

Peer Reviewed Paper **openaccess**

Detection and segmentation of erythrocytes in multispectral label-free blood smear images for automatic cell counting

Solange Doumun,^{a,b,*} Sophie Dabo^a and Jérémie Zoueu^b^aUniversité Lille, CNRS, UMR 8524-Laboratoire Paul Painlevé, INRIA-MODAL, F-59000 Lille, France^bLaboratoire d'Instrumentation, Image et Spectroscopie, Institut National Polytechnique, Félix Houphouët-Boigny, BP 1013, Yamoussoukro, Côte d'Ivoire

Contacts

Solange Doumun: mekapeusolange.doumunepseoulai.etu@univ-lille.fr,
<https://orcid.org/0000-0001-7049-1499>Jérémie Zoueu: jzoueu@yahoo.fr,
<https://orcid.org/0000-0003-4588-1641>Sophie Dabo: sophie.dabo@univ-lille.fr,
<https://orcid.org/0000-0002-4000-6752>

In this work we propose an efficient approach to image segmentation for multispectral images of unstained blood films and automatic counting of erythrocytes. Our method takes advantage of Beer–Lambert's law by using, first, a statistical standardisation equation applied to transmittance images, followed by the local adaptive threshold to detect the blood cells and hysteresis contour closing to obtain the complete blood cell boundaries, and finally the watershed algorithm is used. With this method, image pre-processing is not required, which leads to time savings. We obtained the following results that show that our technique is effective, efficient and fast: Precision of 98.47% and Recall of 98.23%, a degree of precision (F-Measurement) of 98.34% and an Accuracy of 96.75%.

Keywords: multispectral imaging, segmentation, malaria, automatic diagnosis, image analysis

Introduction

Malaria continues to be a common and critical public health problem in developing countries. It is caused by *Plasmodium* parasites that are transmitted through the bite of mosquitoes.¹ In 2018, 228 million cases of malaria were recorded and 405,000 deaths from this disease, mostly children under five years old, and mostly in Sub-Saharan Africa.² Early and reliable detection is needed for quick action and is considered as an effective

way to fight this pandemic. Gold standard diagnostics are made by blood collection and visual inspection by clinic staff using simple, visible light-based microscopes. This manual process is time-consuming and unreliable. In this paper, we discuss detection and segmentation of red blood cells in multispectral and multimodal unstained blood smear images for automatic counting of erythrocytes in an effort to achieve computer-assisted malaria

Correspondence

S. Doumun (mekapeusolange.doumunepseoulai.etu@univ-lille.fr)**Received:** 2 June 2020**Revised:** 4 August 2020**Accepted:** 25 August 2020**Publication:** 9 September 2020**doi:** 10.1255/jsi.2020.a10**ISSN:** 2040-4565

Citation

S. Doumun, S. Dabo and J. Zoueu, "Detection and segmentation of erythrocytes in multispectral label-free blood smear images for automatic cell counting", *J. Spectral Imaging* 9, a10 (2020).
<https://doi.org/10.1255/jsi.2020.a10>

© 2020 The Authors

This licence permits you to use, share, copy and redistribute the paper in any medium or any format provided that a full citation to the original paper in this journal is given, the use is not for commercial purposes and the paper is not changed in any way.



diagnosis. Automatic malaria diagnosis based on image processing is today an active field of research due to the advent of computer-aided diagnosis.³ Generally, the analysis begins with image pre-processing, followed by segmentation and finally classification. According to the state-of-the-art, several works have proposed a number of methods for segmentation in malaria diagnosis. We highlight the following methods.

Otsu⁴⁻⁷ is a well-established technique based on a global thresholding.⁸ The principle is built on the calculation of an optimal threshold (which maximises the inter-class variance) from the image histogram. This approach is simple, fast and well adapted to bimodal histograms (background and foreground) such as blood smear images. However, it does not consider the spatial position of pixels in the image. Uneven illumination images are also missed in the processing, making it difficult to find the optimal threshold.

Watershed segmentation⁹ describes the image as a topographic landscape with ridges and valleys.¹⁰ Considering images as reliefs, the grey levels are identified as altitudes. The image is thus subdivided into watersheds, corresponding to the geographical space in which all the waters converge towards the same minimum. These basins are then delimited by ridge lines or watersheds. They make it possible to identify and extract continuous and homogeneous regions of the image (objects such as cells). Although sensitive to noise, it has shown a tendency towards oversegmentation because each minimum generates a watershed.

The K-means algorithm^{11,12} partitions the points in the data matrix $\mathbf{X}$ into K clusters. This partition is iterative and minimises the sum, over all clusters, of the within-cluster sums of point-to-cluster-centroid distances.¹² The algorithm is initialised by a randomly chosen set of centres. Afterwards, it iteratively searches the optimal partition. Each pixel is assigned to their nearest cluster centroid. After all pixels have been assigned, the new cluster centroids are recalculated by the average of each cluster. The algorithm stops when no further changes are made. This is the steady state. Its simplicity and its ability to process large datasets makes it effective. It works well with uneven illumination images. However, the result is highly dependent on the starting centres. Therefore, it is necessary to run the algorithm several times before reaching the optimal solution. A prior knowledge of the data is also essential. Applied to blood cells, the algorithm leads to a partial segmentation of the blood cells.

In fact, it does not consider the spatial dependencies of the pixels. Moreover, the boundaries of the partitions are different from the cell contours.

Hough's transform^{13,14} is a pattern recognition method that uses the spatial information of the characteristic points of the image. In its implementation, it allows to pass from the space of the acquired image to the space of the parameters that represent the geometrical shapes searched for in the image.¹⁵ Robust to noise, Hough's transform is also a robust curve detector, especially the circular Hough transform for circular shapes such as red blood cells. The drawback of this method is the difficulty in identifying and classifying unwanted contours, as well as discontinuities in contours. Changes in the slope (generally closed) around the area of interest towards an equilibrium position at the edges of the object must be detected under the action of several forces (internal energy, potential energy and external energy). This equilibrium position is the one that minimises the energy. This technique is called active contour¹⁶⁻¹⁸ and is suitable for several imaging applications. But the result depends on the initial setting, which leads to false contours if incorrectly fixed. It also encounters difficulties in detecting several objects in the same image as well as digital instability and does not perform well in segmenting highly overlapping cells.¹⁹

Fuzzy Logic Segmentation²⁰ implements the concept of degree of membership. It defines the membership function which lies at the heart of any fuzzy logic system. The capability of fuzzy rule-based techniques significantly depends upon it. The fuzzy logic is flexible enough to consider inaccuracies and uncertainties. Imitating human reasoning, it is easy to understand even by non-specialists. However, the accuracy of the segmentation is not very high and the number of classes of objects in the image must be known in advance.

Thus, these methods depend on the characteristics of the image, which are affected by a number of factors such as the acquisition system, the red blood staining type etc. As a result of the approaches described above, it is challenging to provide a universal method, especially for the type of images used in this work. Various methods have been proposed with detection rates higher than 98 % compared to manual counting.¹³ However, these techniques are very sensitive to noise and involve pre-processing steps which require the acquisition of additional images, making them time consuming. To counter this, we employ front-end

processing using image standardisation in transmission mode aiming to isolate the red blood cells and facilitate the segmentation.

Materials and methods

Multispectral and multimodal microscope

The image acquisition device consists of a modified Brunel optical microscope. Zoueu *et al.*^{12,21} introduced the concept of unstained blood smear malaria diagnosis using spectral imaging with an initial microscope design and in 2011, Brydegaard *et al.*²² built the model used in this work. The system is composed of three acquisition modes (transmission, reflection and scattering), using a set of 13 wavelengths ranging from ultraviolet to near infrared (375–940 nm). This device has been previously described by Brydegaard *et al.*²² It is a modified commercial metallurgical microscope (Brunel SP80). The ocular was substituted by a 5 Mpx (2592 × 1944) monochromatic CMOS camera (Guppy-503B, Allied Vision Technology, with a MT9P031 sensor from Micron/Aptina) equipped with individual pixel size of 2.2 × 2.2 μm, each having 12-bit depth.²³

The original filament light sources were substituted by a set multiplexing module of LEDs. The whole system was controlled by a PC running a custom-made LabVIEW code [National Instruments (NI)].

