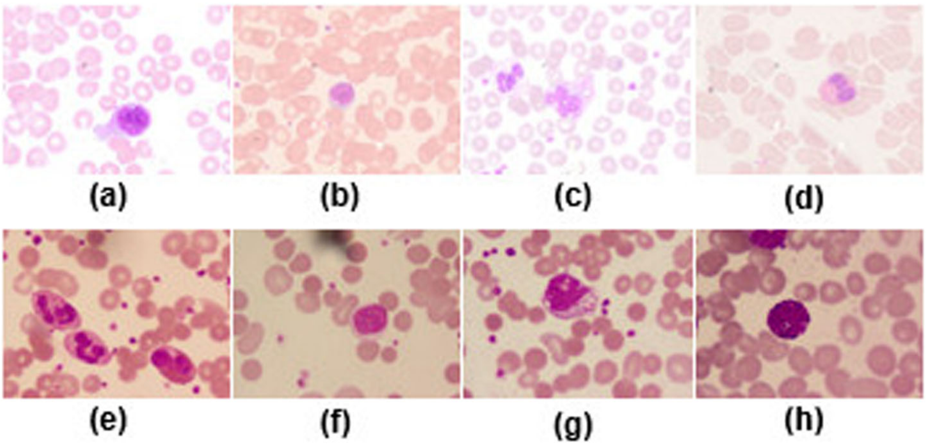


Fig. 3 Sample images of acquired dataset: **a**–**d** images using lab-1 setup, **e**–**h** images using lab-2 setup

Further, to test the robustness of the proposed method we introduced uniform brightness variation into the original images. We added a constant of 10 and 20 to the pixel values which we will be referring as *dataset-2a* and *dataset-2b* respectively. We also subtracted a constant of 10 and 20 from the pixel values which we will be referring as *dataset-2c* and *dataset-2d* respectively. To demonstrate the robustness, we also induced non-uniform brightness variations into the original images which we will be referring to as *dataset-2e*. Details of the datasets are provided in Table 1.

2.2 Development of a robust method to detect nuclei

The block diagram of the proposed method is shown in Fig. 4. Input image is a color image. Pre-processing the input images were considered since this helps in increasing the reliability of nuclei segmentation [12]. To obtain appropriate selection of region of nuclei, we considered a novel image enhancement method. The steps used for image enhancement method are as follows.

1. Input color image I (Fig. 4a)
2. Separate R, G and B components
3. Normalize G component to obtain G_n using the relation $G_n = G/G_{mean}$, where G_{mean} is the mean value of G component
4. Contrast enhance G_n to obtain G' using *imadjust* function of MATLAB
5. Normalize B component to obtain B_n using the relation $B_n = B/B_{mean}$, where B_{mean} is the mean value of B component
6. Combine G', G', B_n to obtain I_{GGB} (Fig. 4b)

In this image enhancement method, we replaced R and G component of original color image by normalized and contrast enhanced G component (G'). We replaced B component by normalized B (B_n) which resulted in GGB color representation of the original image. This enhances the color of nucleus by eliminating red component in images as shown in Fig. 4b. This also increases the contrast between the region of nuclei and the background. The GGB color image is then considered for detection of nucleus using TissueQuant method which is a color image segmentation method as explained in [28]. This method provides accurate color segmentation and hence it can be used for selection of nucleus region in peripheral blood smears [29]. The nuclei of various WBCs present in different color shades. To pick-up all color shades present in region of nuclei, an accurate color segmentation approach is required. It is

Table 1 Details of the datasets

Dataset	Source	Number of Images
Dataset-1	ALL-DB2	160
Dataset-2	Lab-1	80
Dataset-2	Lab-2	80
Dataset-2a	Brightness+10	160
Dataset-2b	Brightness+20	160
Dataset-2c	Brightness-10	160
Dataset-2d	Brightness-20	160
Dataset-2e	Non-uniform Brightness	160

