Int. J. Nonlinear Anal. Appl. 13 (2022) 2, 3251–3259
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.28179.3825

Development of images segmentation using image thresholder and batch processing technique on the blood smears

Suhad Al-Shoukry[a,*], Zalili Binti Musa[b], Duha Amer[a], Safaa Muhsen Kareem[a]

[a]Department of Computer Systems Techniques, Al-Najaf Technical Institute, Al-Furat Al-Awsat Technical University (ATU), Kufa, Al-Najaf, Iraq
[b]University Malaysia Pahang (UMP), Pekan, Pahang, Malaysia

(Communicated by Javad Vahidi)

Abstract

Image segmentation is an important part of image processing, and one of the most common approaches is threshold segmentation. A new segmentation technique with each pixel in the image has its own threshold is developed in response to the fact that standard threshold-based segmentation algorithms only establish one or many thresholds, making it difficult to extract the complex information in an image. This work employs image segmentation tools to examine images of thin blood smears data set. The goal is to explore options for a noniterative-based and automated system for detecting parasites in blood smears. This can be achieved by detecting the presence of a parasite in thin blood smears and quantifying the portion of red blood cells in the sample that are infected. First, we try segmenting the individual red blood cells from the background using the color thresholder. Next, we clean up the obtained cell mask and examine cell properties using the image region analyzer function, which allows quickly filling in region holes and filtering out regions based on their properties such as area dimensions or eccentricity. Then quickly gauge and specify the expected diameter range of the cells in pixels and indicate that the circles are dark relative to the background. Finally, we've combined the code for finding circles matching image histograms and the parasite threshold detection logic into a single function to quickly examine the performance of this function on the other images using the image batch processing technique. The proposed detection function labels the detected cells with blue circles the parasites are marked in red and the infected cells are highlighted in green. The proposed algorithm has appropriately compensated for the variability in image quality.

Keywords: Blood Sample Analysis, image analysis, blood cell segmentation, cell morphology, classification
2020 MSC: 68U10

1 Introduction

Image processing segmentation is basically processed images with a 2d matrix with each coordinate having the value that is intensity value. It is required to use somehow method to separate those pixels into different groups. Many in both cases will be two groups, one will be the object group and the second will be the background group. The object group will be the signal which is the interesting area that is required to be created to separate that area. This is called the mask image, in which the pixels are divided into two groups; the white pixels are normally where the signal is, and the black pixel is the background.

*Corresponding author
Email address: inj.suhadaalzhra2010@atu.edu.iq (Suhad Al-Shoukry)

Received: April 2022 *Accepted:* June 2022

Several studies discussed image segmentation as the main task of image processing tools, especially in medical applications like thin smears inspections. For example, pathologists use hematoxylin and eosin (H&E) to stain body tissue during cancer diagnosis to distinguish between tissue types [7]. They next utilize clustering, an image segmentation technique, to identify the different tissue types in their photos. Clustering is a technique for separating things in a scene into groups. The K-means clustering algorithm discovered separations that keep objects inside every cluster as near together as possible while keeping them as far apart as possible from objects in other clusters [5]. One of the most important instruments in the diagnosis and evaluation of blood disorders is the segmentation of blood cells in digital hematology microscope images. For example, malaria continues to be a major threat to global health, with an estimated 200 million cases and over 400,000 fatalities each year as indicated in [9]. Pathological examinations are the gold standard in several fields of hematology and histopathology, and they also play an important role in an illness diagnosis. Pathologists examine blood cells from a patient's peripheral blood smears during clinical diagnosis. This examination is mostly focused on the forms, sizes, colors, textures, maturity stages, and staining processes of blood cells and their nuclei and cytoplasm, as well as morphological aspects and characteristics of the nuclei and cytoplasm. Computer-Aided Diagnosis methods linked to blood cells, their nuclei, and cytoplasm identification, as well as segmentation and classification algorithms, have been quickly increasing in the digital hematology domain recently. These strategies have played, and will continue to play, an important role in delivering traceable clinical information, consolidating relevant second opinions, and minimizing human intervention in digital hematology image processing [1].