Image acquisition

The program controls camera exposure time and gain, LED multiplexing and current through the NI data acquisition board.²² Bright reference is used to calibrate the exposure times and gains and sample images are captured and saved in 12-bit unsigned integer images in TIFF format (Figure 1). The result is a low-contrast greyscale image. The blood cells cannot obviously be differentiated from the background contribution. The proposed method aims to process these acquired images for detection, segmentation and counting of blood cells.

All samples for all geometries (39 images) were acquired keeping the same region of interest.^{19–21} The spectral image is represented by a 2D matrix:

$$I_{xy} = \int_0^{+\infty} S_{sy\lambda} R_{sy\lambda} D_{sy\lambda} d\lambda \quad (1)$$

where **I** is a spectral image, **S** is the emission spectrum, **R** the sample response and **D** the detector sensitivity.



Figure 1. Original sample image acquired with LED light at 435 nm in transmission mode.

If we consider the spectral dimension, we have a 3D image $I_{sy\lambda}$ for each geometry.

Pre-processing

Initially, the pre-processing is performed using a customised algorithm in MATLAB (MathWorks). Once acquired, the images are saved and backgrounds are subtracted from the sample image, followed by image normalisation to ensure homogeneous illumination. Spectral images are obtained as below:

$$I_{\lambda} = \frac{I_{S_{\lambda}} - I_{B_{\lambda}}}{I_{R_{\lambda}} - I_{B_{\lambda}}} \quad (2)$$

I_{λ} is the greyscale image at wavelength λ , $I_{S_{\lambda}}$ is the sample image, $I_{R_{\lambda}}$ is the bright reference image and $I_{B_{\lambda}}$ is the dark reference image.

To obtain 13 spectral images for each mode, a total of 39 images (13 samples, 13 bright reference images and 13 dark reference images) are taken. The final spectral images of a blood smear sample consists of 39 greyscale images (13 per mode) requiring extra acquisition time for the additional images needed in the normalisation Equation (2). Finally, a 2D-median filter is applied to the normalised images in order to remove the noise.

Beer-Lambert's law in transmission

Beer-Lambert's law establishes a proportionality between the concentration of a chemical entity and the length path (thickness) of light in each media. According to this law the absorbance is:

$$A = \epsilon lc = -\log(T),$$

where l is the optical path length (thickness), c is concentration, ϵ is absorptivity and T is transmittance.

Transmittance is, therefore, inversely proportional to the thickness and the concentration:

$$T = e^{-\epsilon lc} \quad (3)$$

According to this law, in the transmission mode image, the pixel values of the red blood cells are lower than the pixel values of the background image (Figure 1) and, therefore, lower than the average pixels values.

Standardisation

Standardisation allows the reduction of data to have zero mean and a standard deviation of 1. The aim of this process is to overcome the sensitivity of the data to overly large dispersions, while maintaining the shape of joint distributions.

Let X_j be the grey level of the pixel in a given image, the standardised grey level is

$$\frac{x_j - \bar{x}}{\bar{\sigma}} \quad (4)$$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

is the sample grey level mean of all pixels of the image, while the sample standard deviation is:

$$\bar{\sigma} = \sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2}$$

Normalisation

With the same intent as standardisation, normalisation allows scaling data in $[0,1]$.

Let x_{min} and x_{max} be, respectively, the minimum and the maximum values of the pixels' grey level in a given image. A given grey level X_j is normalised as

$$\frac{x_j - x_{min}}{x_{max} - x_{min}}$$

Applying the standardisation equation on Figure 1 gives the result in Figure 2.

The image contrast is then enhanced and the blood cells can be clearly distinguished from the background. However, the background illumination is not homogeneous.

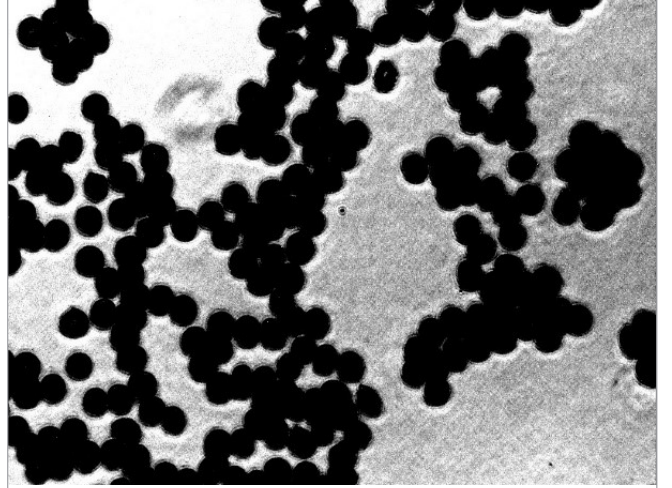


Figure 2. Standardised image of Figure 1.

In the following, the proposed automatic images analysis method is performed through a customised algorithm written in R software v. 3.5.3.

Adaptive thresholding

Unlike global thresholding, the adaptive one allows the threshold to be different depending on the region of the image.²⁴ Therefore, spatial dependencies of the background signal due to spurious signals or uneven illumination can be anticipated. The value of pixel intensity is determined from the local neighbourhood. We obtain a new smoothed image after applying an adaptive thresholding procedure to all pixels. The thresholding is performed by comparing each pixel's intensity with a smoothed background image. For the smoothing, we used the `filter2` function of the image processing package *EBImage*²⁴ of the R language. This function uses a 2D Fast Fourier Transform which is a convolution product equivalent to the following equation:

$$(w \times f)(x, y) = \sum_{s=-\infty}^{+\infty} \sum_{t=-\infty}^{+\infty} w(s, t) f(x + s, y + t)$$

where w is a weighting function whose value is highest in the middle of the window ($s = t = 0$) and decreases at the border.

Hysteresis contour closing

The adaptive thresholding algorithm allows the red blood cells contours to be detected. However,

some of these contours are not closed. Hysteresis contour closing is performed to obtain the complete blood cells' boundaries. It consists of detecting the extreme points of the unclosed contours and their extension.

Two different thresholds are used on **I**, which is a grey-scale image. A high threshold T1 and a low threshold T2 such that:

$$T1 > T2, \min(\mathbf{I}) \leq T1 \text{ and } T2 \leq \max(\mathbf{I})$$

Let **S1** and **S2** be two sets (binary mask images) such as:

$$\begin{aligned} \mathbf{S1} &= \{x \in \mathbb{R}^n / \mathbf{I}(x) \geq T1\}, \\ \mathbf{S2} &= \{x \in \mathbb{R}^n / \mathbf{I}(x) \geq T2\}. \end{aligned}$$

This amounts to finding **C** where $\mathbf{S1} \subset \mathbf{C} \subset \mathbf{S2}$. **C** is the binary image representing the contours, equal to 1 for pixels such as $|\nabla \mathbf{I}(x)| \geq T1$ and 0 elsewhere.²⁵ As far as we are concerned, the set **S1** is our standardised, normalised and segmented image with threshold 0, corresponding to the value for which Equation (4) is zero. The set **S2** is the result of a convolution of the image obtained in **S1** by the magnitude gradient of the Frei-Chen algorithm^{26,27} defined by:

$$\mathbf{M} = \sqrt{\mathbf{D}_x^2 + \mathbf{D}_y^2},$$

where

$$\mathbf{D}_x = \frac{1}{2 + \sqrt{2}} \begin{pmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{pmatrix}$$

and

$$\mathbf{D}_y = \frac{1}{2 + \sqrt{2}} \begin{pmatrix} -1 & \sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & -1 \end{pmatrix}$$

Performance assessment

The effectiveness of the proposed approach was measured using different parameters such as *TP*, True Positive, *FP*, False Positive, and *FN*, False Negative, to evaluate the performance indices which are: Precision, Recall, Accuracy and F-Measurement. In addition, we use a parameter *DR*, Detection Rate, which is the coverage rate of the manually counted cells. A comparison measure, namely, the positive agreement (P_{pos}), an alternative to the well-known kappa index, is also computed to compare the manual method to automatic methods.