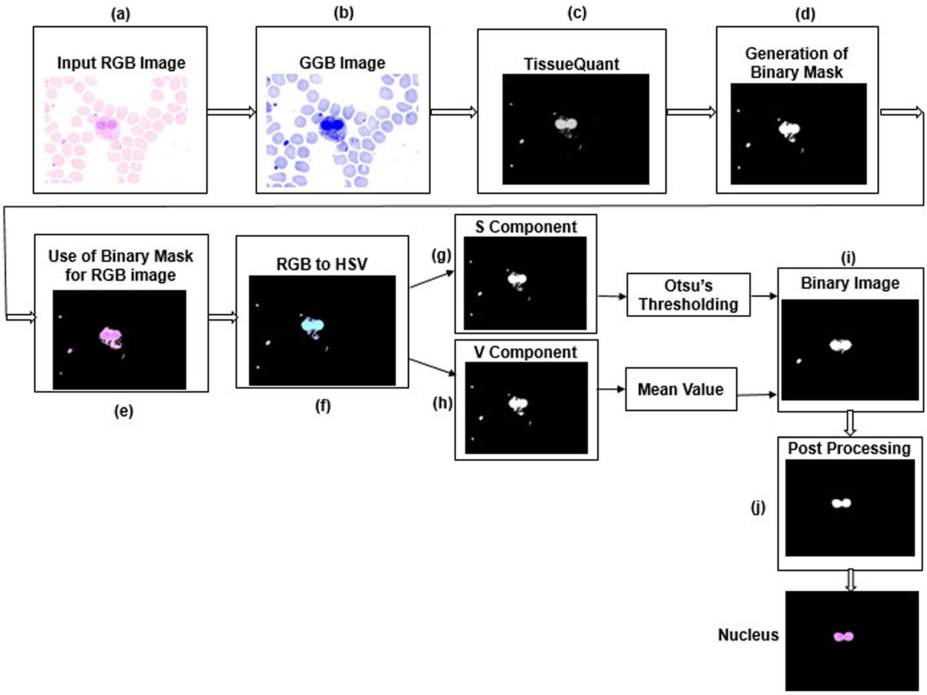


Fig. 4 Block diagram of the proposed method

reported that TissueQant offers efficient color segmentation by selectively choosing color shades appropriately.

2.2.1 TissueQuant algorithm

TissueQuant is a color image segmentation method which selects a particular color and its shades in an efficient manner. It converts given color image to HSI color space and Gaussian weighting functions are applied on each component of the color space in scoring colors based on how much they differ from the selected color. Gaussian weighting function assigns higher score to the color shades which are closer to the selected color. Lower scores are assigned to color shades which are distant from the selected color based on the width of the Gaussian weighing functions and other color shades get score of zero. The color score equations for calculating the score are as follows.

$$I_H(i,j) = e^{-(H(i,j)-H_c-1)^2/{H_w}^2} + e^{-(H(i,j)-H_c+1)^2/{H_w}^2} \tag{1}$$

$$I_S(i,j) = e^{-(S(i,j)-S_c)^2/{S_w}^2} \tag{2}$$

$$I_I(i,j) = e^{-(I(i,j)-I_c)^2/{I_w}^2} \tag{3}$$

Where H_c, S_c and I_c are hue, saturation and intensity values of selected color, H_w, S_w, and I_w are the width of Gaussian weighing functions and the LHS of Eqs. 1, 2 and 3 are the scores closer to the selected color of H, S and I components respectively. The final score is obtained as given in Eq. 4.

$$I\ (i,j) = I_H\ (i,j).I_S\ (i,j).I_I\ (i,j) \tag{4}$$

Where I(i, j) is grayscale image representing color scores of pixels of image.

2.2.2 The proposed method

A bluish color representing the shade of nucleus was as reference color to detect the region of nucleus appropriately using TissueQuant method. The TissueQuant algorithm scores all shades of selected color. Hence, the algorithm selects blue and its shades thereby appropriately detecting nucleus region even in the presence of brightness and color shade variations. A few platelets and staining artifacts present in the image may also be detected at this level of nuclei detection. The gray scale image obtained from TissueQuant method was converted into binary image and used as mask on the original RGB image. The threshold value for this operation was determined experimentally, so as to approximately select the nuclei region. The threshold value of 25 was found to be suitable for all the images in the dataset. The resultant color image represents approximate ROIs of nuclei. To evaluate the brightness level around the nuclei, we converted the color image representation of the ROIs to HSV representation and extracted S and V components. We used this step to adaptively select suitable threshold values for segmentation of nucleus. Algorithm steps for threshold selection are as follows.

1. Input color image I (Fig. 4a)
3. Obtain GGB image I_{GGB} (Fig. 4b)
4. Apply TissueQuant on I_{GGB} to obtain I_{score} using Gaussian weighting functions as given in Eqs. (4–7) (Fig. 4c)
5. Convert I_{score} to binary image I_b using threshold value $T = 25$ (Fig. 4d)
7. Obtain I_m using I_b as mask on I
8. Convert I_m to HSV image I_{hsv} (Fig. 4f)
9. Extract S and V components of I_{hsv} (Fig. 4g and h)
10. Obtain threshold value *thresh* of S component using Otsu's threshold method
11. Obtain mean value *meanV* of V component
12. Select threshold value $T1$ based on *thresh* and *meanV*
13. Apply $T1$ on I_{score} image to obtain binary image I_{bw}
14. Apply area filter and morphological closing operation to obtain nucleus (Fig. 4j)

To assess the brightness level at the neighborhood of each nucleus, we used threshold value thresh of S component and mean value meanV of V component to categorize images based on brightness levels. Further, threshold value T1 was selected depending on thresh value and meanV. Morphological closing operation with disk shaped structuring element of size 10 was used to connect lobulated nucleus and area filter was used to remove platelets and staining artifacts. Results of the proposed method was validated by comparing with the ground truth obtained from an expert.

2.3 Introducing brightness variations into images

To demonstrate robustness of the proposed method we subjected images of dataset-2 to uniform and non-uniform brightness variations. Methods of brightness variations are explained in subsections 2.3.1 and 2.3.2.