Image segmentation is a fundamental problem in computer vision to extract meaningful subjects. The smears' images want to be partitioned into regions or segments and want these segments to be meaningful in some sense. In the context of binary images, it is possible to compute a histogram of such images, find an appropriate threshold, and then by thresholding the original image, a binary image can be obtained. This binary image is the segmentation that has been looking for. Initializing with active contour is very useful in many different domains, but doesn't solve the general segmentation problem. To appreciate how hard segmentation is in general, one needs to look at natural images. So the question here is, how does one segment this image into meaningful regions? For that matter, what is a meaningful segment, in this case? So we're going to develop some very simple methods for image segmentation, and our approach is going to be to group pixels together in an image that has similar visual attributes or characteristics. So first, we're going to start by looking at how humans perform segmentation. What kind of rules, principles do use to guide us during the process of segmentation?. We're going to treat images as a clustering problem. At each pixel in the image, a feature vector describes that pixel, which can be seen as a point in high-dimensional feature space. In this feature space, we're going to cluster points together, and those clusters are going to represent the meaningful regions or segments.

This research could help researchers and pathologists better understand the challenges of performing a thorough analysis of blood cell microscope pictures, which could aid in the identification of blood disorders and help researchers and pathologists in the future. This approach has the effect of increasing the accuracy and efficiency of pathologists' decision-making, which benefits patients by allowing for faster and more accurate diagnoses. The research on intelligent systems is significant because it lays out future ideas for overcoming overlapping blood cell identification and other issues with microscopic images. Better blood cell segmentation and identification could increase the accuracy of the cell counting approach for sickness diagnosis in the future.

2 Literature Review

Detection and segmentation techniques of blood smears, their cytoplasm, and nuclei from digital hematology microscope pictures have been discussed in the study [1], which presented a comparison and significant trends for the performance of existing approaches taking into account the databases used, the constraints, and the number of photos. The article [9] also discussed the current breakthroughs in image analysis and machine learning for microscopic malaria detection and provides an overview of these techniques. However, both these studies didn't provide research technical results and a survey on state-of-art studies has been presented instead. The study [6] developed an effective method for leukemia recognition using deep learning convolutional neural network (CNN) classifier to perform the preprocessing, segmentation, feature extraction, and classification over blood smear images. Although this work achieved an accuracy of 98.7%, the performance of the presented Chronological SCA-based classifier was complicated and multiple algorithms were used.

The paper [2] proposed an automated blood cell of bone marrow smear images based on two stages CNN to classify 10 classes of the erythroid and myeloid maturation sequence and obtained an accuracy of 97.06%. Although the proposed technique showed high classification performance for the blood differential count system, this method is also

quite complicated and time-consuming. These studies suffered from the large variations in cell appearance, such as size, color, and shape of cells, the adhesion between erythrocytes (red blood cells, RBCs) and leukocytes (white blood cells, WBCs), and the emergence of substantial dyeing impurities in blood smear images, pose a major challenge for robust and accurate leukocyte identification and segmentation in blood smear images. The study [3] proposed an end-to-end leukocyte localization and segmentation method, in which pixel-level prior information is used for supervisor training of a CNN, which is then used to locate the region of interests (ROI) of leukocytes, and finally, a segmentation mask of leukocyte is obtained based on the extracted ROI by forwarding propagation of the network. The proposed approach, however, was limited to only being applied on LeukocyteMask. To cap the mentioned limitations, the paper [11] used adaptive thresholding to offer a new segmentation approach for microscopic pictures of malaria parasites derived from thick blood smears. Low-pass filtering and contrast stretching were utilized in the enhancement procedure on 253 parasite candidates, cropped from 22 thick blood smear microphotographs. The proposed algorithm's average segmentation accuracy was 95.2 percent, and cytoplasm was successfully removed, but this method is effective in the detection of malaria parasites only.

