

Automatic Segmentation of Red Blood Cells from Microscopic Blood Smear Images Using Image Processing Techniques



K. T. Navya , Subhraneil Das, and Keerthana Prasad 

Abstract Human blood is a very effective parameter to detect, diagnose and rectify ailments of the human body. Complete blood count (CBC) is a method to clinically obtain a statistical measure of blood and its related parameters, i.e., red blood cells (RBCs), white blood cells (WBCs), platelets, hemoglobin concentration to name a few. This helps to determine the physical state of the subject. For further diagnosis, peripheral blood smear, a thin layer of blood smeared on a microscope slide and stained using various staining methods is examined for the morphology of the cells by the pathologists. However, manual inspection of smear images is tedious, time-consuming, and laboratorian-dependent. Although there are certain software-based approaches to tackle the problem, most of them are not robust for all staining methods. Thus, the need is to create an automated algorithm that will work for different staining types, thereby alleviating both the aforementioned drawbacks. This work aims to create an automatic method of segmenting and counting RBCs from blood smear images using image processing techniques to help diagnose RBC-related disorders. In the proposed method, the images are first preprocessed, i.e., standardized to a uniform color and illumination profile using contrast enhancement, adaptive histogram equalization followed by Reinhard stain normalization algorithms. WBCs and platelets are extracted in HSI color space and subtracted from the original image to retain only RBCs. Thereafter using morphological operations and active contour segmentation algorithms, a count of total RBCs were obtained even for overlapped cells in the microscopic blood smear image. The proposed method achieved counting accuracy of 89.6% for 150 images.

K. T. Navya · S. Das

Department of Electronics and Communication, Manipal Institute of Technology, Manipal, India

K. Prasad

Manipal School of Information Sciences, Manipal, India

e-mail: keerthana.prasad@manipal.edu

K. T. Navya (✉) · S. Das · K. Prasad

Manipal Academy of Higher Education, Manipal 576104, India

e-mail: navya.kt@manipal.edu

Keywords Red blood cells · Color normalization · Image processing · Segmentation · Active contour

1 Introduction

Blood sampling and diagnosis of diseases based on the blood count and characteristics are an important aspect of medicinal science and engineering. Human blood is an essential source for developing diagnosis patterns and antigens for numerous diseases. Blood is a specialized body fluid and has four main components [1].

- *Plasma*, liquid component produced in the bone marrow to transport blood cells throughout the body
- *Erythrocytes or red blood cells* is a biconcave disk with a central pallor, 6–8 μm in diameter with pink stain helps carry oxygen from the lungs to the rest of the body.
- *Leukocytes or white blood cells*, largest cells, varies between 8 and 20 μm in diameter, a dark purple nucleus with pale cytoplasm produce antibodies to fight off infection.
- *Thrombocytes or platelets*, small cell fragments without nuclei, 2–3 μm in diameter, bluish in color helps in blood clotting.

Figure 1 shows a typical blood film or peripheral blood smear stained using the Romanowsky staining method to examine the blood cells microscopically. Staining plays a very important role in the analysis of blood smear images as the color profile, contrast, and feature details in a microscopic blood smear image are determined initially by this aspect. Popular staining methods under Romanowsky staining include Giemsa, Wright-Giemsa, Leishman stain, etc. The blood smear thus taken is analyzed in the feather edge of the slide, and hence, subsequent conclusion helps in knowing the current state of health of the patient as well as diagnosing a specific ailment. Considering the manual techniques followed previously for analyzing blood smear images, it was very tedious and not reliable. Thus, this work is aimed at developing

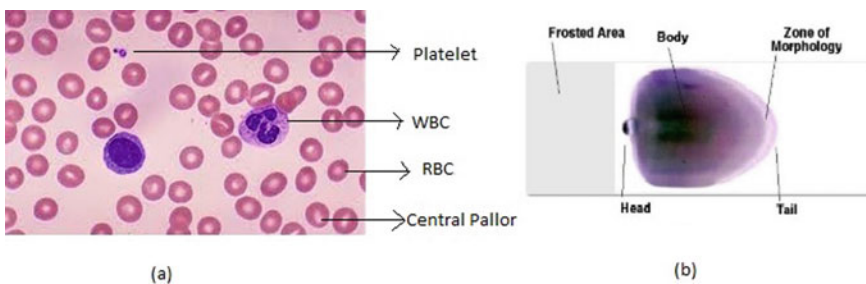


Fig. 1 Blood smear images [2] **a** microscopic view of blood smear **b** blood smear slide

an autonomous and robust system for image preprocessing followed by segmentation of red blood cells in a given blood smear image using image processing techniques in MATLAB to help diagnose RBC-related disorders.