2.3.1 Uniform variation of brightness

We subjected images to different brightness levels by adding or subtracting a constant value as given in Eq. 5. Adding or subtracting a constant value increases or decreases the pixel intensity uniformly throughout the image. The constant value 'c' was varied in the range of −20 to +20 in steps of 10. This resulted in images with either decreased or increased brightness levels as shown in Fig. 5.

$$I_m uniform(i,j) = I_m original(i,j) + c \tag{5}$$

2.3.2 Non-uniform variation of brightness

Non-uniformity in brightness was introduced by using a scaling profile *S(k)* which was multiplied to each row of the original image as given in Eq. 6.

$$I_m non\text{−}uniform = I_m original * S(k) \tag{6}$$

Where *S(k)* is a scaling profile with linearly spaced values. We considered variations of scaling profile values in horizontal direction as shown in Fig. 6. The algorithm steps to generate *S(k)* are as follows:

1. Obtain the original image *I*
2. Obtain the size of the original image
3. Define scaling factor a = 0.25
4. Define start = 1-a and stop = 1 + a
5. For index = 1: number of columns, compute *S(column)* as given in Eq. (7)

$$S\ (column) = start + (index-1) * \frac{(end-start)}{(number\ \ of\ \ columns-1)} \tag{7}$$

This algorithm produces scaling profile of linearly spaced values between 0.75 and 1.25. Each row of the image is multiplied by this scaling profile to introduce non-uniform brightness variations in the image as shown in Fig. 6. The gradual change in brightness levels from darker to brighter shades can be observed in the images as a result of multiplication of original image with scaling profile.

3 Results and discussions

We attempted to address brightness and color shade variations by considering a novel image enhancement method and by using TissueQuant method. We considered images from ALL-

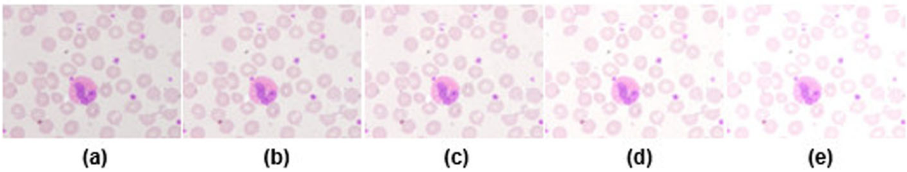


Fig. 5 Images with different brightness levels **a** c = −20, **b** c = −10, **c** original image, **d** c = +10, **e** c = +20

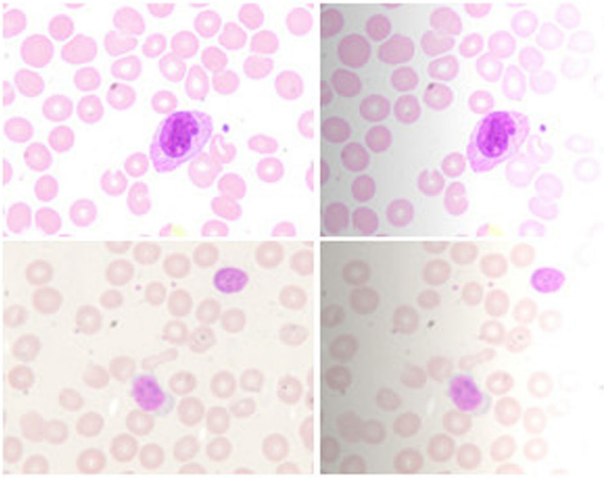


Fig. 6 Images with non-uniform brightness variation **a** original image **b** brightness varied image

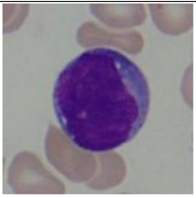
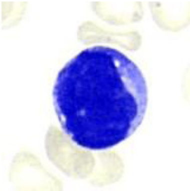
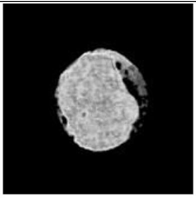
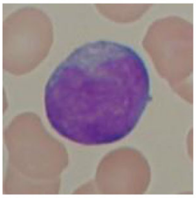
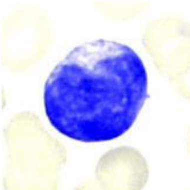
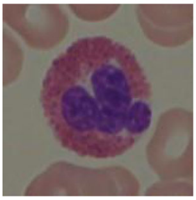
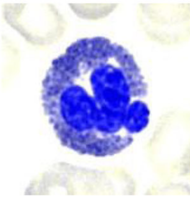
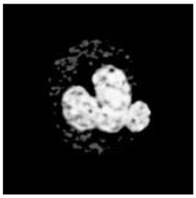
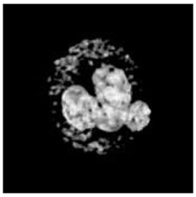
IDB2 dataset, images from two different laboratories, and images with uniform and non-uniform brightness levels to demonstrate robustness of the proposed method. Results obtained are discussed in the following sub-sections. Results of image enhancement method are demonstrated in sub-section 3.1. We provide the results of nuclei detection for *dataset-1* and *dataset-2* in sub-sections 3.2 and 3.3 respectively. Performance of the proposed method for the datasets with introduced brightness variations are demonstrated in sub-section 3.4. Comparison of the results for all the datasets are also provided in the sub-section. Also, we compared the results of the proposed method with results of nine state-of-the-art methods.

3.1 Results of image enhancement method

We considered a novel image enhancement method to minimize the color and brightness variations of the original images. The results of the image enhancement method for *dataset-1* are shown in Table 2. It can be observed from the table that, color variation of region of nuclei is minimized compared to the variation in original images. Also, TissueQuant performs better for GGB images compared to original images which is evident in Table 1 from column 3 and column 4 of row 1 and row 4 respectively. It can be observed that the contrast between the nucleus and the region around the nucleus is high in column 3 compared to that of column 4. Though there is a small variation in color and brightness levels, image enhancement method (column 2) clearly differentiates the region of nuclei from the background in the images. The color variation is managed by TissueQuant method. Also use of adaptive threshold helped in managing small brightness variations.