These indices are calculated according to the following equations:

$$P_{pos} = \frac{2 \times P_{11}}{P_{.1} + P_{1.}}$$

$$DR = 100 \times \frac{MC - \text{CelDel}}{MC}$$

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

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

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F\text{-Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

where $TP = AC - FP$, $FP = CES + FD$, $FN = ME + \text{CelDel}$, $P_{.1} = P_{21} + P_{11}$, $P_{1.} = P_{11} + P_{12}$ and $P_{11} = \text{Min}(\text{Rate A}, \text{Rate B})$.

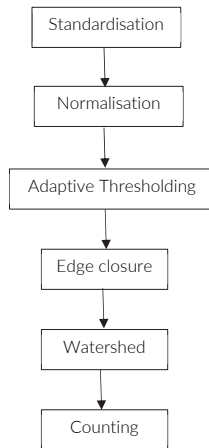
$$P_{12} = \begin{cases} \text{Rate B} - \text{Rate A} & \text{if Rate B} > \text{Rate A} \\ 0, & \text{otherwise} \end{cases}$$

$$P_{21} = \begin{cases} 0 & \text{if Rate B} > \text{Rate A} \\ \text{Rate B} - \text{Rate A}, & \text{otherwise} \end{cases}$$

$P_{22} = 1 - \text{Max}(\text{Rate A}, \text{Rate B})$, P_{pos} is the index of positive agreement between two cell counting methods (*Rate A* and *Rate B*), *CelDel* is the number of cells deleted by segmentation, *CES* is the counting error due to cell segmentation, *FD* is the out-of-cell detection, *FN* is the number of cells omitted from the automatic counting, *ME* is the merging error (counting error due to overlapping cells), *MC* is the number of cells manually counted, *AC* is the number of cells automatically counted and *CE* is the counting error:

$$EC = 100 \times \left| \frac{MC - AC}{MC} \right|$$

Results and discussions



Segmentation and counting process

The proposed segmentation approach used in this work can be divided into three basic steps, which are

summarised in Figure 3. Figure 3a is the original sample image from which the processing starts. It is a low contrast image where the blood cells are not clearly identified.

- The first step consists of standardisation (Figure 2) and normalisation.
- Step 2 realises an adaptive thresholding which aims to separate the erythrocytes from the background (Figure 3b).
- In step 3, the inhomogeneity of segmented cells is overcome by post-treatment which consists of contour closing by hysteresis and cell filling (Figure 3c). The watershed algorithm is thus performed to isolate the red blood cells. This leads to the counting process (Figure 3d).

Figure 3 reflects the whole process using LED light at a wavelength of 435 nm in transmission mode.

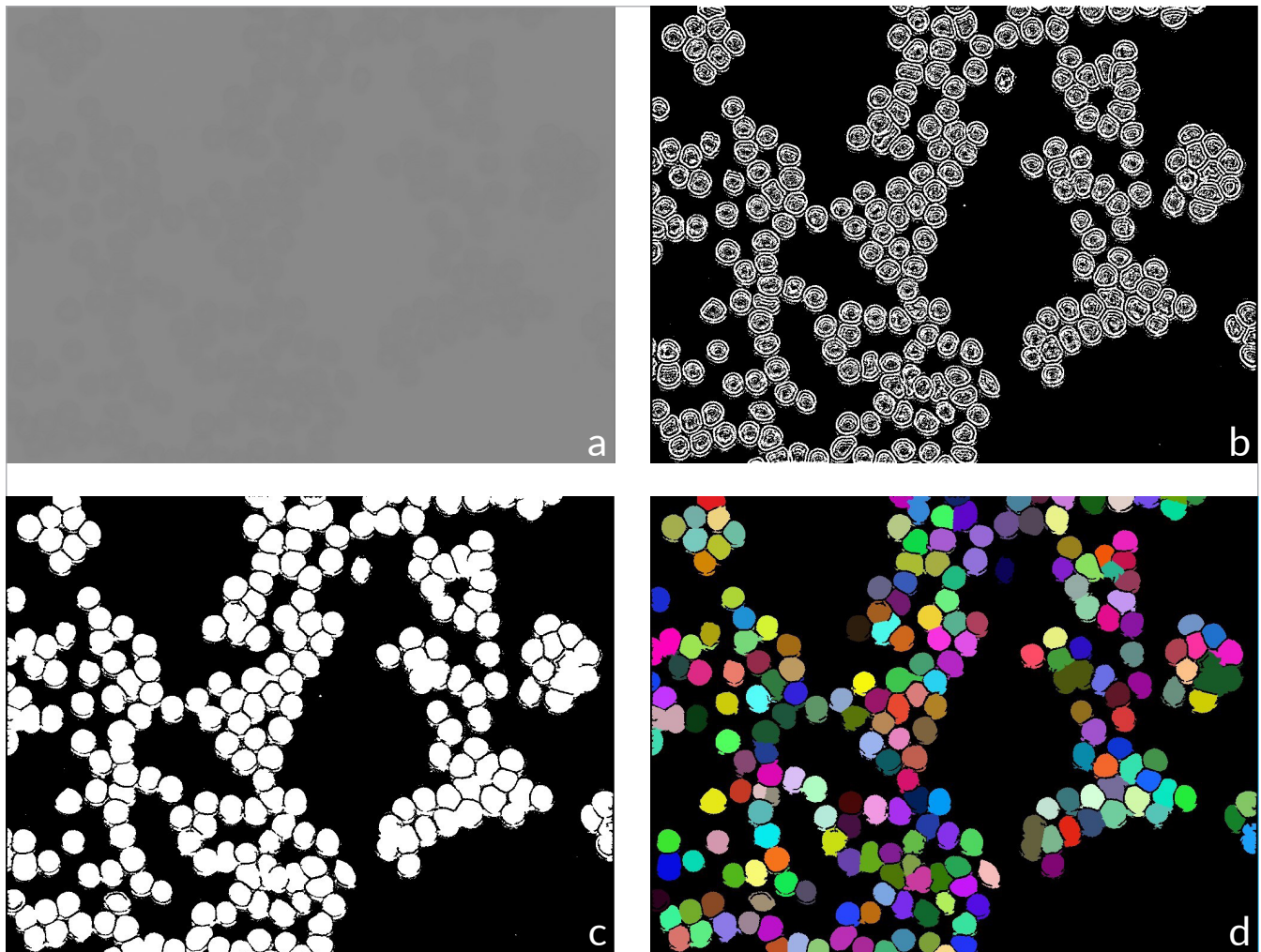


Figure 3. Segmentation process illustrated with a 435 nm image; (a) sample image, (b) adaptive thresholding, (c) post-processing and (d) counting.

Geometry selection

As indicated previously, the image used in this work is acquired thanks to three modalities and 13 wavelengths ranging from 370 nm to 940 nm. This section aims to explore both geometries and wavelengths, using the proposed methodology, to determine the best erythrocyte contrast over the background to achieve suitable segmentation. Figure 4 shows images acquired in transmission mode using wavelengths of 435 nm, 525 nm, 590 nm and 700 nm. The choice of these wavelengths reflects the main absorption peaks of heme and their potential for best contrast.

The original sample images were segmented using adaptive segmentation applied to Figure 2 followed by normalisation.

In each image (Figure 4), we observe a black background which corresponds to both glass slide and the

plasma contribution; the cells with higher pixel intensities correspond to the erythrocytes. The contrast is obtained from the variability of hemoglobin concentration, the absorbancies and the relative thicknesses of the plasma and erythrocytes. Contrast variation from one wavelength image to another is due to absorptivity.

In Figure 5, the mode is switched to reflection, while maintaining both region of interest (ROI) and wavelength at the same value. Figure 5a is the image with 435 nm, which represents the highest absorption peak for hemoglobin and gives the highest contrast of the blood smear images. For reflection, the contrast derives from the relative reflection coefficients between the background and the erythrocytes which are composed predominantly of hemoglobin.