The hematological counter equipment is both fast and accurate, but it is also highly expensive. In the paper [8], these issues prompted the development of a low-cost Computer-Aided System (CAS) for analyzing blood-smeared images produced from a microscope. The most critical phase of CAS is segmentation, and any failure here will result in inaccuracies. Recognizing the importance of segmentation, this study looked at various color band thresholding-based segmentation algorithms for counting white blood cells (WBCs). It was discovered that the S color component of HSV produced the highest WBC segmentation accuracy of 96.92 percent, but this method was required images with limited resolutions.

Due to differences in the appearance of cells across the slide, segmenting leukocytes is a difficult task, as stated in the study [4]. With color and light fluctuations, an automated approach for detecting nuclei and extracting leukocytes from peripheral blood smear images is provided. Nuclei are detected using arithmetic and morphological processes, while leukocytes are detected using the active contours method. Although the findings of this work showed that the suggested approach successfully detects nuclei and leukocytes with Dice scores of 0.97 and 0.96, respectively, the method's overall sensitivity was low and roughly obtained. The study [10] divided the WBC into two major components: cytoplasm and nucleus. The segmentation was carried out with the help of a proposed segmentation framework that combines numerous digital image processing methods applied on twenty microscopic blood pictures. The proposed framework achieved nucleus segmentation accuracy of 92 percent but the segmentation was with accuracy of 78 percent for cytoplasm.

The findings of these researches were crucial for the identification of different diseases as well as post-chemotherapy care. Manual, labor-intensive procedures for determining the differential count, on the other hand, result in inter-and intra variations among hematologists' results. As a result, an automated method to do the blood cell differential count is extremely desirable, but the advancement is hampered by several issues. Because of the diverse acquisition and staining techniques, each maturation stage has different blood cells, minor inter-class differences within each stage, and different images. Furthermore, for bone marrow smear analysis, a significant number of classes must be identified, and the high density of contacting cells smears makes single-cell segmentation challenging, which is critical in traditional image processing.

3 Objectives

The aim of the study is to explore options for a noniterative-based and automated system for detecting parasites in blood smears.

To achieve this aim, the following objectives are accomplished:

- To detect the presence of a parasite in thin blood smears.
- To quantify the portion of red blood cells in the sample that are infected.

MATLAB-based image processing techniques are used to do this task.

4 Method and materials

Image segmentation is the process of turning a picture into a series of pixel sections that are represented by a mask or a labeled image. It is feasible to process only the important parts of a picture rather than the complete image by separating it into segments.

4.1 Data

The blood smears we'll be using were obtained from the centers for disease controls DPDX website. About 186 digitized photos of May Grünwald-Giemsa (MGG)-stained peripheral blood (PB) smears are included in his dataset. They were generated from five individuals with hereditary spherocytosis during everyday work in the Core Laboratory at the Hospital Clinic of Barcelona. A 1,000x magnification microscope (Olympus BX43) and a digital camera were used to create the digital photographs (Olympus DP73). The photos in this dataset are in the JPG format (RGB, 2400 x 1800 pixels). Six samples of these images are shown in Fig. 1.