Color and brightness variations in *dataset-2* are more compared to *dataset-1*. Hence to bring out the role of the image enhancement method, we also demonstrated the results of TissueQuant method for *dataset-2*. The results of TissueQuant method for original images and GGB images are shown in Table 3. Color of region of nuclei vary in original images as shown in Table 3 (column 1). Region of nuclei show less color variation after converting original images to GGB representation as shown in Table 3 (column 2). TissueQuant results for GGB images are shown in Table 3 (column 3). It can be observed from the table that, the

Table 2 Results of TissueQuant on original and GGB images for *dataset-1*

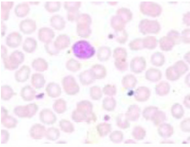
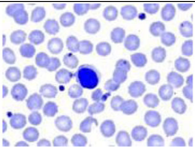
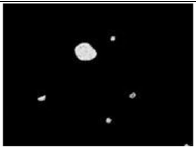
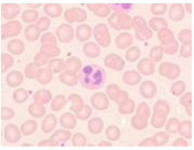
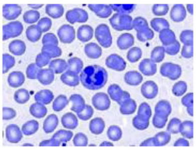
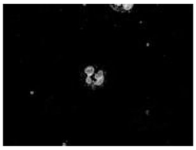
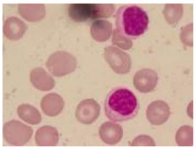
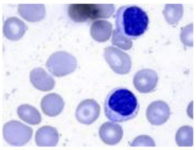
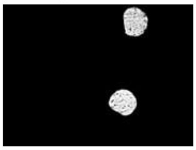
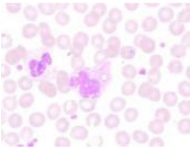
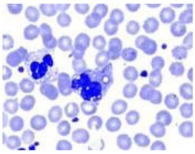
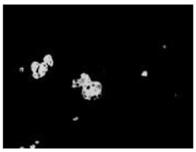
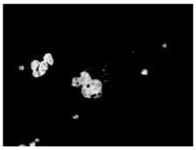
No.	Original Image 1	GGB Images 2	TissueQuant Output for GGB Images 3	TissueQuant Output for Original Image 4
1				
2				
3				
4				

performance of TissueQuant for GGB images is superior compared to the results obtained using original images. This is due to the fact that image enhancement method enhances the contrast between region of nuclei and the background region. The selected single reference color using TissueQuant method is suitable for all the images of both the datasets. Hence, the image enhancement method facilitates automated detection nuclei with single reference color. TissueQuant method failed to detect the region of nuclei with single reference color for original images as shown in Table 3 (row 3) whereas it is not so for *dataset-1*. This is due to the fact that, color variation of region of nuclei is more in *dataset-2*. A single reference color for TissueQuant method is suitable for all the images of both the datasets. Thus GGB enhancement facilitates fully automated approach of nuclei detection even in the presence of color and brightness variations.

3.2 Performance evaluation of the proposed method for *dataset-1*

The results of the proposed automated nuclei detection method are compared with the expert annotated images for performance evaluation. The results of nuclei detection are shown in Fig. 7, images a1-a4 are expert annotated images and images b1-b4 are results of the proposed

Table 3 Results of TissueQuant on original and GGB images for *dataset-2*

No.	Original Image 1	GGB Images 2	TissueQuant Output for GGB Images 3	TissueQuant Output for Original Image 4
1				
2				
3				
4				

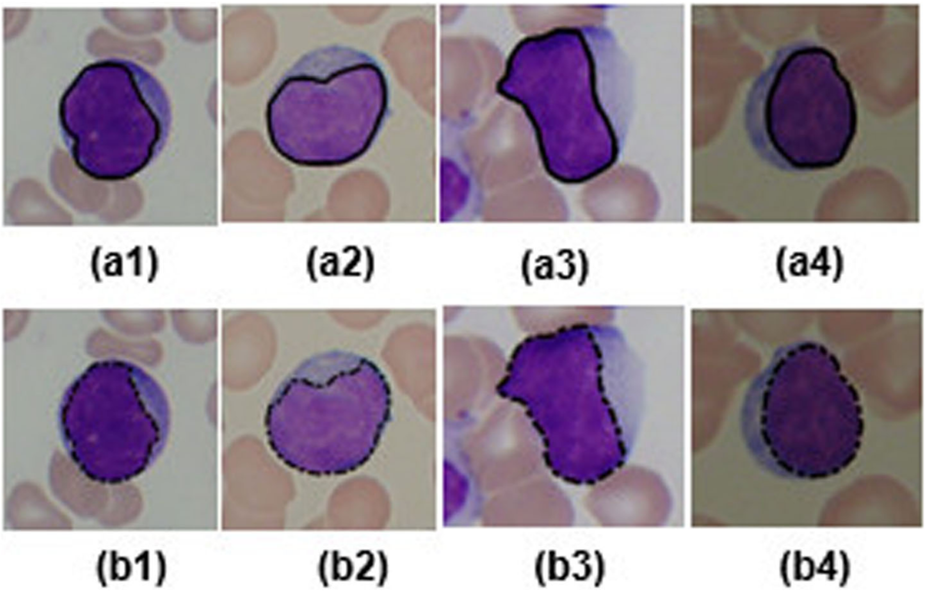


Fig. 7 Nuclei detection results for dataset-1; a1-a4 (upper row): expert annotated images, b1-b4 (bottom row): results of the proposed method