By comparison, the segmentation with transmission images leads to higher contrast and reveals more

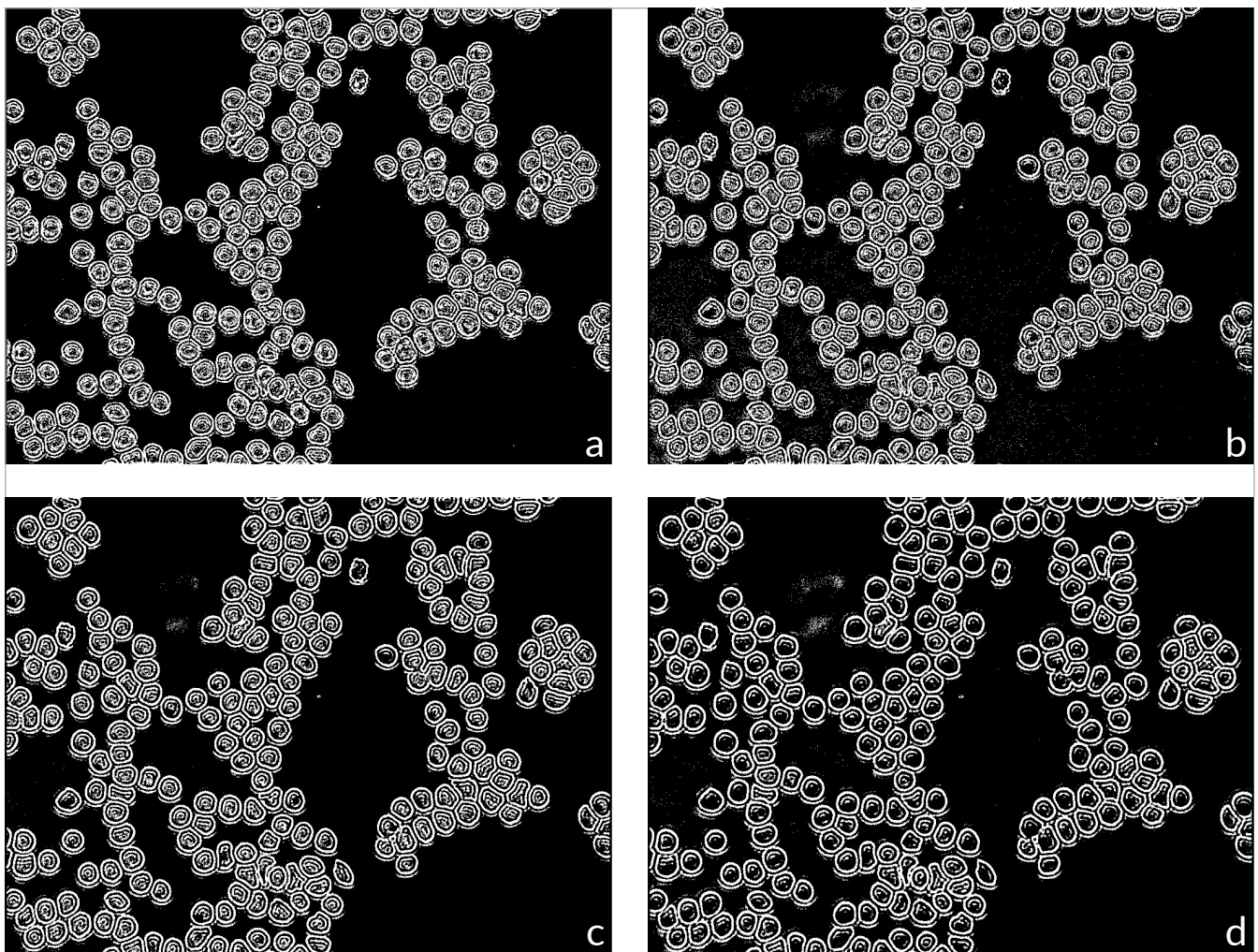


Figure 4. Adaptive thresholding in transmission according to wavelengths: (a) image at 435 nm, (b) image at 525 nm, (c) image at 590 nm and (d) image at 700 nm.

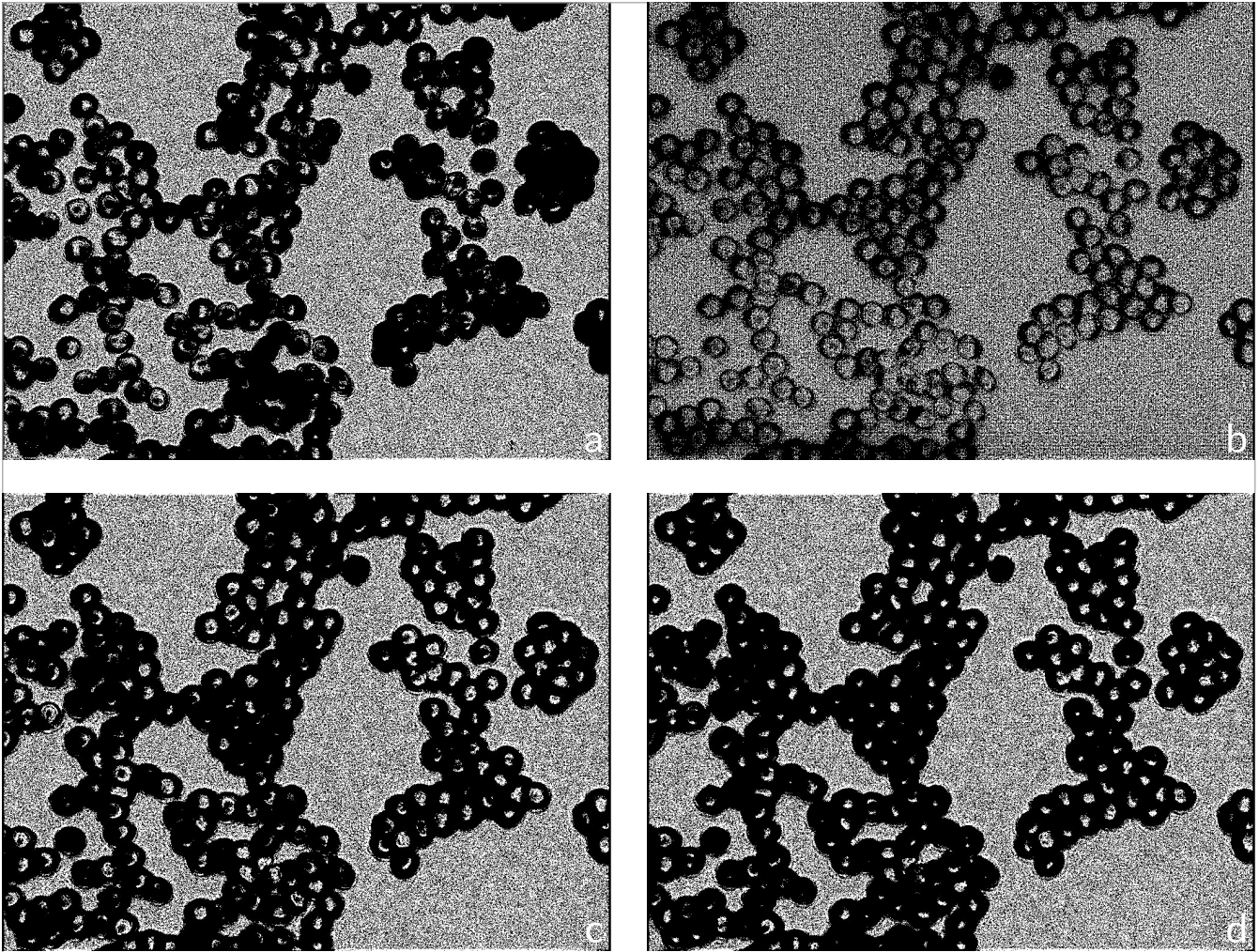


Figure 5. Adaptive thresholding in reflection according to wavelengths: (a) image at 435 nm, (b) image at 525 nm, (c) image at 590 nm and (d) image at 700 nm.

clearly erythrocyte edges, while resulting in simple post-processing by the cell filling algorithm, edge closing algorithm and watershed closing algorithm. The cells' isolation and counting processing is thus shortened. Transmission mode provides the best results. The system is then switched to the scattering mode for above-mentioned ROI and wavelengths (Figure 6).