2 Related Work

This section attempts at giving a brief overview of various past developments in the field of automated blood counting and its related parameters such as color normalization, preprocessing, RBC segmentation, background, and foreground differentiation using various segmentation methods based on underlying principles of image processing. Sharif et al. [3] elaborated a robust methodology for RBC segmentation using masking and morphological operation. This paper has a combination of pixel-based, region-based, and morphological segmentation. YCbCr color has been chosen for illumination issues. A marker-controlled watershed algorithm was used to handle overlapping cells from 20 images. Mazalan et al. [4] proposed an automated method to count RBCs in microscopic images using circular Hough transform (CHT) for 10 sample images and obtained 91.9% accuracy. This method is cost-effective and provides an alternative way to recognize and count circular cells. Alomari et al. [5] proposed an iterative structured circle method to segment WBC and RBC, and average accuracy of 95.3% for RBCs and 98.4% for WBCs was achieved for 100 images. However, this method depends on the number of iterations. Abbas et al. [6] presented a method for RBC segmentation in YCbCr color space. K-means clustering was used to identify cells from 90 Giemsa stained images. Tomari et al. [7] proposed a Hough transform (HT) method to count overlapped RBCs and obtained 94% average accuracy for four sample images. However, this method involved many parameters. Alam et al. [8] presented the YOLO algorithm to identify and count blood cells and obtained accuracy 96.09% RBC, 86.89% WBC, and 96.36% for 364 annotated images. Wei et al. [9] proposed K-means clustering-based method to segment and count overlapped RBCs. The method used the H and S components to differentiate between WBC and RBC. The author of the work obtained 92.9% accuracy for 100 Wright-Giemsa stained images. Acharya et al. [10] presented a method to identify and count RBC using K-Medoids and geometric features. This method achieved 98% accuracy for 1000 Wright stain images. Ejaz et al. [11] proposed HT-based method to segment and count RBCs and obtained 94.9% accuracy for 500 subjects. Berge et al. [12] proposed RBC segmentation method using boundary extraction and curvature calculation. The Delaunay triangulation method was used to split overlapped RBCs of any shapes. This method obtained 2.8% absolute error for 49 Giemsa stained images. Hegde et al. [13] proposed active contour method to segment WBCs. G'G'B channel representation for handling illumination and color variations and obtained 96% sensitivity for 54 images from the ALL-IDB2 dataset. Adagale et al. [14] proposed an overlapped red blood cell counting algorithm using template matching and pulse coupled neural network and achieved 90% average accuracy for 40 images. However, accuracy decreases due to overlapped RBCs. Cruz et al. [15] presented RBC count-

ing method using blob analysis and watershed transform in the HSV component and obtained 96% average accuracy for 10 blood samples. Loddo et al. [16] proposed a blood cell counting method using nearest neighbor and SVM techniques by cropping each cell manually. Clumped cells were counted using CHT. This method used 368 images from the ALL-IDB dataset and obtained an average accuracy of 99.2% for WBCs and 98% for RBCs. Yeldhos et al. [17] implemented FPGA-based RBC counting system. Watershed transform and CHT segmentation method were used in YCbCr color space and obtained 90.98% accuracy for 108 blood smear images from the ALL-IDB database. Tran et al. [18] presented deep learning semantic segmentation method for RBC and WBC segmentation and counting. SegNet architecture was utilized to segment blood cells by labeling each pixel. The segmentation accuracy for WBCs and RBCs was 94.93% and 91.11%, respectively. For cell counting, Euclidean distance transform and binary dilation are used and obtained 93.3% for RBC and 97.29% for WBC for 42 ALL-IDB database images. From the past studies, different methods such as circular Hough transform, watershed transform, morphological operations, thresholding-based methods, K-means clustering, ANN, DNN have been used for RBC segmentation and counting. However, these methods lack robustness in handling blood smear images with multiple stains. Hence, there is a need for developing a robust system to handle images taken from various laboratory settings.