Table 4 Mathematical relation of similarity measures

Parameters	Relation
DSC	$2TP/(2TP + FP + FN)$
Accuracy	$(TP + TN)/(TP + TN + FP + FN)$
Precision rate	$TP/(TP + FP)$
Recall rate	$TP/(TP + FN)$

method. It can be observed from the figure that the results of the proposed method match well with the expert annotated images. Though there was brightness and color variations in the images, the proposed method performed well. This is due to the combination of image enhancement, TissueQuant and multiple threshold methods used in the proposed method. We computed Dice Similarity Coefficient (DSC), accuracy, precision rate and recall rate of detected nuclei to evaluate the results of the proposed method. The description of the similarity measures is given in Table 4.

Where TP is true positive, TN is true negative, FP is false positive and FN is false negative. Dice score measures the similarity between the segmented region of nucleus and the groundtruth. It approaches 1 if the segmented region and groundtruth are similar. Precision and recall rates were considered because they are sensitive to region boundaries. Low precision rate indicates over-segmentation whereas low recall rate indicates under-segmentation. High value of both the precision and recall rates indicate that the segmented region of nucleus and the groundtruth are similar.

The mean and Standard Deviation (SD) values of Similarity Measures (SM) for dataset-1 are tabulated in Table 5. It can be observed from the table that the average DSC is around 0.97, which indicates the region of nuclei detected using the proposed method and the expert annotated region of nuclei are nearly equal. Also, accuracy of 0.99 indicates the there is less mismatch between the results of the proposed method and the expert annotated regions. This could be due to assessment of color shades around the nucleus using TissueQuant method and also due to multiple threshold values selected depending on brightness variations around the nucleus region. Threshold based methods can be used for detection of nuclei which is a fast and reliable method [32].

3.3 Performance evaluation of the proposed method for *dataset-2*

To test the robustness of the proposed method, we evaluated the proposed method by considering *dataset-2* which consists of images acquired from two different setups as mentioned in section 2.1. Results of the proposed method were compared with the groundtruth images to evaluate the performance of the proposed method. The nuclei detection results for

Table 5 Similarity measures of *dataset-1*

SM	*Dataset-1* mean	SD
DSC	0.97	0.009
Accuracy	0.99	0.002
Precision	0.96	0.02
Recall	0.97	0.03

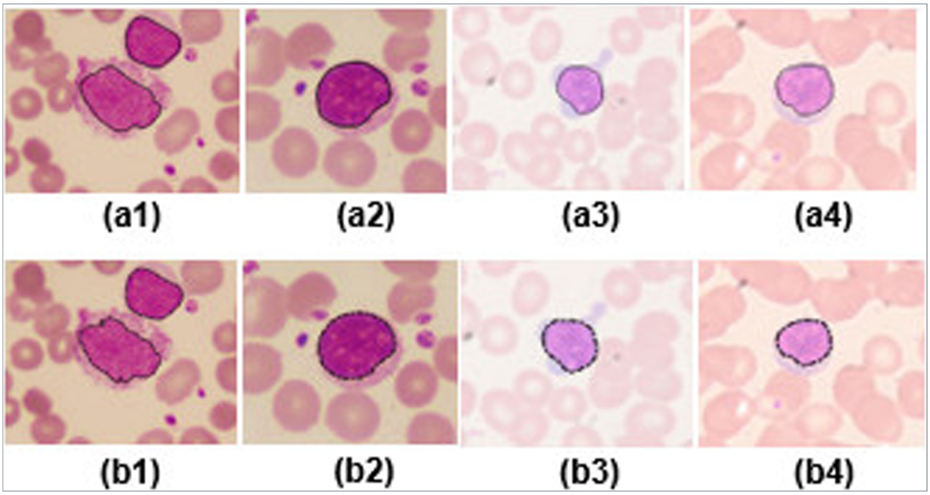


Fig. 8 Nuclei detection results for *dataset-2*; a1-a4(upper row): expert annotated images, b1-b4(bottom row): results of the proposed method

this dataset are shown in Fig. 8. Images a1-a4 in the figure are expert annotated images and b1b4 are results of the proposed method. It can be observed from the figure that, the results of the proposed method match well with the expert annotated images even in presence of brightness variations. This indicates that, the proposed method manages the color and brightness variations present in the images.

To evaluate the performance of the proposed method for this dataset we computed the similarity measures as in case of *dataset-1*. The mean and SD of similarity measures for the *dataset-2* are listed in Table 6. It can be observed from the table that, the results obtained are as good as the results obtained for *dataset-1*. This implies that, the proposed method is suitable for detection of nuclei even in presence of color and brightness variations.

The variation of DSC value for the *dataset-1* and *dataset-2* are shown in Fig. 9. It can be observed from the figure that, the DSC values remain above 0.94 for majority of the images of both the datasets. The variation is less for *dataset-1* compared to that of *dataset-2*. This is due to the fact that, the *dataset-2* consists of images of two different resolutions, images of different color shades and images of brightness variations whereas the images in *dataset-1* are cropped sub-images with less variations in color and brightness levels. Hence less variation can be observed for *dataset-2*. Though the variation is more for *dataset-1*, the average DSC values for both the datasets remain almost the same. Thus the proposed method can efficiently handle the color shade variations in PBS images.