Automatic counting results from the proposed segmentation approach applied to the 13 wavelengths in the scattering, reflection and transmission modes are presented in Table 1. The examined ROI contains 225 erythrocytes according to a manual count.

The automatic counting deviation from the manual one is estimated as 8.58%, 43.59% and 77.88% in transmission, reflection and scattering modes, respectively. These

findings lead to the selection of the transmission mode, which is compatible with the Beer-Lambert's law equation.

By comparing the counting accuracy of all wavelengths in transmission mode, the best result is obtained with 435 nm, which is associated with an automatic counting error of 1.98% and an accuracy of 98%. Furthermore, the images at 525 nm, 590 nm, 625 nm, 660 nm, 700 nm, 750 nm and 810 nm also introduce smaller automatic counting errors (less than 10%) and an accuracy higher than 90%, as shown in Table 2. Also, all these wavelengths detect 100% of the cells.

Results on non-pretreated images

Here we describe the results of our method on images acquired from the unstained blood smear without the

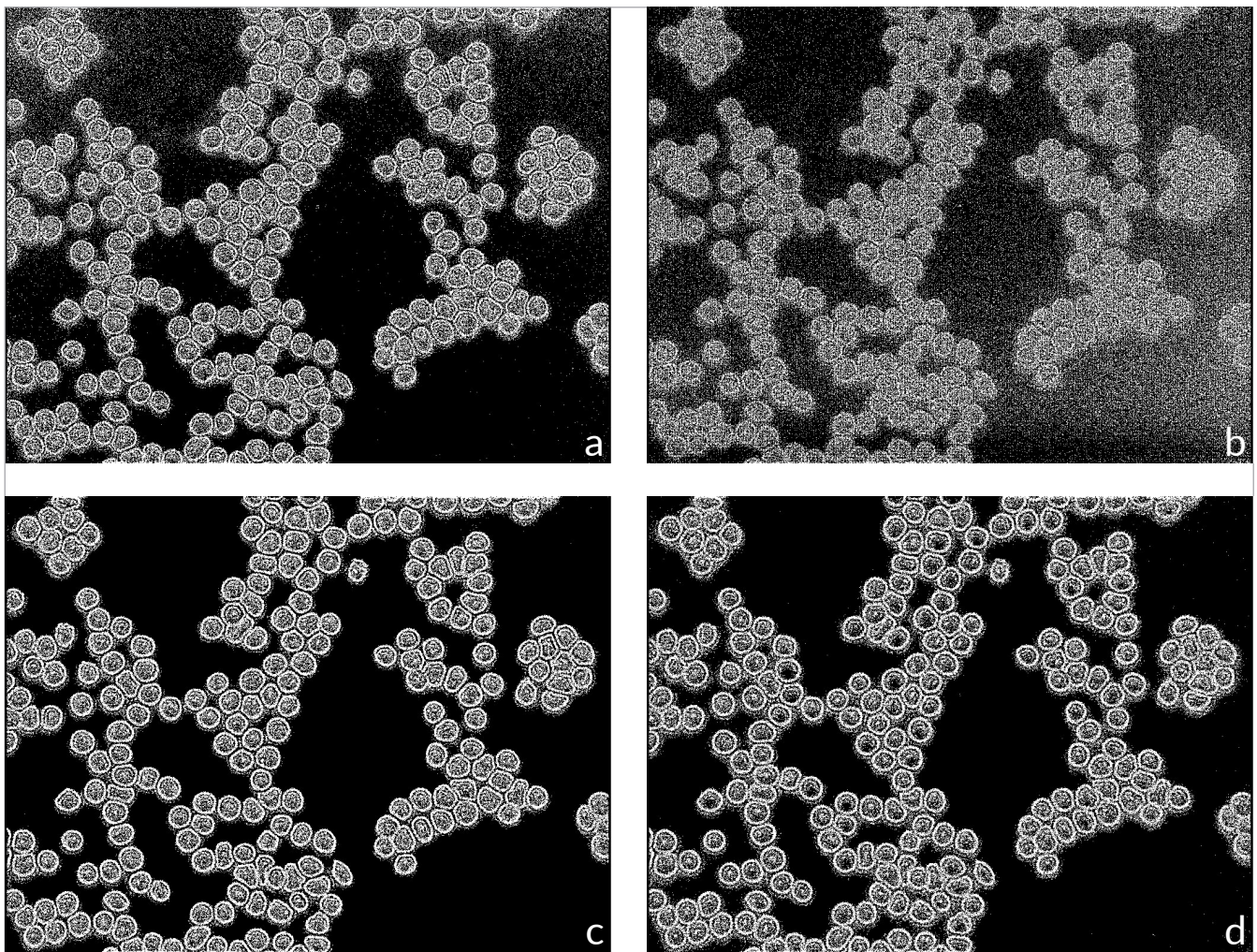


Figure 6. Adaptive thresholding in scattering according to wavelengths: (a) image at 435 nm, (b) image at 525 nm, (c) image at 590nm and (d) image at 700 nm.

normalisation pre-processing step which requires the acquisition of dark and bright reference images [Equation (2)]. As illustration is given with the 435nm wavelength image, see Figure 3a.

By applying the standardisation Equation 4, we get negative values for the red blood cell pixels, due to Beer-Lambert's law [Equation (3)]. This is consistent with the inverted standardised image as depicted in Figure 7.

All the cells are effectively identified. As stated above, the counting effectiveness depends on the cell isolation algorithm.

Adaptive thresholding is used to isolate red blood cells. However, depending on the wavelength, some features (contour details) can clearly be seen and others not. In this way, the blood cells uncorrectly segmented are removed.

All cells are detected by the combination of several wavelengths. We can obtain and improve this result by using a proposed hysteresis contour closing algorithm applied to a single image (one of these wavelengths) as shown in Figures 3c and 3d.

Result with pre-processed image

Applying the standardisation equation to the pre-processed image produces a different result. In fact, only the contours are shown in Figure 8a. This is due to the pre-processing algorithm used for the image smoothing. Because of the concave shape of the red blood cells, the effect of Beer-Lambert's law in transmission [Equation (3)] is accentuated on the contours which are thick and less affected by the smoothing

Table 1. Geometry comparison.

Wavelength (nm)	Scattering		Reflection		Transmission	
	AC	Error rate (%)	AC	Error rate (%)	AC	Error rate (%)
375	34	84.89	80	64.44	160	28.89
400	25	88.89	182	19.11	196	12.89
435	28	87.56	216	4.00	221	1.78
470	154	31.56	165	26.67	200	11.11
525	18	92	110	51.11	225	0
590	48	78.67	203	9.78	224	0.44
625	49	78.22	110	51.11	217	3.56
660	33	85.33	80	64.44	227	0.89
700	58	74.22	150	33.33	211	6.22
750	98	56.44	109	51.56	227	0.89
810	44	80.44	130	42.22	211	6.22
850	29	87.11	65	71.11	190	15.56
940	29	87.11	50	77.78	173	23.11
Average	49.8	77.88	126.92	43.59	206.31	8.58

algorithm. However, the cells are segmented by closing the contours, as shown in Figure 8.

Discussion

In the summary Table 3, we compare the performance of the proposed automatic counting method with the manual one on 30 raw images. The proposed algorithm is applied to images with two different sizes. A first batch of

7 samples of size 1020×768 and a second batch of 23 images of size 2048×1088 .

The thresholding algorithm reveals erythrocytes that can be overlapped or attached and cannot be isolated spatially by the isolation algorithm, which leads to counting error. However, the result can be improved by considering the magnification parameter of the acquisition system in order to define the relative size of the erythrocyte.

As shown in Table 3, the acquisition time is one-third of that of the method previously used in our laboratory.

Table 2. Wavelength comparison in transmission mode.