3 Methodology

This section gives a sequence of steps that are to be followed to obtain the desired results. A methodology of the work is depicted in Fig. 2.

3.1 Image Acquisition

The required Leishman stained blood smear images are acquired from Kasturba Medical College (KMC), Hematology department, Manipal, with a 100x lens objective. Also, smear images from Isfahan MISP online database[19] are gathered for the process.



Fig. 2 Methodology of proposed work

3.2 Preprocessing

Due to the varied color profile and contrast of the microscopic smear images, there is a need to standardize the image to obtain a consistent color profile. To make an image ready for the automatic segmentation under various image settings and stains, illumination correction and color normalization methods have been implemented. The various preprocessing methods used are discussed further [20].

- *Linear Contrast Enhancement using Adaptive Histogram Equalization*

This method is helpful to reduce illumination variation. The images are first converted from RGB to LAB color space where a consistent normalization of the L parameter (Luminosity) is performed followed by grayscale conversion and subsequent adaptive histogram equalization.

- *Gamma Correction*

is a nonlinear method used to correct the image's luminance; i.e., it amplifies the shadows or the bright regions of the image as per the requirement. It is used to correct uneven illumination by encoding luminance in video or still image systems. Gamma correction is defined as:

$$V_{out} = AV_{in}^{\Gamma} \quad (1)$$

Powers larger than 1 make the shadows darker and smaller than 1 make dark regions lighter.

- *Gray World Assumption method*

It is a color correction function based on the principle that on average the scene is neutral gray. The algorithm produces an estimate of illumination by computing the mean of each channel of the image. The average pixel value of an unsigned 8-bit integer image is 127.5, so by calculating the real average pixel value, the scaling value is computed to scale the entire image linearly. In practice, the average of each individual channel is used to calculate a separate scaling value for each channel.

- *Histogram Matching*

It adjusts histogram of the 2D image to match histogram of reference image in order to normalize the color. The reference RGB image is converted into HSV, and hue component is extracted and used for histogram matching with the input image to segment WBC and platelets. The HSV image is then converted into RGB image for further processing.

- *Reinhard's Stain Normalization*

The algorithm thus maps the color distribution of an image to that of a well-stained target image thereby solving the problem of inconsistency. Technically it matches the mean and the standard deviation of each color channel in the two images in that color space [21]. The statistical approach to color mapping can be shown as in equations

$$\begin{aligned}
 l_{\text{mapped}} &= \frac{l_{\text{original}} - \tilde{l}_{\text{original}}}{\hat{l}_{\text{original}}} \hat{l}_{\text{target}} + \tilde{l}_{\text{target}} \\
 \alpha_{\text{mapped}} &= \frac{\alpha_{\text{original}} - \tilde{\alpha}_{\text{original}}}{\hat{\alpha}_{\text{original}}} \hat{\alpha}_{\text{target}} + \tilde{\alpha}_{\text{target}} \\
 \beta_{\text{mapped}} &= \frac{\beta_{\text{original}} - \tilde{\beta}_{\text{original}}}{\hat{\beta}_{\text{original}}} \hat{\beta}_{\text{target}} + \tilde{\beta}_{\text{target}}
 \end{aligned} \tag{2}$$

where the superscripted variables \tilde{l} and \hat{l} represent the mean and standard deviation of luminosity parameter, respectively, in the LAB color space. Similarly, the subscripted variables l_{original} , l_{target} , and l_{mapped} indicate whether the parameter concerned is either of the original, target or the mapped image of the LAB color space parameters. An image with acceptable color profile, contrast, and illumination was chosen as the target image, and the source image was subjected to Reinhard staining normalization.

- *Noise Removal*

The blood smear image post color normalization is subjected to noise removal and deblurring using Wiener filter. The filter processes a two-dimensional image and eliminates granular noise from the image. It is also useful in removing motion blur from the image, thereby making the image ready for further processing.