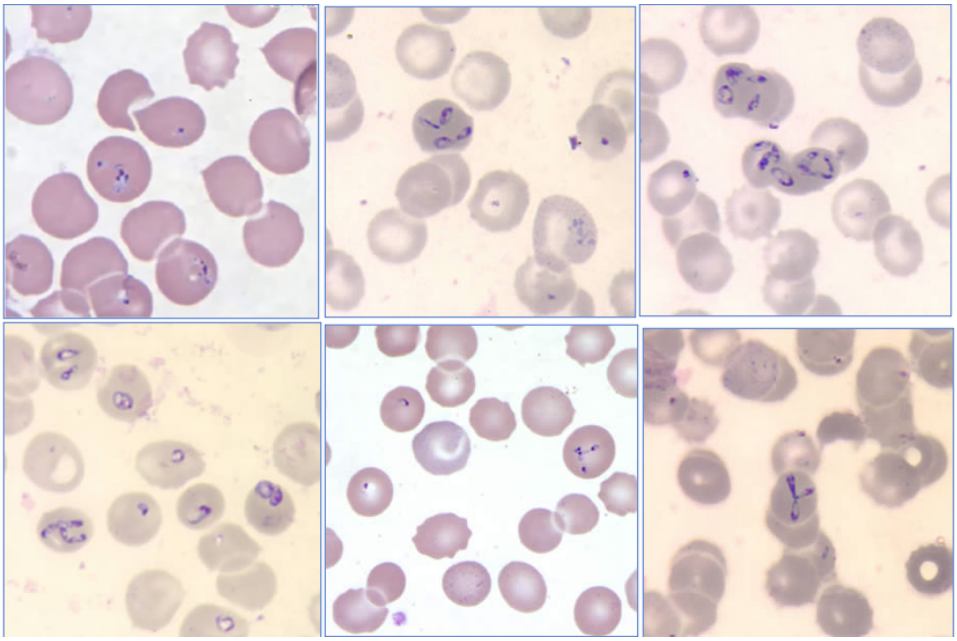


Figure 1: Six samples of thin blood smears

4.2 Procedure

MATLAB-based image processing techniques are used to examine thin blood smears. To perform the developed approach on thin blood smears, the following steps have been considered:

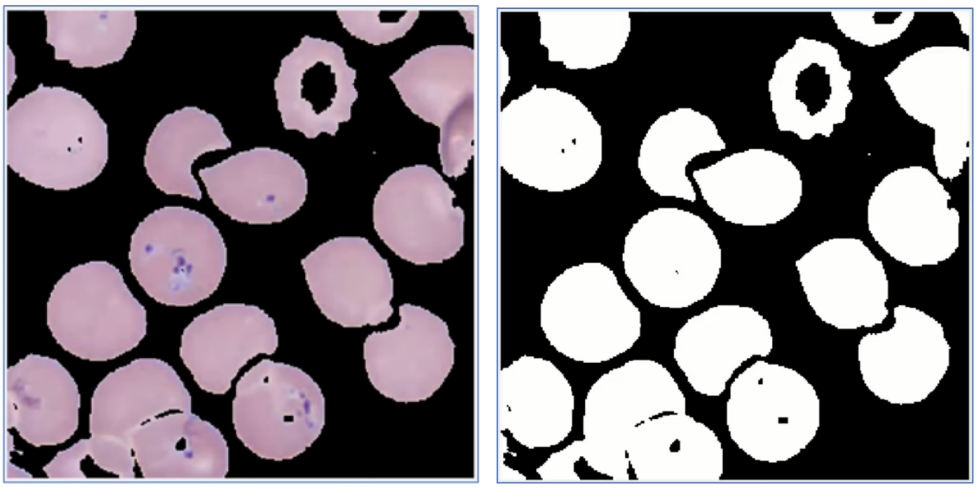
1. We try segmenting the individual red blood cells from the background using the color thresholder.
2. Since the blood cells are purple against a light background, we use the green intensity values of the image pixels to determine which pixels correspond to cells and which are background.
3. Next, we clean up the obtained cell mask and examine cell properties using the image region analyzer function, which allows quickly filling in region holes and filtering out regions based on their properties such as area dimensions or eccentricity.
4. Alternatively, since most of the cells in the blood smears are roughly circular, we instead pursue an algorithm based on finding circles using the image segmenter function. Then quickly gauge and specify the expected diameter range of the cells in pixels and indicate that the circles are dark relative to the background.
5. With the cells isolated, we can then identify parasites within the cells based on their darker pixel intensity in a grayscale version of the image.