Table 6 Similarity measures of *dataset-2*

SM	*Dataset-2* mean	SD
DSC	0.96	0.017
Accuracy	0.998	0.008
Precision	0.94	0.04
Recall	0.96	0.03

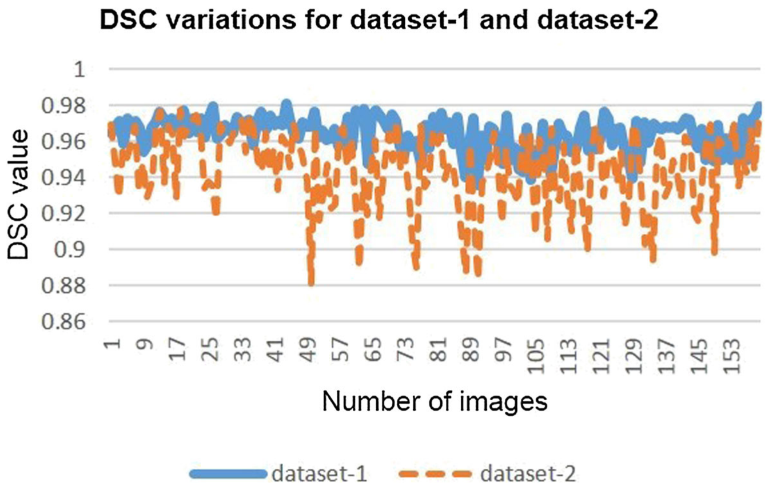


Fig. 9 DSC variations between *dataset-1* and *dataset-2*

3.4 Performance evaluation of the proposed method for introduced brightness variations

To demonstrate the robustness of the proposed method we also considered images of different brightness levels as mentioned in section 2.1. Details of the method used to introduce brightness variations into the images are given in section 2.3. We evaluated the results of the proposed method by computing similarity measures as in the case of *dataset-1* and *dataset-2*. The mean and SD values of the similarity measures for all the considered brightness variations are listed in Table 7. It can be observed from the table that, the similarity values are almost the same for the considered variations. A small variation in DSC values can be observed. The brightness variations affect the edges of the regions of nuclei, which result in loss of a few edge pixels. Hence, the variation in DSC values can be observed in Table 7. Further, non-uniform variation of brightness levels results in region of nuclei to appear in different shades as shown in Fig. 6. This also results in loss of edge pixels. Though there was color shade and brightness variations, the proposed method performed equally well for all the considered variations. Thus it is robust to color and brightness variations.

We compared the DSC and recall rate of nuclei detection for all the datasets. The bar plots for the average DSC and recall rate are shown in Fig. 10. The DSC and recall rate values remain almost the same for all the datasets except (*dataset-1*). The average DSC and recall rate of detection of nuclei is 0.96 and 0.95 for the combined dataset as shown in the figure. The DSC and recall rate values are higher for *dataset-1* compared to the other datasets. This is due

Table 7 Similarity measures for brightness varied images of *dataset-2*

SM	*dataset*	*-2a*	*dataset*	*-2b*	*dataset*	*t-2c*	*dataset*	*-2d*	*dataset*	*-2e*
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
DSC	0.956	0.017	0.957	0.017	0.956	0.017	0.957	0.017	0.956	0.01
Accuracy	0.998	0.001	0.998	0.001	0.998	0.002	0.998	0.002	0.998	0.001
Precision	0.94	0.05	0.94	0.06	0.94	0.05	0.937	0.05	0.937	0.05
Recall	0.955	0.03	0.947	0.03	0.947	0.03	0.947	0.03	0.948	0.03

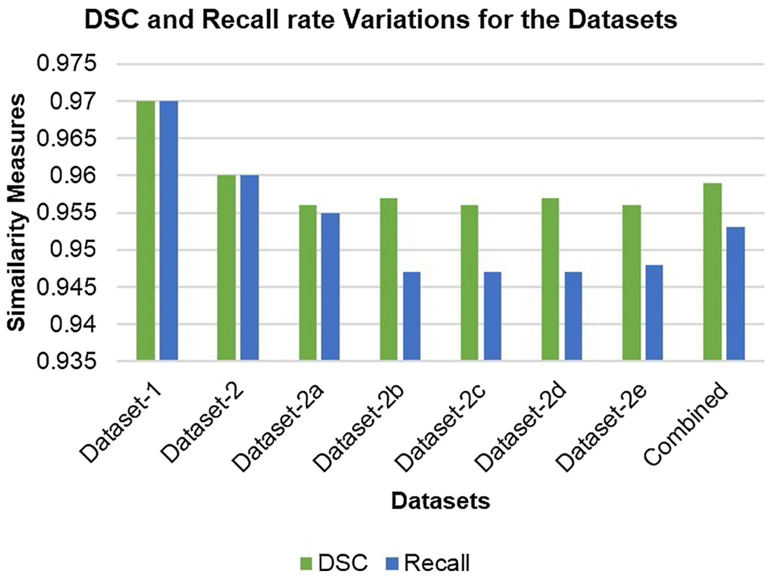


Fig. 10 Plots for DSC and recall rate for the datasets

to the fact that, the *dataset-1* consists of images with less variations and also the images are cropped images of the whole PBS images.