The image enhancement algorithms on being implemented on a set of images gave a series of results based on their resolution, contrast standardization, and processing method as stated in this section. In Fig. 3, the original faded Leishman stained microscopic image of a normal blood smear is shown which is almost devoid of sharpness, contrast, and color and the reference image for color normalization. The contrast-enhanced reference image is exposed to gamma correction for illumination correction and then passed through a Wiener filter with a filtering window of [10]. The resultant images are smoothed and exhibit a much less extent of jagged edges in the features of the image. The input image is subjected to grayworld color normalization in order to correct uniformity of color, contrast, and illumination variation. However, results obtained show consistency in image regions with proper contrast and deviation for other regions thereby proving to be inefficient for the required purpose. The images are then subjected to the statistical means of color matching in order to achieve a fixed level of uniform tone and contrast which can then be applied and fixated to any image which is fed as an input to the algorithm. The target image ideally should possess uniformity in regard to the factors of illumination, color tone, contrast, warmth,

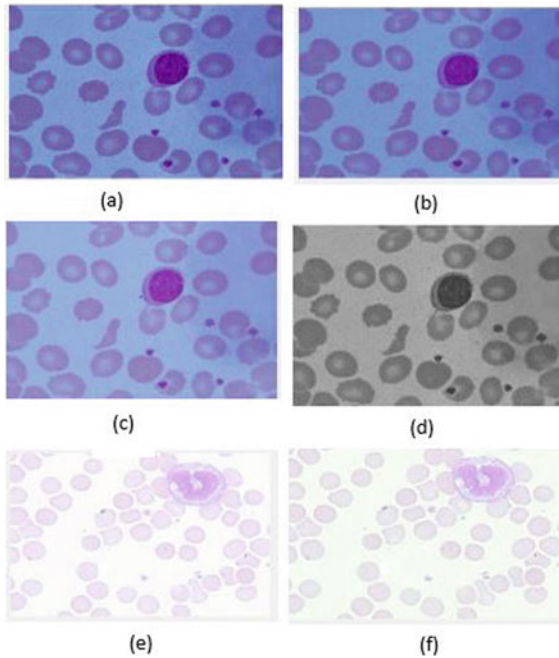


Fig. 3 Preprocessed blood smear images **a** Leishman reference image **b** contrast-enhanced image **c** gamma correction for image **d** Wiener filtered image **e** Leishman stained input image **f** grayworld normalized image

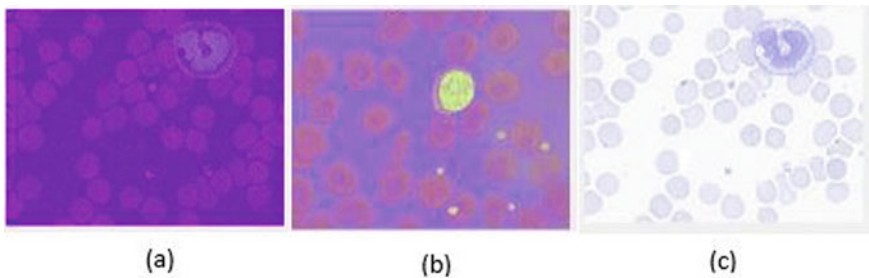


Fig. 4 **a** H component of input image **b** H component of preprocessed reference image **c** histogram matched RGB image

and sharpness. Histogram matching technique is used for color normalization and illumination correction where hue component of the reference image is matched to hue component of the source image to extract WBCs and platelets from blood smear images, and the matched RGB image is as shown in Fig. 4. In order to compare the histogram matched image with another color normalization method, Reinhard stain normalization was applied for source image with reference to the target image. The normalized results obtained with different stain input and target images taken in 100x