AC	Error rate(%)	Accuracy (%)
160	28.89	82.47
196	12.89	90.74
221	1.98	98.22
200	11.11	90.91
225	0.00	90.68
224	0.44	96.89
217	3.56	94.71
227	0.89	81.60
211	6.22	94.95
227	0.89	93.91
211	6.22	97.22
190	15.56	94.97
173	23.11	89.58
206.31	8.58	92.07

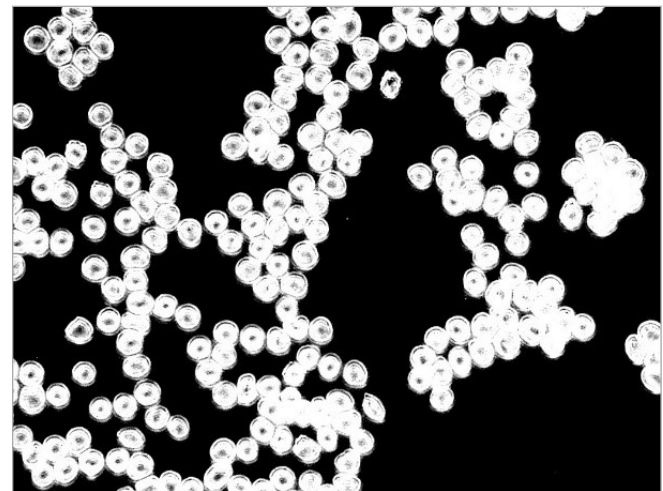


Figure 7. Standardised and inverted image of Figure 2.

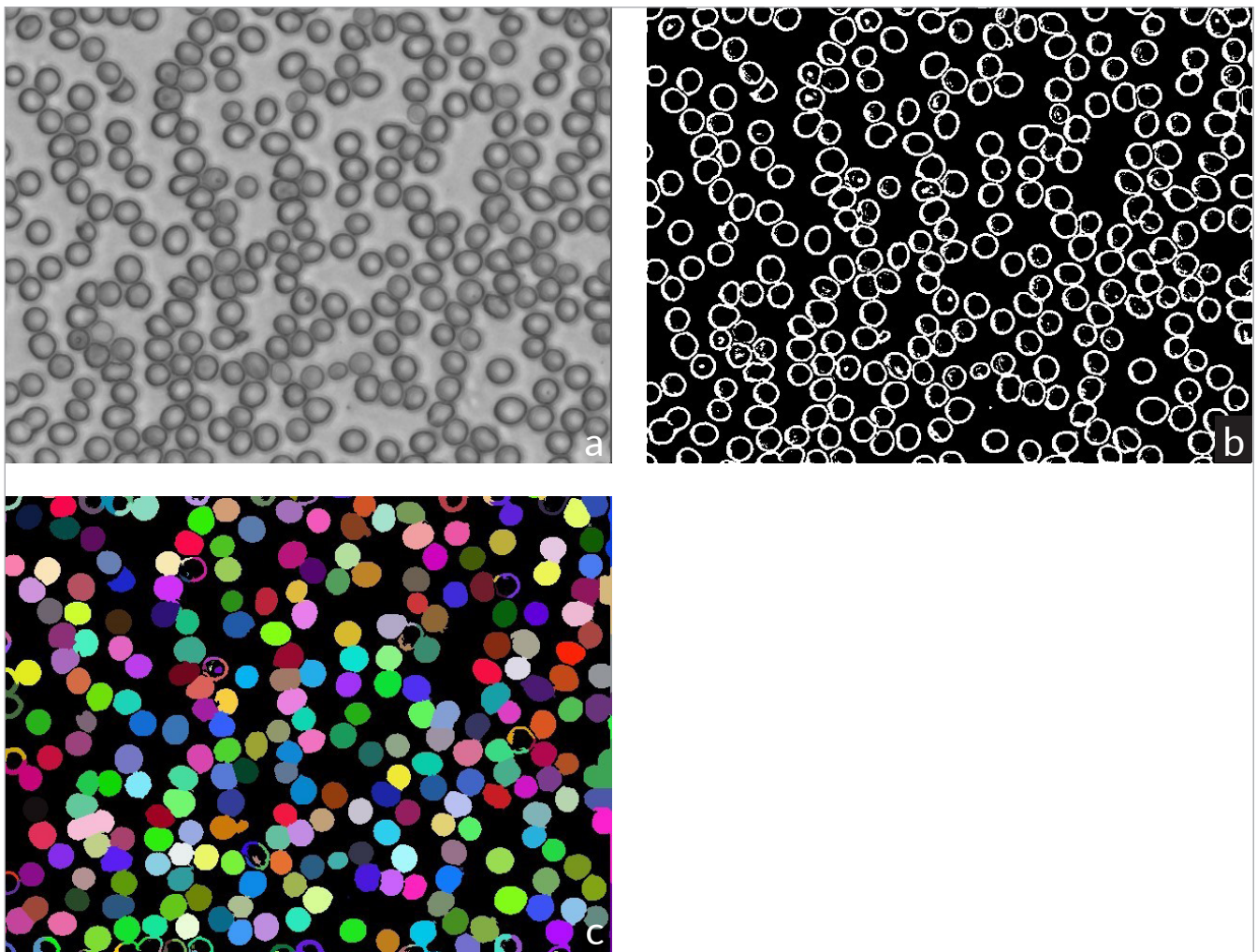


Figure 8. Pre-processed image segmentation; (a) image at 435 nm after normalisation and filtering, (b) image after adaptive thresholding and (c) counting.

Second, we capitalise on the execution time of the normalisation and pre-processing algorithms that are usually performed.^{11,19,21,26–28} Finally, the speed average is 30s for 1020×768 size images and less than 2 min for 2048×1088 size images. This meets the demands of real-time processing.

The numerical results show also the detection rates of 100% for the first batch and 99.65 % for the second, with an average of 99.82%, indicating the capacity that can be obtained if the isolation and counting post-processing is improved.

Figure 9 gives a determination coefficient R^2 of 0.9998 (when performing a simple linear regression) which shows the closeness between the automatic and manual counting methods. We also obtained an error rate of 1.6 % (Table 3) and root-mean-square error (RMSE) of

5.86 (Table 4). These results show a clear improvement over those of Bagui *et al.*¹³

Our method was compared with those of five other studies which reported on the segmentation and counting of red blood cells.³¹

Analysis of Table 5 gives us an accuracy of 96.8 %, which is not as good as Loddo.³² This is due to inadequacy of our cell isolation algorithm, which could be improved by considering the magnification features of the acquisition system in order to parameterise the relative size of the erythrocyte. However, we have a better result for precision, 98.5 %, and Recall, 98.2 %. We obtained a degree of precision (F-Measurement) of 98.3 %, which is a key indicator of performance for a diagnostic test. The positive agreement index of 0.99 shows the high agreement^{33–36} level of the

Table 3. Counting synthesis table.