We also compared the results of the proposed method with the state-of-the art methods. The similarity measure for nuclei detection is given in Table 8. It can be observed from the table that, the performance of the proposed method is better compared to the methods proposed in [5, 8, 10, 20, 26, 31, 33–35]. We obtained average DSC around 0.97 for *dataset-1*. Also the average DSC value is considered for a total of 1120 images with color and brightness variations. This indicates that, use of multiple threshold values as in case of the proposed method and the method proposed in [8] perform better for images with color and illumination variations. Madhloom et al. [20] used Otsu's thresholding method for detection of nuclei and they reported the accuracy of segmentation between 85 and 98%. Though Otsu's method gives the threshold value based on the image intensity level, use of global and single threshold value for nuclei segmentation for the dataset may not be sufficient to handle the variations present in the images. Multiple threshold values offer appropriate detection of region of nuclei. The

Table 8 Similarity measure of the state-of-the-art methods and the proposed method

Method	Average Similarity Measure
The proposed method	0.965
Gram-schmidt [33]	0.93
Arithmetic operation [20]	0.92
Multi-level thresholding [8]	0.96
Ellipse-curve-fitting [10]	0.93
SDM-based clustering [26]	0.86
Active contour [34]	0.923
GVF snake [35]	0.92
Image fusion [31]	0.894
K-means clustering [5]	0.95

proposed method gives promising results for detection of nuclei in presence of possible variations found in microscopic images.

Further, segmentation of nuclei can be explored using visual attention techniques [4] in which region of interest should have different contrast compared to the background. Since the contrast of region of nucleus is different from the other regions visual attention method can be explored for detection of nucleus. Also use of Convolution Neural Network can be explored for segmentation of region of nucleus. It is reported that, CNN can learn the features by itself and it can provide greater accuracy [17].

4 Conclusion

In this paper we presented a robust method for detection of nuclei of WBCs. We considered a novel image enhancement method to minimize the color and brightness variations present in the images. We used TissueQuant method for color segmentation to detect approximate region of nuclei. This region was used to assess brightness level around the nucleus and identified an appropriate threshold value. The robustness of the proposed method was evaluated by considering images from ALL-IDB2 dataset, a dataset obtained from two lab setups, and introducing uniform and non-uniform brightness variations into the images. The similarity measure of the proposed method was compared with nine state-of-the-art methods and was found to be better. The average DSC of the proposed method for all the images considered including color and brightness variations is around 0.97.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Adnan K (2008) IBCIS: intelligent blood cell identification system. Prog Nat Sci 18:1309–1314
2. Amin MM, Kermani S, Talebi A, Oghli MG (2015) Recognition of acute lymphoblastic leukemia cells in microscopic images using kmeans clustering and support vector machine classifier. J Med Sign Sens 5:49–58
3. Azam B, Qureshi RJ, Jan Z, Khattak TA (2014) Color based segmentation of white blood cells in blood photomicrographs using image quantization. Res J Recent Sci 3(4):34–39
4. Chen J, Ma B, Cao H, Chen J, Fan Y, Xia T, Li R (2017) Attention region detection based on closure prior in layered bit planes. Neurocomputing 251:16–25
5. Congcong Z, Xiaoyan X, Xiaomei L, Ying-Jie C, Wu Z, Jun C, Chengyun Z, Zhi L (2014) White blood cell segmentation by color-space-based k-means clustering. Sensors (Basel) 14(9):16128–16147
6. Devi SS, Roy A, Singha J, Sheikh SA, Laskar RH (2017) Malaria infected erythrocyte classification based on a hybrid classifier using microscopic images of thin blood smear. Multimed Tools Appl 77(1):631–660
7. Dorini LB, Minetto R, Leite NJ (2007) White blood cell segmentation using morphological operators and scale-space analysis. Proc IEEE XX Braz Sym Comput Graph Image Process: 294–304
8. Huang D-C, Hung KD, Chan YK (2012) A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images. J Syst Softw 85:2104–2118