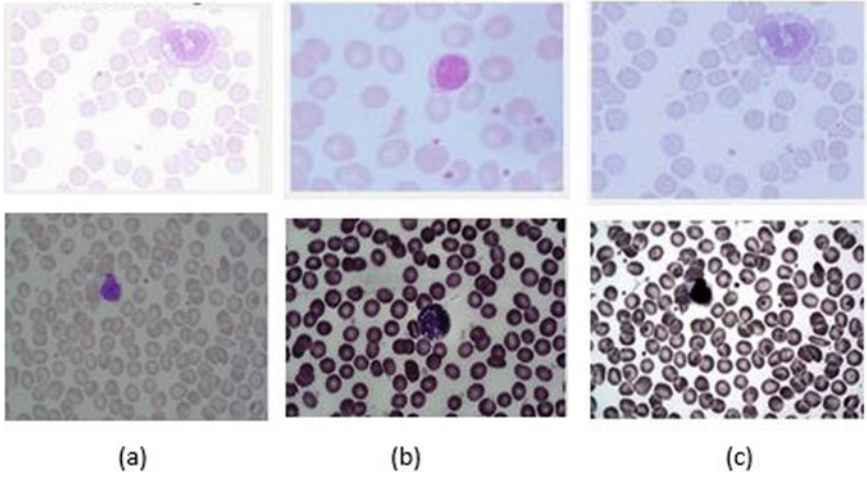


Fig. 5 **a** Leishman and Wright stain input image **b** 100 \times and 40 \times reference image **c** color normalized image

and 40X magnification are shown in Fig. 5. On visualization, Reinhard normalized image combat both illumination and color variation and show a prominent difference between the blood cells.

3.3 Segmentation of Blood Cells

The normalized image is processed further for WBC, platelets, and RBC differentiation. Thus, the Reinhard normalized image is subjected to unsharp masking for sharpening the object boundaries to distinct with respect to the background, especially at the edges of the RBCs. WBCs are extracted using the S-channel of HSV that highlights the nucleus and green channel of RGB color space which highlights RBCs and cytoplasm of WBC. Then by combining both the channel outputs, WBCs are removed. Further, traces of WBCs are removed using morphological opening via a disk-shaped morphological element with the radius range of 80–90, and then similarly, platelets are removed, subtracted, and smoothed to obtain the image with only RBCs. The appropriate results are shown in Fig. 6.

We can observe from the RBC segmented image that due to morphological operations, boundaries of RBCs are altered. For further RBC disorder analysis, the morphology of the cell is very important. So to improve on this diverging active contour segmentation algorithm is used. The basic premise of active contours is energy minimizing models; i.e., the snake tends to minimize its energy, thereby shrinking onto the boundaries of the image objects[22]. The entire process of fitting onto the bound-

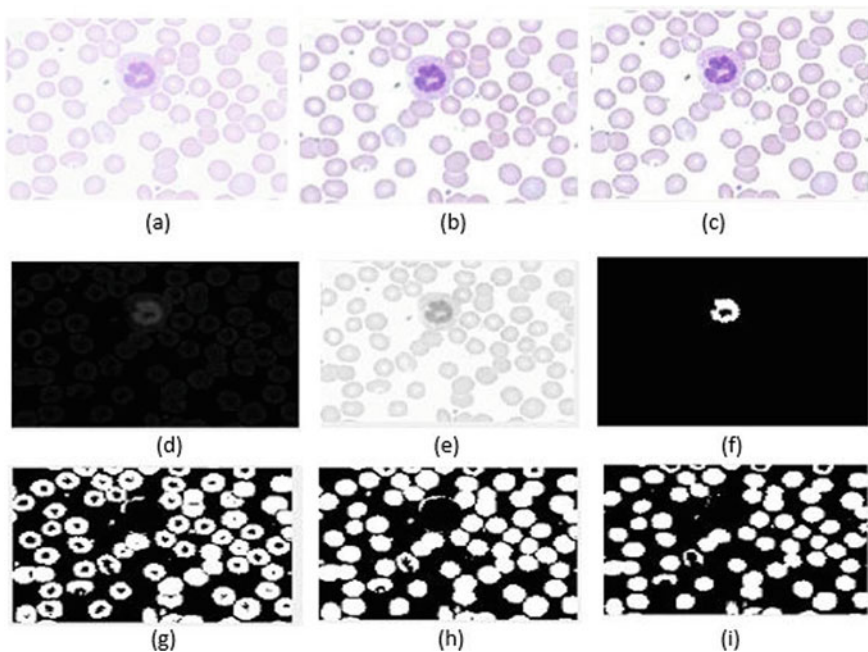


Fig. 6 a Input image b reinhard normalized image c sharpened image d S-channel of original image e green channel extraction f extracted WBC image g subtracted image h image smoothed by morphological dilation i resultant image after erosion