The image segmental function helps to determine the appropriate threshold value. However, to determine a threshold value that is meaningful across all of our images, we will first need to account for the varying image quality by adjusting the histogram of a 2d image to match the histogram of a reference image.

5 Results of segmenting the thin blood smears

5.1 Detection of the presence of a parasite in thin blood smears

This stage can be expressed by two images (a) and (b) of Fig. 2.



(a) (b)

Figure 2: Using the green intensity values of the image pixels to determine which pixels correspond to cells and which are background: (a) adjusting green intensity values of the image pixels, (b) BW image

Next, we clean up the obtained cell mask and examine cell properties, filling in region holes and filtering out regions based on their properties. A sample of the produced image is shown in Fig. 3.

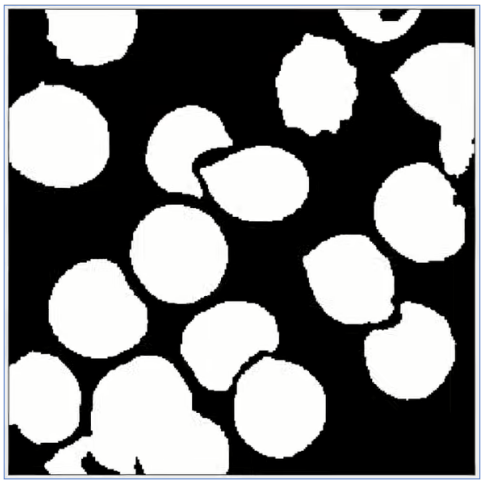
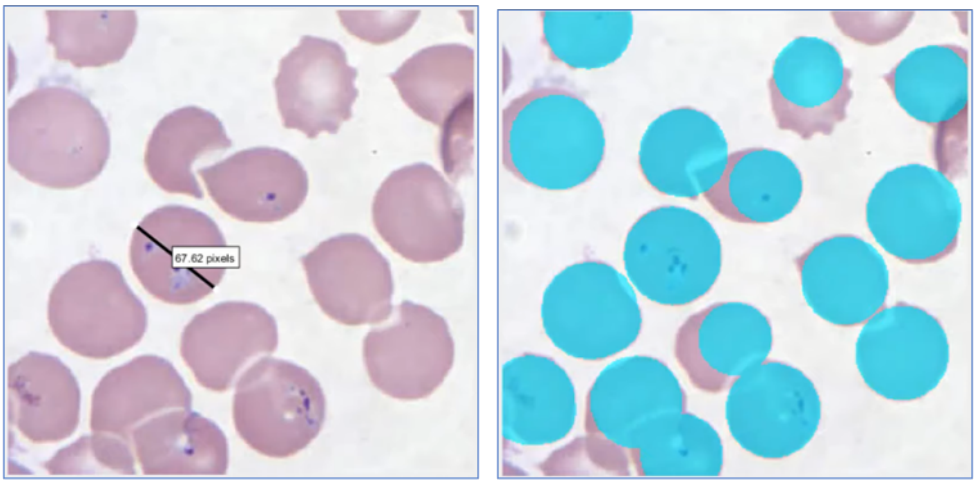


Figure 3: Filling in the region of holes and filtering out regions.