9. Jan Z, Khan A, Sajjad M, K. Muhammad, S. Rho, I. Mehmood, (2017) A review on automated diagnosis of malaria parasite in microscopic blood smears images, Multimed Tools Appl: 1–26
10. Jaroonrut P, Charnchai P (2015) Segmentation of white blood cells and comparison of cell morphology by linear and nave bayes classifiers. Biomed Eng Online 14(63):1–19
11. Joshi MD, Karode AH, Suralkar SR (2013) Detection of acute leukemia using white blood cells segmentation based on blood samples. Int J Electron Commun Eng Technol 4:148
12. Kalinathan L, Kathavarayan RS, Nagendram D, Vij M, Rela M (2018) Segmentation of hepatocellular carcinoma and dysplastic liver tumors in histopathology images using area based adaptive expectation maximization. Multimed Tools Appl 77(2):1761–1782
13. Labati RD, Piuri V, Scotti F (2011) ALL-IDB: the acute lymphoblastic leukemia image database for image processing. Proc IEEE Int Conf Image Process: 2045–2048
14. Li Y, Ziu R, Mi L, Cao Y, Di Yao D (2016) Segmentation of white blood cell from acute lymphoblastic leukemia images using dual-threshold method. Comput Math Methods Med 2016:12
15. Li Y, Li Y, Kim H, Serikawa S (2017) Active contour model-based segmentation algorithm for medical robots recognition, Multimed Tools Appl: 16
16. Lin Y, Meer P, Foran DJ (2005) Unsupervised segmentation based on robust estimation and color active contour models. IEEE Trans Inf Technol Biomed 9(3):475–486
17. Liu N, Han J, Zhang D, Wen S, Liu T (2015) Predicting eye fixations using convolutional neural networks. Proc IEEE Conf Comput Vision Pattern Recogn. https://doi.org/10.1109/CVPR.2015.7298633
18. Loffler H, Rastetter J, Haferlach T (2005) Atlas of clinical hematology, sixth edn. Springer
19. Longo D, Harrison TR (2012) Atlas of hematology and analysis of peripheral blood smears. McGraw- Hill
20. Madhloom HT, Kareem SA, Ariffin H, Zaidan AA, Alanazi HO, Zaidan BB (2010) An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold. J Appl Sci 10(11):959–966
21. Madhloom HT, Kareem SA, Ariffin H (2012) An image processing application for the localization and segmentation of lymphoblast cell using peripheral blood images. J Med Syst 36:2149–2158
22. Marzukia NIC, Mahmoodb NH, Razakb MAA (2015) Segmentation of white blood cell nucleus using active contour. J Teknologi 74(6):115–118
23. Mathur A, Tripathi AS, Kuse M (2013) Scalable system for classification of white blood cells from leishman stained blood stain images. J Pathol Inform 4(2):15
24. Nagabhushana RM, Rajeshwari P (2014) Geometrical analysis of leukocyte nucleus to detect lymphoblast from microscopic blood images. Int J Technol Eng Sci 2(6):1947–1951
25. Nemane JB, Chakkarwar VA, Lahoti PB (2013) White blood cell segmentation and counting using global threshold. Int J Emerg Technol Adv Eng 3:639–643
26. Neoh SC, Srisukkham W, Zhang L, Todryk S, Greystoke B, Lim CP, Hossain MA, Aslam N (2015) An intelligent decision support system for leukaemia diagnosis using microscopic blood images. Sci Rep 5:14
27. Othman MZ, Alaa AB (2014) Segmentation and feature extraction of lymphocytes WBC using microscopic images. Int J Eng Res Technol 3:696–701
28. Prasad K, Bhagath K, Chakravarthy M, Prabhu G (2012) Applications of 'TissueQuant' - a color intensity quantification tool for medical research. Comput Methods Prog Biomed 106:27–36
29. Prasad K, Winter J, Bhat UM, Acharya RV, Prabhu GK (2012) Image analysis approach for development of a decision support system for detection of malaria parasites in thin blood smear images. J Digit Imaging 25: 542–549
30. Ramin S, Hossein R, Ardeshir T (2012) Extraction of nucleolus candidate zone in white blood cells of peripheral blood smear images using curvelet transform. Comput Math Methods Med 2012:12
31. Rawat J, Singh A, Bhadauria HS (2014) An approach for leukocytes nuclei segmentation based on image fusion, Proc IEEE Int Sym Signal Process Inform Technol (ISSPIT): 456–461
32. Rawat J, Singh A, Bhadauria HS, Virmani J, Devgun JS (2017) Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers. Multimed Tools Appl 76(18):19057–19085
33. Rezatofighi SH, Soltanian-Zadeh H (2011) Automatic recognition of five types of white blood cells in peripheral blood. Comput Med Imaging Graph 35:333–343
34. Rong C, Xiaoqin Z, Lixin H, Min W, (2015) Subimage cosegmentation in a single white blood cell image. Proc IEEE 7th Int Conf Comput Intell Commun Syst Netw: 152–157
35. Sadeghian F, Seman Z, Ramli AR, Kahar A, Hisham B, Saripan MI (2009) A frame work for white blood cell segmentation in microscopic blood images using digital image processing. Biol Proc Online 11:196–206
36. Taneja A, Ranjan P, Ujlayan A (2017) Multi-cell nuclei segmentation in cervical cancer images by integrated feature vectors. Multimed Tools Appl: 1–20
37. Yiping Y, Yiping C, Wenxian S (2014) A method of leukocyte segmentation based on S component and B component images. J Innov Optic Health Sci 7(1):1–8

Roopa B Hegde has been working as assistant professor in NMAM Institute of Technology, Nitte since past 12 years. Currently, she is a research scholar in School of Information Sciences, Manipal. She is working on automation of blood smear analysis using image processing. Her area of interest are image processing and digital electronics.

Dr. Keerthana Prasad has been working on applications of image processing and analysis in the area of medical diagnostics and research. She has developed many in-house analysis tools for medical researchers of the university. She has won many awards such as Rural Innovation award, Innovation Day award at Philips, IDF Innovation award and best paper award. She has submitted three patent applications in India which are pending examination and published around 20 research papers in indexed international journals.

Dr. Harishchandra Hebbar received his PhD at Manipal University, India. He is a professor in School of Information Science. He has published more than five research papers in international journals. He has undertaken five funded projects. His area of interest are embedded systems design, content based image retrieval techniques, computer networks and product design.

Dr. Brij Mohan Kumar Singh completed M.D. in Pathology at Kasturba Medical College (KMC), Manipal University. Currently, he is assistant professor in the department of Immunohematology and Blood Transfusion in KMC. He has published more than 17 papers in national and international journals.