aries occurs in several iterations to get the best fit possible. Now, considering the snake as a continuous parametric variable, we can define its position in the image as

$$v(t) = (x(t), y(t)) \quad (2)$$

where $v(t)$ is the active contour and $x(t)$, $y(t)$ are the continuous contour coordinates. Thus, the energy equation associated with active contours is as follows:

$$E = \int_0^1 [E_{\text{int}}(v(t)) + E_{\text{img}}(v(t)) + E_{\text{con}}(v(t))] dt \quad (3)$$

where E_{int} is internal energy, E_{img} is image forces and E_{con} is external constraint force. The divergence of the contour was achieved by exercising a certain degree of *contraction bias*. This parameter ranges from -1 to 1 , and a negative value indicates expanding contour, whereas a positive value indicates a shrinking contour. For the initial contour mask, centroids of the morphologically segmented RBCs are taken as a seed point to diverge. Figure 7 depicts an initial mask and contour detected RBC binary image as an overlay mask to present the detected cells clearly.

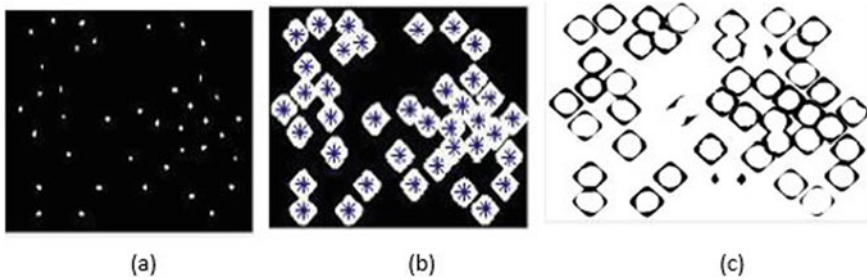


Fig. 7 a Initial contour mask b RBC segmented image c Detected RBCs along with their shape enclosed within binary mask

4 Results and Discussion

In the case of any algorithm, a reference or ground truth is required to measure the accuracy of the used algorithm. The segmented RBCs are counted manually and compared with the ground truth. Here for ground truth formation, we make use of ImageJ Tool to segment the microscopic blood smear images to get a clear count of the RBCs present in the image. Clustered or ambiguous cell boundaries are drawn using a freehand line tool to obtain clear cell boundaries and count. Now using the count obtained from the ground truth of the images and the count obtained via our proposed algorithm we determine the accuracy of the method. In this work, accuracy is determined as the percentage of detected count divided by the actual count of RBCs present in that image. A total of 150 images, 75 Leishman stained images from KMC and 75 images from the online database [19] were used in the study. The active contour algorithm yields an overall accuracy of 89.6% for 150 images. Though it resolves a significant amount of overlapped cells present in an image, some overlapped cells and clusters still remain unresolved. Still, there is a need for a robust segmentation method and accurate ground truth for RBC count to achieve higher accuracy.

5 Conclusion

The methods implemented to meet the objective of the work has been achieved so far. Preprocessing of the images proved to be a cumbersome task and after several trials, Reinhard's method demonstrated to be the most effective method of color normalization. Similarly, post-WBC and platelet removal, the active contour model provided an RBC count closest to the actual values as obtained from the ground truth and achieved an overall accuracy of 89.6%. However, to overcome overlapping and densely clustered RBCs, a robust segmentation algorithm has to be implemented to get a much more accurate count. Further, segmentation processes using convolutional

neural networks and artificial intelligence toolboxes can be used to create a shape classifier for identifying cell edges, overlapped cells, and cell clusters accurately in order to detect specific diseases and anomalies.