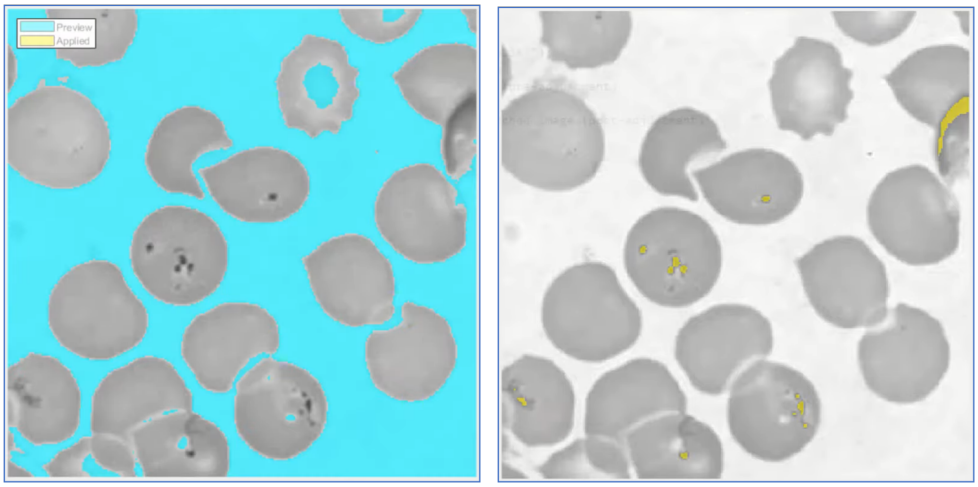
Since most of the cells in the blood smears are generally circular, we pursued an algorithm based on finding circles using an image segmenter function to indicate that the circles are dark relative to the background. This is demonstrated in Fig. 4.



(a) (b)

Figure 4: Produced image with:- (a) gauge and specify the expected diameter range, (b) the circles are dark relative to the background.

With the cells isolated we can then identify parasites within the cells based on their darker pixel intensity in a grayscale version of the image. The image segmenter function helps to determine the appropriate threshold value as shown in Fig. 5(a), while the image after thresholding, creating a mask, and inverting mask are shown in Fig. 5(b).



(a) (b)

Figure 5: Produced image based on: - (a) its darker pixel intensity in a grayscale image, (b) image after thresholding, creating a mask and inverting mask.

However, to determine a threshold value that is meaningful across all of the images, we first accounted for the varying image quality by adjusting the histogram of a 2D image to match the histogram of a reference image as shown in Fig. 6.

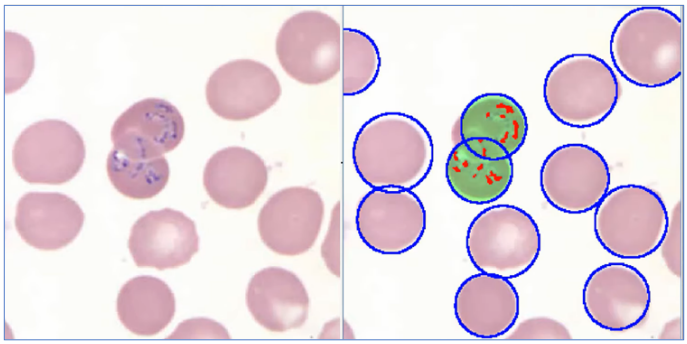


Figure 6: The original image vs. matched image

This figure indicates that this thin blood smear is infected by 17% with parasites. Furthermore, there are two diagnosed cells (with green color) from 12 detected cells (blue circles).

5.2 Quantifying the portion of red blood cells in the infected samples

In this stage, we've combined the code for finding circles matching image histograms and the parasite threshold detection logic into a single function. The results are shown in Fig. 7.