Image no.	MC	AC	NOC	CES	FD	ET (s)	DR (%)	CE (%)	Precision	Recall	Accuracy	F-measure	P _{pos}
1	225	221	5(12)	1	0	26.34	100	1.77	99.55	96.92	96.49	98.21	0.99
2	150	156	7(14)	1	0	22.06	100	4.00	99.36	95.54	95.09	97.48	1.00
3	149	150	5(10)	2	3	33.59	100	0.67	96.67	96.67	93.55	96.67	0.99
4	204	200	5(10)	1	0	26.38	100	1.96	99.50	97.55	97.07	98.51	0.99
5	262	257	5(10)	0	0	15.11	100	1.90	100	98.09	98.09	99.04	0.99
6	119	120	3(6)	2	0	44.78	100	0.84	98.33	97.52	95.93	97.93	1.00
7	119	121	3(6)	3	0	48.34	100	1.68	97.52	97.52	95.16	97.52	1.00
Partial average 1						30.94	100	1.89	98.70	97.12	95.91	97.91	0.99
8	1045	1017	20(44)	0	0	234.27	99.61	2.68	100	97.31	97.31	98.64	0.99
9	139	139	1(4)	4	0	101.76	99.28	0.00	97.12	97.12	94.41	97.12	0.99
10	155	160	0	3	0	136	99.37	3.23	98.13	99.38	97.53	98.75	1.00
11	178	177	0	3	0	124.18	100	0.56	98.31	100	98.31	99.15	0.99
12	182	182	1(2)	2	0	102.92	100	0.00	98.90	99.45	98.36	99.17	0.99
13	171	171	0	0	0	76.7	100	0.00	100	100	100	100	1.00
14	183	181	0	4	0	86.09	100	1.09	97.79	100	97.79	98.88	0.98
15	185	189	0	4	0	68.93	100	2.16	97.88	100	97.88	98.93	1.00
16	181	183	0	4	0	93.22	98.87	1.10	97.81	98.87	96.73	98.34	0.99
17	172	175	0	4	0	90.41	99.41	1.74	97.71	99.41	97.15	98.55	1.00
18	164	163	0	3	0	107.81	99.39	0.61	98.16	99.38	97.56	98.77	0.99
19	153	156	0	5	0	114.72	100	1.96	96.79	100	96.79	98.37	0.99
20	175	177	0	3	0	133.12	98.28	1.14	98.31	98.30	96.66	98.30	1.00
21	180	185	0	5	0	114.62	100	2.78	97.30	100	97.30	98.63	1.00
22	182	183	0	2	0	108.53	99.45	0.55	98.91	99.45	98.37	99.18	1.00
23	156	159	0	3	0	94.83	100	1.92	98.11	100	98.11	99.05	1.00
24	165	165	0	0	0	89.73	100	0.00	100.00	100	100	100	1.00
25	159	162	0	3	0	84.06	100	1.89	98.15	100	98.15	99.07	1.00
26	160	161	1(2)	3	0	87	99.37	0.63	98.14	98.75	96.93	98.44	0.99
27	164	166	0	3	0	83	100	1.22	98.19	100	98.19	99.09	1.00
28	175	177	1(3)	5	0	105.55	99.42	1.14	97.18	98.28	95.55	97.72	0.99
29	174	177	1(2)	3	0	90.94	100	1.72	98.31	99.43	97.75	98.86	1.00
30	181	185	0	3	0	95.89	99.42	2.21	98.38	99.43	97.82	98.90	1.00
Partial average 2						105.40	99.65	1.32	98.24	99.33	97.59	98.78	1.00
Total average							99.82	1.60	98.47	98.23	96.75	98.34	0.99

MC: number of cells manually counted; AC: number of cells automatically counted; NOC: Number of overlapping cells; CES: Counting error due to cell segmentation; FP: False positive; the number of cells which should not be included in the counting; ET (s): Execution Time in seconds; DR: Detection rate; the manual counting cells coverage rate;

CE: Counting error; $CE = 100 \times \frac{|MC - CA|}{MC}$;

Partial average 1: 1020×768 size images average; Partial average 2: 2048×1088 size images average; Total average: average of all 30 images; RMSE: Root Mean Square Error; MAE: Mean Absolute Error; MAPE: Mean Absolute Percentage Error; R^2 adjust: determination coefficient

proposed method with manual counting compared to others.

Conclusion and perspectives

In this study, a method was introduced for detection and counting of RBCs from spectral images without using

reference images. By using this approach, the time saved can be critical in an emergency situation and when the peak demand of the test is highest: each extra batch of 13 images will add 2.6 min of acquisition time and about the same for pretreatment. Moreover, the novel algorithm runs in less than 2 min for large images (2048 × 1088 px).

Evaluation of the performance of the proposed approach results in a Precision of 98.47% and Recall of 98.23%, which gives a degree of precision (F-Measurement) of 98.34% and Accuracy of 96.75%. These results show that the proposed method is efficient and can lead to real-time diagnosis.

In perspective, we plan to improve the counting rate by merging over-segmented cells and disconnecting the ones that are overlapping.

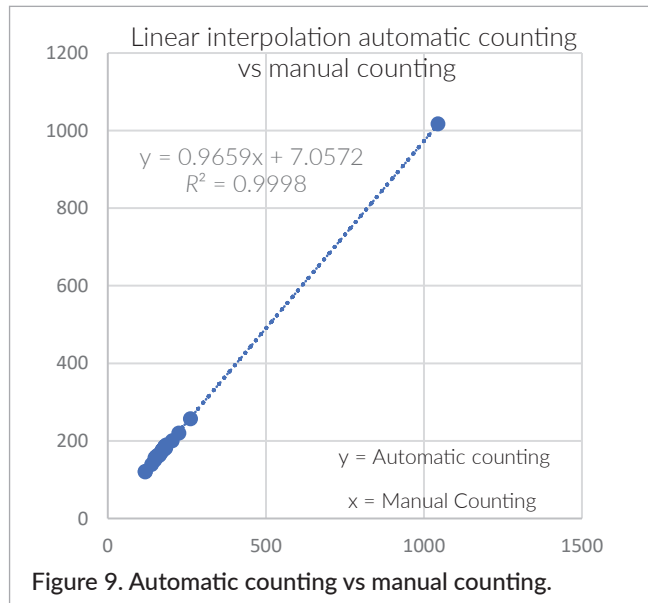


Table 4. Performance indicator.

Performance indicator	Value
RMSE	5.86
MAE	3.03
MAPE %	1.30
R^2 adjust	0.9998

Table 5. Blood cells counting performance compared with the state-of-the-art.

	Alomari <i>et al.</i> ³⁷	Loddo <i>et al.</i> ³²	Moallem <i>et al.</i> ¹⁸	Poostchi <i>et al.</i> ¹⁹	Cecilia <i>et al.</i> ³¹	Our approach
Images type	Stained thin blood smears RGB colour images from Dhruv Pathology Lab	Stained thin blood smears JPG 24-bit colour depth (2592 × 1944) images from ALL-IDB dataset (Canon Powershot G5 camera)	Stained thin blood smears RGB colour images from regular light microscope by smartphone	Stained thin blood smears RGB colour images acquired by a Zeiss Axio imager	Stained thin blood smears JPG 24-bit colour depth (2592 × 1944) images from ALL-IDB and MP-IDB dataset (Canon Powershot G5 camera)	Unstained blood smears greyscale 12-bit depth images acquired by a 5Mpx (2592 × 1944) monochromatic CMOS camera
Methods used	Iterative structured circle detection algorithm	Nearest neighbour and support vector machine	Adaptive meanshift clustering and snake algorithm	Laplacien of Gaussian filter and Geodesic active contour	Modified edge boxes	Standardisation, adaptive thresholding and the watershed algorithm
Accuracy	95.3	98		92.5	95.6	96.8
Precision	95	89	98	99.3	98.4	98.5
Recall	98	98	93	90.3	95	98.2
F-Measure	96	93	95	94.5	96.6	98.3
P_{pos}	98.99	98.99	96.37	94.90	97.44	99.09