References

1. H. Mohan, *Textbook of Pathology* (Medical Publishers Pvt. Limited, Jaypee Brothers, 2018)
2. K.W. Jones, Evaluation of cell morphology and introduction to platelet and white blood cell morphology. *Clin. Hematol. Fundam. Hemost.* 93–116 (2009)
3. J.M. Sharif, M. Miswan, M. Ngadi, M.S.H. Salam, M.M. bin Abdul Jamil, Red blood cell segmentation using masking and watershed algorithm: a preliminary study, in *2012 International Conference on Biomedical Engineering (ICoBE)* (IEEE, 2012), pp. 258–262
4. S.M. Mazalan, N.H. Mahmood, M.A.A. Razak, Automated red blood cells counting in peripheral blood smear image using circular Hough transform, in *2013 1st International Conference on Artificial Intelligence, Modelling and Simulation* (IEEE, 2013), pp. 320–324
5. Y.M. Alomari, S. Abdullah, S.N. Huda, R. Zaharatul Azma, K. Omar, Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection algorithm. *Comput. Math. Methods Med.* **2014** (2014)
6. N. Abbas, D. Mohamad et al., Microscopic RGB color images enhancement for blood cells segmentation in YCBCR color space for k-means clustering. *J. Theor. Appl. Inf. Technol.* **55**(1), 117–125 (2013)
7. R. Tomari, W.N.W. Zakaria, R. Ngadengon, M.H.A. Wahab, Red blood cell counting analysis by considering an overlapping constraint 2006–2015. *Asian Res. Publishing Netw. (ARPNet)* **10**(3) (2015)
8. M.M. Alam, M.T. Islam, Machine learning approach of automatic identification and counting of blood cells. *Healthcare Technol. Lett.* **6**(4), 103–108 (2019)
9. X. Wei, Y. Cao, G. Fu, Y. Wang, A counting method for complex overlapping erythrocytes-based microscopic imaging. *J. Innov. Optical Health Sci.* **8**(06), 1550033 (2015)
10. V. Acharya, P. Kumar, Identification and red blood cell automated counting from blood smear images using computer-aided system. *Med. Biol. Eng. Comput.* **56**(3), 483–489 (2018)
11. Z. Ejaz, A. Hassan, H. Aslam, Automatic red blood cell detection and counting system using Hough transform. *Indo Am. J. Pharm. Sci.* **5**(7), 7104–7110 (2018)
12. H. Berge, D. Taylor, S. Krishnan, T.S. Douglas, Improved red blood cell counting in thin blood smears, in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, 2011), pp. 204–207
13. R.B. Hegde, K. Prasad, H. Hebbar, B.M.K. Singh, Image processing approach for detection of leukocytes in peripheral blood smears. *J. Med. Syst.* **43**(5), 114 (2019)
14. S. Adagale, S. Pawar: Image segmentation using PCNN and template matching for blood cell counting, in *2013 IEEE International Conference on Computational Intelligence and Computing Research* (IEEE, 2013), pp. 1–5
15. D. Cruz, C. Jennifer, L.C. Castor, C.M.T. Mendoza, B.A. Jay, L.S.C. Jane, P.T.B. Brian et al., Determination of blood components (WBCs, RBCs, and platelets) count in microscopic images using image processing and analysis, in *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (IEEE, 2017), pp. 1–7
16. A. Loddo, L. Putzu, C. Di Ruberto, G. Fenu, A computer-aided system for differential count from peripheral blood cell images, in *2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)* (IEEE, 2016), pp. 112–118
17. M. Yeldhos, Red blood cell counter using embedded image processing techniques. *Res. Rep.* **2** (2018)

18. T. Tran, O.H. Kwon, K.R. Kwon, S.H., Lee, K.W. Kang, Blood cell images segmentation using deep learning semantic segmentation, in *2018 IEEE International Conference on Electronics and Communication Engineering (ICECE)* (IEEE, 2018), pp. 13–16
19. M. Kashefpor, R. Kafieh, S. Jorjandi, H. Golmohammadi, Z. Khodabande, M. Abbasi, N. Teifuri, A.A. Fakharzadeh, M. Kashefpoor, H. Rabbani, Isfahan MISP dataset. *J. Med. Signals Sens.* **7**(1), 43 (2017)
20. Image enhancement techniques. <https://in.mathworks.com>. Accessed May 2020
21. E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images. *IEEE Comput. Graph. Appl.* **21**(5), 34–41 (2001)
22. M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models. *Int. J. Comput. Vis.* **1**(4), 321–331 (1988)