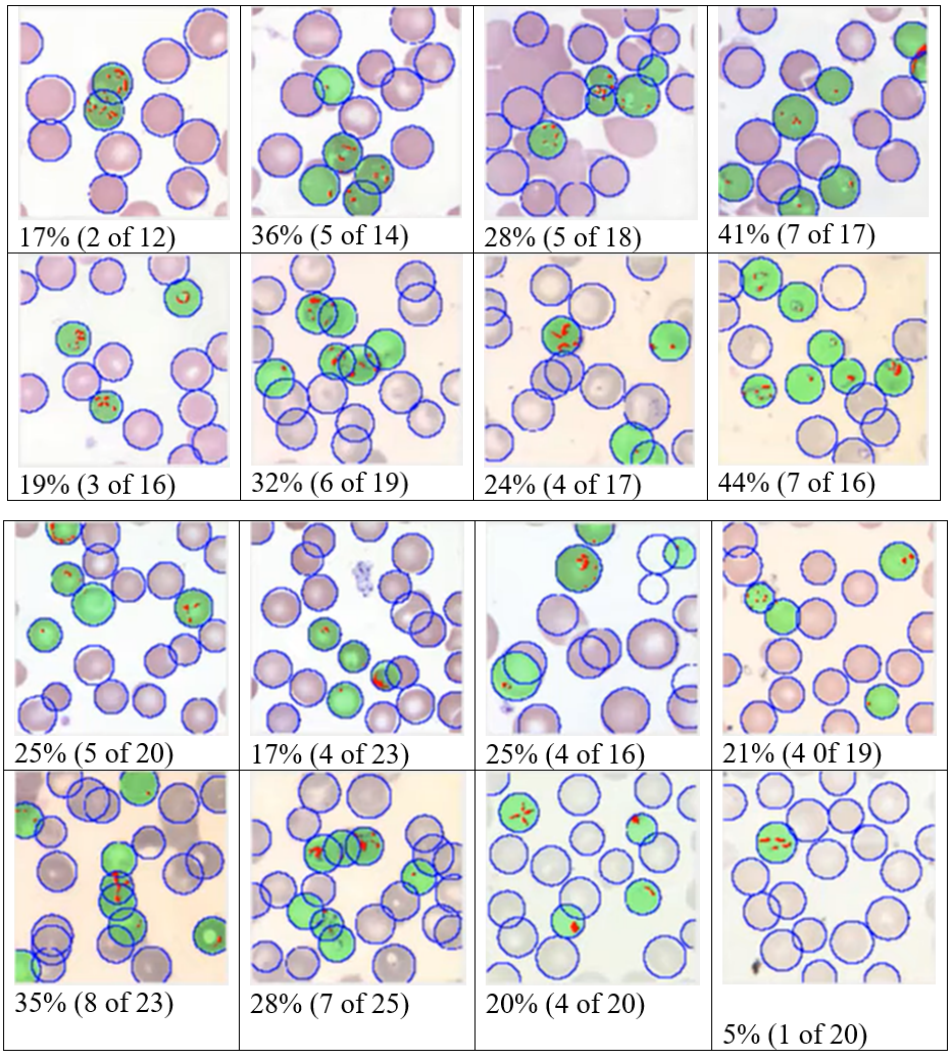


Figure 7: Percent of infected cells (of those detected).

This leads to quickly examining the performance of this function on the other images using the image batch processing technique.

6 Discussion of the results of the thin red blood smears

First, we try segmenting the individual red blood cells from the background using the color thresholder. Since the blood cells are purple against a light background, we adjusted the green intensity values of the image pixels to determine which pixels correspond to cells and which are the background. This was noticed in Fig. 4. Next, (see Fig. 5) we clean up the obtained cell mask and examine cell properties using MATLAB-based image region analyzer function, which allows quickly filling in region holes and filtering out regions based on their properties such as area dimensions or eccentricity. Since most of the cells in the blood smears are approximately circular, we instead pursued an algorithm based on finding circles using an image segmenter function. Then, quickly gauge and specify the expected diameter range of the cells in pixels and indicate that the circles are dark relative to the background as demonstrated in Fig. 6. With the cells isolated, it is possible to identify parasites within the cells based on their darker pixel intensity in a grayscale version of the image. The image segmenter function helps to determine the appropriate threshold value as shown in Fig. 7. By accounting for varied image quality and modifying the histogram of a 2D image to match the histogram of a reference image as illustrated in Fig. 8, a threshold value that is relevant across all of the photos can be determined.

We've merged the code for finding circles that match picture histograms with the parasite threshold detection logic into a single function at this point. Fig. 9 depicts the results.