References

1. Paludisme. <https://www.who.int/fr/news-room/fact-sheets/detail/malaria> [Accessed: 5 November 2019].
2. WHO | World Malaria Report 2019. WHO. <http://www.who.int/malaria/publications/world-malaria-report-2019/en/> [Accessed: 30 July 2020].
3. M. Poostchi, K. Silamut, R.J. Maude, S. Jaeger and G. Thoma, "Image analysis and machine learning for detecting malaria", *Transl. Res.* **194**, 36 (2018). <https://doi.org/10.1016/j.trsl.2017.12.004>
4. S.S. Savkare and S.P. Narote, "Automated system for malaria parasite identification", in 2015 *International Conference on Communication, Information Computing Technology (ICCICT)*, pp. 1–4 (2015). <https://doi.org/10.1109/ICCICT.2015.7045660>
5. S.S. Devi, S.A. Sheikh, A. Talukdar and R.H. Laskar, "Malaria infected erythrocyte classification based on the histogram features using microscopic images of thin blood smear", *Indian J. Sci. Technol.* **9(45)**, 1–10 (2016). <https://doi.org/10.17485/ijst/2016/v9i45/94119>
6. S.S. Devi, R.H. Laskar and S.A. Sheikh, "Hybrid classifier based life cycle stages analysis for malaria-infected erythrocyte using thin blood smear images", *Neural Comput. Appl.* **29(8)**, 217 (2018). <https://doi.org/10.1007/s00521-017-2937-4>
7. G.P. Gopakumar, M. Swetha, G. Sai Siva and G.R.K. Sai Subrahmanyam, "Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner", *J. Biophotonics* **11(3)**, e201700003 (2018). <https://doi.org/10.1002/jbio.201700003>
8. N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Trans. Syst. Man Cybern.* **9(1)**, 62–66 (1979). <https://doi.org/10.1109/TSMC.1979.4310076>
9. J.-D. Kim, K.-M. Nam, C.-Y. Park, Y.-S. Kim and H.-J. Song, "Automatic detection of malaria parasite in blood images using two parameters", *Technol. Health Care Off. J. Eur. Soc. Eng. Med.* **24 Suppl 1**, S33 (2015). <https://doi.org/10.3233/THC-151049>
10. S. Derivaux, S. Lefevre, C. Wemmert and J. Korczak, "Segmentation par ligne de partage des eaux basée sur des connaissances texturales". <https://icube-publis.unistra.fr/5-DLWK07> [Accessed: 6 February 2020].
11. A. Nanoti, S. Jain, C. Gupta and G. Vyas, "Detection of malaria parasite species and life cycle stages using microscopic images of thin blood smear", in 2016 *International Conference on Inventive Computation Technologies (ICICT)* **1**, 1–6 (2016). <https://doi.org/10.1109/INVENTIVE.2016.7823258>
12. J.T. Zoueu, S. Ouattara, A. Toure, S. Safi and S.T. Zan, "Spectroscopic approach of multispectral imaging of *Plasmodium falciparum*-infected human erythrocytes", in 2009 *3rd ICTON Mediterranean Winter Conference (ICTON-MW)*, pp. 1–7 (2009). <https://doi.org/10.1109/ICTONMW.2009.5385598>
13. O.K. Bagui and J.T. Zoueu, "Red blood cells counting by circular Hough transform using multispectral images", *J. Appl. Sci.* **14**, 3591–3594 (2014). <https://doi.org/10.3923/jas.2014.3591.3594>
14. Z. Zhang, L.L.S. Ong, K. Fang, A. Matthew, J. Dauwels, M. Dao and H. Asada, "Image classification of unlabeled malaria parasites in red blood cells", in 2016 *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3981–3984 (2016). <https://doi.org/10.1109/EMBC.2016.7591599>
15. H. Maitre, *Survey on Hough Transform* (1985). <http://hdl.handle.net/2042/2334>
16. D. Bibin, M.S. Nair and P. Punitha, "Malaria parasite detection from peripheral blood smear images using deep belief networks", *IEEE Access* **5**, 9099–9108 (2017). <https://doi.org/10.1109/ACCESS.2017.2705642>
17. D.K. Das, A.K. Maiti and C. Chakraborty, "Automated system for characterization and classification of malaria-infected stages using light microscopic images of thin blood smears", *J. Microsc.* **257(3)**, 238 (2015). <https://doi.org/10.1111/jmi.12206>
18. G. Moallem, H. Sari-Sarraf, M. Poostchi, R.J. Maude, K. Silamut, M.A. Hossain, S. Antani, S. Jaeger and G. Thoma, "Detecting and segmenting overlapping red blood cells in microscopic images of thin blood smears", in *Medical Imaging 2018: Digital Pathology* **10581**, 105811F (2018). <https://doi.org/10.1117/12.2293762>
19. M. Poostchi, I. Ersoy, K. McMenamin, E. Gordon, N. Palaniappan, S. Pierce, R.J. Maude, A. Bansal, P. Srinivasan, L. Miller, K. Palaniappan, G. Thoma and S. Jaeger, "Malaria parasite detection and cell counting for human and mouse using thin blood smear

- microscopy", *J. Med. Imaging* **5(4)**, 044506 (2018). <https://doi.org/10.1117/1.JMI.5.4.044506>
20. J.T. Zoueu and S.T. Zan, "Trophozoite stage infected erythrocyte contents analysis by use of spectral imaging led microscope", *J. Microsc.* **245(1)**, 90 (2012). <https://doi.org/10.1111/j.1365-2818.2011.03548.x>
 21. S. Dabo-Niang and J.T. Zoueu, "Combining kriging, multispectral and multimodal microscopy to resolve malaria-infected erythrocyte contents", *J. Microsc.* **247(3)**, 240 (2012). <https://doi.org/10.1111/j.1365-2818.2012.03637.x>
 22. M. Brydegaard, A. Merdasa, H. Jayaweera, J. Ålebring and S. Svanberg, "Versatile multispectral microscope based on light emitting diodes", *Rev. Sci. Instrum.* **82(12)**, 123106 (2011). <https://doi.org/10.1063/1.3660810>
 23. A.J. Merdasa, M. Brydegaard, S. Svanberg and J.T. Zoueu, "Staining-free malaria diagnostics by multispectral and multimodality light-emitting-diode microscopy", *J. Biomed. Opt.* **18(3)**, 036002 (2013). <https://doi.org/10.1117/1.JBO.18.3.036002>
 24. G. Pau, F. Fuchs, O. Sklyar, M. Boutros and W. Huber, "EBImage—an R package for image processing with applications to cellular phenotypes", *Bioinformatics* **26(7)**, 979 (2010). <https://doi.org/10.1093/bioinformatics/btq046>
 25. M. Bergounioux, *Introduction au Traitement Mathématique des Images - Méthodes Déterministes*. Springer (2015). <https://doi.org/10.1007/978-3-662-46539-4>
 26. D. Apdilah, M.Y. Simargolang and R. Rahim, "A study of Frei-Chen approach for edge detection", *Int. J. Sci. Res. Sci. Eng. Technol. IJSRSET* **3(1)**, 59 (2017). <https://doi.org/10.5281/zenodo.239730>
 27. M. Al-khassaweneh, "Robust and invisible watermarking technique based on Frei-Chen bases", in *2019 IEEE International Conference on Electro Information Technology (EIT)*, p. 1 (2019). <https://doi.org/10.1109/EIT.2019.8833875>
 28. J.T. Zoueu and S. Zan, "Trophozoite stage infected erythrocyte contents analysis by use of spectral imaging led microscope", *J. Microsc.* **245(1)**, 90 (2012). <https://doi.org/10.1111/j.1365-2818.2011.03548.x>
 29. O.K. Bagui and J.T. Zoueu, "Automatic malaria diagnosis by the use of multispectral contrast imaging", *PCN J.* **75**, 86–98 (2015). https://www.pcnjournal.com/157513_2053.htm
 30. O.K. Bagui, W. Yavo, D. Tano and J.T. Zoueu, "Etude de l'effet de l'amodiaquine sur les globules rouges infectés par le paludisme dans les images multispectrales", *Africa Sci.* **10(4)**, (2014). <http://www.afriques-science.info/document.php?id=3968>
 31. C. Di Ruberto, A. Loddo and L. Putzu, "Detection of red and white blood cells from microscopic blood images using a region proposal approach", *Comput. Biol. Med.* **116**, 103530 (2020). <https://doi.org/10.1016/j.compbiomed.2019.103530>
 32. A. Loddo, L. Putzu, C.D. Ruberto and 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)*, p. 112 (2016). <https://doi.org/10.1109/SITIS.2016.26>
 33. M.L. McHugh, "Interrater reliability: the kappa statistic", *Biochem. Medica* **22(3)**, 276 (2012). <https://doi.org/10.11613/BM.2012.031>
 34. J.R. Landis and G.G. Koch, "The measurement of observer agreement for categorical data", *Biometrics* **33(1)**, 159 (1977). <https://doi.org/10.2307/2529310>
 35. A.R. Feinstein and D.V. Cicchetti, "High agreement but low kappa: I. the problems of two paradoxes", *J. Clin. Epidemiol.* **43(6)**, 543 (1990). [https://doi.org/10.1016/0895-4356\(90\)90158-L](https://doi.org/10.1016/0895-4356(90)90158-L)
 36. D.V. Cicchetti and A.R. Feinstein, "High agreement but low kappa: II. resolving the paradoxes", *J. Clin. Epidemiol.* **43(6)**, 551 (1990). [https://doi.org/10.1016/0895-4356\(90\)90159-M](https://doi.org/10.1016/0895-4356(90)90159-M)
 37. Y.M. Alomari, S.N.H. Sheikh Abdullah, R. Zaharatul Azma and K. Omar, "Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection algorithm", *Computational and Mathematical Methods in Medicine* **2014**, 979302 (2014). <https://doi.org/10.1155/2014/979302>