This approach ultimately provides a statistic of how much the percentage of parasites have infected those thin blood smear samples as well as the number of diagnosed/infected cells (green in color) among the discovered cells (blue circles). As a result, utilizing the pictured batch processing technique, it is possible to easily check the performance of this function on different thin blood smears.

This approach is limited to investigating the thin red blood smears and can be expanded to cover the thick blood smears by using other thresholder and segmentation techniques.

7 Conclusions

1. The proposed approach was able to detect the presence of a parasite in thin blood smears by gauging the cells and specifying the expected diameter range of those cells in pixels, and indicating those circles are dark relative to the background. This statistic of a sample implies that parasites have infected 17% of this thin blood smear and there are two diagnosed cells among the 12 discovered cells.
2. This study also quantified the proportion of infected red blood cells in the sample by providing statistics of how much the percentage of parasites have infected those thin blood smear samples as well as the number of diagnosed/infected cells among the discovered cells.

References

[1] Kh. AL-Dulaimi, J. Banks, K. Nugyen, A. Al-Sabaawi, I. Tomeo-Reyes and V. Chandran, *Segmentation of white blood cell, nucleus and cytoplasm in digital haematology microscope images: A review–challenges, current and future potential techniques*, IEEE Rev. Biomed. Engin.**14** (2021), 290–306.

[2] J.W. Choi, Y. Ku, B.W. Yoo, J.Ah Kim, D.S. Lee, Y.J. Chai, H.-J. Kong, and H.Ch. Kim, *White blood cell differential count of maturation stages in bone marrow smear using dual-stage convolutional neural networks*, PLOS ONE **12** (2017), no. 12, e0189259.

[3] H. Fan, F. Zhang, L. Xi, Z. Li, G. Liu and Y. Xu, *LeukocyteMask: An automated localization and segmentation method for leukocyte in blood smear images using deep neural networks*, J. Biophoton. **12** (2019), no. 7.

[4] R.B. Hegde, K. Prasad, H. Hebbar and B.M. Kumar Singh, *Image Processing Approach for Detection of Leukocytes in Peripheral Blood Smears*, J. Med. Syst. **43** (2019), no. 5, 114.

[5] Z.R. Hussein, A.I. Mahameed and J.A. Fadhil, *Distinguishing of different tissue types using K-Means clustering of color segmentation*, Eastern-Eur. J. Enterprise Technol. **5** (2021), no. 2 (113), 22–28.

[6] K.K. Jha and H.S. Dutta, *Mutual information based hybrid model and deep learning for Acute Lymphocytic Leukemia detection in single cell blood smear images*, Comput. Meth. Prog. Biomed. **179** (2019), 104987.

[7] A. Kar, K. Pattnaik, T. Kar, P. Biswal, Ch. Mishra and L. Guru, *Clear cell lesions in pathology: Histomorphologic approach to diagnosis*, Indian J. Pathol. Microbio. **63** (2020), no. 2, 177.

[8] S.N. Mohd Safuan, M.R. Md Tomari and W.N. Wan Zakaria, *White blood cell (WBC) counting analysis in blood smear images using various color segmentation methods*, Measurement **116** (2018), 543–555.

[9] M. Poostchi, K. Silamut, R.J. Maude, S. Jaeger and G. Thoma, *Image analysis and machine learning for detecting malaria*, Translation. Res. **194** (2018), 36–55.

[10] F. Sadeghian, Z. Seman, A.R. Ramli, B.H. Abdul Kahar and M.-I. Saripan, *A framework for white blood cell segmentation in microscopic blood images using digital image processing*, Bio. Proced. Online **11** (2009), no. 1, 196.

[11] U. Salamah, R. Sarno, A.Z. Arifin, S. Sarimuddin, A.S. Nugroho, I.E. Rozi, and P.B. Setia Asih, *Segmentation of malaria parasite candidate from thickblood smear microscopic images using watershed and adaptive thresholding*, J. Telecommun. Electronic Comput. Engin. **10** (2018), no. 2-4, 113–117.